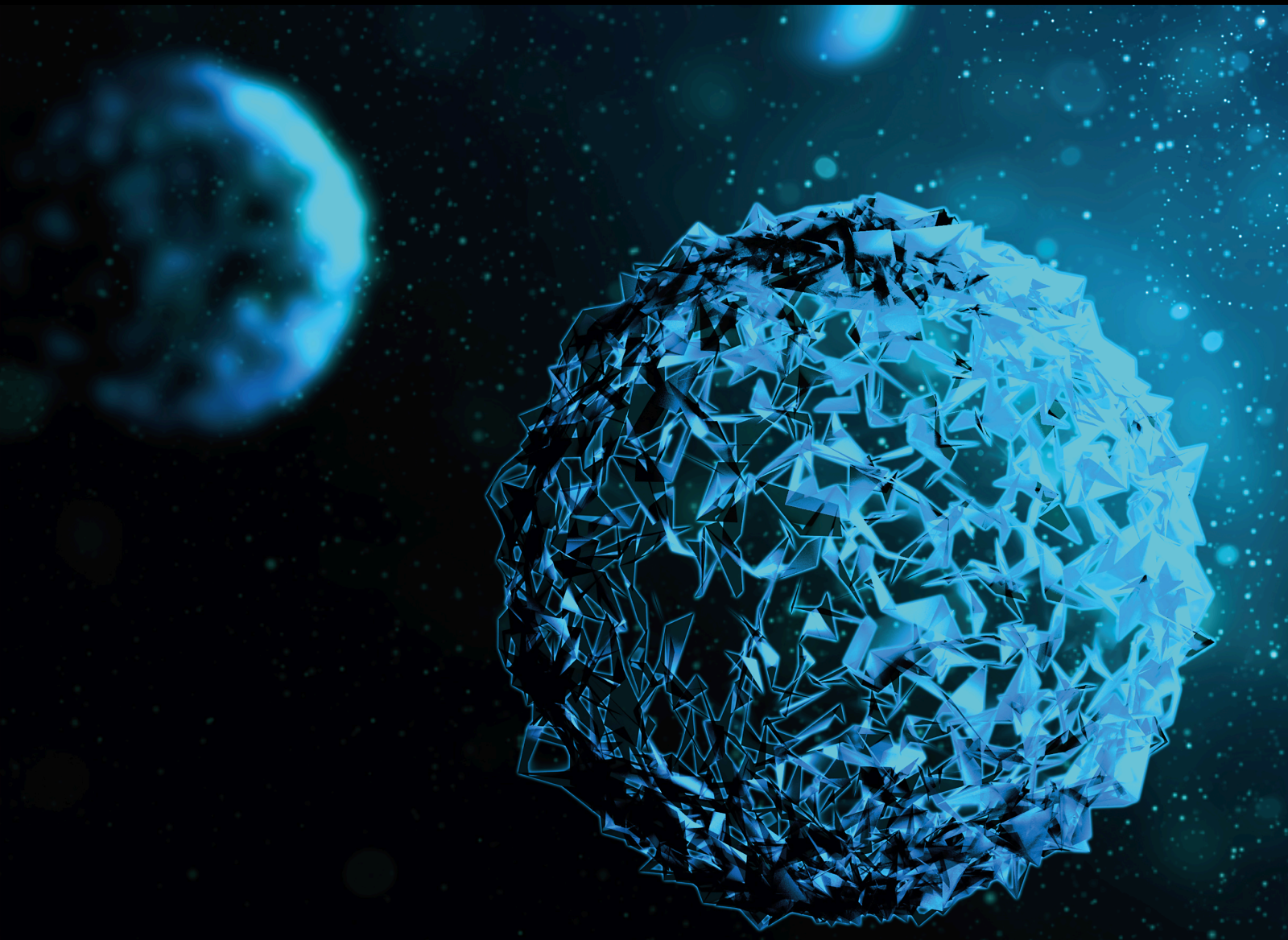


Applications of Bioinformatics and Systems Biology in Precision Medicine and Immuno-Oncology 2020

Lead Guest Editor: Yudong Cai
Guest Editors: Tao Huang and Shijia Zhu





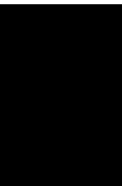
**Applications of Bioinformatics and Systems
Biology in Precision Medicine and Immuno-
Oncology 2020**

BioMed Research International

**Applications of Bioinformatics and
Systems Biology in Precision Medicine
and Immuno-Oncology 2020**

Lead Guest Editor: Yudong Cai

Guest Editors: Tao Huang and Shijia Zhu



Copyright © 2024 Hindawi Limited. All rights reserved.

This is a special issue published in "BioMed Research International." 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.

Section Editors

Penny A. Asbell, USA
David Bernardo , Spain
Gerald Brandacher, USA
Kim Bridle , Australia
Laura Chronopoulou , Italy
Gerald A. Colvin , USA
Aaron S. Dumont, USA
Pierfrancesco Franco , Italy
Raj P. Kandpal , USA
Fabrizio Montecucco , Italy
Mangesh S. Pednekar , India
Letterio S. Politi , USA
Jinsong Ren , China
William B. Rodgers, USA
Harry W. Schroeder , USA
Andrea Scribante , Italy
Germán Vicente-Rodríguez , Spain
Momiao Xiong , USA
Hui Zhang , China

Academic Editors

Bioinformatics


Contents

Retracted: The Relationship between Urinary Stones and Gut Microbiome by 16S Sequencing

BioMed Research International


Retraction (1 page), Article ID 9760313, Volume 2024 (2024)

Corrigendum to “Construction of Protein-related Risk Score Model in Bladder Urothelial Carcinoma”

Qizhan Luo and Xiaobo Zhang 


Corrigendum (1 page), Article ID 9758361, Volume 2021 (2021)

Identification and Verification of Molecular Subtypes with Enhanced Immune Infiltration Based on m6A Regulators in Cutaneous Melanoma

Yitong Lin, Shu Wang, Shirui Liu, Sha Lv, Huayang Wang, and Fuqiu Li 


Research Article (19 pages), Article ID 2769689, Volume 2021 (2021)

Nomogram Personalizes and Visualizes the Overall Survival of Patients with Triple-Negative Breast Cancer Based on the Immune Genome

Peipei Wang, Yang Fu, Yueyun Chen, Qing Li, Ye Hong, Ting Liu, and Zhenyu Ding 


Research Article (16 pages), Article ID 4029062, Volume 2020 (2020)

Solanine Inhibits Immune Escape Mediated by Hepatoma Treg Cells via the TGF β /Smad Signaling Pathway

Juwei Gao , Yinyin Ying, Jue Wang, and Yiyi Cui

Research Article (10 pages), Article ID 9749631, Volume 2020 (2020)

Immune Infiltration Landscape in Lung Squamous Cell Carcinoma Implications

Jungang Zhao, Wenming Bao, and Weiyang Cai 


Research Article (21 pages), Article ID 5981870, Volume 2020 (2020)

The Regulation of circRNA RNF13/miRNA-1224-5p Axis Promotes the Malignant Evolution in Acute Myeloid Leukemia

Rong Zhang, Yingchun Li, Hongtao Wang, Ke Zhu, and Guojun Zhang 


Research Article (11 pages), Article ID 5654380, Volume 2020 (2020)

JAK3 and TYK2 Serve as Prognostic Biomarkers and Are Associated with Immune Infiltration in Stomach Adenocarcinoma

Lingkai Meng, Ling Ding, Yue Yu, and Wang Li 



Research Article (15 pages), Article ID 7973568, Volume 2020 (2020)

[Retracted] The Relationship between Urinary Stones and Gut Microbiome by 16S Sequencing

Chenhao Zhou, Kai Li, Lun Zhao, Wei Li, Zongbao Guo, Jingyao Xu, Xiaofei Qi, and Hexing Yuan 


Research Article (7 pages), Article ID 1582187, Volume 2020 (2020)

Identification of Latent Oncogenes with a Network Embedding Method and Random Forest



Ran Zhao, Bin Hu, Lei Chen , and Bo Zhou 

Research Article (11 pages), Article ID 5160396, Volume 2020 (2020)



Inhibition Mir-92a Alleviates Oxidative Stress and Apoptosis of Alveolar Epithelial Cells Induced by Lipopolysaccharide Exposure through TLR2/AP-1 Pathway

Jian Cui, Huanhuan Ding, Yongyuan Yao, and Wei Liu 
Research Article (8 pages), Article ID 9673284, Volume 2020 (2020)

circHIPK3 Acts as Competing Endogenous RNA and Promotes Non-Small-Cell Lung Cancer Progression through the miR-107/BDNF Signaling Pathway

Weijun Hong, Yajuan Zhang, Jingyi Ding, Qilian Yang, Haixiang Xie , and Xiwen Gao 
Research Article (9 pages), Article ID 6075902, Volume 2020 (2020)




SRCIN1 Regulated by circCCDC66/miR-211 Is Upregulated and Promotes Cell Proliferation in Non-Small-Cell Lung Cancer

Weijun Hong, Suyun Yu, Yaqing Zhuang, Qingqing Zhang, Jiqin Wang , and Xiwen Gao 
Research Article (11 pages), Article ID 5307641, Volume 2020 (2020)



A Novel RNA-Seq-Based Model for Preoperative Prediction of Lymph Node Metastasis in Oral Squamous Cell Carcinoma

Bo Qiao , Min Zhao, Jing Wu, Huan Wu, Yiming Zhao, Fanhao Meng, Yu Tian, Situo Wang, Jinlong Shi , and Haizhong Zhang 
Research Article (13 pages), Article ID 4252580, Volume 2020 (2020)

Development of an Immune Infiltration-Related Eight-Gene Prognostic Signature in Colorectal Cancer Microenvironment

Beilei Wu , Lijun Tao , Daqing Yang , Wei Li , Hongbo Xu , and Qianggui He 
Research Article (43 pages), Article ID 2719739, Volume 2020 (2020)

Comprehensive Characterization of Prognostic Long Noncoding RNAs in Osteosarcoma

Hua Gao, Yuanyuan Guo , Miaomiao Zhang, and Zuqiang Yi 
Research Article (12 pages), Article ID 6725753, Volume 2020 (2020)


Construction of Protein-related Risk Score Model in Bladder Urothelial Carcinoma

Qizhan Luo and Xiaobo Zhang 
Research Article (13 pages), Article ID 7147824, Volume 2020 (2020)

miR-1258 Regulates Cell Proliferation and Cell Cycle to Inhibit the Progression of Breast Cancer by Targeting E2F1







Xianbao Zhao 
Research Article (14 pages), Article ID 1480819, Volume 2020 (2020)

Identification of Differentially Expressed Genes and miRNAs Associated with Esophageal Squamous Cell Carcinoma by Integrated Analysis of Microarray Data

Lemeng Zhang , Jianhua Chen, Tianli Cheng, Hua Yang, Changqie Pan, and Haitao Li
Research Article (16 pages), Article ID 1980921, Volume 2020 (2020)


Contents

Prognostic Value of Complement Component 2 and Its Correlation with Immune Infiltrates in Hepatocellular Carcinoma

Gang Ning , Yan-Lin Huang , Li-Min Zhen, Wen-Xiong Xu , Xue-Jun Li , Li-Na Wu, Ying Liu, Chan Xie , and Liang Peng 

Research Article (12 pages), Article ID 3765937, Volume 2020 (2020)

Profiles of Immune Infiltration and Prognostic Immunoscore in Lung Adenocarcinoma

Yanyan Li, Liping Tao, and Weiyang Cai 

Research Article (15 pages), Article ID 5858092, Volume 2020 (2020)

Effect of Different Expression of Immune-Related lncRNA on Colon Adenocarcinoma and Its Relation to Prognosis

Meiwei Mu, Yi Tang, Zheng Yang , Yuling Qiu, Xiaohong Li , Wuning Mo , and Qisheng Su 

Research Article (9 pages), Article ID 6942740, Volume 2020 (2020)

Retraction

Retracted: The Relationship between Urinary Stones and Gut Microbiomeby 16S Sequencing

BioMed Research International

Received 12 March 2024; Accepted 12 March 2024; Published 20 March 2024

Copyright © 2024 BioMed Research International. 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.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] C. Zhou, K. Li, L. Zhao et al., "The Relationship between Urinary Stones and Gut Microbiomeby 16S Sequencing," *BioMed Research International*, vol. 2020, Article ID 1582187, 7 pages, 2020.

Corrigendum

Corrigendum to “Construction of Protein-related Risk Score Model in Bladder Urothelial Carcinoma”

Qizhan Luo^{1,2} and Xiaobo Zhang¹

¹Xiangya International Medical Center, Department of Geriatrics, Xiangya Hospital, Central South University, Changsha, China

²Department of Urology, RWTH Aachen University, Pauwelsstrasse 30, 52074 Aachen, Germany

Correspondence should be addressed to Xiaobo Zhang; 549651566@qq.com

Received 23 May 2021; Accepted 23 May 2021; Published 7 June 2021

Copyright © 2021 Qizhan Luo and Xiaobo Zhang. 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.

In the article titled “Construction of Protein-related Risk Score Model in Bladder Urothelial Carcinoma” [1], the Acknowledgments section should read as follows:

Acknowledgments

We would like to acknowledge the supportive technique by the Xiangya International Medical Center, Department of Geriatrics, Xiangya Hospital, Central South University, and Guangjun, and the China Scholarship Council and Xiaobo Zhang. Also, the authors acknowledge Dr. Jun Guang from Sukebiota for assisting with data analysis, gene selection, and the construction and for verification of the risk model.

References

- [1] Q. Luo and X. Zhang, “Construction of protein-related risk score model in bladder urothelial carcinoma,” *BioMed Research International*, vol. 2020, Article ID 7147824, 13 pages, 2020.

Research Article

Identification and Verification of Molecular Subtypes with Enhanced Immune Infiltration Based on m6A Regulators in Cutaneous Melanoma

Yitong Lin,¹ Shu Wang,² Shirui Liu,¹ Sha Lv,¹ Huayang Wang,³ and Fuqiu Li¹ 

¹Department of Dermatology, The Second Hospital of Jilin University, Changchun, Jilin Province, 130041, China

²Department of Radiotherapy, The Second Hospital of Jilin University, Changchun, Jilin Province, 130041, China

³Department of Dermatology, Xiang'an Hospital of Xiamen University, Xiamen, Fujian Province, 361101, China

Correspondence should be addressed to Fuqiu Li; lifuqiu1234@126.com

Received 8 July 2020; Revised 13 November 2020; Accepted 22 December 2020; Published 4 January 2021

Academic Editor: Shijia Zhu

Copyright © 2021 Yitong Lin 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. As the most aggressive type of skin cancer, cutaneous melanoma (CM) is experiencing a rapidly rising mortality in recent years. Exploring potential prognostic biomarkers or mechanisms of disease progression therefore has a great significance for CM. The purpose of this study was to identify genetic markers and prognostic performance of N6-methyladenosine (m6A) regulators in CM. **Method.** Gene expression profiles, copy number variation (CNV), and single nucleotide polymorphism (SNP) data of patients were obtained from The Cancer Genome Atlas (TCGA) database. **Results.** Genomic variation and association analysis of gene expressions revealed a high degree of genomic variation in the presence of m6A-regulated genes. m6A patients with high-frequency genomic variants in the regulatory gene tended to develop a worse prognosis ($p < 0.01$). Unsupervised cluster analysis of the expression profiles of m6A-regulated genes identified three clinically distinct molecular subtypes, including degradation-enhanced subgroup and immune-enhanced subgroup, with significant prognostic differences ($p = 0.046$). A novel prognostic signature, which was established according to m6A-related characteristic genes identified through genome-wide expression spectrum, could effectively identify samples with poor prognosis and enhanced immune infiltration, and the effectiveness was also verified in the dataset of the chip. **Conclusion.** We identified genetic changes in the m6A regulatory gene in CM and related survival outcomes. The findings of this study provide new insights into the epigenetic understanding of m6A in CM.

1. Introduction

As the most aggressive type of skin cancer, cutaneous melanoma (CM) resulted in 287,723 new cases and 60,712 deaths worldwide in 2018 [1]. Although it accounts for only 4% of all skin cancer cases, it is far more dangerous than other skin cancers. The treatment and control of CM still remain less effective, though great efforts have been devoted to the improvements in managing advanced regional and metastatic CM [2]. UV exposure is known to be closely associated with the development of melanoma [3]. The incidence of CM shows regional variations such as ethnic skin phenotypes and differences in sun exposure [4]. A study demonstrated that the mortality of melanoma patients is rapidly increasing

in recent years [5] and also pointed out the management of melanoma patients remains a complex and evolving subject [6]. Thus, exploring potential prognostic biomarkers or the mechanism of disease development is highly necessary for CM.

RNA methylation has become a common phenomenon and a key regulator of transcript expressions [7]. N6-Methyladenosine (m6A), which is methylated at the N6 position of adenosine, is considered to be the most common conserved internal transcriptional modifications of messenger RNAs (mRNAs), microRNA (miRNAs), and long noncoding RNAs (lncRNAs) [8]. The deposition of m6A is encoded by a methyltransferase complex involving three homologous factors, including methyltransferases (termed as “writers”),

demethylases (termed as “erasers”), and recognition from m6A binding proteins (termed as “readers”). The m6A dysregulation consists of multiple cellular processes even in human cancers. It has been demonstrated that m6A mRNA demethylase FTO could regulate melanoma tumorigenicity and react to anti-PD-1 blockade [9]. The m6A methyltransferase METTL3 promotes osteosarcoma progression by regulating the m6A level of LEF1 [10]. Furthermore, METTL3 (an oncogene) maintains SOX2 expression via the m6A-IGF2BP2-dependent mechanism in colorectal carcinoma cells and therefore could serve as a potential biomarker for cancer prognosis prediction [11].

So far, however, little is known about the relationship between m6A-related genes and CM. The Cancer Genome Atlas (TCGA) database allows an easy access to human cancer data of gene expressions, copy number variation (CNVs), and single nucleotide polymorphism (SNPs), which all play important roles in the development and progression of human cancers [12]. However, the CNV and SNP of m6A-related genes remain unknown to us.

This retrospective study developed a novel gene signature based on m6A regulators in CM with data acquired from TCGA database and also analyzed the performance of the signature. By analyzing genomic variations and gene expression associations, we found that these m6A regulatory genes showed high genomic variations; interestingly, patients with m6A regulatory genes of high-frequency genomic variation tended to develop a worse prognosis ($p < 0.01$). Three clinical subtypes with different molecular characteristics and significant prognostic differences were identified by performing unsupervised clustering analysis on the expression profiles of m6A regulatory genes. m6A-related characteristic genes were determined based on genome-wide expression profiles. A new prognostic signature was then built for identifying subgroups with poor prognosis and enhanced immune infiltration, and its performance was validated in the training set and validation set. In summary, we determined genetic changes in the m6A regulatory genes in CM and patient survival outcomes. The findings of this study provide new insights into the epigenetic understanding of m6A in CM.

2. Materials and Methods

2.1. Data Resource and Processing. All CM clinical data, copy number variation (CNV), single nucleotide polymorphism (SNP), and mRNA expression data were retrieved from TCGA-Assembler of the TCGA website [13] and downloaded in May 2019. For transcriptome data, we obtained 472 samples and downloaded data of read counts. Data was normalized by R package *DESeq*. For SNP data, we obtained a total of 469 samples, and the downloaded data was processed by MuTect [14]. For CNV data, there were 940 samples acquired by R package *RTCGA*. Here, the “deletion” was defined as $\text{segment_mean} < -0.2$, and the “amplification” was defined as the $\text{segment_mean} > 0.2$. For clinical information data, initially, there were a total of 389 clinical samples; after integrating the data and excluding samples with a survival time shorter than 90 days, a total of 250 CM samples were finally decided for further analysis.

The expression spectrum dataset GSE65904 [15] with 214 melanoma samples from Illumina HumanHT-12 V4.0 Expression BeadChip platform was downloaded from the Gene Expression Omnibus (GEO) database. The samples with a survival time shorter than 90 days were selected, resulting in a final of 189 samples. Multiple probes corresponding to a gene were retained and shown as the median of the gene expression level, while probes corresponding to multiple genes were eliminated. Finally, the expression profile of genes was obtained. The work flow chart is shown in Figure 1.

2.2. Univariate Cox Survival Analysis. Here, the prognostic performance of m6A regulatory genes was examined. Based on the expression profile of m6A regulatory genes and clinical follow-up data of the samples, each m6A regulatory gene was evaluated using R software package *survival*, and the forest map was plotted applying R software package *forestplot*.

2.3. Clustering Analysis. For a better classification of the patients, an unsupervised clustering method was employed in hierarchical clustering analysis of the expression profile of CNV-related m6A regulatory genes. The clinical subtypes were selected according to the minimum intragroup variance and the maximum intergroup variance of the method.

2.4. Differential Expression Gene Analysis. Gene differences between each clinical subgroup were analyzed using R software package *DESeq2*, with $|\log_2(\text{fold change})| > 2$ and $\text{FDR} < 0.05$ serving as the thresholds.

2.5. Analysis of Immune Cells and Immune Infiltration. The ssGSEA (single-sample gene set enrichment analysis) algorithm was applied to quantify the relative abundance of cell invasion in each tumor microenvironment (TME) of a CM patient. The gene sets labeled as TME-infiltrating immune cell type were obtained from Charoentong et al.’s study [16], which investigated a variety of human immune cell subtypes, including activated CD8 T cells, activated dendritic cells, macrophages, natural killer T cells, and regulatory T cells. The enrichment fraction calculated by ssGSEA was used to represent the relative abundance of TME-infiltrating cells in the samples. Patient’s infiltration score was assessed by the R package *Estimate*.

2.6. KEGG Pathway Enrichment Analysis of Different Clinical Subgroups. ssGSEA KEGG pathway analysis was performed on each sample using R software package *GSEA* [17] to calculate the difference of enrichment score of each pathway in different clinical subgroups. Moreover, the relationship between enrichment scores of each sample pathway and clinical subgroup was analyzed by performing Pearson correlation. $p < 0.05$ was the threshold value in determining the KEGG pathway that was the most relevant to the clinical subgroup.

2.7. Principal Component Analysis. Principal component analysis was performed according to the expression profile of m6A-related gene markers, with the first principal component being considered the scoring coefficient. A scoring

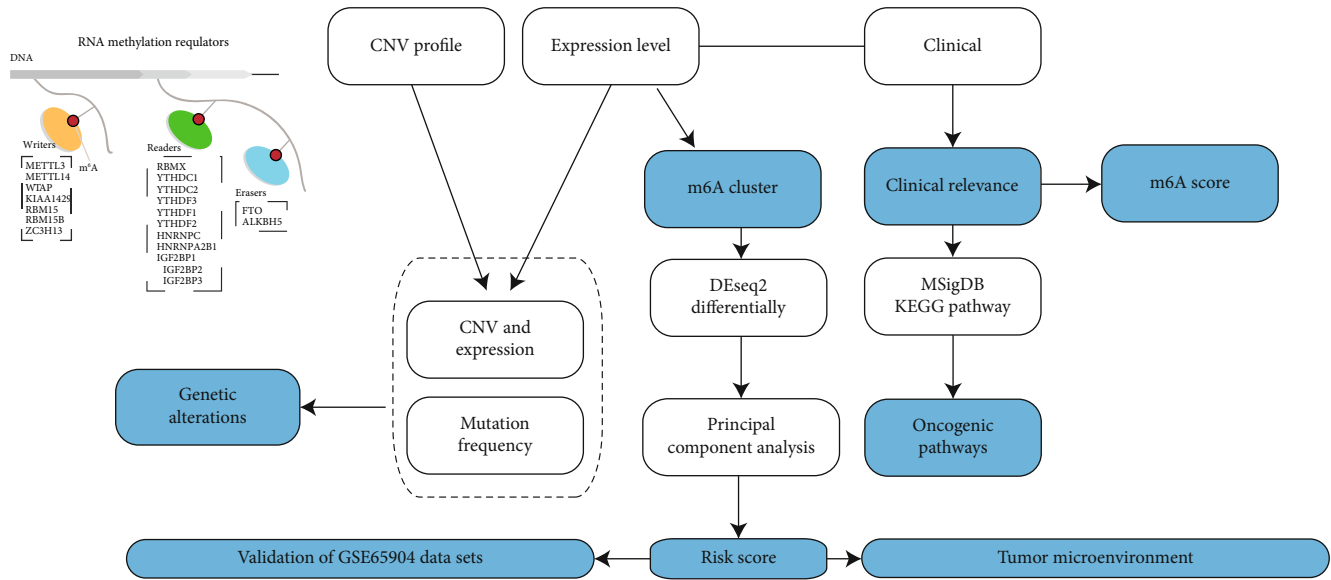


FIGURE 1: The work flow chart.

model was established to calculate the risk score of each sample according to the model. Samples scored above 75% were assigned into the high-risk group with immunorejection phenotypes, while those scored below 75% were in the low-risk group.

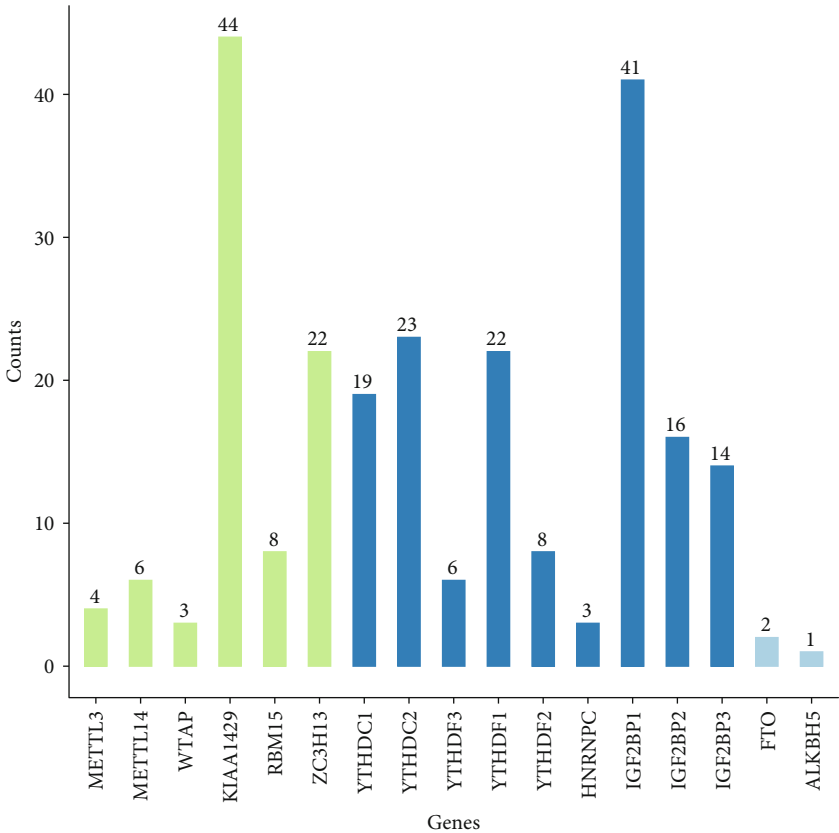
2.8. Statistical Analysis. All statistical data were analyzed by SPSS 23 (IBM, Chicago, USA) and R language. The association between CNV and SNP of m6A regulatory genes and clinicopathological characteristics was examined by the chi-squared test. The association between three CM key genes and CNV and SNP of m6A regulatory genes was analyzed by the chi-squared test. The Kaplan-Meier curve and log-rank test were applied to evaluate the prognostic performance of the alterations in m6A regulatory genes. All statistical results with a p value ≤ 0.05 were considered to be significant.

3. Results

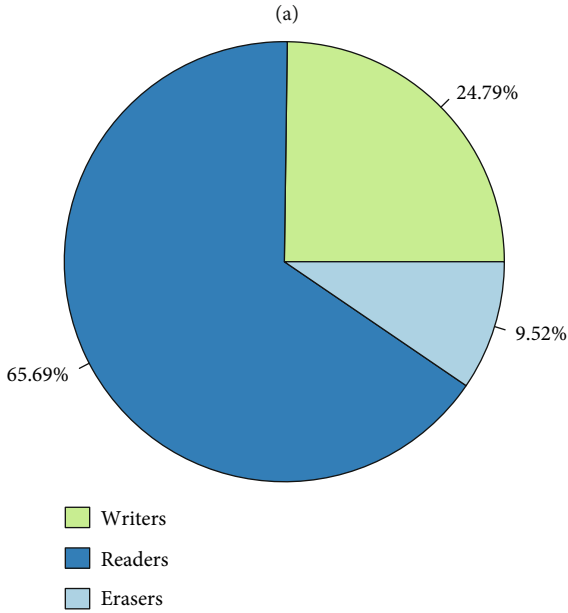
3.1. The m6A Regulatory Genes Showed a High Frequency of Genomic Variation in CM Patients. A total of 17 m6A regulatory genes were recruited in this study. 136 out of the 242 SNP of m6A regulatory genes with sequencing data appeared as independent samples (Table 1). Among them, the “writer” gene KIAA1429 in 44 samples and the “reader” gene IGF2BP1 in 41 samples demonstrated a higher SNP, with all the “reader” genes showing a higher SNP frequency than that of the “writer” or “eraser” genes; noticeably, the frequency of SNP of the “eraser” gene was the lowest (Figure 2(a)). We also observed that in all 469 CM samples with CNV data, the m6A regulatory genes had a high frequency of CNV (Figure 2(b)). For example, the “writer” gene WTAP showed the highest frequency of 57.08% of its CNV events, and the frequency of the “reader” gene IGF2BP3 was 52.1% (Table 2). These results indicated a high frequency of genomic SNP in the m6A regulatory gene in CM, and

these abnormalities may affect gene expression and lead to different clinical outcomes in CM patients.

3.2. CNV of m6A Regulatory Genes Was Associated with Adverse Clinical Outcomes. Considering that changes in CNV can affect gene expression levels through dosage compensation effects, to this end, the effect of m6A-regulated gene changes on mRNA expression was evaluated. In 472 KM samples, mRNA expression levels showed a significant correlation with CNV patterns. For all 17 regulatory genes, except for IGF2BP3 and IGF2BP1, increased CNV of the rest 15 genes was related to a higher mRNA expression and deletion in mRNAs with decreased expression (Figure 3(a)). Furthermore, the prognostic differences between samples of m6A regulatory genes with the high-frequency and low-frequency CNV abnormalities were analyzed. We observed that patients with a high CNV tended to show worse clinical outcomes ($p < 0.001$), directly indicating that abnormal copy number of m6A regulatory genes was related to poor clinical outcomes (Figure 3(b)). Finally, by analyzing the relationship between the expressions of m6A regulatory genes and patients’ prognosis, we observed that RBM15, YTHDF1, WTAP, and METTL14 genes were significantly associated with CM prognosis (Figure 3(c)). Specifically, high-expressed YTHDF1 was indicative of a poor prognosis and could therefore be considered a risk factor, at the same time, low-expressed RBM15, WTAP, and METTL14 were also related to a poor prognosis and were all seen as protective factors. In order to further determine the prognostic value of m6A regulators, a risk score model (m6AScore) was established based on a multivariate regression method to calculate the risk score of each patient in TCGA cohort, and the prognostic value of m6AScore was determined by univariate and multivariate Cox survival analyses. The results of univariate analysis demonstrated that m6AScore and a variety of clinical features such as T, N, and stages were strongly correlated with CM patients’ prognosis (Figure 3(d)), while multivariate



Functions
Erasers
Readers
Writers



(a)

FIGURE 2: SNP and CNVs of m6A regulatory genes in CM patients: (a) frequency of SNP of different m6A regulatory genes in CM samples; (b) the CNV data of m6A regulatory genes in CM samples.

TABLE 2: Different CNV patterns occur in CM patients.

| Types | Genes | Diploid | Deletion | Amplification | CNV sum | Amplification (%) | Deletion (%) | Percentage |
|---------|----------|---------|----------|---------------|---------|-------------------|--------------|------------|
| Writers | METTL3 | 338 | 93 | 40 | 133 | 30.08% | 69.92% | 28.24% |
| | METTL14 | 354 | 76 | 41 | 117 | 35.04% | 64.96% | 24.84% |
| | WTAP | 203 | 246 | 24 | 270 | 8.89% | 91.11% | 57.08% |
| | KIAA1429 | 268 | 10 | 194 | 204 | 95.10% | 4.90% | 43.22% |
| | RBM15 | 236 | 70 | 75 | 145 | 51.72% | 48.28% | 38.06% |
| | ZC3H13 | 303 | 76 | 92 | 168 | 54.76% | 45.24% | 35.67% |
| | YTHDC1 | 361 | 68 | 52 | 120 | 43.33% | 56.67% | 24.95% |
| | YTHDC2 | 324 | 100 | 47 | 147 | 31.97% | 68.03% | 31.21% |
| Readers | YTHDF3 | 282 | 14 | 175 | 189 | 92.59% | 7.41% | 40.13% |
| | YTHDF1 | 271 | 3 | 198 | 201 | 98.51% | 1.49% | 42.58% |
| | YTHDF2 | 346 | 55 | 70 | 125 | 56.00% | 44.00% | 26.54% |
| | HNRNPC | 338 | 93 | 40 | 133 | 30.08% | 69.92% | 28.24% |
| | IGF2BP1 | 3342 | 43 | 90 | 133 | 67.67% | 32.33% | 3.83% |
| | IGF2BP2 | 358 | 51 | 62 | 113 | 54.87% | 45.13% | 23.99% |
| Erasers | IGF2BP3 | 228 | 18 | 230 | 248 | 92.74% | 7.26% | 52.10% |
| | FTO | 329 | 116 | 36 | 152 | 23.68% | 76.32% | 31.60% |
| | ALKBH5 | 336 | 209 | 26 | 235 | 11.06% | 88.94% | 41.16% |
| Total | | 8217 | 1341 | 1492 | 2833 | 52.67% | 47.33% | 25.64% |

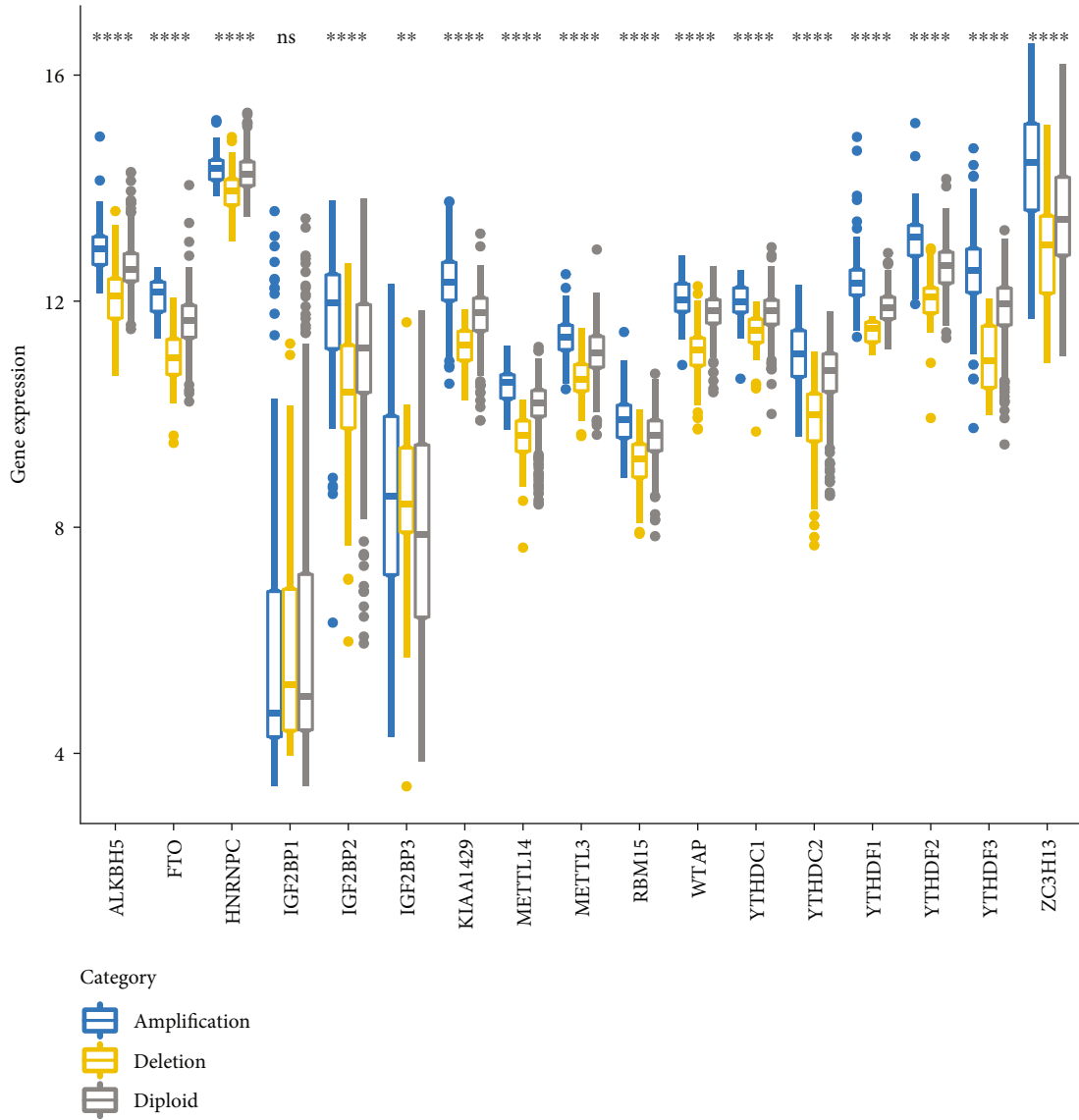
analysis showed that only m6AScore, N3, N2, and T4 had a significant prognostic correlation (Figure 3(e)). In addition, HR of N3, N2, and T4 was abnormally high. These results indicated that m6AScore has a more stable prognostic performance than T, N, M, and stage or other clinical features.

3.3. Different Molecular Subgroups Identified by Unsupervised Cluster Analysis Based on Expressions of m6A Genes. Given that m6A regulatory genes are associated with CM prognosis, unsupervised hierarchical cluster analysis was conducted based on the expression profiles of 15 CNV-related m6A regulatory genes. We observed that these 15 genes were divided into three main categories (Figure 4(a)) with different expression patterns of m6A genes. Further analysis showed that 12 genes (80%) out of the 15 genes in the three subtypes had significant expression differences (Figure 4(b)). Subsequently, prognostic differences among the three types of samples were examined, and the data demonstrated that m6A Cluster3 was related to a significantly better prognosis than m6A Cluster1 and m6A Cluster2 (Figures 4(c) and 4(d)). These results suggest that the three m6A Clusters have different molecular characteristics, which will lead to different clinical outcomes.

3.4. The Regulatory Characteristics of m6A Cluster. To explore the relationship between the biological behaviors of CM and m6A Cluster, the enrichment scores of each sample in the KEGG pathway were calculated by ssGSEA to further determine the differences of the enrichment scores (Figure 5(a)). Significant differences in 44 KEGG pathways can be observed; interestingly, 40 pathways showed significant positive correlations with m6A Cluster2, and they were mainly important signaling pathways involved in tumor development, for example, VEGF_signaling_pathway, basal_cell_carcinoma, and drug_metabolism_cytochrome.

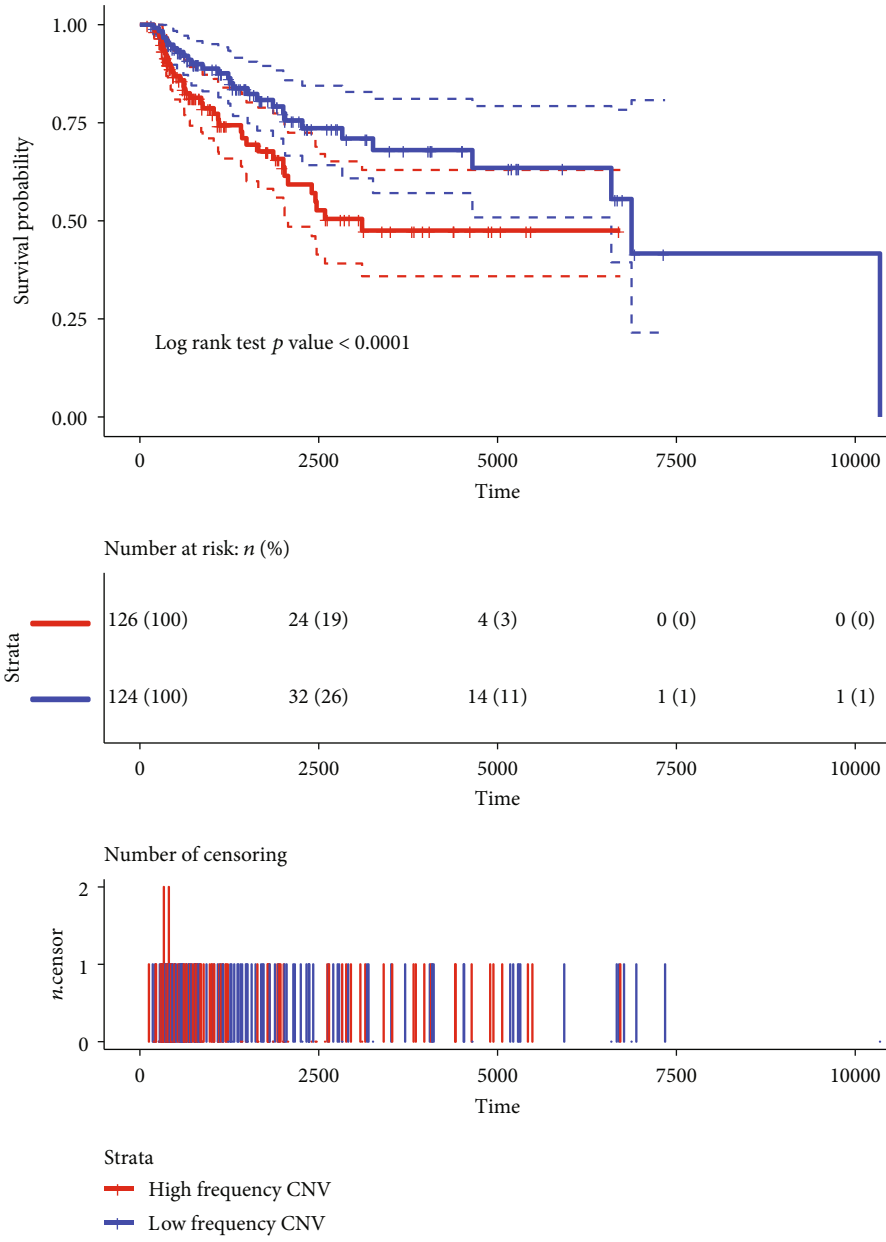
In addition, the remaining 4 KEGG pathways, namely, ubiquitin_mediated_proteolysis, non_homologous_end_joining, basal_transcription_factors, and RNA_degradation, were significantly upregulated in m6A Cluster3 and are mainly related to the degradation process. 28 immune cell components were calculated for each patient by the ssGSEA, and a significant difference in the distribution of 19 immune cells (67.9%) was detected (Figure 5(b)). Among these significant immune cells, memory B cell and type 2 T helper cell showed the highest scores in m6A Cluster1, while the remaining 17 immune cells all had the highest scores in m6A Cluster2. These results suggested that m6A Cluster3 is related to immune processes and a better CM prognosis, while m6A Cluster2 is related to a stable immune microenvironment and a worse prognosis. We also compared the relationships between the 3 m6A-related molecular subtypes and the previously reported genomic subtypes (BRAF, RAS, NF1, and triple-WT) [18]). The data showed that m6A Cluster3 had the largest intersection with the RAS subtype, whereas m6A Cluster1 and m6A Cluster2 had the largest intersection with triple_WT (Figure 5(c)). In our study, the results that m6A Cluster1 and m6A Cluster2 were indicative of a poorer prognosis and higher frequency of m6A copy number are consistent with triple-WT with significantly more copy number fragments.

3.5. Identification of m6A Gene Markers and Molecular Characteristics. To determine the gene expression differences among the m6A cluster samples, we identified 5,533 differentially expressed genes based on the gene expression profile. It has been found that m6A Cluster2 and m6A Cluster1/m6A Cluster3 had the most differentially expressed genes, as compared with m6A Cluster1 (Figure 6(a)). Gene expression profiles of these 5533 genes were extracted for cluster analysis,



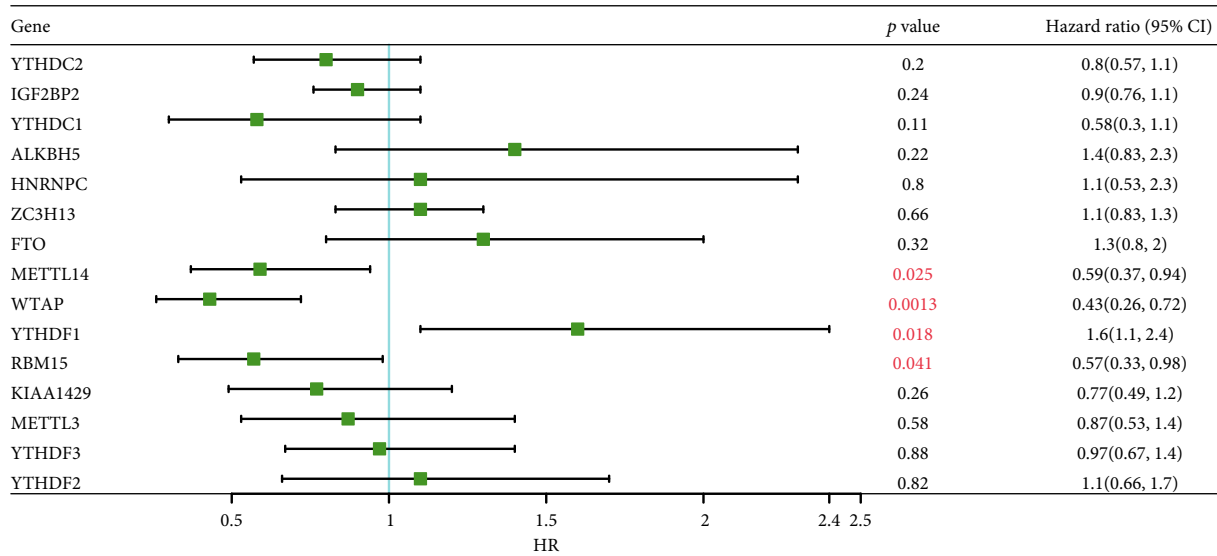
(a)

FIGURE 3: Continued.

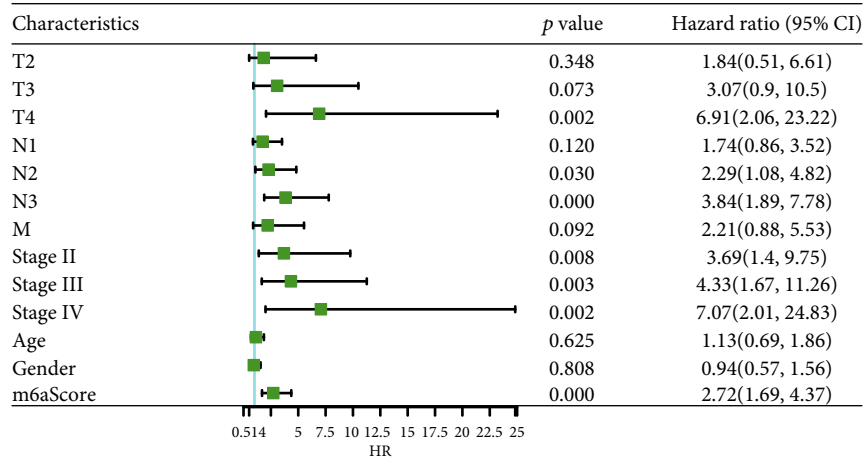


(b)

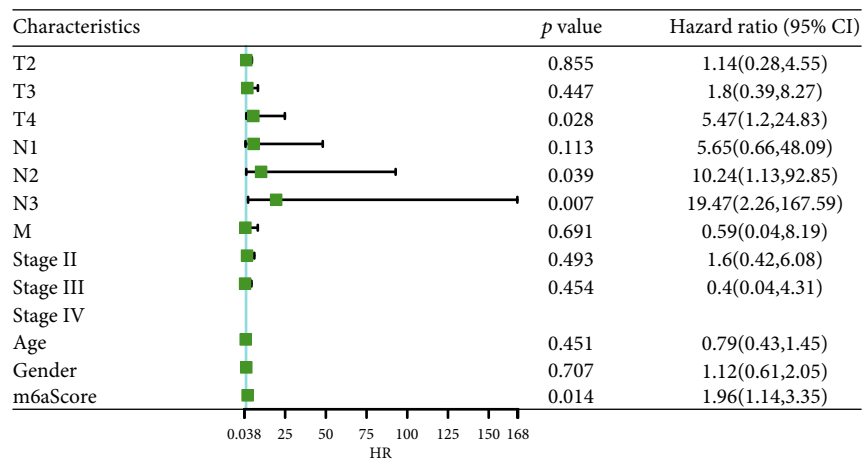
FIGURE 3: Continued.



(c)

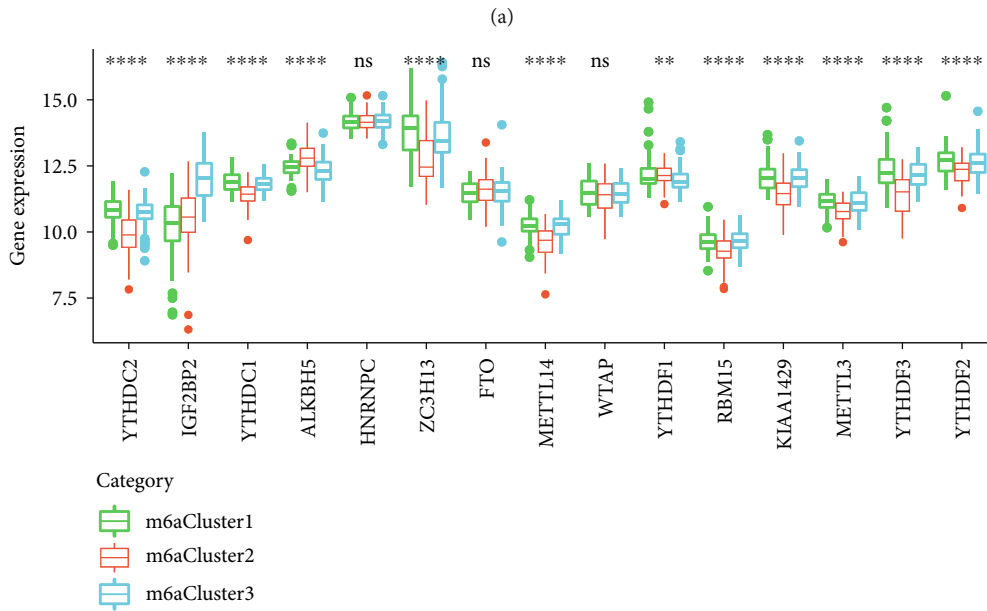
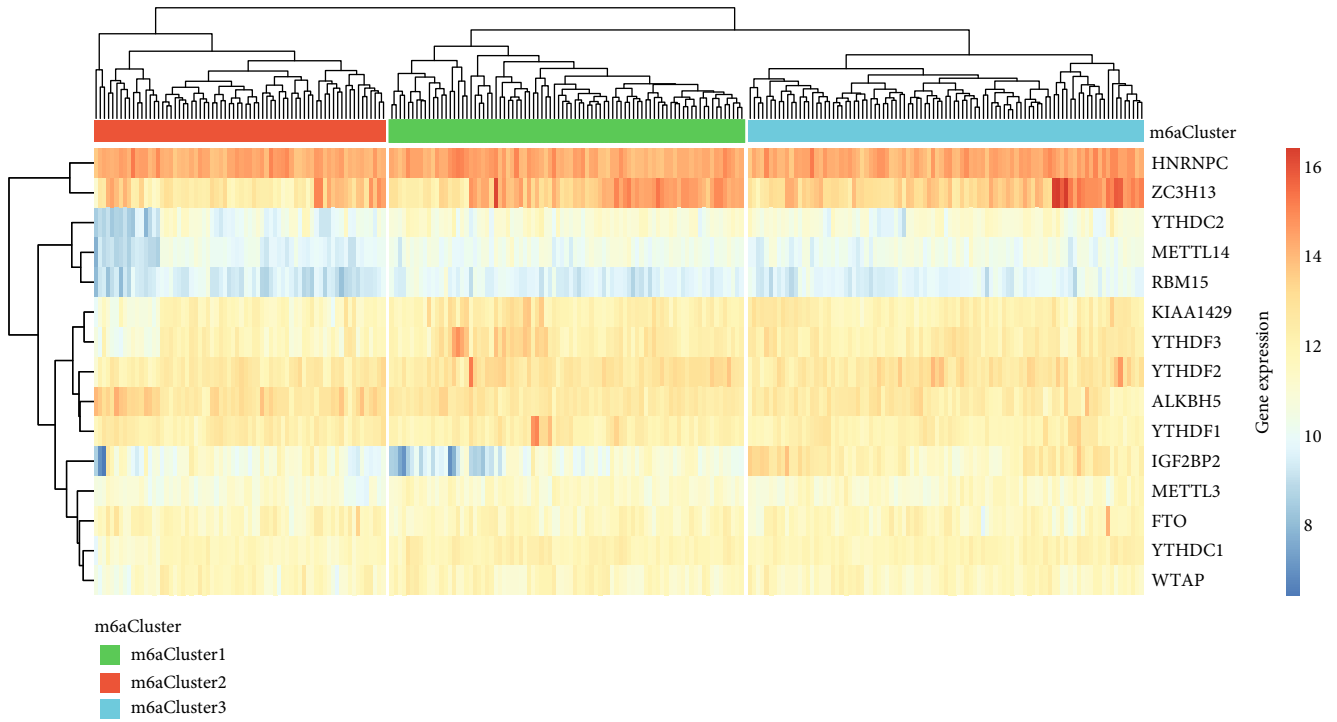


(d)



(e)

FIGURE 3: Relationship between m6A regulatory genes and CNV: (a) the relationships between CNV and expression levels of m6A regulatory genes. Kruskal test was used to detect the differences between different groups; (b) the Kaplan-Meier curves of CNV and prognosis in SCKM patients; (c) relationship between m6A regulatory genes and prognosis; (d) survival analysis by univariate Cox analysis on m6AScore and clinical features; (e) survival analysis by multivariate survival Cox on m6AScore and clinical features.



(b)

FIGURE 4: Continued.

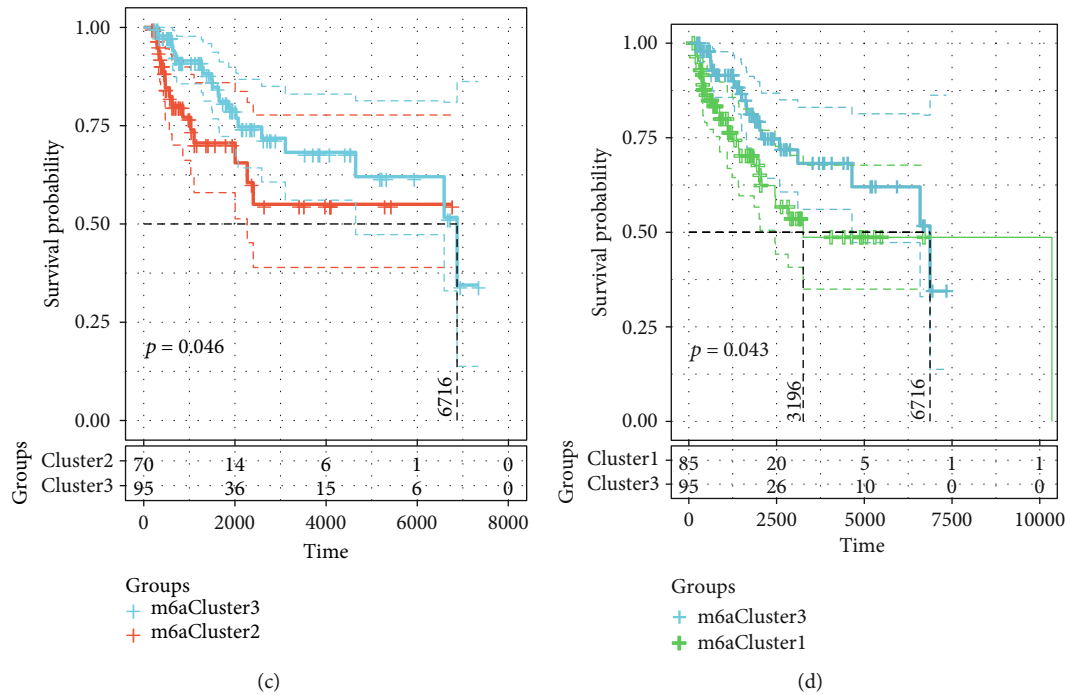


FIGURE 4: Different molecular subgroups identified by unsupervised cluster analysis based on m6A gene expression: (a) identification of molecular subtypes; (b) heat map of differentially expressed genes; (c) KM prognosis curve of m6A Cluster2 and m6A Cluster3; (d) KM prognosis curve of m6A Cluster1 and m6A Cluster3.

and the data demonstrated that these genes divided the samples into two categories (Cluster1 and Cluster2). Specifically, Cluster2 mainly consisted of m6A Cluster2 samples and a small part of m6A Cluster3 samples, Cluster1 contained almost all the m6A Cluster1 samples, and Cluster1 contained most of the m6A Cluster3 samples (Figure 6(b)). Further analysis demonstrated that Cluster2 was related to a significantly worse prognosis than Cluster1 (Figure 6(c)). Moreover, we also detected the expression differences of 15 m6A regulatory genes with CNV abnormalities in Cluster1 and Cluster2 (Figure 6(d)), and 12 (80%) out of the 15 genes showed significant expression differences, and among the 12 genes, ALKBH5 and YTHDF1 were significantly highly expressed in Cluster2, while the remaining 10 genes were significantly highly expressed in Cluster1. These results indicated that a class of molecular subgroups (m6A Cluster1/Cluster2) with unfavorable prognosis which resulted from abnormalities of immune microenvironment could be identified by m6A regulatory genes or m6A-related gene markers.

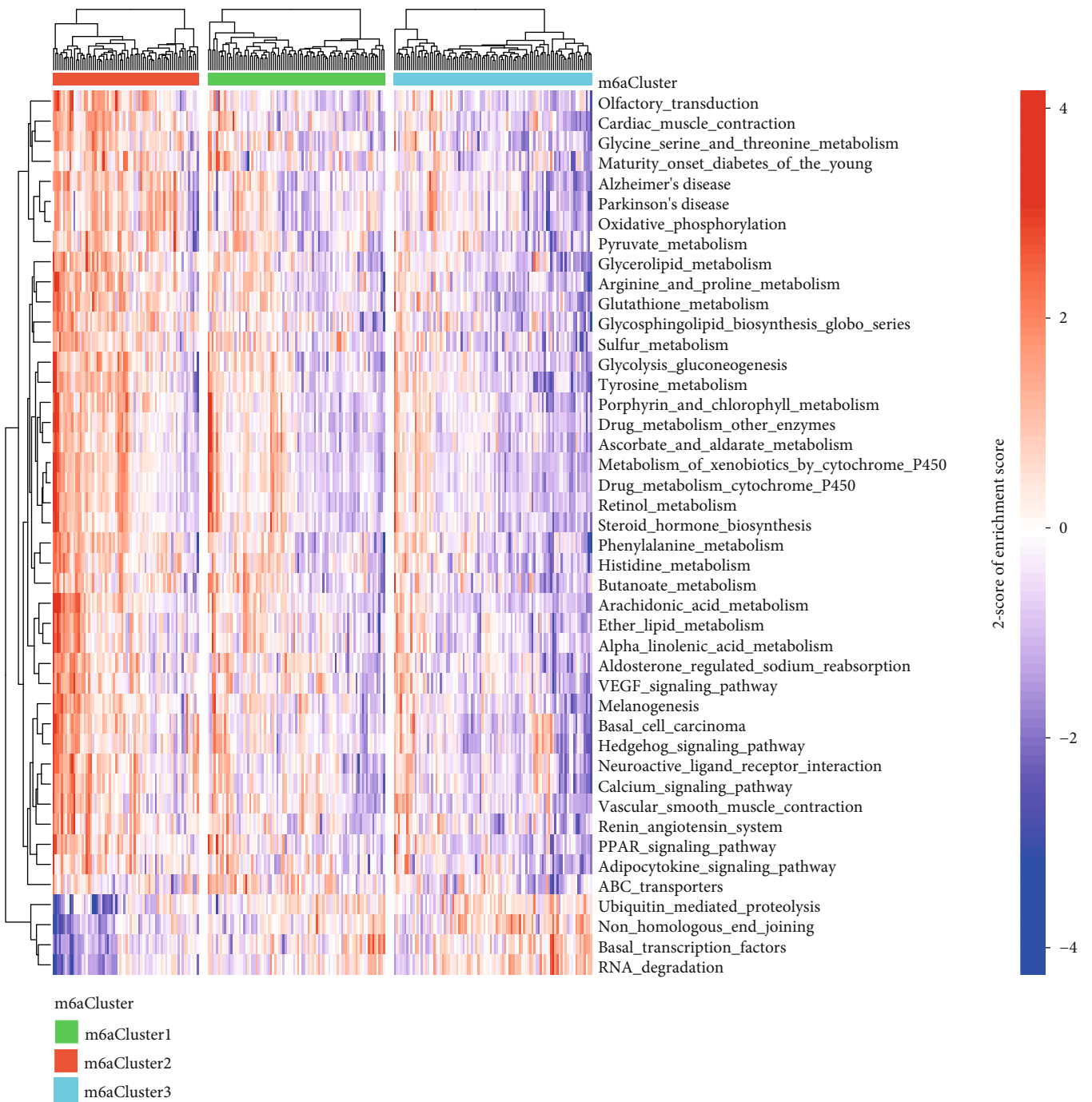
3.6. Establishment and Verification of Prognostic Signature Based on m6A Regulatory Genes. The subgroup of immune-related prognosis was determined by performing the principal component analysis based on the expression profiles of m6A-related gene markers to divide the samples into a high-risk group and low-risk group. A careful analysis of the prognostic differences between the two groups showed that the prognosis in the high-risk group was significantly worse than the low-risk group (Figure 7(a)). Moreover, a higher immune score was observed in the high-risk group, as compared with the low-risk group (Figure 7(b)). Further-

more, the patients were accordingly group based on the chip platform GSE65904. Here, the data showed that patients' prognosis in the high-risk group sample was greatly more unfavorable than that in the low-risk group (Figure 7(c)). Immunomicroenvironment analysis demonstrated that the samples from the high-risk group had higher immune microenvironment scores (Figure 7(d)). In addition, the Swegene Center for Integrative Biology at Lund University melanoma cohort (GSE22153) was introduced for verification. It has been found that the prognosis of the high-risk group was significantly worse than that of the low-risk group samples (Figure 7(e)). Immunomicroenvironment analysis showed that the high-risk group samples had a higher immune microenvironment score (Figure 7(f)). These results suggested that the expression profiles of m6A-related gene markers could serve as prognostic markers to evaluate the prognosis of CM patients.

4. Discussion

The present study developed a novel gene signature based on m6A regulators to CM using TCGA database and also assessed its predictive performance. Genetic alterations of the m6A regulatory genes have been found to be closely related to CM patients' survival outcomes. Our findings provide new insights into the epigenetic understanding of m6A in CM.

As the most common internal chemical modification of mRNA, m6A is widely involved in many pathological processes of cancer development. The deposition of m6A is encoded by a methyltransferase complex involving three homologous factors including methyltransferases (such as



(a)

FIGURE 5: Continued.

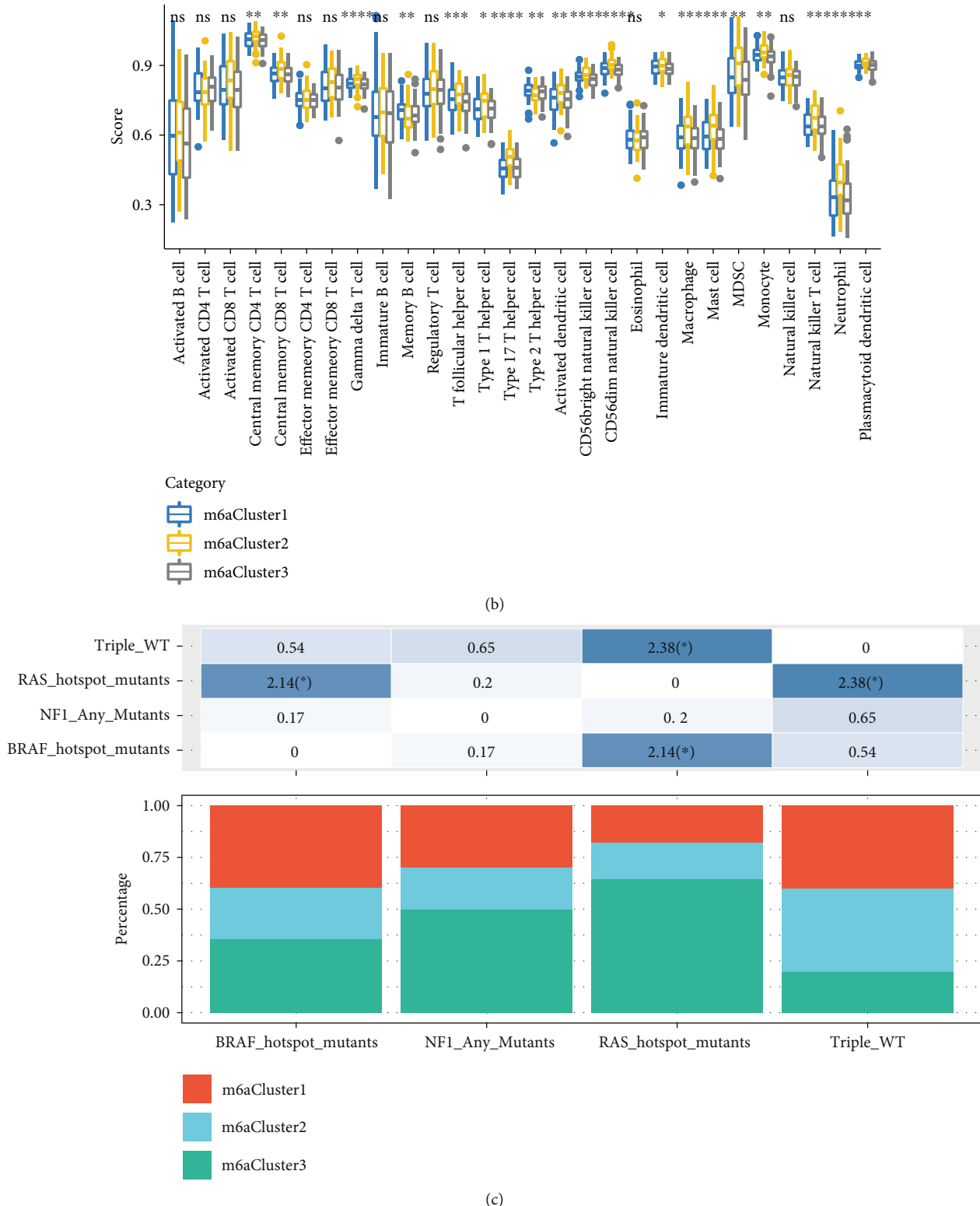
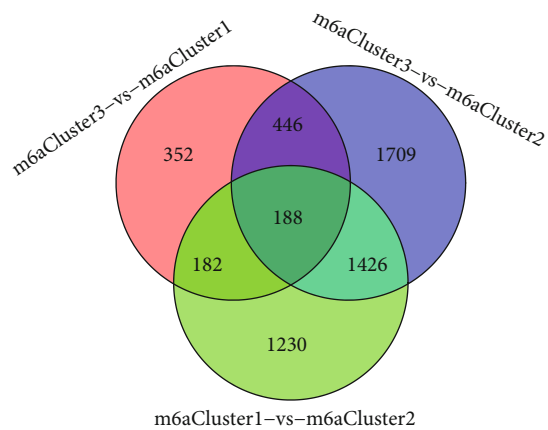


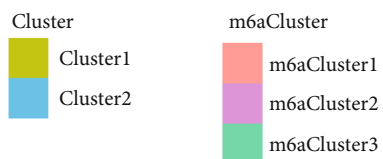
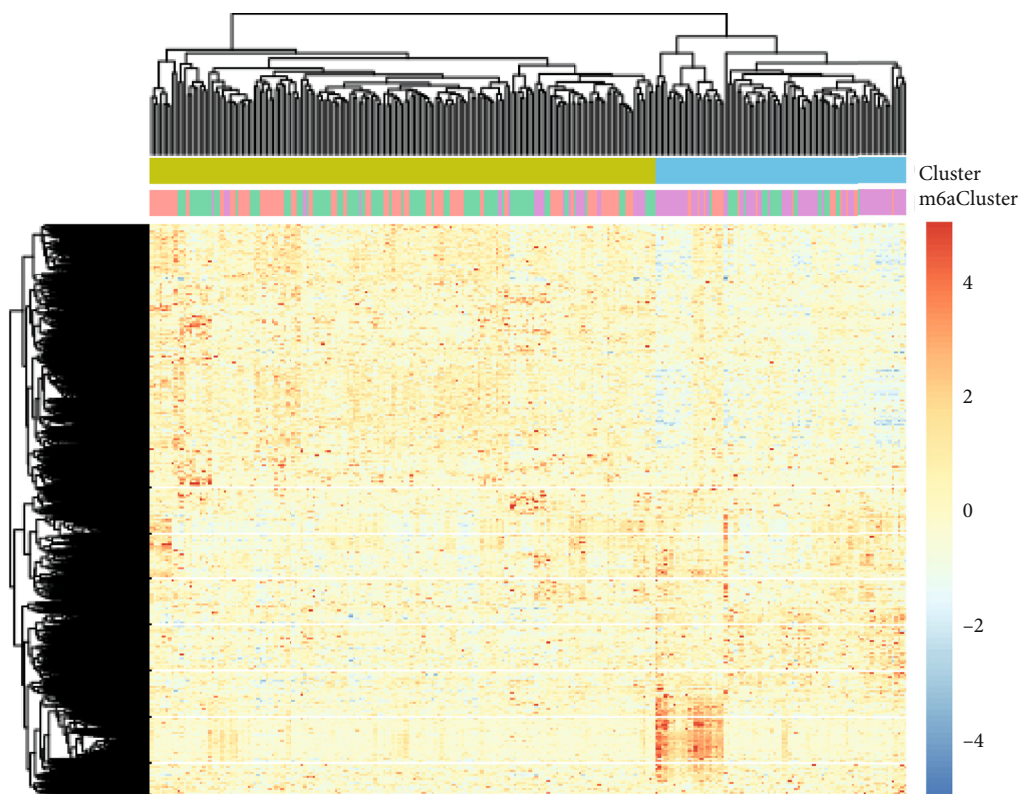
FIGURE 5: The regulatory characteristics of m6A cluster: (a) enrichment fraction heat map of the KEGG pathway in each sample with significant enrichment; (b) the score distribution of 28 immune cells in different m6A clusters; (c) the intersection relationship between the three m6A clusters and the four genomic isoforms.

METTL3/14, WTAP, RBM15/15B, and KIAA1429, termed as “writers”), demethylases (such as FTO and ALKBH5, termed as “erasers”) and recognition from m6A binding pro-

teins (such as YTHDF1/2/3, IGF2BP1 and HNRNPA2B1, termed as “readers”) [8]. Specifically, METTL3 silencing reduces m6A methylation and downregulates the total

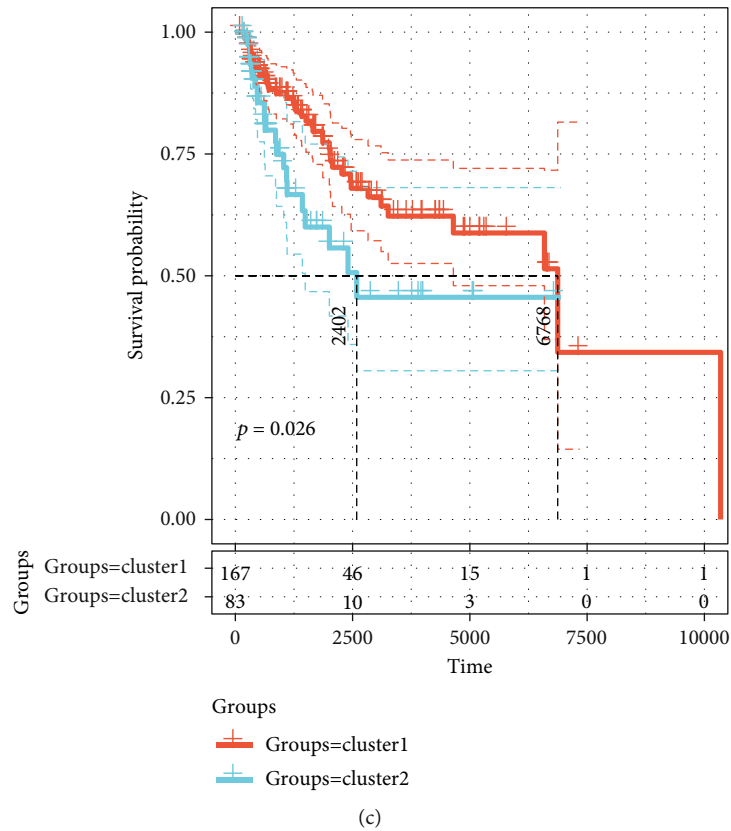


(a)

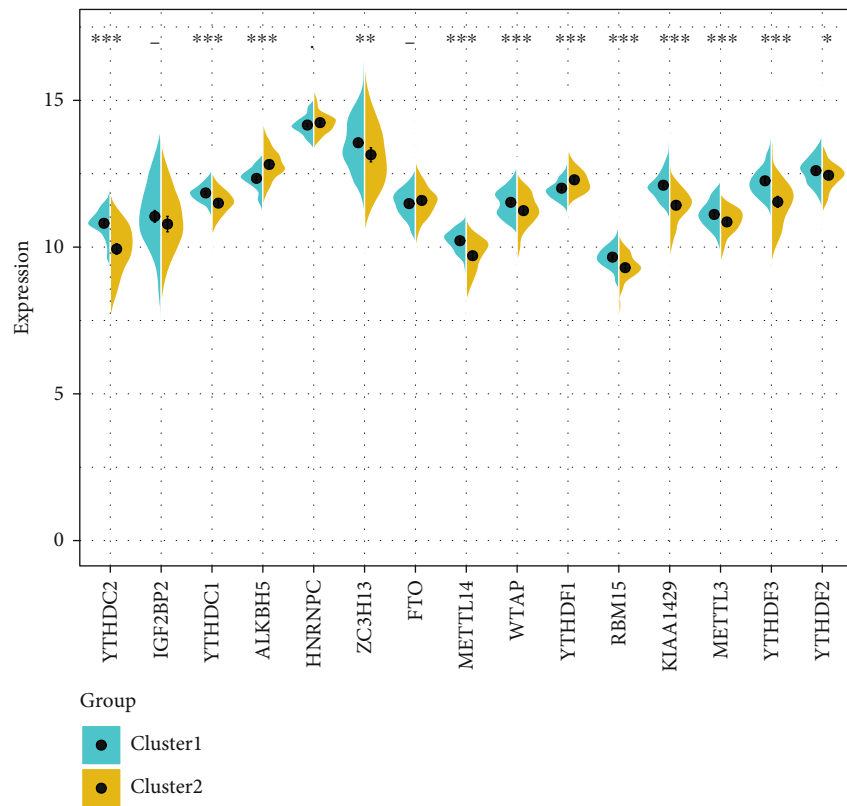


(b)

FIGURE 6: Continued.



(c)



(d)

FIGURE 6: Identification of m6A gene markers and molecular characteristics: (a) Wayne diagram of differential genes in m6AClusters; (b) heat map of differentially expressed genes in m6AClusters; (c) prognostic KM curve of Cluster1 and Cluster2; (d) the expression differences of 15 CNV-related m6A regulatory genes in Cluster1 and Cluster2.

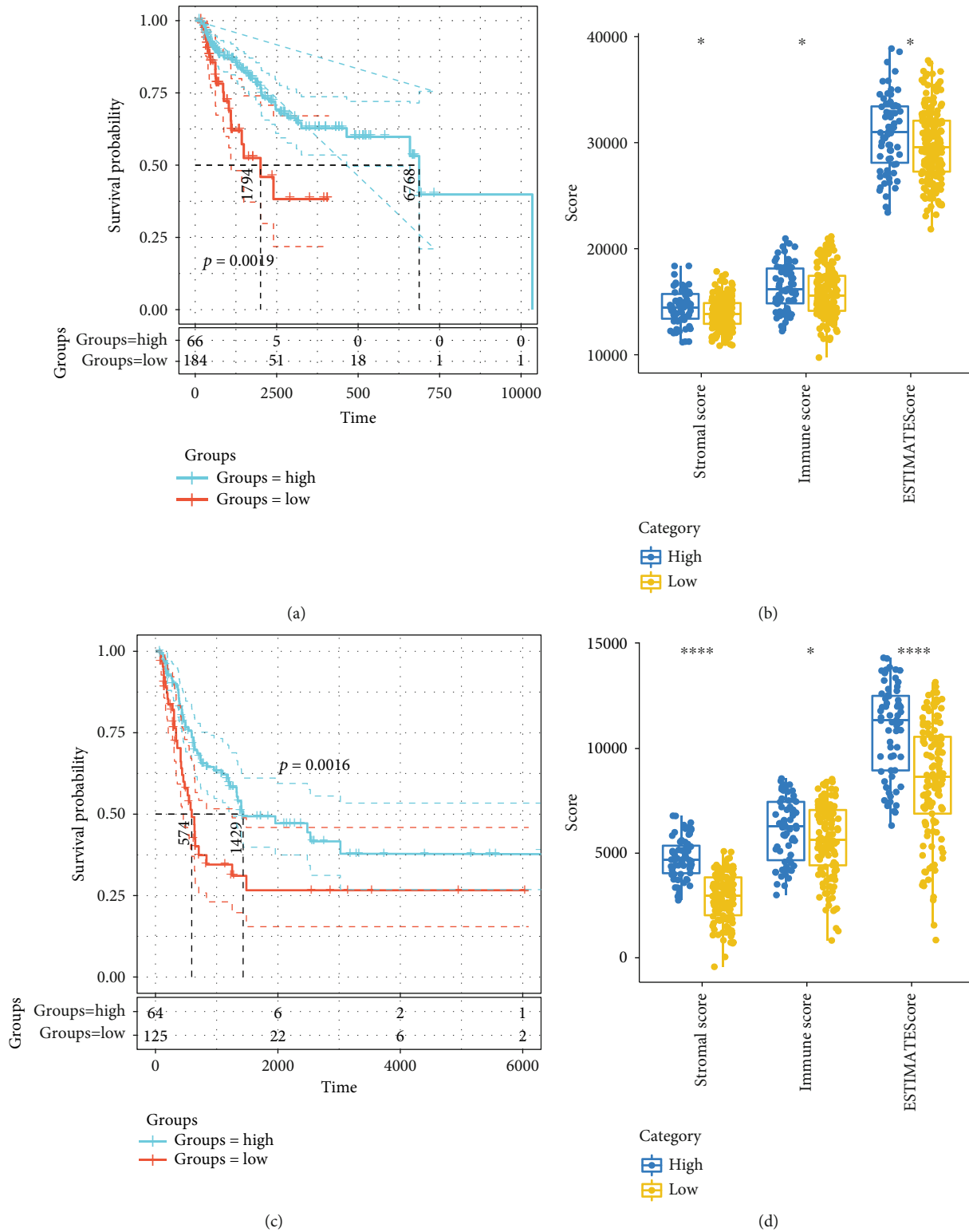
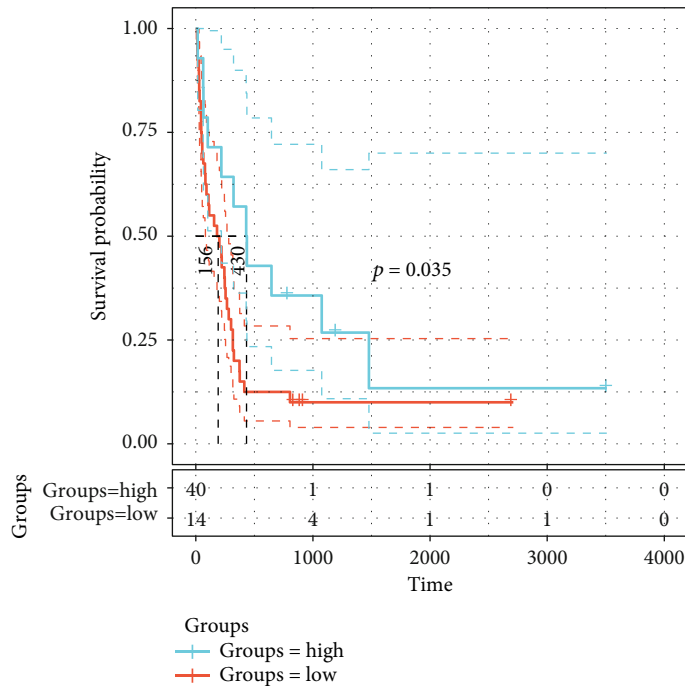
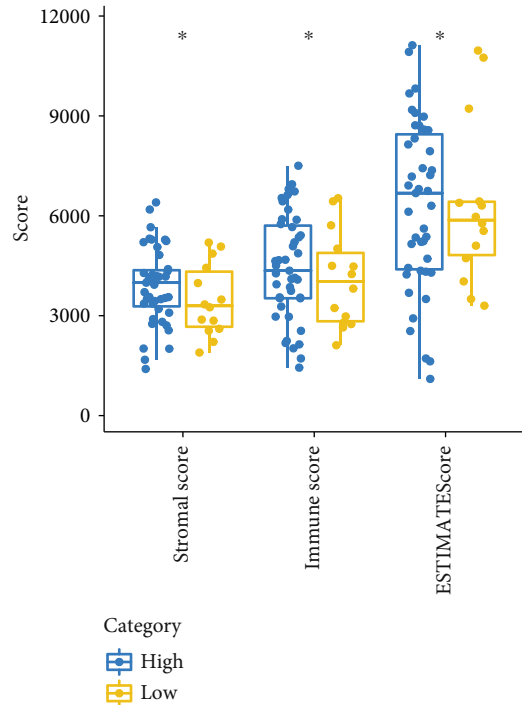


FIGURE 7: Continued.



(e)



(f)

FIGURE 7: Establishment and validation of risk scores based on m6A-related genetic markers: (a) KM curves for prognostic differences between high-risk and low-risk groups; (b) differences in immune microenvironment scores between high-risk and low-risk groups; (c) KM curves for prognostic differences between high-risk and low-risk groups in the GSE65904 dataset; (d) differences in immune microenvironment scores between high-risk and low-risk groups in the GSE65904 dataset; (e) KM curves for prognostic differences between high-risk and low-risk groups in the GSE22153 dataset; (f) differences in immune microenvironment scores between high-risk and low-risk groups in the GSE22153 dataset.

mRNA level of lymphoid enhancer binding factor 1 (LEF1), thereby inhibiting the activity of the Wnt/ β -catenin signaling pathway. Moreover, METTL3 promotes osteosarcoma cell progression by regulating the m6A level of LEF1 and activating the Wnt/ β -catenin signaling pathway [10]. Previously, upregulated expression of METTL3 has been observed in human hepatocellular carcinoma and multiple solid tumors [19]. Knockdown of METTL3 can reduce hepatocellular carcinoma cell proliferation, migration, and colony formation, and more interestingly, knockout of METTL3 remarkably suppresses tumorigenicity and lung metastasis. In colorectal cancer, METTL3 acts as a tumor suppressor on cell proliferation, migration, and invasion via the p38/ERK pathway [20]. In addition, reducing m6A methylation could activate oncogenic Wnt/PI3K-Akt signaling and promote malignant phenotypes of gastric cancer cells [21]; moreover, METTL3 knockdown could inhibit the level of total RNA methylation of m6A, cell proliferation, and migration of gastric cancer cells [22]. Noticeably, m6A methylation has also been reported in bladder cancer [23], renal cancer [24], acute myeloid leukemia [25], breast cancer [26], and hepatocellular cancer [27].

The regulators of m6A RNA methylation function critically in the malignant progression of glioma and may be use-

ful for the development of prognostic stratification and therapeutic strategies [28]. In this study, 15 m6A regulatory genes with abnormal copies of the genome were identified based on TCGA gene expression data and CNV data. Our analysis showed that a higher frequency of the CNV of m6A regulatory genes was related to a worse the prognosis of CM patients. Three clinical subgroups with different molecular characteristics were detected after performing unsupervised cluster analysis. Here, m6aCluster2 was associated with abnormal immune microenvironment and poor prognosis. A risk assessment model was established with the m6A-related gene markers, and subgroups with high immune scores and poor prognosis were identified in both sequencing data and chip data. Our results indicated that m6A regulatory genes play an important role in CM, as they could influence immune-related processes during tumor development. Zhou et al. [29] proved that a low expression of the writer gene METTL3 is related to activations of adipogenesis and mTOR pathways. However, m6A regulators in CM have not been reported before.

Some cancer-related pathways are dysregulated during CM development. In this study, We determined a molecular subgroup with favorable prognosis based on m6A regulatory genes and found that degradation pathways such as ubiquitin

pathways, mediated proteolysis, and RNA degradation were active in the subgroups, and interestingly, these pathways play critical roles in regulating cell growth and proliferation by controlling the abundance of key cyclins [30]. Evidences increasingly showed that abnormal proteolysis of cell cycle regulators significantly promotes tumorigenesis, while enhanced degradation of cell cycle regulators (i.e., proto-oncoproteins) can inhibit tumor metastasis. The driving force of the cell cycle is the activation of cyclin-dependent kinases (CDKs), whose activity is controlled by the proteolysis of ubiquitin-mediated key regulators (such as cyclins and CDK inhibitors), and disorder of proteolytic system may lead to uncontrolled proliferation, genomic instability, and cancer initiation [31]. Selective RNA degradation has been shown to be able to strongly resist disease development and is therefore believed to have the potential to treat cancer [32]. In antitumor therapy, activation of peroxisomal proliferator activator (PPAR) upregulates RNA degradation pathways, leading to reprogramming of tumor metabolism [33]. These results suggest that degradation-enhancing molecular subgroup identified by the m6A regulatory genes could be explored as a novel effective strategy for CM treatment.

5. Conclusion

In summary, this study first detected the genetic alterations of m6A regulatory genes in CM and identified molecular subsets with enhanced degradation and immune enhancement based on genomic mutagenesis of m6A regulatory genes. A m6A signature with a high predictive accuracy in predicting the prognosis of CM patients was constructed in this study. Our findings provide new insights into the epigenetic understanding of m6A in CM.

Data Availability

The datasets analyzed in this study are available from the GEO repository (<https://www.ncbi.nlm.nih.gov/geo/>).

Conflicts of Interest

The authors stated they have no conflicts of interest.

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] P. Gorayski, B. Burmeister, and M. Foote, "Radiotherapy for cutaneous melanoma: current and future applications," *Future Oncology*, vol. 11, no. 3, pp. 525–534, 2015.
- [3] A. Sample and Y. Y. He, "Mechanisms and prevention of UV-induced melanoma," *Photodermatology, Photoimmunology & Photomedicine*, vol. 34, no. 1, pp. 13–24, 2018.
- [4] G. C. Leonardi, L. Falzone, R. Salemi et al., "Cutaneous melanoma: from pathogenesis to therapy (review)," *International Journal of Oncology*, vol. 52, no. 4, pp. 1071–1080, 2018.
- [5] A. M. Eggermont, A. Spatz, and C. Robert, "Cutaneous melanoma," *The Lancet*, vol. 383, no. 9919, pp. 816–827, 2014.
- [6] S. N. Pavri, J. Clune, S. Ariyan, and D. Narayan, "Malignant melanoma: beyond the basics," *Plastic and Reconstructive Surgery*, vol. 138, no. 2, pp. 330e–340e, 2016.
- [7] Q. Lan, P. Y. Liu, J. Haase, J. L. Bell, S. Huttelmaier, and T. Liu, "The critical role of RNA m6A methylation in cancer," *Cancer Research*, vol. 79, no. 7, pp. 1285–1292, 2019.
- [8] X. Y. Chen, J. Zhang, and J. S. Zhu, "The role of m6A RNA methylation in human cancer," *Molecular Cancer*, vol. 18, no. 1, p. 103, 2019.
- [9] S. Yang, J. Wei, Y. H. Cui et al., "m(6)A mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade," *Nature Communications*, vol. 10, no. 1, article 2782, 2019.
- [10] W. Miao, J. Chen, L. Jia, J. Ma, and D. Song, "The m6A methyltransferase METTL3 promotes osteosarcoma progression by regulating the m6A level of LEF1," *Biochemical and Biophysical Research Communications*, vol. 516, no. 3, pp. 719–725, 2019.
- [11] T. Li, P. S. Hu, Z. Zuo et al., "METTL3 facilitates tumor progression via an m6A-IGF2BP2-dependent mechanism in colorectal carcinoma," *Molecular Cancer*, vol. 18, no. 1, p. 112, 2019.
- [12] C. O. Gigeck, D. Q. Calcagno, L. T. Rasmussen et al., "Genetic variants in gastric cancer: risks and clinical implications," *Experimental and Molecular Pathology*, vol. 103, no. 1, pp. 101–111, 2017.
- [13] J. N. Weinstein, The Cancer Genome Atlas Research Network, E. A. Collisson et al., "The Cancer Genome Atlas pan-cancer analysis project," *Nature Genetics*, vol. 45, no. 10, pp. 1113–1120, 2013.
- [14] K. Cibulskis, M. S. Lawrence, S. L. Carter et al., "Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples," *Nature Biotechnology*, vol. 31, no. 3, pp. 213–219, 2013.
- [15] R. Cabrita, M. Lauss, A. Sanna et al., "Tertiary lymphoid structures improve immunotherapy and survival in melanoma," *Nature*, vol. 577, no. 7791, pp. 561–565, 2020.
- [16] 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.
- [17] S. Hanzelmann, R. Castelo, and J. Guinney, "GSVA: gene set variation analysis for microarray and RNA-seq data," *BMC Bioinformatics*, vol. 14, no. 1, p. 7, 2013.
- [18] R. Akbani, K. C. Akdemir, B. A. Aksoy et al., "Genomic classification of cutaneous melanoma," *Cell*, vol. 161, no. 7, pp. 1681–1696, 2015.
- [19] 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.
- [20] R. Deng, Y. Cheng, S. Ye et al., "m(6)A methyltransferase METTL3 suppresses colorectal cancer proliferation and migration through p38/ERK pathways," *Oncotargets and Therapy*, vol. 12, pp. 4391–4402, 2019.
- [21] C. Zhang, M. Zhang, S. Ge et al., "Reduced m6A modification predicts malignant phenotypes and augmented Wnt/PI3K-

- Akt signaling in gastric cancer,” *Cancer Medicine*, vol. 8, no. 10, pp. 4766–4781, 2019.
- [22] T. Liu, S. Yang, J. Sui et al., “Dysregulated N6-methyladenosine methylation writer METTL3 contributes to the proliferation and migration of gastric cancer,” *Journal of Cellular Physiology*, vol. 235, no. 1, pp. 548–562, 2019.
- [23] J. Han, J. Z. Wang, X. Yang et al., “METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m6A-dependent manner,” *Molecular Cancer*, vol. 18, no. 1, p. 110, 2019.
- [24] D. Gong, J. Zhang, Y. Chen et al., “The m6A-suppressed P2RX6 activation promotes renal cancer cells migration and invasion through ATP-induced Ca²⁺ influx modulating ERK1/2 phosphorylation and MMP9 signaling pathway,” *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, p. 233, 2019.
- [25] Z. Ianniello, A. Paiardini, and A. Fatica, “N6-Methyladenosine (m6A): a promising new molecular target in acute myeloid leukemia,” *Frontiers in Oncology*, vol. 9, p. 251, 2019.
- [26] L. Wu, D. Wu, J. Ning, W. Liu, and D. Zhang, “Changes of N6-methyladenosine modulators promote breast cancer progression,” *BMC Cancer*, vol. 19, no. 1, p. 326, 2019.
- [27] X. Cheng, M. Li, X. Rao et al., “KIAA1429 regulates the migration and invasion of hepatocellular carcinoma by altering m6A modification of ID2 mRNA,” *Oncotargets and Therapy*, vol. 12, pp. 3421–3428, 2019.
- [28] R. C. Chai, F. Wu, Q. X. Wang et al., “m6A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas,” *Aging (Albany NY)*, vol. 11, no. 4, pp. 1204–1225, 2019.
- [29] J. Zhou, J. Wang, B. Hong et al., “Gene signatures and prognostic values of m6A regulators in clear cell renal cell carcinoma - a retrospective study using TCGA database,” *Aging (Albany NY)*, vol. 11, no. 6, pp. 1633–1647, 2019.
- [30] T. Bashir and M. Pagano, “Aberrant ubiquitin-mediated proteolysis of cell cycle regulatory proteins and oncogenesis,” *Advances in Cancer Research*, vol. 88, pp. 101–144, 2003.
- [31] K. I. Nakayama and K. Nakayama, “Ubiquitin ligases: cell-cycle control and cancer,” *Nature Reviews Cancer*, vol. 6, no. 5, pp. 369–381, 2006.
- [32] S. K. Dey and S. R. Jaffrey, “RIBOTACs: small molecules target RNA for degradation,” *Cell Chemical Biology*, vol. 26, no. 8, pp. 1047–1049, 2019.
- [33] N. W. Chang and Y. P. Huang, “The RNA degradation pathway is involved in PPAR α -modulated anti-oral tumorigenesis,” *Biomedicine*, vol. 9, no. 4, p. 27, 2019.

Research Article

Nomogram Personalizes and Visualizes the Overall Survival of Patients with Triple-Negative Breast Cancer Based on the Immune Genome

Peipei Wang, Yang Fu, Yueyun Chen, Qing Li, Ye Hong, Ting Liu, and Zhenyu Ding 

Department of Biotherapy, Cancer Center, West China Hospital, West China Medical School, State Key Laboratory of Biotherapy, Sichuan University, Chengdu, China

Correspondence should be addressed to Zhenyu Ding; dingzhenyu@scu.edu.cn

Peipei Wang and Yang Fu contributed equally to this work.

Received 21 July 2020; Revised 2 October 2020; Accepted 12 November 2020; Published 24 November 2020

Academic Editor: Shijia Zhu

Copyright © 2020 Peipei Wang 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. Triple-negative breast cancer (TNBC) is usually poorly differentiated, highly invasive, susceptible to distant metastasis, and less responsive to endocrine and targeted therapy. However, immunotherapy is a promising treatment for TNBC patients recently. **Methods.** The prognostic value of immune-related genes (IRGs) was explored by using RNA sequencing and microarray data of 123 and 107 TNBC patients from TCGA and GEO databases, respectively. **Results.** In TCGA database, GO and KEGG pathway analysis of 119 differential IRGs indicated that they actively participate in the interaction of cytokines and receptors. A nomogram model constructed by the prognosis-related CCL25, IL29, TDGF3, GPR44, and GREM2 in the IRGs could personalize and visualize the 1-, 2-, 3-, 4-, and 5-year overall survival (OS) of TNBC patients. Moreover, TNBC patients could be defined as low-risk (risk score < 194) or high-risk (risk score \geq 194) cohorts based on the risk score derived from the nomogram model. The results could be validated by the GSE58812 dataset. Furthermore, the risk score was an independent risk factor for TNBC patients (HR = 1.019, 95% CI 1.012-1.027, $p < 0.001$) and was positively related to stage ($p = 0.017$). Interestingly, the risk score could reflect the infiltration of B cells, CD4+ T cells, CD8+ T cells, dendritic cells, and neutrophils. **Conclusion.** These findings provided a reference for personalized OS prediction in TNBC patients and might be potential immune biomarkers for designing novel therapy.

1. Introduction

Breast cancer has high incidence in women. Different types of breast cancer have obvious differences in morphology, molecular pathological characteristics, clinical features, and responses to tumor treatment [1]. As a special subtype of breast cancer, triple-negative breast cancer (TNBC) lacks the expression of estrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor 2 (HER2), which severely limits the clinical usage of endocrine and targeted therapy.

TNBC is usually poorly differentiated, highly invasive, susceptible to distant metastasis, and less responsive to treat-

ment than other hormone receptor-positive breast cancers, so it has a higher risk of early relapse [2]. However, due to its unstable genome and high mutation rate, TNBC is highly immunogenic [3]. At present, some progress has been made in immunotherapy for TNBC. Immunotherapy stimulates the immune response of TNBC patients through active immunity, such as cancer vaccines, or passive immunity, such as adoptive T cell therapy, tumor-specific antibodies, and immune checkpoint inhibitors [4]. A study has fused TNBC cells with peripheral blood monocyte-derived dendritic cells (DCs) to generate DC vaccines, which stimulate the proliferation of T lymphocytes and enhance the cytotoxic effect on breast cancer cells [5]. Emerging immune

checkpoint inhibitors have received increasing attention in numerous TNBC study. TNBC patients with higher PD-L1 expression levels and more tumor-infiltrating lymphocytes (TILs) have higher immunogenicity, which plays a crucial part in regulating the immune response [6]. FDA has approved PD-L1 blockade combined with chemotherapy for patients with PD-L1-positive TNBC [7].

Increasing evidence suggests that immune gene expression and TIL may be prognostic for TNBC. Rody et al. [8] analyzed the RNA sequencing data of 579 TNBC patients and find that the expression of immune cell metagenes is closely related to prognosis. A study found that TNBC patients with higher TILs have better overall survival (OS) [9]. Yeong et al. [10] found that high PD-1+ cell infiltration significantly improved disease-free survival in TNBC patients. Similarly, PD-1, IFNG, and IFN signaling genes are positively correlated with the improvement of clinical outcomes of TNBC patients. These studies suggest that immune genes and TILs play an important role in TNBC. However, there are currently no immune-related genes that individually and visually predict OS and TILs for patients with TNBC.

In this study, we analyzed the RNA-seq and microarray data in The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases to comprehensively evaluate immune-related genes (IRGs) expression levels and predict prognosis and immune cell infiltration in TNBC patients. The cell functions involved in IRGs were also explored. These explorations are especially important for the individual assessment of the prognosis of TNBC patients and the discovery of targeted immunization methods.

2. Materials and Methods

2.1. Acquisition of Datasets and Patients' Information. UCSC Xena (<https://xenabrowser.net/datapages/>) [11] was used to download the RNA-seq data of breast cancer in the form of $\log_2(\text{norm_count} + 1)$ from TCGA database (<http://www.tcg.org/>), including 123 newly diagnosed PR-, ER-, and Her2-negative breast cancers and 13 normal tissues adjacent to tumor. At the same time, the clinical information of patients was obtained, involving age, gender, tumor invasion depth (T), lymph node metastasis (N), distant metastasis (M), TNM stage, survival time, and status. The 123 TNBC patients in TCGA database served as a training cohort, and the clinical characteristics are listed in Table S1. In addition, a microarray dataset and clinical characteristics from 107 TNBC cases obtained from the GSE58812 dataset [12] in the GEO database were used as a validation cohort.

2.2. Differential Analysis of Genes. The limma package in R (version 3.6.1) was used to select differential genes between 123 cases of primary TNBC and 13 normal tissues adjacent to tumor. A total of 20,530 genes were included into the differential analysis, setting the adjusted p value < 0.05 , \log_2 |fold change| > 1 , and expression level > 0.2 . And a total of 1076 genes with significant differences were finally determined. We focused on the role of IRGs in TNBC. The ImmPort database (<https://www.immport.org/>) [13] provides and

updates 2498 IRGs for cancer research, all of which have been identified as being involved in the biological processes of immunity. Therefore, we downloaded the list of IRGs from the ImmPort database. Finally, 4.8% of IRGs were identified as those (a total of 119) differentially expressed genes and were included in subsequent analysis. Functional enrichment analysis of KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology) was performed on 119 IRGs through the Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/>) to explore potential molecular mechanisms.

2.3. Construction and Validation of Nomogram Model. The foreign and rms packages in R (version 3.6.1) were applied for establishing a nomogram model to analyze the role of IRGs in TNBC patients, which was described in detail in our previous study [14]. First, according to the nomogram, a point is given for each patient's IRG expression level. Then, a total risk score is obtained by gathering the given points of all the IRGs of the patients, which can predict OS. The concordance index (C-index) of 1000-sample bootstrap and the receiver operating characteristic (ROC) curve were used to evaluate the prognostic prediction ability of the nomogram model, and the judgment criterion was Area Under the Curve (AUC) or C-index > 0.5 . Model performance was assessed through both the internal and external calibration curve of 1000-sample bootstrap.

2.4. The Clinical Value of Risk Score. The risk score was divided into high or low based on the median value. Risk distribution, survival status, and IRG expression distribution of high- and low-risk TNBC were plotted by the heat map package in R (version 3.6.1). Survival and survminer packages were used to draw survival curves. We also explored the predictive value of risk score in TNBC for immune cell infiltration. The TIMER database (<http://cistrome.org/TIMER/>) [15] analyzed the TILs of 32 cancers in TCGA. Therefore, we obtained TIL abundances in TNBC patients from the TIMER and evaluated the correlation between the risk score and TILs.

2.5. Statistical Analysis. All statistical analysis was conducted in R (version 3.6.1, <https://www.r-project.org/>). Uni- and multivariate Cox regression analyses were used to screen prognostic variables. A log-rank test was used to compare the difference between survival curves. Two sets of quantitative data were compared by the Wilcoxon test. The correlation between the two sets of quantitative data was expressed by Spearman coefficient. A two-tailed p value < 0.05 was considered statistically significant.

3. Results

3.1. Identification of Differential IRGs. Differential analysis found that a total of 1076 genes were differentially expressed in TNBC, including 323 high expression and 753 low expression genes (Figure 1). Among them, a total of 119 differential IRGs, including 36 high expression and 83 low expression IRGs, were found. Their positions on the chromosome are shown in Figure 2. The KEGG showed that the enrichment

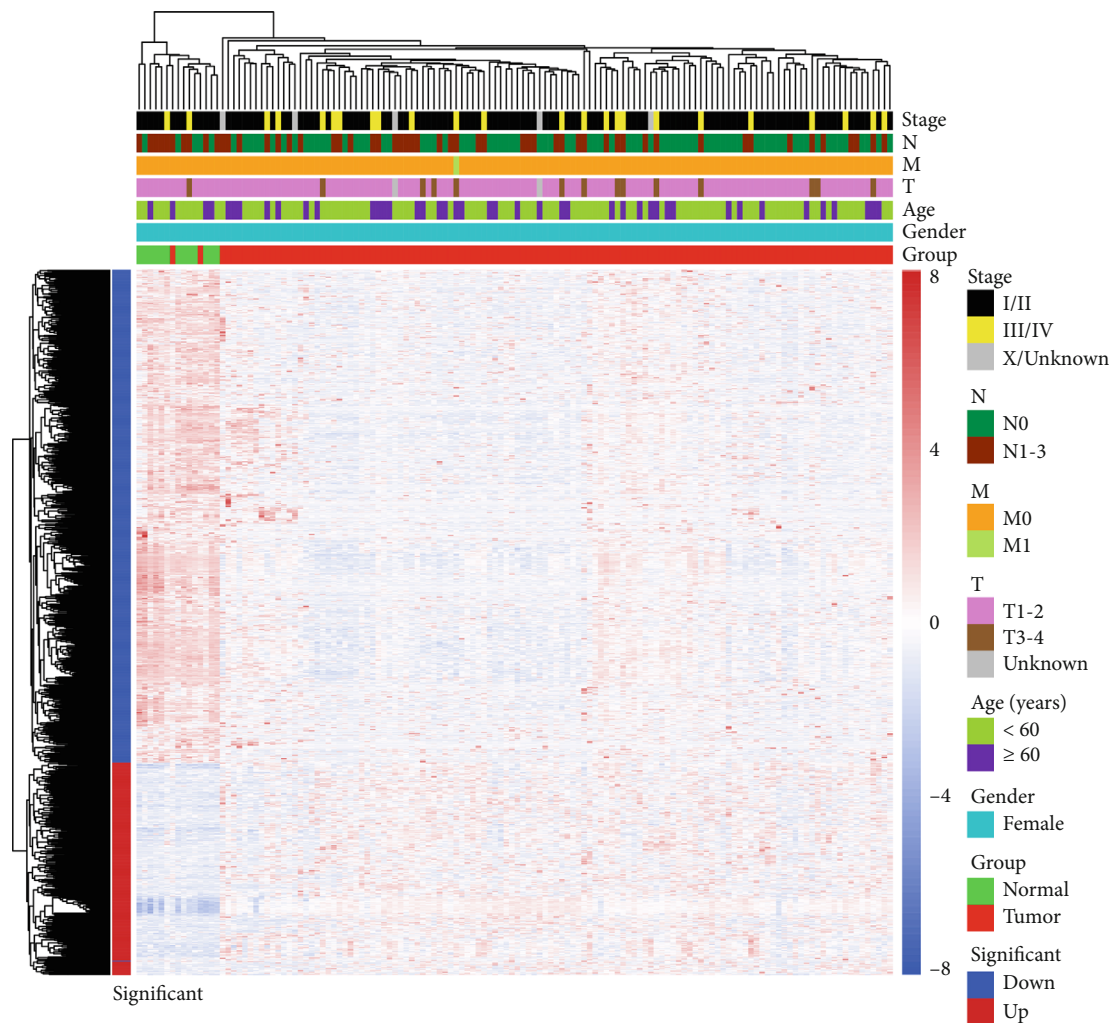


FIGURE 1: The heat map shows the differential genes between 123 cases of primary TNBC and 13 normal tissues adjacent to tumor. The color scale from green to red represented the difference in gene expression from low to high.

of differentially expressed IRGs was primarily in the “neuroactive ligand-receptor interaction” and “cytokine-cytokine receptor interaction” (Figure 3(a)). Biological processes, cellular components, and molecular functions are primarily enriched in “cell-cell signaling,” “extracellular region,” and “growth factor activity,” respectively (Figures 3(b)–3(d)). These findings suggested that cytokines and receptor pathways were most frequently implicated.

3.2. Identification of Prognosis-Related IRGs. Because prognostic molecular biomarkers are important for guiding treatment and disease monitoring, we focus on the impact of IRGs on the OS in TNBC. Prognostic analysis revealed that a total of 6 IRGs had significant impact on the OS of TNBC patients. Among them, high expression of C-C motif chemokine ligand 25 (CCL25), interleukin 29 (IL29), teratocarcinoma-derived growth factor 3 (TDGF3), and killer cell immunoglobulin like receptor, two Ig domains and long cytoplasmic tail 4 (KIR2DL4) predicted a favorable OS in TNBC patients (Figures 4(a), 4(d), 4(e), and 4(f)). Conversely, high expression of G protein-coupled receptor 44 (GPR44) and gremlin

2, DAN family BMP antagonist (GREM2) predicted a poor OS in TNBC patients (Figures 4(b) and 4(c)).

3.3. Establishment of Nomogram Model. CCL25, IL29, TDGF3, KIR2DL4, GPR44, and GREM2 obtained from the analysis of TCGA database were included into the establishment of a nomogram model. According to the expression levels of these IRGs, we got the total risk score of each individual, which could predict 1-, 2-, 3-, 4-, and 5-year OS (Figure S1). However, we could observe from Table S2 that the upregulated KIR2DL4 accounted for only 5 risk scores in the model. Compared with the risk scores of the other five IRGs, KIR2DL4 had a lower contribution to predicting the OS of TNBC patients, so we excluded KIR2DL4 in the establishment of the nomogram model. Finally, five prognosis-related IRGs including CCL25, IL29, TDGF3, GPR44, and GREM2 were included into the establishment of the nomogram model (Figure 5). To assess the predictive effect of the nomogram model on the OS of TNBC patients, we used the AUC of the ROC curve and the C-index of 1000-sample bootstrap for evaluation. When KIR2DL4 was

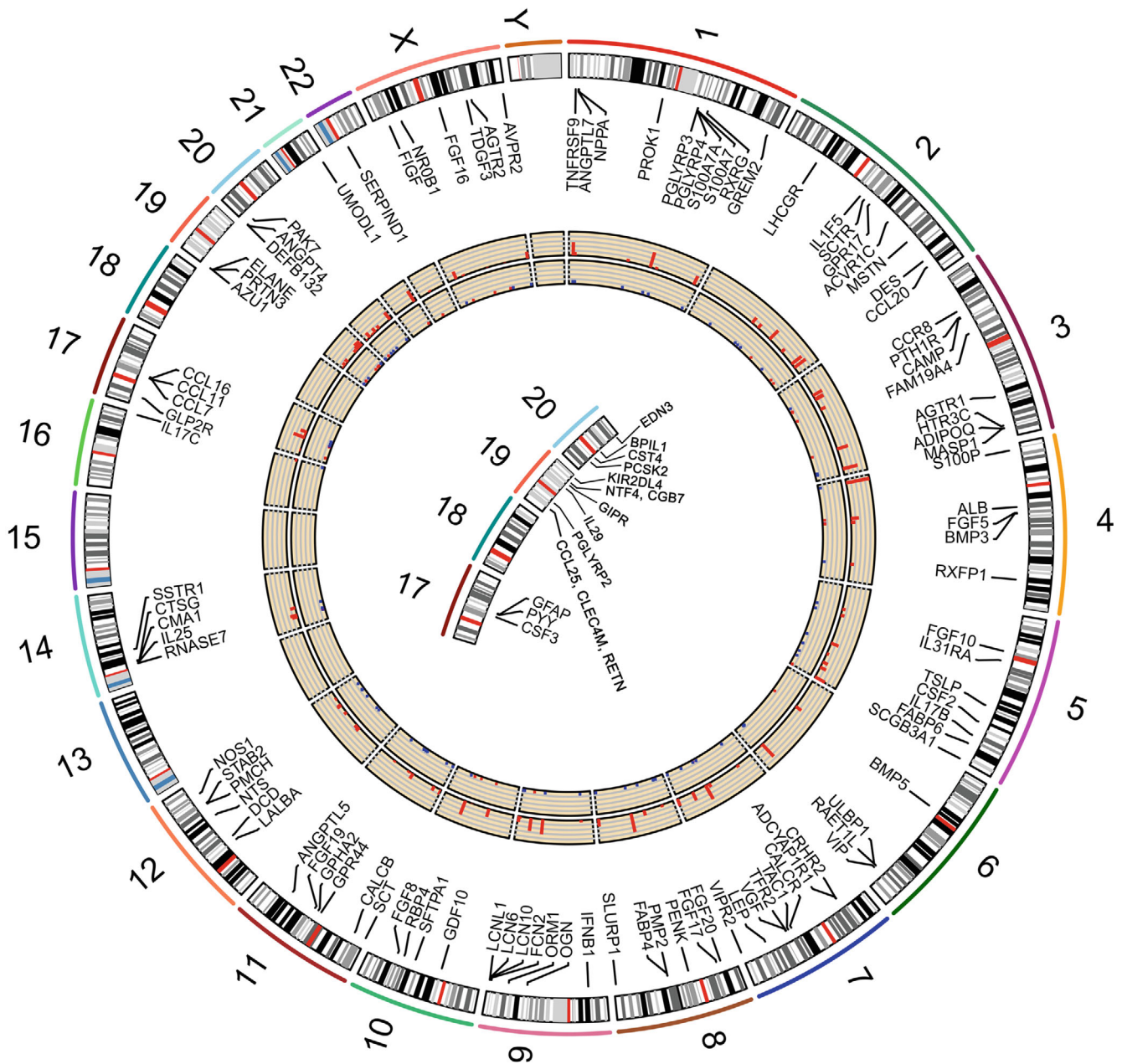
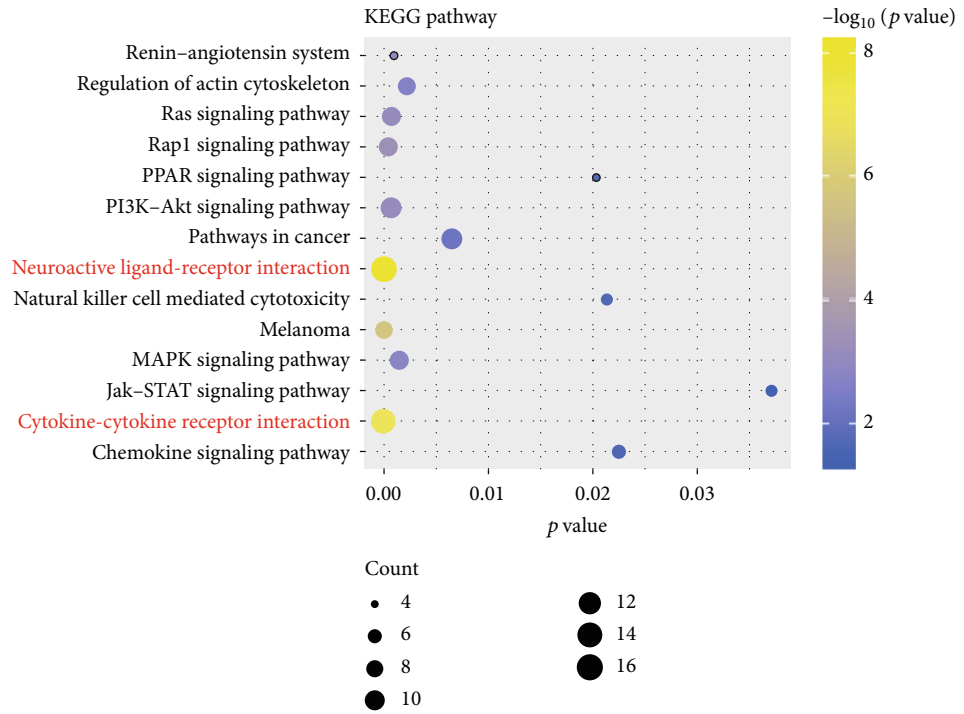


FIGURE 2: Differentially immune-related genes (IRGs) and their chromosomal locations. The number in the outermost circle is the name of the chromosome. The second circle is the positions of the genes in the chromosome. The black and white bars represent the chromosome bands, and the red bars represent the centromeres. The third circle is the gene names. The fourth circle is the average expression levels of the genes, and the bars from low to high represent the gene expression from low to high. Blue represents fold change < -2 , while red represents fold change > 2 . In the center of the circle diagram is the positions of genes that cannot be fully displayed on the second circle.

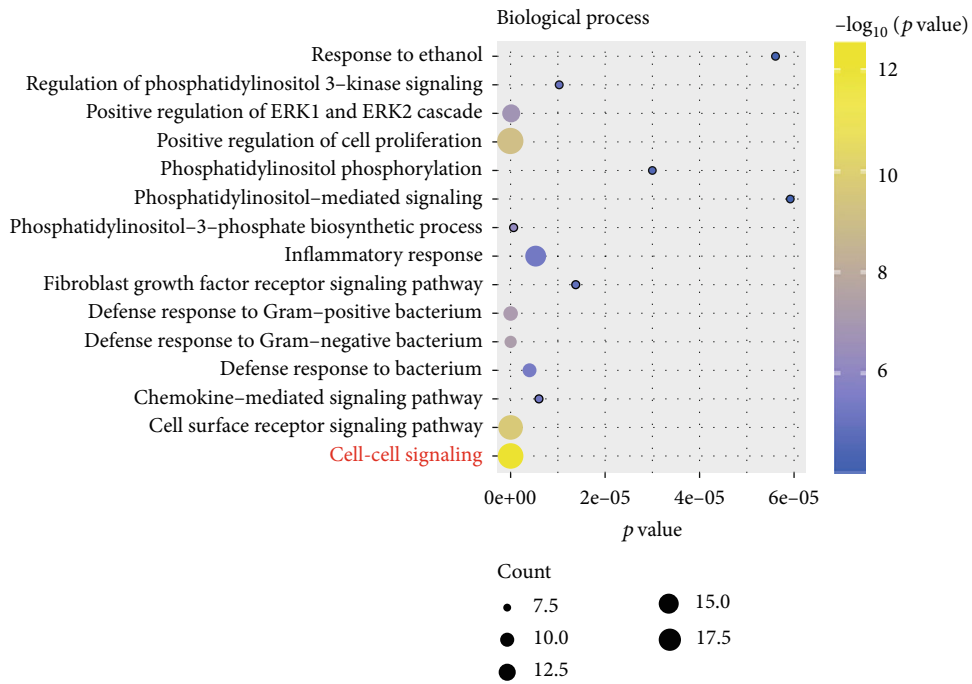
not removed, the AUC was 0.839 and C-index was 0.878 (Figure S2). When KIR2DL4 was excluded, AUC was 0.852 and C-index was 0.879 (Figure 6(a)). Further external validation of the nomogram model with the GSE58812 dataset in the GEO database showed that the AUC of the ROC curve was 0.619 and the C-index is 0.615 (Figure 6(b)). We also used the calibration curves to further validate the nomogram model. The calibration curves for the training group showed a good consistency between the predicted and actual 1-, 2-, 3-, 4-, and 5-year OS of the

nomogram model (Figure 7(a)). At the same time, the calibration curves in the validation group were also well identified (Figure 7(b)). Through internal and external verification, it was proved that the nomogram model in this study could conduct relatively accurate prediction of the OS of TNBC patients.

3.4. Risk Stratification. In order to analyze the clinical value of the risk score, we compare the risk scores of different clinical information and found that compared with stage I/II



(a)



(b)

FIGURE 3: Continued.

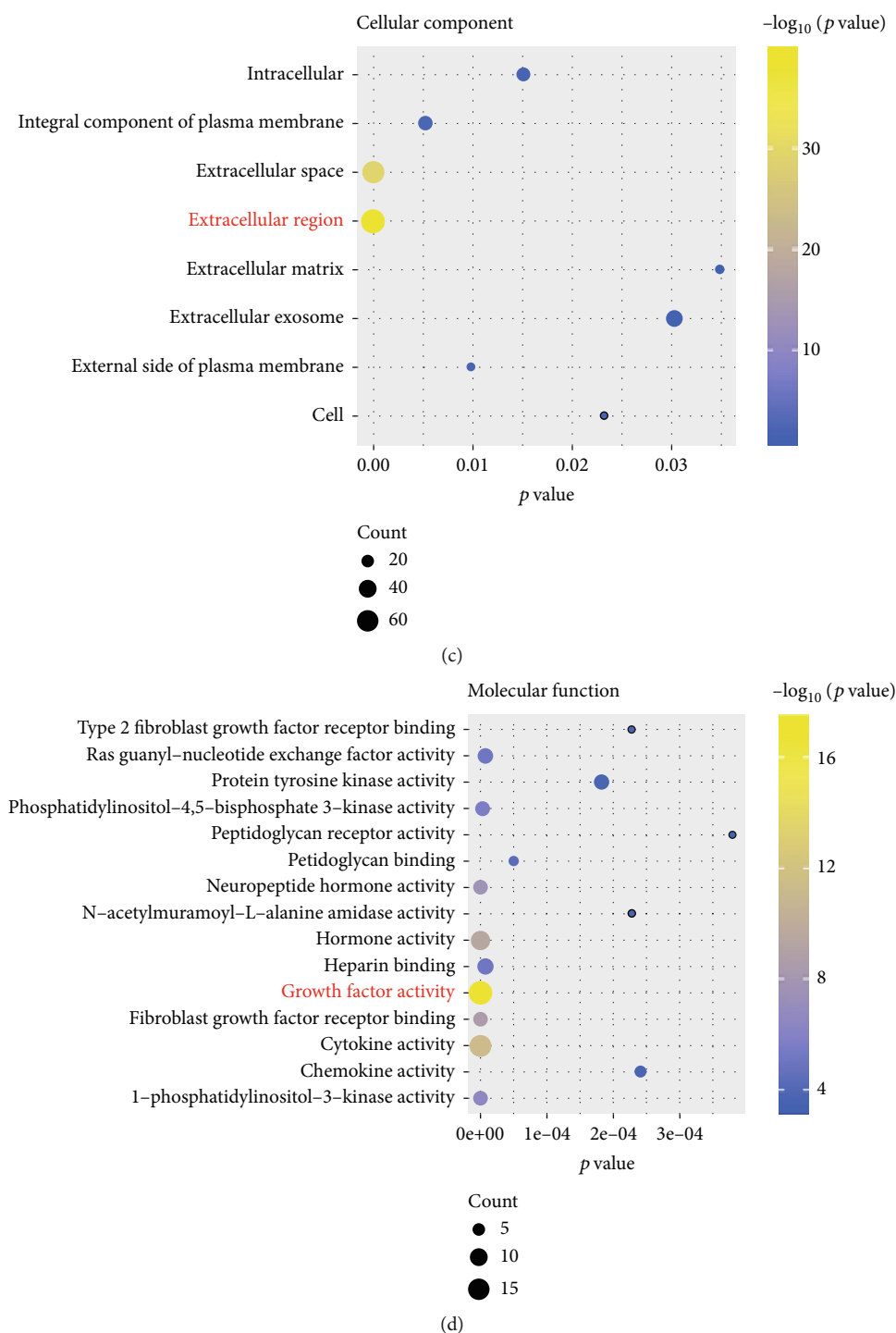
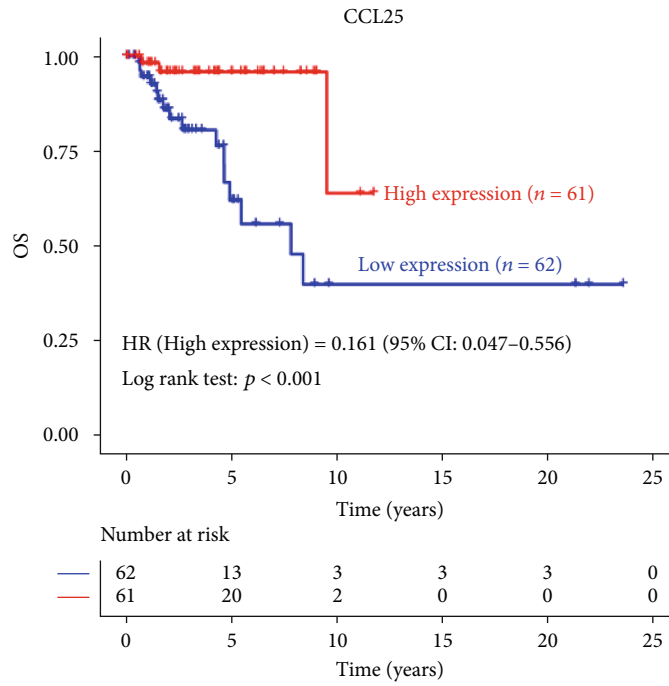


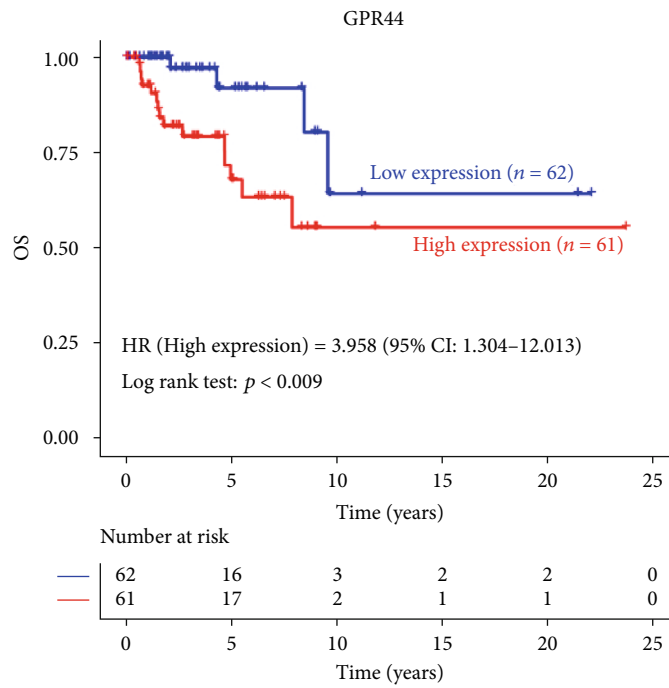
FIGURE 3: Functional enrichment of differential IRGs, and the top 15 of items are displayed. (a) KEGG pathway analysis; (b) biological process; (c) cellular component; (d) molecular function. The bubbles in the figures from small to large indicate that the number of enriched genes is from small to large. The color scale from blue to red indicates that the p value is from large to small.

TNBC patients with stage III/IV had higher risk scores ($p = 0.017$), and there was no significant difference between different ages, T, N, and M ($p > 0.05$) (Figure 8). After uni- and multivariate Cox regression analyses, we found that the risk score was an independent risk factor (HR = 1.019, 95% CI 1.012-1.027, $p < 0.001$, Table 1). The risk score was highly

correlated with TNM stage, indicating that risk score could be used to construct risk stratification. TNBC patients can be defined as low-risk (risk score < 194) or high-risk (risk score ≥ 194) cohort based on the risk score (Figures 9(a)–9(d), left panel). As shown in Figure 9(d), left panel, the survival curve of low- and high-risk TNBC patients

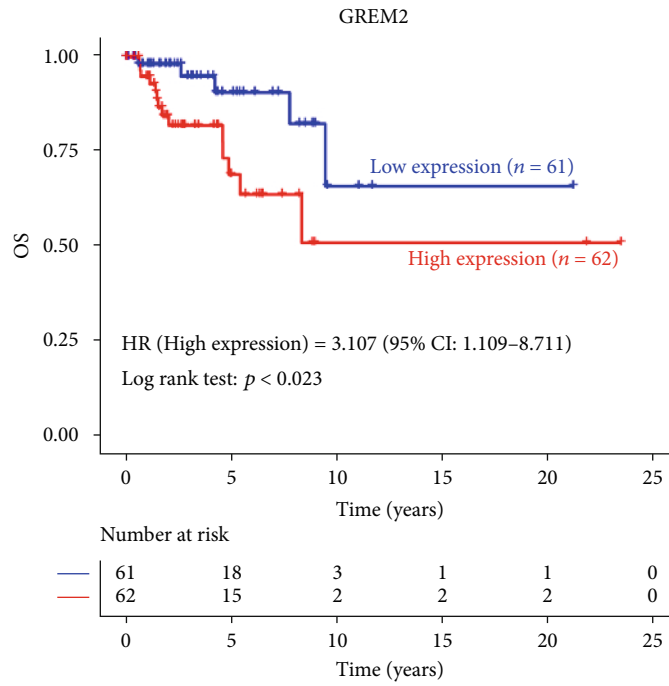


(a)

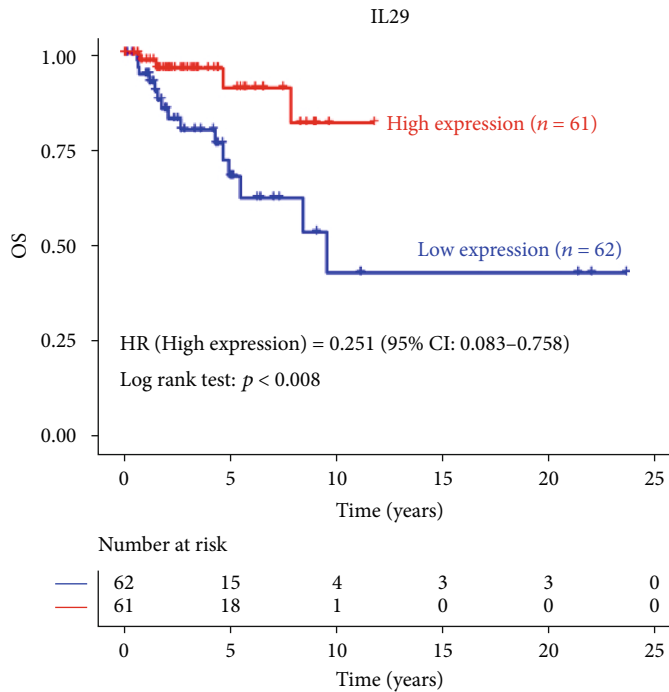


(b)

FIGURE 4: Continued.

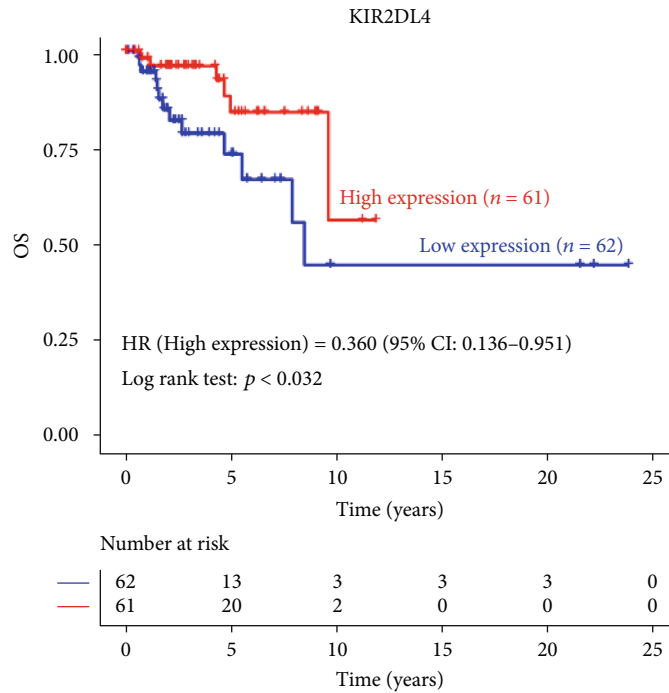


(c)

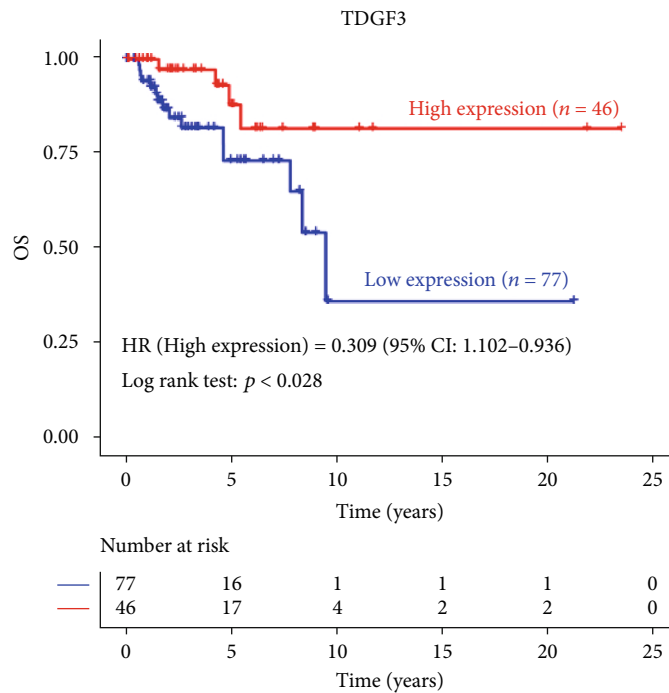


(d)

FIGURE 4: Continued.



(e)



(f)

FIGURE 4: Survival curve analysis of prognostic IRGs. Survival analysis based on the expression levels of 119 IRGs shows that 6 genes are closely related to overall survival (OS).

showed significant differences ($p < 0.001$). This finding could be validated by the external dataset GSE58812 ($p = 0.032$, Figures 9(a)–9(d), right panel).

3.5. Correlation between Risk Score and TILs. In order to investigate whether the risk score reflected by immune genes could accurately assess the state of TNBC immune microen-

vironment, we performed the correlation analysis between risk score and TILs (Figure 10) and found that risk score was negatively related to B cells ($R = -0.26$, $p = 0.005$), CD4 + T cells ($R = -0.21$, $p = 0.019$), CD8+ T cells ($R = -0.19$, $p = 0.034$), dendritic cells (DCs) ($R = -0.25$, $p = 0.005$), and neutrophils ($R = -0.27$, $p = 0.002$). However, the risk score had no significant correlation with macrophages ($p = 0.3$).

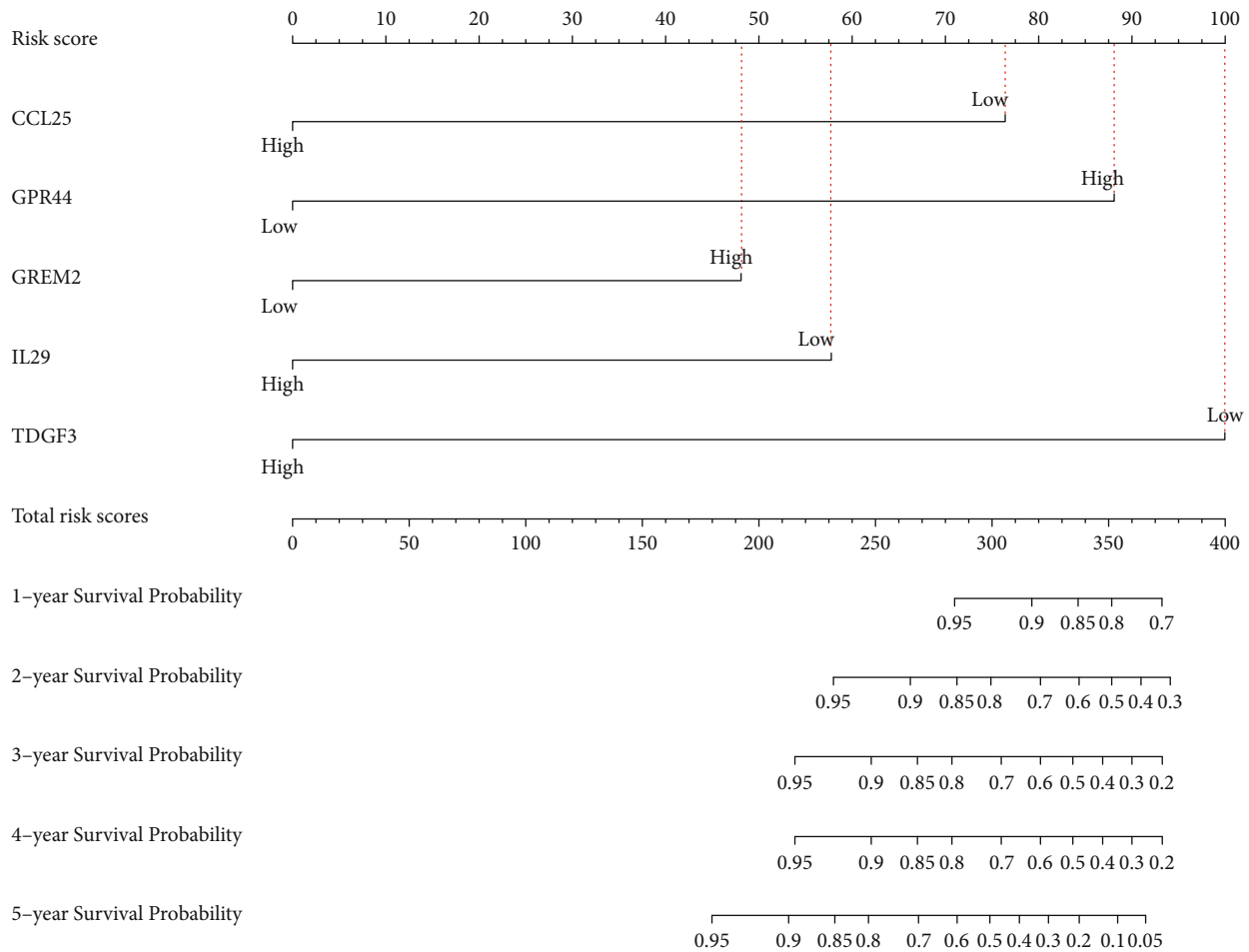


FIGURE 5: Construction of nomogram model with prognostic IRGs for predicting 1-, 2-, 3-, 4-, and 5-year OS in TNBC patients. A higher risk score meant a lower survival probability.

4. Discussion

The clonal proliferation and metastasis characteristics of cancer cells depend on genome changes. We focus our research on changes in the immune genome to reveal its relationship with the immune microenvironment. In this study, we extracted differential IRGs from two large TNBC cohorts in the GEO and TCGA databases and analyzed the underlying immune mechanism. Pathway and GO analysis found that TNBC patients primarily function through interactions between cytokines and receptors. Various studies have showed that cytokines and receptors are involved in the growth, invasion, and metastasis of breast cancer, and corresponding immune inhibitors against cytokines and receptors have been applied to the treatment of breast cancer [16, 17]. In addition, fibroblast growth factor receptors (FGFRs) are highly expressed in TNBC patients, and inhibitors against FGFRs have been tested in clinical trials [18]. Bioinformatics analysis provided clues that changes in the immune system of TNBC patients were initiated through cytokine and receptor pathways, which laid the foundation for in-depth immune-related research.

The nomogram model has been widely applied to systematically assess the outcome of cancer patients [14, 19]. At the same time, IRGs can provide personalized immune signature to assess the prognosis of lung cancer patients [20]. The prognosis-related CCL25, IL29, TDGF3, GPR44, and GREM2 in IRGs were used to construct a nomogram model to evaluate its clinical value in TNBC patients. The nomogram model we constructed can individualize and visualize 1-, 2-, 3-, 4-, and 5-year OS for TNBC patients. Evidence suggests that blocking the CCR9/CCL25 axis can promote tumor progression and distant metastasis [21, 22]. IL29 appears to inhibit tumor growth in a variety of cancers [23]. TDGF3, also known as TDGF1P3 or CRIPTO3, is expressed in cancer [24]. Our study found that high expression of CCL25, IL29, and TDGF3 predicted a good prognosis in TNBC patients. Study has shown that GPR44 is involved in the release of cytokines from immune cells and the development of inflammation [25]. Silencing GREM2 can inhibit the JNK signaling pathway in gastric cancer, which inhibits tumor growth [26]. In this study, high expressions of GPR44 and GREM2 could predict adverse outcomes in TNBC

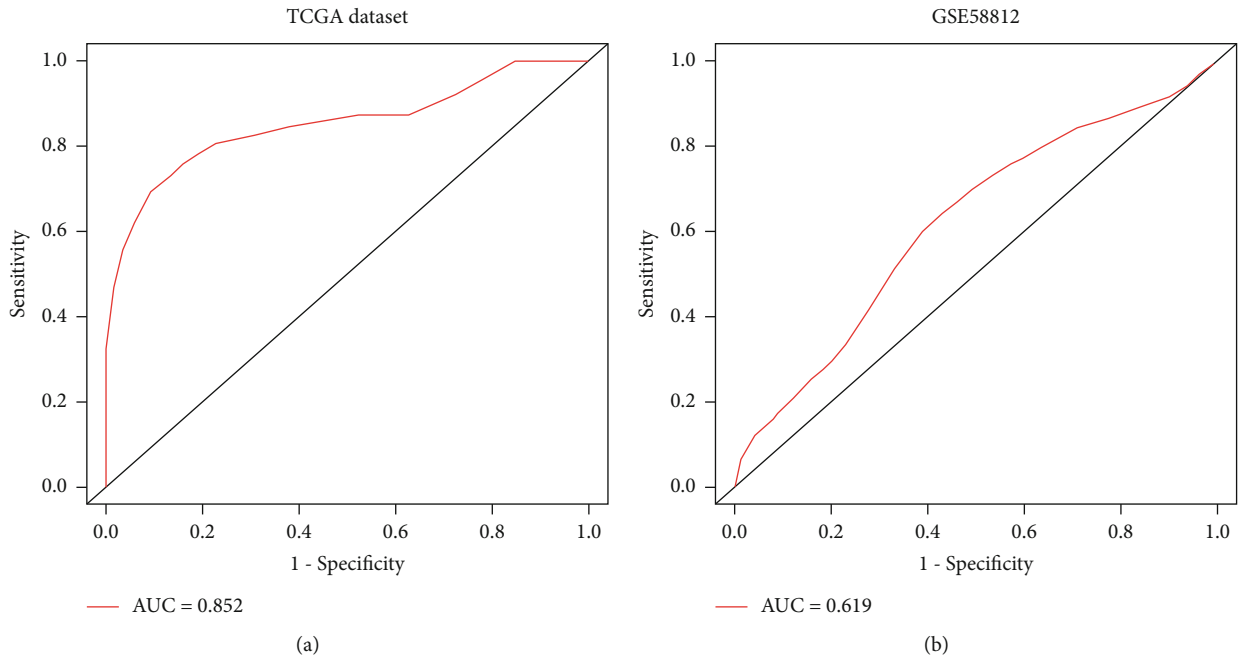


FIGURE 6: The ROC curves for the nomogram models. (a) ROC curve verified the nomogram model constructed by CCL25, IL29, TDGF3, GPR44, and GREM2. (b) ROC curve verified the nomogram model constructed by prognostic IRGs (CCL25, GPR44, GREM2, IL29, and TDGF3) in the dataset GSE58812 of the GEO database.

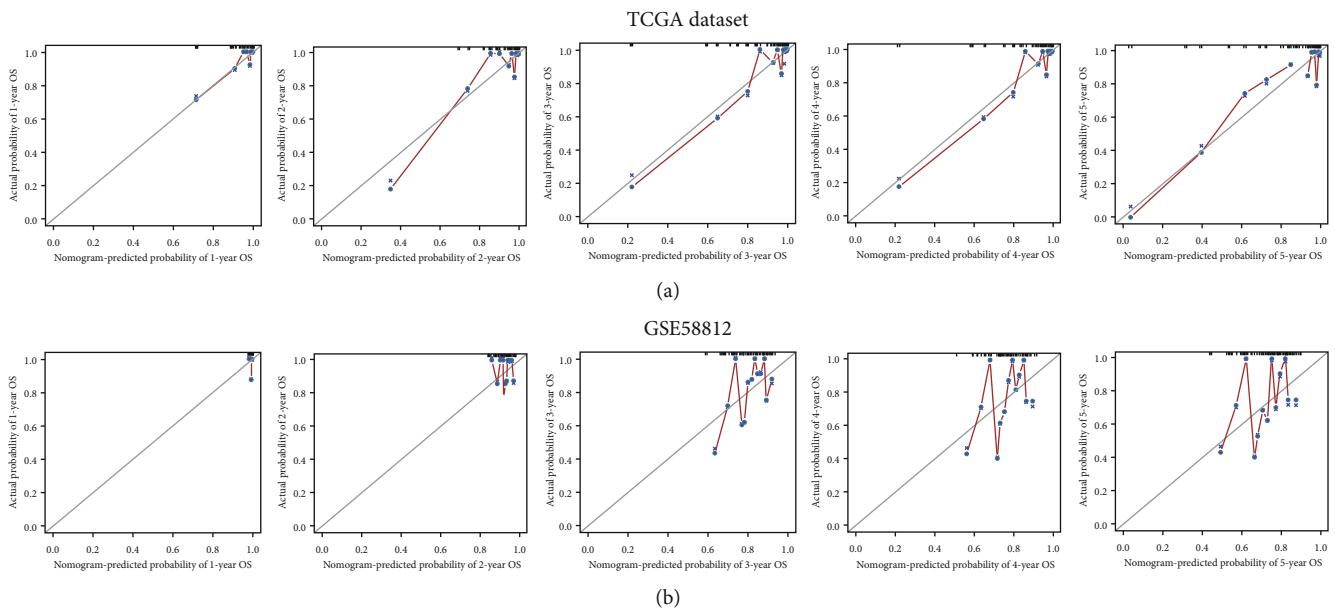


FIGURE 7: Internal and external calibration curve validation of the nomogram model. (a) Calibration curves of internal validation in TCGA database. (b) Calibration curves of external validation in the GSE58812 dataset.

patients. The establishment of risk stratification can provide a reference for clinicians to more rationally manage cancer patients and individualized treatment plans [27]. We conducted risk stratification for TNBC patients based on the risk score and found that patients with a risk score ≥ 194 belonged to a high-risk cohort, while patients with a risk score < 194 belonged to a low-risk cohort.

The risk score showed good clinical practicability. The risk score was an independent risk factor for TNBC patients and was positively related to TNM stage. In addition, the risk score was negatively correlated with B cells, CD4+ T cells, CD8+ T cells, dendritic cells, and neutrophils, which could reflect the level of TILs and provide a reference for clinicians to adjust the treatment plan.

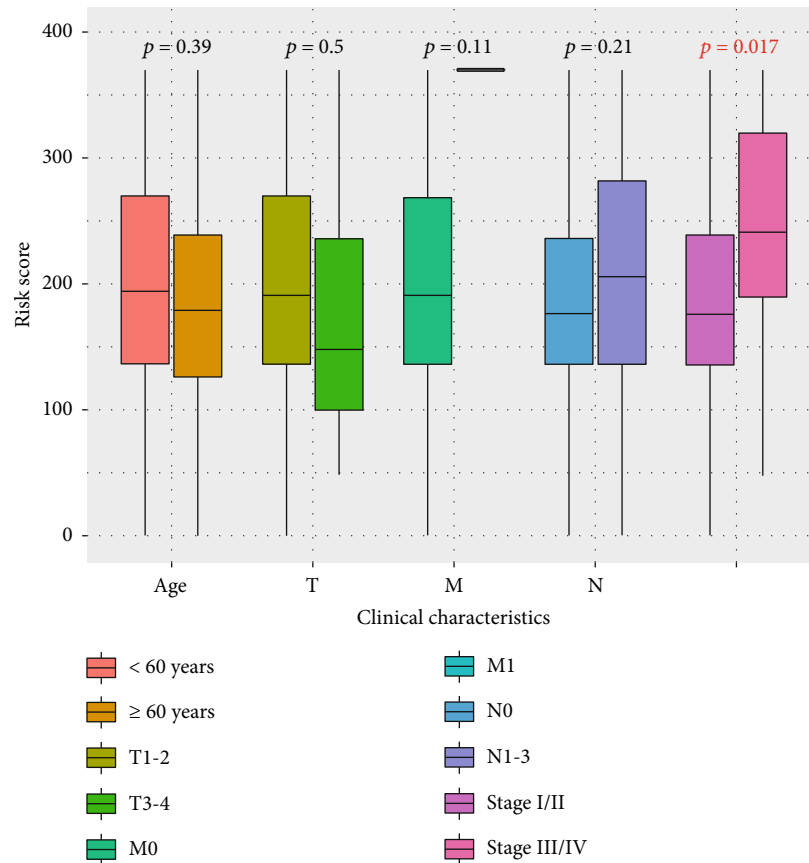


FIGURE 8: The distribution of risk scores in TNBC patients among different age, tumor invasion depth (T), distant metastasis (M), lymph node metastasis (N), and TNM stage.

TABLE 1: Uni- and multivariate Cox regression analyses of 123 TNBC patients in TCGA database.

| Variables | Univariate Cox | | Multivariate Cox | |
|-----------------------|------------------------|-----------|----------------------|-----------|
| | HR (95% CI) | p value | HR (95% CI) | p value |
| Risk score | 1.022 (1.014-1.031) | <0.001*** | 1.019 (1.012-1.027) | <0.001*** |
| Age (years) | 1.004 (0.969-1.040) | 0.842 | | |
| Tumor invasion depth | | | | |
| T1-2 | Reference | | | |
| T3-4 | 2.986 (0.976-9.140) | 0.055 | | |
| Distant metastasis | | | | |
| M0 | Reference | | Reference | |
| M1 | 54.325 (4.926-599.140) | 0.001** | 3.099 (0.272-35.283) | 0.362 |
| Lymph node metastasis | | | | |
| N0 | Reference | | Reference | |
| N1-3 | 3.785 (1.437-9.967) | 0.007** | 1.778 (0.483-6.548) | 0.387 |
| Stage | | | | |
| I/II | Reference | | Reference | |
| III/IV | 5.441 (2.113-14.012) | <0.001*** | 4.273 (1.043-17.515) | 0.044* |

HR: hazard ratio; CI: confidence interval; *p < 0.05; **p < 0.01; ***p < 0.001.

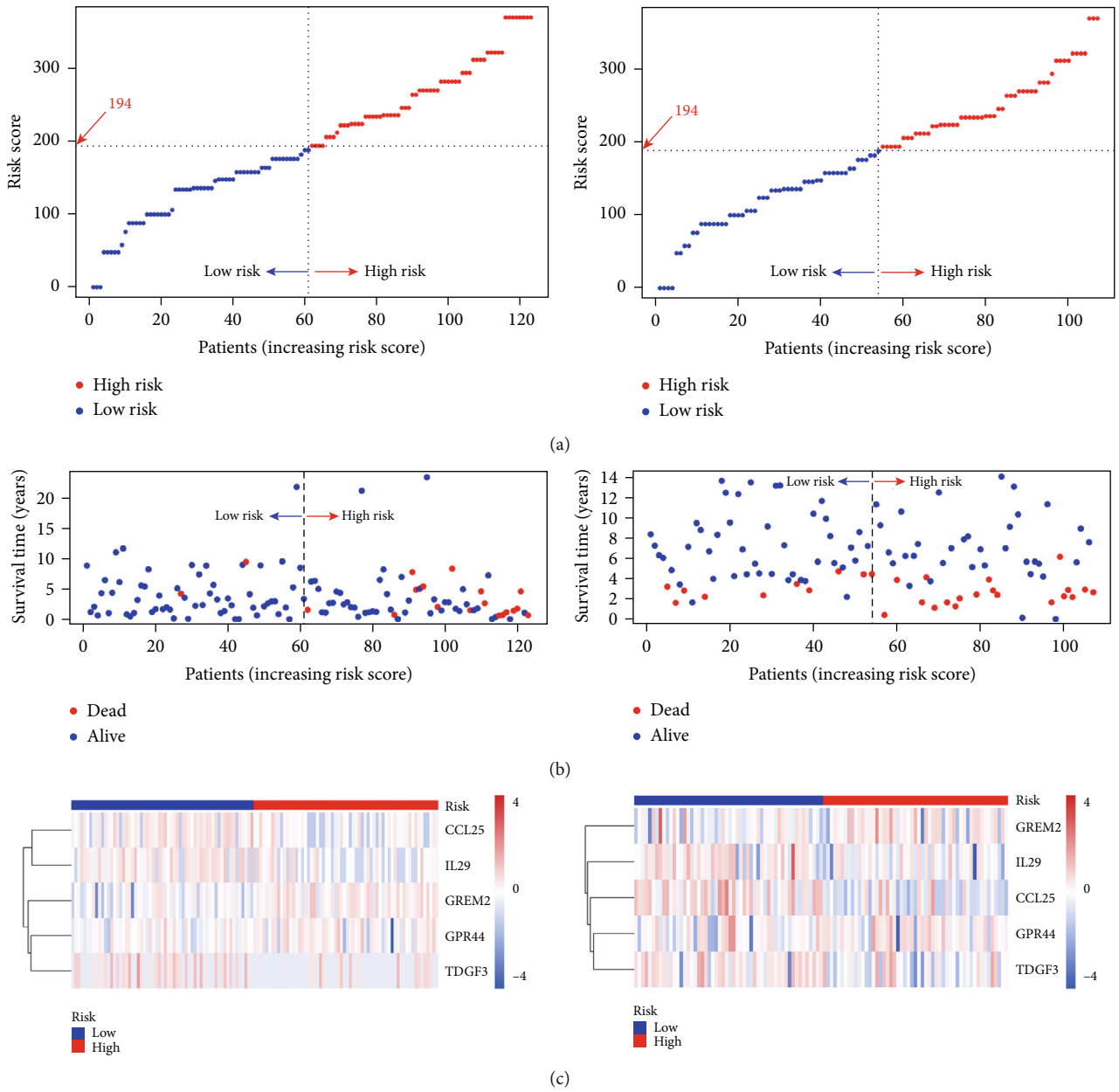
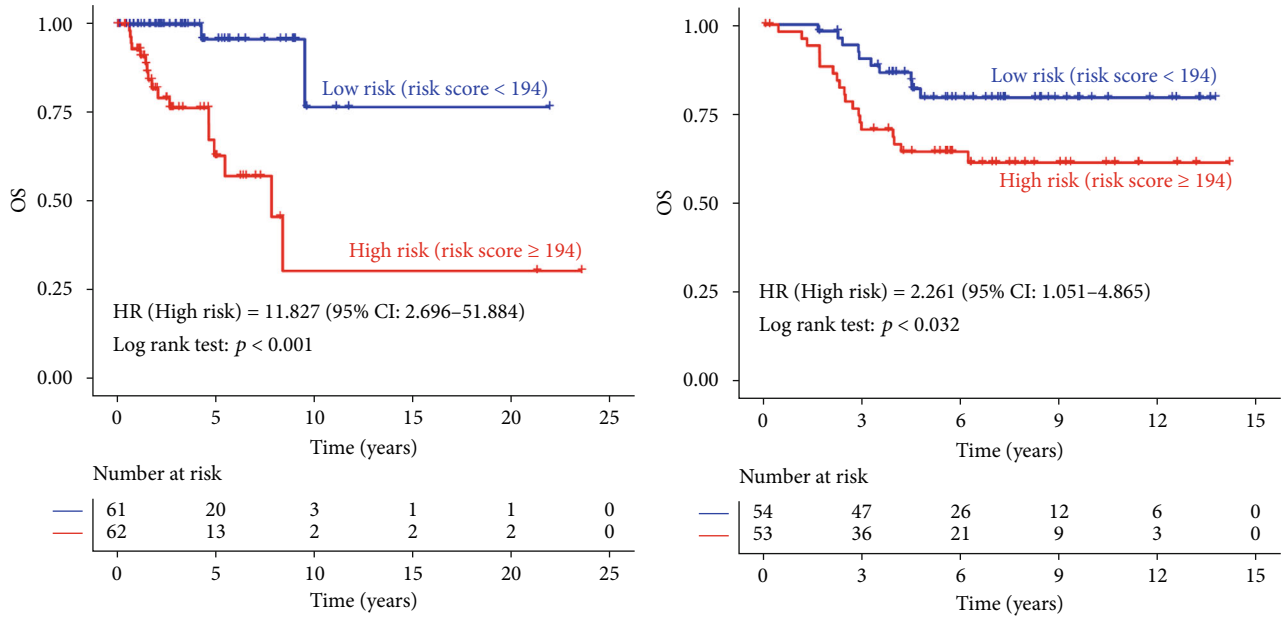


FIGURE 9: Continued.



(d)

FIGURE 9: Risk stratification of TNBC patients by risk score. (a) Risk score, (b) survival time, and (c) gene expression distribution of low- and high-risk groups in TCGA dataset (left panel) and GSE58812 (right panel). (d) Plotting survival curves according to low- and high-risk scores in TCGA dataset (left panel) and GSE58812 (right panel).

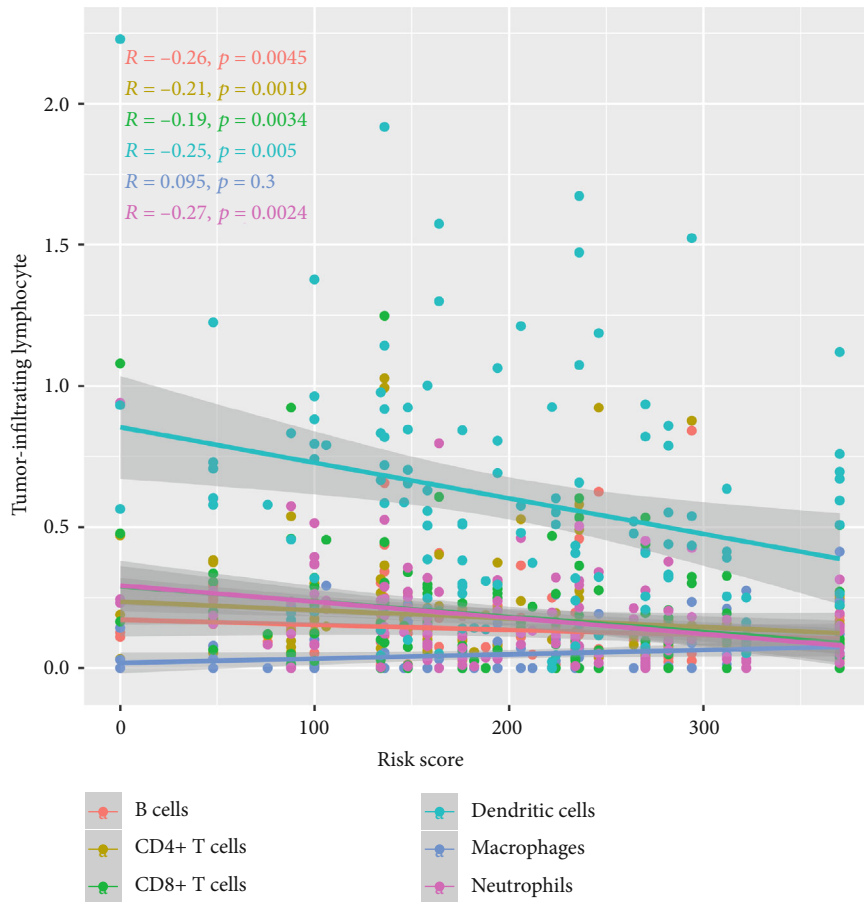


FIGURE 10: Spearman correlation between risk score and tumor-infiltrating lymphocytes (TILs).

Studies indicate that a favorable prognosis is observed in TNBC patients with high B cells and CD4+ T cells [28, 29]. Furthermore, it has been reported that activated CD8+ T cells have been shown to kill cancer cells through various mechanisms [30]. Not surprisingly, less infiltration of CD8+ T cells into tumors in TNBC patients is often related to worse outcomes [31]. In addition, a study suggests that neutrophil infiltration is a favorable prognostic factor for colon cancer [32], which is consistent with our findings. As we know that DCs take a vital part in the tumor microenvironment, infiltration of activated DCs into tumors can improve the antitumor effect of immune cells [33]. In this study, high-risk patients might have the lower infiltration levels of B cells, CD4+ T cells, CD8 + T cells, dendritic cells, and neutrophils, which was associated with a poor OS.

The limitation of this study was that no clinical samples and corresponding clinical information were used to validate the nomogram model and risk stratification constructed by CCL25, IL29, TDGF3, GPR44, and GREM2. Additionally, the reliability of the results was still challenged because we also lacked validation *in vitro* and *in vivo*.

5. Conclusion

A nomogram model constructed by CCL25, IL29, TDGF3, GPR44, and GREM2 could predict the 1-, 2-, 3-, 4-, and 5-year OS of TNBC patients and perform risk stratification. The risk score derived from the nomogram model could also predict the level of immune cell infiltration in tumors. These findings provided a reference for personalized prognosis prediction in TNBC patients and might be potential immune biomarkers for designing novel therapy.

Data Availability

The data used and analyzed during the current study are available from UCSC Xena platform (<https://xenabrowser.net/datapages/>) and GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that no competing interest exists.

Authors' Contributions

ZYD contributed to the concept development and study design. PPW analyzed the data, performed bioinformatics training, and wrote the manuscript. YF analyzed the data and helped to write the manuscript. YYC, QL, and YH carried out data management and statistical analysis. TL helped with cohort identification and data management. All authors read and approved the final manuscript. Peipei Wang and Yang Fu contributed equally to this work.

Supplementary Materials

Supplementary 1. Figure S1: construction of a nomogram model with CCL25, IL29, TDGF3, KIR2DL4, GPR44, and GREM2 for predicting 1-, 2-, 3-, 4-, and 5-year OS in TNBC patients.

Supplementary 2. Figure S2: ROC curve verified the nomogram model constructed by CCL25, IL29, TDGF3, KIR2DL4, GPR44, and GREM2.

Supplementary 3. Table S1: the clinical information of TNBC patients.

Supplementary 4. Table S2: the risk scores for variables in nomograms.

References

- [1] J. Ferlay, M. Colombet, I. Soerjomataram et al., "Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods," *International Journal of Cancer*, vol. 144, no. 8, pp. 1941–1953, 2019.
- [2] R. Dent, M. Trudeau, K. I. Pritchard et al., "Triple-negative breast cancer: clinical features and patterns of recurrence," *Clinical Cancer Research*, vol. 13, no. 15, pp. 4429–4434, 2007.
- [3] C. Denkert, C. Liedtke, A. Tutt, and G. von Minckwitz, "Molecular alterations in triple-negative breast cancer—the road to new treatment strategies," *Lancet*, vol. 389, no. 10087, pp. 2430–2442, 2017.
- [4] N. M. Ayoub, K. M. al-Shami, and R. J. Yaghan, "Immunotherapy for HER2-positive breast cancer: recent advances and combination therapeutic approaches," *Breast Cancer-Target*, vol. 11, pp. 53–69, 2019.
- [5] P. Zhang, S. Yi, X. Li et al., "Preparation of triple-negative breast cancer vaccine through electrofusion with day-3 dendritic cells," *PLoS One*, vol. 9, no. 7, p. e102197, 2014.
- [6] M. J. Kwa and S. Adams, "Checkpoint inhibitors in triple-negative breast cancer (TNBC): where to go from here," *Cancer*, vol. 124, no. 10, pp. 2086–2103, 2018.
- [7] "Atezolizumab combo approved for PD-L1-positive TNBC," *Cancer Discovery*, vol. 9, no. 5, 2019.
- [8] A. Rody, T. Karn, C. Liedtke et al., "A clinically relevant gene signature in triple negative and basal-like breast cancer," *Breast Cancer Research*, vol. 13, no. 5, p. R97, 2011.
- [9] H. Mori, M. Kubo, M. Kai et al., "T-bet(+) lymphocytes infiltration as an independent better prognostic indicator for triple-negative breast cancer," *Breast Cancer Research and Treatment*, vol. 176, no. 3, pp. 569–577, 2019.
- [10] J. Yeong, J. C. T. Lim, B. Lee et al., "Prognostic value of CD8 + PD-1+ immune infiltrates and PDCD1 gene expression in triple negative breast cancer," *Journal for Immunotherapy of Cancer*, vol. 7, no. 1, p. 34, 2019.
- [11] C. Chen, C. Liang, S. Wang et al., "Expression patterns of immune checkpoints in acute myeloid leukemia," *Journal of Hematology & Oncology*, vol. 13, no. 1, p. 28, 2020.
- [12] P. Jézéquel, D. Loussouarn, C. Guérin-Charbonnel et al., "Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response," *Breast Cancer Research*, vol. 17, no. 1, p. 43, 2015.
- [13] S. Bhattacharya, S. Andorf, L. Gomes et al., "ImmPort: disseminating data to the public for the future of immunology," *Immunologic Research*, vol. 58, no. 2-3, pp. 234–239, 2014.

- [14] P. P. Wang, S. H. Liu, C. T. Chen et al., "Circulating tumor cells as a new predictive and prognostic factor in patients with small cell lung cancer," *Journal of Cancer*, vol. 11, no. 8, pp. 2113–2122, 2020.
- [15] T. W. Li, J. Y. Fan, B. B. Wang et al., "TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells," *Cancer Research*, vol. 77, no. 21, pp. E108–E110, 2017.
- [16] P. Bedano, B. Schneider, G. Sledge, and K. Miller, "Antiangiogenic therapy for breast cancer," *Breast Cancer Res*, vol. 12, 2007.
- [17] M. Katoh, "FGFR inhibitors: effects on cancer cells, tumor microenvironment and whole-body homeostasis (review)," *International Journal of Molecular Medicine*, vol. 38, no. 1, pp. 3–15, 2016.
- [18] M. Katoh and H. Nakagama, "FGF receptors: cancer biology and therapeutics," *Medicinal Research Reviews*, vol. 34, no. 2, pp. 280–300, 2014.
- [19] C. T. Chen, P. P. Wang, W. J. Mo et al., "Expression profile analysis of prognostic long non-coding RNA in adult acute myeloid leukemia by weighted gene co-expression network analysis (WGCNA)," *Journal of Cancer*, vol. 10, no. 19, pp. 4707–4718, 2019.
- [20] B. L. Li, Y. Cui, M. Diehn, and R. Li, "Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer," *JAMA Oncology*, vol. 3, no. 11, pp. 1529–1537, 2017.
- [21] N. Jacquelot, D. P. Enot, C. Flament et al., "Chemokine receptor patterns in lymphocytes mirror metastatic spreading in melanoma," *The Journal of Clinical Investigation*, vol. 126, no. 3, pp. 921–937, 2016.
- [22] H. J. Chen, J. Sun, Z. L. Huang et al., "Comprehensive models of human primary and metastatic colorectal tumors in immunodeficient and immunocompetent mice by chemokine targeting," *Nat Biotechnol*, vol. 33, no. 6, pp. 656–660, 2015.
- [23] N. E. Kelm, Z. W. Zhu, V. A. Ding et al., "The role of IL-29 in immunity and cancer," *CRIT REV ONCOL HEMATOL*, vol. 106, pp. 91–98, 2016.
- [24] C. Sun, O. Orozco, D. L. Olson et al., "CRIPTO3, a presumed pseudogene, is expressed in cancer," *Biochemical and Biophysical Research Communications*, vol. 377, no. 1, pp. 215–220, 2008.
- [25] N. Nagata, H. Iwanari, H. Kumagai et al., "Generation and characterization of an antagonistic monoclonal antibody against an extracellular domain of mouse DP2 (CRTH2/GPR44) receptors for prostaglandin D2," *PLoS One*, vol. 12, no. 4, p. e0175452, 2017.
- [26] A. Ran, L. Guan, J. N. Wang, and Y. Wang, "GREM2 maintains stem cell-like phenotypes in gastric cancer cells by regulating the JNK signaling pathway," *Cell Cycle*, vol. 18, no. 19, pp. 2414–2431, 2019.
- [27] M. S. Tallman, E. S. Wang, J. K. Altman et al., "Acute myeloid leukemia, version 3.2019, NCCN clinical practice guidelines in oncology," *Journal of the National Comprehensive Cancer Network*, vol. 17, no. 6, pp. 721–749, 2019.
- [28] H. Matsumoto, A. A. Thike, H. Li et al., "Increased CD4 and CD8-positive T cell infiltrate signifies good prognosis in a subset of triple-negative breast cancer," *Breast Cancer Res Tr.*, vol. 156, no. 2, pp. 237–247, 2016.
- [29] J. Yeong, J. C. T. Lim, B. Lee et al., "High densities of tumor-associated plasma cells predict improved prognosis in triple negative breast cancer," *Frontiers in Immunology*, vol. 9, p. 1209, 2018.
- [30] L. Martinez-Lostao, A. Anel, and J. Pardo, "How do cytotoxic lymphocytes kill cancer cells?," *Clinical Cancer Research*, vol. 21, no. 22, pp. 5047–5056, 2015.
- [31] H. Vihervuori, T. A. Autere, H. Repo et al., "Tumor-infiltrating lymphocytes and CD8(+) T cells predict survival of triple-negative breast cancer," *Journal of Cancer Research and Clinical Oncology*, vol. 145, no. 12, pp. 3105–3114, 2019.
- [32] M. L. Wikberg, A. Ling, X. R. Li, Å. Öberg, S. Edin, and R. Palmqvist, "Neutrophil infiltration is a favorable prognostic factor in early stages of colon cancer," *Human Pathology*, vol. 68, pp. 193–202, 2017.
- [33] J. M. Tran Janco, P. Lamichhane, L. Karyampudi, and K. L. Knutson, "Tumor-infiltrating dendritic cells in cancer pathogenesis," *Journal of Immunology*, vol. 194, no. 7, pp. 2985–2991, 2015.

Research Article

Solanine Inhibits Immune Escape Mediated by Hepatoma Treg Cells via the TGF β /Smad Signaling Pathway

Juwei Gao , Yinyin Ying, Jue Wang, and Yiyi Cui

Department of Oncology, The Third Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, 310012 Zhejiang, China

Correspondence should be addressed to Juwei Gao; tuohun890878@163.com

Received 24 July 2020; Revised 5 September 2020; Accepted 15 October 2020; Published 3 November 2020

Academic Editor: Tao Huang

Copyright © 2020 Juwei 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.

Objective. To observe the inhibitory effect of solanine on regulatory T cells (Treg) in transplanted hepatoma mice and to study the mechanism of solanine inhibiting tumor growth. **Methods.** The levels of Treg cells and IL-2, IL-10, and TGF β in the blood of patients with liver cancer were detected by flow cytometry and ELISA, respectively. A mouse hepatocellular carcinoma (HCC) graft model was established and randomly divided into four groups: control group, solanine group, TGF β inhibitor group (SB-431542), and solanine +TGF β inhibitor combined group. Tumor volume of each group was recorded, tumor inhibition rate was calculated, and tumor metastasis was counted. The proportion of CD4⁺CD25⁺Foxp3⁺ Treg in transplanted tumor tissues was detected by flow cytometry. The expression levels of Foxp3 and TGF β in transplanted tumor tissues were detected by quantitative fluorescence PCR. **Results.** Compared with healthy people, Treg cells and IL-2, IL-10, and TGF β contents in peripheral blood of liver cancer patients were increased. The results of the transplanted tumor model in mice showed that the tumor volume of the transplanted mice in the solanine group and the TGF β inhibitor mice was reduced compared with the control group. The combined group had the smallest tumor volume. The proportion of CD4⁺CD25⁺Foxp3⁺ Treg in the transplanted tumor tissues of mice in the solanine treatment group was significantly lower than that in the control group. The expressions of Foxp3 and TGF β in the transplanted tumor tissues of mice in the solanine group were significantly lower than those in the control group. **Conclusion.** Solanine may enhance the antitumor immune response by downregulating the proportion of CD4⁺CD25⁺ Treg and the expression of Foxp3 and TGF β in tumor tissues.

1. Introduction

The mechanism of HCC development and distant metastasis is very complex, among which immune dysfunction is an important factor [1–5]. Various immune mechanisms can target tumor cells to stimulate the body's immune response to tumor cells, including cellular immunotherapy, cytokines, tumor vaccines, and immuncheckpoint inhibitors [6–8].

CD4⁺CD25⁺ Treg cells are usually divided into two categories according to the differences in cell development, biological characteristics, and mechanism of action [9–11]. These include naturally occurring regulatory T cells (nTreg) and induced (or acquired) regulatory T cells (iTreg) from the thymus [12]. iTreg cells are induced, immunomodulatory

cells. Recent studies have shown that in addition to CD4⁺CD25⁺ Treg cells produced by the thymus, peripheral common CD4⁺CD25⁺ T cells can express Foxp3 in vivo and transform into iTreg cells under the conditions of chronic antigen stimulation and cytokines [13, 14]. iTreg cells are derived from peripheral CD4⁺ T cells under the influence of various cytokines and environmental factors after receiving antigen stimulation [15]. CD4⁺CD25⁺ Treg cells are closely related to immune evasion and tolerance of tumors. A large number of studies have found that CD4⁺CD25⁺ Treg cells can be detected in peripheral blood and tumor-infiltrating lymph nodes of patients with gastric cancer, liver cancer, breast cancer, and other cancers [16]. CD4⁺CD25⁺ Treg cells may play an immunosuppressive role by activating

them after recognizing tumor antigens and inhibit the production of antitumor immune responses in the body [9–11], to leave the body in a state of low or no response to the tumor. Foxp3, a member of the forkhead transcription factor family, is a specific morphological and functional indicator of human CD4⁺CD25⁺ Treg cells. Foxp3 plays an important role in the development, differentiation, maturation, and function maintenance of Treg cells, and is considered to be the most sensitive marker of CD4⁺CD25⁺ Treg cells [17–19]. As a member of the TGF β family of growth factors, TGF β can regulate the proliferation and differentiation of immune cells and suppress the immune response. TGF β is a growth regulator of a variety of human epithelial-derived cells and plays an important role in the occurrence, development, and metastasis of tumors.

Solanine is a weakly basic glycoside [20, 21]. Studies have shown that solanine has a wide range of anticancer effects in liver cancer, breast cancer, liver cancer, pancreatic cancer, colorectal cancer, and other tumors [22]. Solanine can inhibit cell proliferation, induce cell apoptosis, block cell cycle, induce autophagy, enhance chemoradiotherapy, inhibit epithelial-mesenchymal transformation, inhibit tumor metastasis, and inhibit angiogenesis [21, 23]. However, there are few studies on the effect of solanine on tumor immune function.

In this study, the effects of solanine on CD4⁺CD25⁺ Treg cells in tumor microenvironment of transplanted hepatocellular carcinoma mice were observed, to study the antitumor effect of solanine and to explore the possibility of solanine in immunotherapy for tumor.

2. Materials and Methods

2.1. Collection of Patient Specimens. Fifteen cases of hepatocellular carcinoma were collected from our hospital from December 2018 to April 2020. Preoperative chemotherapy and immunotherapy were not performed to exclude immune-related diseases. The age ranged from 36 to 68, with a median age of 47. Fifteen healthy subjects excluded tumor and immune-related diseases in hospital physical examination were taken as the control group. All patients signed informed consent forms, and the study was reviewed by the ethics committee of our hospital.

2.2. Serum Samples Were Detected by ELISA. Cytokines IL-2, IL-6, and IL-10, TGF β , and enzyme-linked immunosorbent assay (ELISA) kits were purchased from Wuhan Baodu Company. In the detection of TGF β , the reagent was purchased from Jingmei Bioengineering Co., Ltd. and was detected by Denley automatic labeling analyzer in the United States. Serum TGF β was detected by ELISA, and the operation was strict according to the instructions.

2.3. Cell Culture and Transfection. RPMI1640 cell culture medium, fetal bovine serum, and lymphocyte separation solution were purchased from Gibco, USA. Human hepatocytes and hepatocellular carcinoma cells were purchased from American Type Culture Collection (ATCC, Manassas, VA, USA). H22 cells were inoculated in culture bottles at appropriate concentrations and added to culture medium

containing 10% FBS RPMI1640, cultured in an incubator with constant temperature of 37°C, 5% CO₂, and saturated humidity. Adherent cell growth, passage every 2~3 days once. At the time of administration, human liver cancer cells at exponential growth stage were taken and digested by 0.25% trypsin. Cell suspension with 10% fetal bovine serum was adjusted to a concentration of 3×10^5 /mL and inoculated into 6-well plates with 1 mL per well. The 6-well plate was placed in a carbon dioxide incubator at 37°C and 5% CO₂. After 24 h of culture, solanine was added, and the final concentrations were 0.4, 2, and 10 μ mol/L, respectively.

2.4. Xenograft Tumor Model. Eighty female nude mice aged 4 to 5 weeks, with SPF grade, were purchased from the Laboratory Animal Center of Peking University and kept in SPF room. Liver cancer cell H22 was diluted with RPMI1640 cell culture medium to a suspension containing 1×10^7 cells per 200 μ L and injected subcutaneously into 3 mice. After continued feeding for 1 week, the subcutaneous eminence of mice was visible, indicating the formation of tumors containing liver cancer tumor cells. After the mice were put to death, the subcutaneous skin was cut off in the superclean workbench after sterilization with medical alcohol, and the disinfected scalpel was cut into small pieces. The cells were washed with sterile PBS solution twice. The nude mice were randomly divided into 4 groups: the tumor blocks were inoculated subcutaneously under the right armpit to about 2 cm above the midaxillary line of the mice, and tumor formation in the mice was observed. When the diameter of the subcutaneous cervical tumor block was 5 mm, the tumor-bearing model was successful.

2.5. Administration Method and Dose. In the model control group, 0.2 mL/d of normal saline was intraperitoneally injected. In the solanine group, solanine was intraperitoneally injected at a dose of 37.5 mg/kg, 0.2 mL/day. In the combined TGF β inhibitor group, solanine was intraperitoneally injected at a dose of 37.5 mg/kg, 0.2 mL per day, and TGF β inhibitor SB-431542 was added on the 6th, 7th, and 13th and 14th days, at a dose of 0.2 mL per 10 mg/kg. In the TGF β inhibitor group, 0.2 mL/d of normal saline was intraperitoneally injected. TGF β inhibitor SB-431542 was added on the 6th, 7th, and 13th and 14th days, and 0.2 mL of 10 mg/kg dose was intraperitoneally injected.

2.6. Flow Cytometry to Detect Treg in Peripheral Blood. Reagents were purchased from Jingmei Bioengineering Co., Ltd. and tested using Coulter Epics XL flow cytometer (Beckman Coulter, USA). 40 μ L of EDTA-K2 anticoagulated blood and 5 μ L of CD4-PE and CD25-FITC antibodies were added. In the same type of control tube, another 40 μ L of blood was added to 5 μ L of IgG1-PE and 5 μ L of IgG2a-FITC, mixed well, and incubated in the dark at room temperature for 20 min. After hemolysis, add 3 mL of PBS buffer, centrifuge at 1800 r/min for 5 min, and discard the supernatant. Then, add 500 μ L of PBS buffer solution and check by flow cytometry. The results are expressed as a percentage.

2.7. Treg Cells in Mouse Spleen Lymphocytes Were Detected by Flow Cytometry. The mouse spleen was isolated to make cell

suspension. After filtration, the cells were transferred to the new centrifuge tube. Centrifuge at $800 \times g$ for 5 min and remove the supernatant. The cells were washed with PBS solution 2 times and then centrifuged to remove supernatant. An appropriate amount of cell culture medium was added to adjust the cell density to 1×10^7 /mL. The cells were stratified by the treatment of lymphocyte separation solution, and the lymphocyte layer was the cloud layer between the stratified solution and the supernatant. Draw the needle into the new centrifuge tube. After washing with sterile PBS solution for 2 times, the supernatant was centrifuged to obtain the mononuclear cells of spleen tissue. $100 \mu\text{L}$ of mononuclear cell suspension of spleen tissue was added into the flow tube with PERCP-labeled CD4 antibody and the second tube with PERCP-labeled CD4 and FITC-labeled CD25. After incubation at room temperature in dark for 20 min, $500 \mu\text{L}$ of intracellular membrane breaker was added for 10 min, and the supernatant was removed by centrifugation. PE-labeled Foxp3 antibody was added, was incubated in the dark for 20 min, was washed with PBS once, and was suspended. The proportion of $\text{CD4}^+\text{CD25}^+\text{Foxp3}^+/\text{CD4}^+$ (Treg) cells in mouse lymphocytes in each group was measured by upflow cytometry.

2.8. Immunofluorescence. The slides were fixed with 4% paraformaldehyde for 15 min, and the slides were soaked with PBS for 3 times, 3 min each. Triton X-100 (PBS) permeable cells at room temperature for 20 min. Normal goat serum was added to the slides and sealed at room temperature for 30 min. Each slide was dripped with enough diluted primary antibody (Foxp3, Abcam, 1:100) and placed in a wet box, incubated at 4°C overnight. Add fluorescent secondary antibody. After the absorbent paper was dried and the excess liquid was added to the slide, the diluted fluorescent secondary antibody was added and incubated at 37°C in a wet box for 1 h. PBST was soaked and sliced for 3 times, 3 min each. DAPI was added for incubation in dark for 30 min, and the specimen was stained. The tablets were sealed with a sealing solution containing an antifluorescent quenching agent. Images were observed and collected under a fluorescence microscope (Nikon, Japan).

2.9. Immunohistochemical. The IHC EnVision two-step method was used to detect the expression of Ki-67 in tumor tissue. All fresh specimens were fixed with 10% formaldehyde and sampled within 48 h, paraffin-embedded, and consecutively sectioned, with $4 \mu\text{m}$ thick. Mouse anti-human Ki-67 monoclonal antibody was purchased from Abcam Company. The working fluid concentration was 1:100, and PBS was used as the negative control instead of two primary antibodies. EnVision Kit was purchased from Dako. The sodium citrate buffer (pH = 6.0) was used to repair the antigen by microwave. The rest of the steps strictly follow the instructions.

2.10. qRT-PCR. The reverse transcription kit and qPCR kit were purchased from Dalian Takara Company. The total RNA of spleen lymphocytes in each group was extracted by the TRIzol method. Total RNA concentration was determined by spectrophotometry. A cDNA template was

obtained by using reverse transcription kit to extract 500 ng of total RNA from mouse lymphocytes. qPCR primers for target genes were designed and synthesized by Shanghai Sangon Co., Ltd.: Foxp3 Sense: $5'$ -CACCTATGCCACCC TTATCCG- $3'$, Foxp3 Anti-sense: $5'$ -CATGCGAGTAA ACCAATGGTAGA- $3'$; GAPDH Sense: $5'$ -AGGTCGGTG AACGGATTG- $3'$, GAPDH Anti-Sense: $5'$ -GGGGTC GTTGATGGCAACA- $3'$. Procedures were performed as described in SYBR Premix Ex TaqTM II and tested by the U.S. Bio-Rad iQ5 qPCR system. Reaction conditions are as follows: 94°C , 30 s; 58°C , 30 s; and 72°C , 30 s, with a total of 40 cycles. As an internal reference gene, GAPDH was repeated for 3 times, and the mRNA expression of each group was calculated by the $2^{-\Delta\Delta\text{CT}}$ method.

2.11. Statistical Analysis. SPSS 20.0 (SPSS Inc., Chicago, IL, USA) statistical software was used for relevant data analysis. All data were expressed as the mean \pm standard deviation. Comparisons between the two groups were performed using the *t*-test. Comparisons between groups were performed using one-way ANOVA. $P < 0.05$ or $P < 0.01$ means that the difference was statistically significant.

3. Results

3.1. Treg Cells Increased in Peripheral Blood of Tumor Patients, and the Levels of IL-2, IL-10, and TGF β Increased. In order to study the content of immune cells and the changes of tumor immune-related factors in peripheral blood of patients with liver cancer, we collected the peripheral blood of 15 healthy volunteers and 15 patients with liver cancer. The experimental results showed that compared with healthy volunteers, the proportion of Treg in peripheral blood of liver cancer patients in CD4^+ T cells was analyzed and cell surface markers were detected on flow cytometry, and the number of Treg in the liver cancer group was higher than that in the control group ($P < 0.01$, Figure 1(a)). Further, ELISA was used to detect the levels of IL-2, IL-10, and TGF β in serum. The results showed that the levels of IL-2, IL-10, and TGF β in the serum of liver cancer patients were significantly increased compared with healthy volunteers ($P < 0.01$, Figures 1(b)–1(d)).

3.2. Liver Cancer Cells Release More TGF β than Normal Stem Cells. Furthermore, the expression of TGF β in liver cancer cell lines and normal liver cells was determined. The experimental results showed that compared with normal liver cells, TGF β expression was higher in HCC cell lines, with H22 having the highest TGF β expression (Figure 2(a)). In the previous experimental results, we confirmed the inhibitory effect of solanine on liver cancer. In this study, we examined the inhibitory effect of solanine on TGF β of hepatocellular carcinoma cell line H22. The results showed that solanine could significantly reduce the content of TGF β in H22 with a dose-dependent relationship (Figures 2(b) and 2(c)).

3.3. Solanine Inhibits the Growth and Metastasis of Transplanted Hepatoma in Mice. After inoculation of liver cancer cells in mice for 5-7 days, subcutaneous tumor

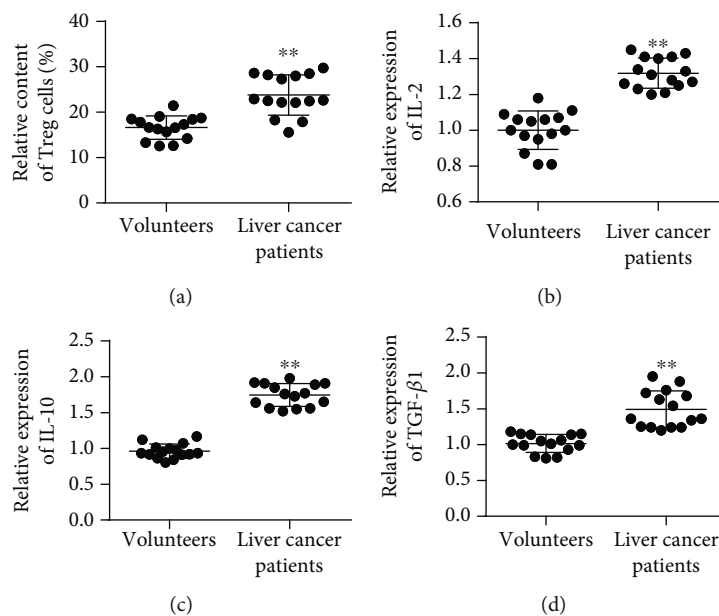


FIGURE 1: Detection of Treg cell content and TGF β in peripheral blood of tumor patients. (a) Treg cell content in serum. (b) IL-2 content in serum. (c) IL-10 content in serum. (d) TGF β content in serum. * $P < 0.05$ and ** $P < 0.01$.

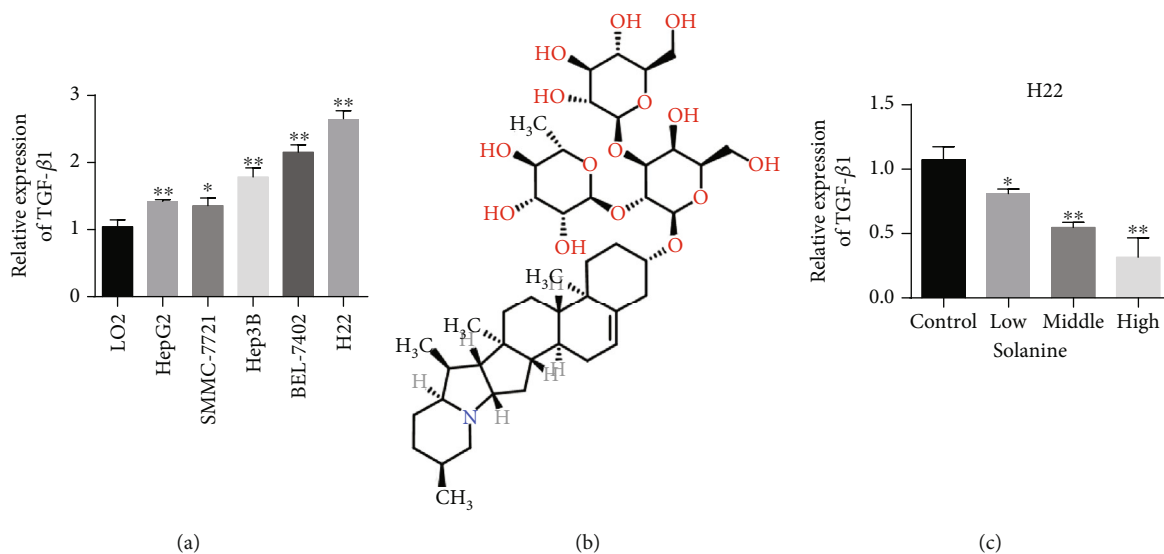


FIGURE 2: Liver cancer cells release more TGF β than normal stem cells. (a) Compared with normal liver cells (LO2) and liver cancer cells (HepG2, SMMC-7721, Hep3B, BEL-7402, and H22), TGF β content was detected. (b) Structural formula of solanine. (c) Solanine treats liver cancer cells and inhibits TGF β . * $P < 0.05$ and ** $P < 0.01$.

formation could be felt, and tumor lesions could be observed by naked eye on the 7th day or so. The transplanted tumors in the PBS group grew rapidly, and the growth curve was steep (Figures 3(a) and 3(b)). There was no significant difference in the growth rate of transplanted tumors in the solanine group, the TGF β inhibitor group, and the combination group compared with the control group at the initial stage of treatment. With the extension of treatment time, the growth of transplanted tumors in mice in the sola-

nine group, the TGF β inhibitor group, and the combination group was slow. The tumor weight test results showed that the transplanted weight of mice in the solanine group was significantly lower than that in the control group (Figure 3(c)). The tumor metastasis detection results showed that compared with the control group, tumor metastasis was significantly reduced in the solanine group, the TGF β inhibitor group, and the combination group (Figure 3(d)). Further, we detected the cell proliferation index Ki-67 staining

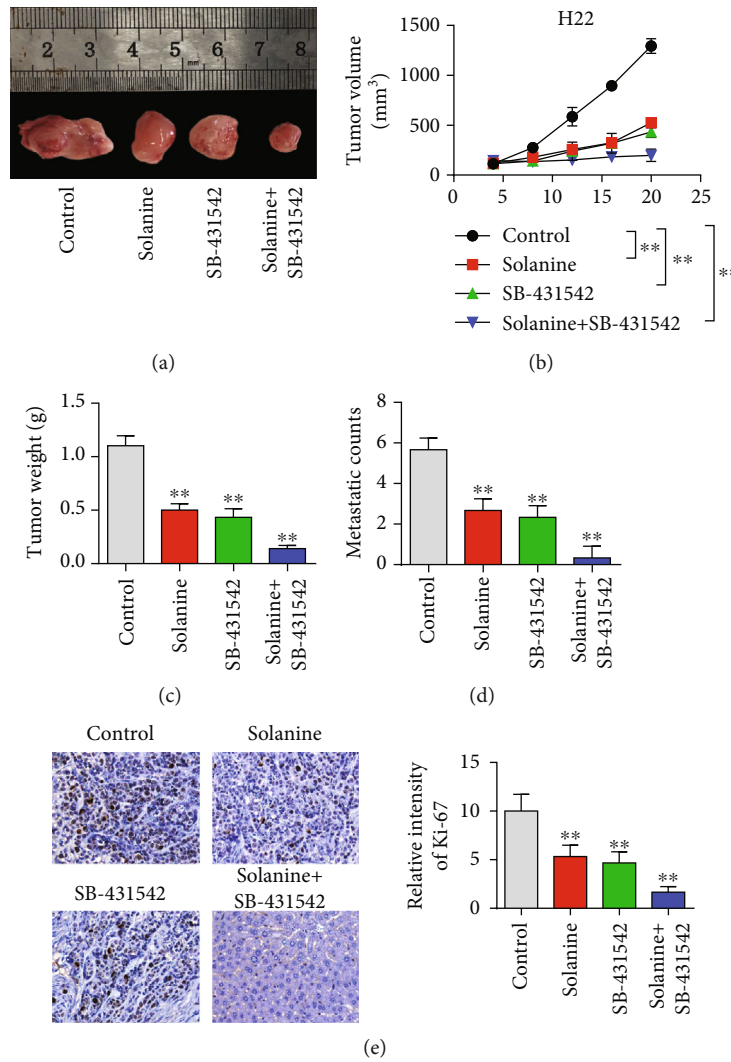


FIGURE 3: Animal level study of solanine antitumor. (a) Pictures of tumors in a mouse subcutaneous tumor-bearing model. (b) Mouse tumor volume curve. (c) Tumor weight detection in mice. (d) Detection of liver metastasis in mice. (e) Ki-67 immunohistochemical staining of mouse tumor tissue. * $P < 0.05$ and ** $P < 0.01$.

by immunohistochemistry. The experimental results showed that compared with the control group, Ki-67 staining of tumor cells was significantly decreased in the solanine group, the TGF β inhibitor group, and the combination group, indicating that the solanine group, the TGF β inhibitor group, and the combination group could significantly reduce the proliferation ability of tumor cells (Figure 3(e)).

3.4. Solanine Treatment Reduced the Proportion of CD4⁺CD25⁺Foxp3⁺ Treg in Transplanted Tumor Tissues. In order to observe the effect of solanine on Treg in hepatocellular carcinoma-bearing mice, CD4⁺CD25⁺Foxp3⁺ was used as the marker of Treg in this study. CD4 coil gates were first used; then, CD25 coil gates were used to detect the proportion of CD4⁺CD25⁺Foxp3⁺ Treg in CD4⁺ T cells in each group. The proportion of CD4⁺CD25⁺Foxp3⁺ Treg in CD4⁺ T cells in the solanine group, the TGF β inhibitor group, and the combination group was significantly lower than that in the control

group (Figures 4(a) and 4(b)). At the same time, immunofluorescence staining results showed that the expression of Foxp3 in Treg cells of the solanine group, the TGF β inhibitor group, and the combination group was also significantly decreased (Figure 4(c)). At the same time, we also found that solanine treatment reduced the contents of IL-2, IL-10, and TGF β in the serum of tumor-bearing mice (Figures 5(a)–5(c)). The Treg cells in the control group, the solanine group, the TGF β inhibitor group, and the combined group were separated by flow cytometry and then cultured to detect the contents of IL-2, IL-10, and TGF β in the cell supernatant. The results showed that the levels of IL-2, IL-10, and TGF β in Treg cells isolated from the solanine group, the TGF β inhibitor group, and the combined treatment group were significantly lower than those in the control group (Figures 6(a)–6(c)).

3.5. Solanine Inhibited the Expression of Foxp3 and TGF β mRNA in Transplanted Tumor Tissues. In order to further

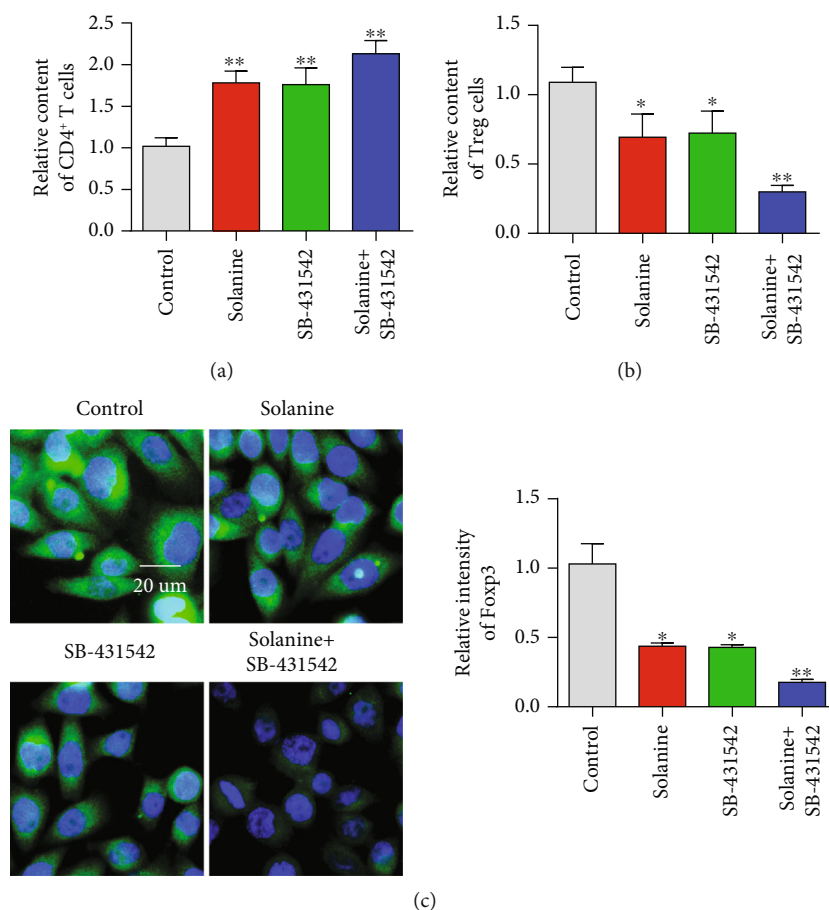


FIGURE 4: Expression levels of CD4⁺ cells, D4⁺CD25⁺Foxp3⁺ Treg, and Foxp3⁺ immunofluorescence intensity in tumor tissues treated with solanine. (a) Detect the content of CD4⁺ T cells by flow cytometry. (b) Detect the content of CD4⁺CD25⁺Foxp3⁺ Treg by flow cytometry. (c) Foxp3⁺ immunofluorescence intensity of tumor tissue sections. * $P < 0.05$ and ** $P < 0.01$.

investigate the molecular mechanism of solanine in the treatment of liver cancer, we detected the changes of the TGF β /Smad signaling pathway in tumor tissues after treatment with solanine. The experimental results showed that compared with the control group, TGF β expression was significantly reduced in the solanine group, the TGF β inhibitor group, and the combination group, with the maximum reduction in the combination group (Figure 7(a)). Further test results showed that the overall expression of Smad2 did not change significantly, but the p-Smad2 expression decreased significantly (Figures 7(b) and 7(c)). The detection results of Smad3 were consistent with those of Smad2. The overall expression level of Smad2 in the solanine group, the TGF β inhibitor group, and the combination group did not change significantly, but the expression level of p-Smad2 was significantly decreased, and the reduction degree of the combination group was the largest (Figures 7(d) and 7(e)).

4. Discussion

Hepatocellular carcinoma (HCC) is one of the most common malignancies and ranks the second leading cause of cancer

death worldwide [24–26]. Immunotherapy is one of the important means of tumor therapy. In recent years, it has been found that the occurrence, development, and prognosis of tumors are closely related to CD4⁺CD25⁺ Treg cells. Treatment measures against Treg in vivo are obviously beneficial to antitumor immunity [27–30]. Ideal tumor therapeutic effects can be achieved by removing Treg cells in tumor patients and regulating the migration, distribution, and function of Treg. Tumor cells retain the antigenicity of the original cells and produce tumor antigens during carcinogenesis. Therefore, the immune escape of the tumor results in the body's unresponsiveness to the tumor or low immune function. Treg cells can inhibit the development and activation of effector cells that recognize their own tumor cells, and play an important role in mediating tumor immune tolerance. Therefore, inhibition of Treg cells is of great significance for tumor immunotherapy [31–34].

Treg cells are involved in tumor escape by secreting inhibitory cytokines. Treg cells mainly secrete TGF β and IL-10 to inhibit the body's antitumor immunity. IL-10 promotes tumor immune escape by inhibiting the function of APC, Th1 cells, NK cells, and macrophages. TGF β also induces excessive IL-10 secretion in tumors, leading to

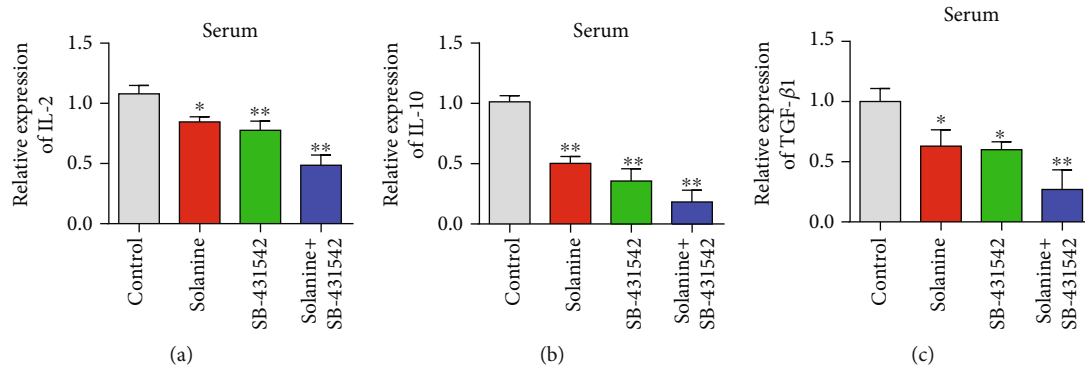


FIGURE 5: Detection of IL-2, IL-10, and TGFβ levels in serum of tumor-bearing mice. (a) Detection of IL-2 in serum by qPCR. (b) Detection of IL-10 in serum by qPCR. (c) Detection of TGFβ in serum by qPCR. *P < 0.05 and **P < 0.01.

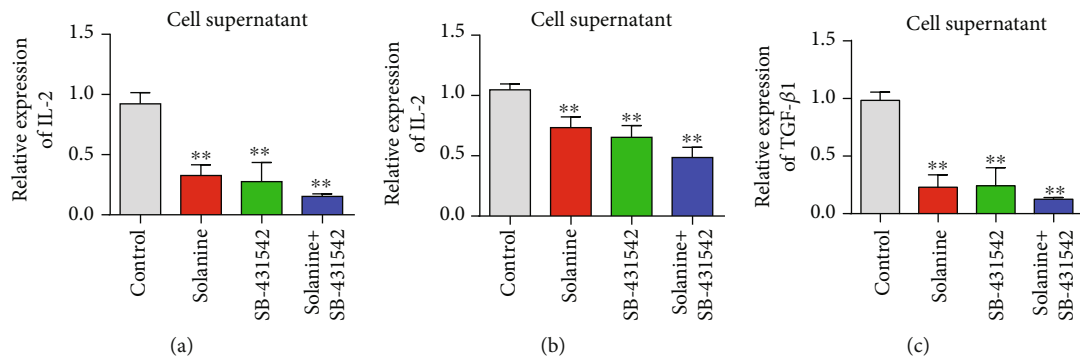


FIGURE 6: Levels of IL-2, IL-10, and TGFβ in spleen CD4⁺ cell culture supernatant of tumor-bearing mice. (a) Detection of IL-2 content in the supernatant of CD4⁺ T cells by qPCR. (b) Detection of IL-10 in CD4⁺ T cell supernatant by qPCR. (c) qPCR detection of TGFβ content in CD4⁺ T cell supernatant. *P < 0.05 and **P < 0.01.

immunosuppression of the antitumor response. Chen et al. [35] found that CD4⁺CD25⁺ Treg cells could inhibit the tumor clearance effect mediated by CD8⁺ T cells. TGFβ plays a key role in this inhibitory process. Treg cells promote tumor immune escape by interacting with immune cells in the tumor microenvironment. Treg cells inhibit tumor immunity by inhibiting tumor-infiltrating T cells (TIL), dendritic cells, and macrophages in the tumor microenvironment. Treg cells inhibit proliferation and induce immune incompetence by competing with effector T cells to bind IL-2 or acting directly on them. IL-2 is a signal of cell proliferation. CD4⁺CD25⁺ Treg cells can upregulate the expression of IL-2 receptor chain, while interfering with the expression of IL-2 receptor chain in effector T cells. Finally, IL-2 receptor can be used competitively, so that effector cells cannot get growth signal and cannot proliferate. In this study, it was found that solanine and TGFβ inhibitors could significantly inhibit Treg cell content and reduce IL-2 and IL-10 levels.

In this study, a mouse model of hepatocellular carcinoma was successfully established, and it was found that solanine could significantly inhibit the growth of hepatocellular carcinoma in vivo. With the extension of time, the

rate of tumor inhibition increased. The results of animal experiments verified the inhibitory effect of solanine on liver cancer. Foxp3 is the most specific marker of Treg cells and plays a key role in the differentiation and function regulation of Treg cells. In this study, qPCR was used to detect gene expression changes of key transcription factors of Treg cells in spleen lymphocytes of hepatoma-bearing mice. The results showed that the expression level of Foxp3 mRNA in lymphocytes of mice in the tumor-bearing model group increased significantly, while the expression level of Foxp3 mRNA in the solanine treatment group decreased significantly, with a time effect [36, 37]. This suggests that solanine may play an immunomodulatory role by inhibiting the expression of Foxp3, a key transcription factor in Treg cells [38, 39]. At the same time, it was also found in this study that the levels of Treg-related immune cytokines (IL-2, IL-10, and TGFβ) were significantly increased in mice in the hepatoma-bearing model group. However, Treg cell-related immune cytokines decreased after treatment with solanine. This indicated that solanine could reduce Treg cell content in spleen lymphocytes of hepatocellular carcinoma-bearing mice by inhibiting Treg cell-related immune cytokines.

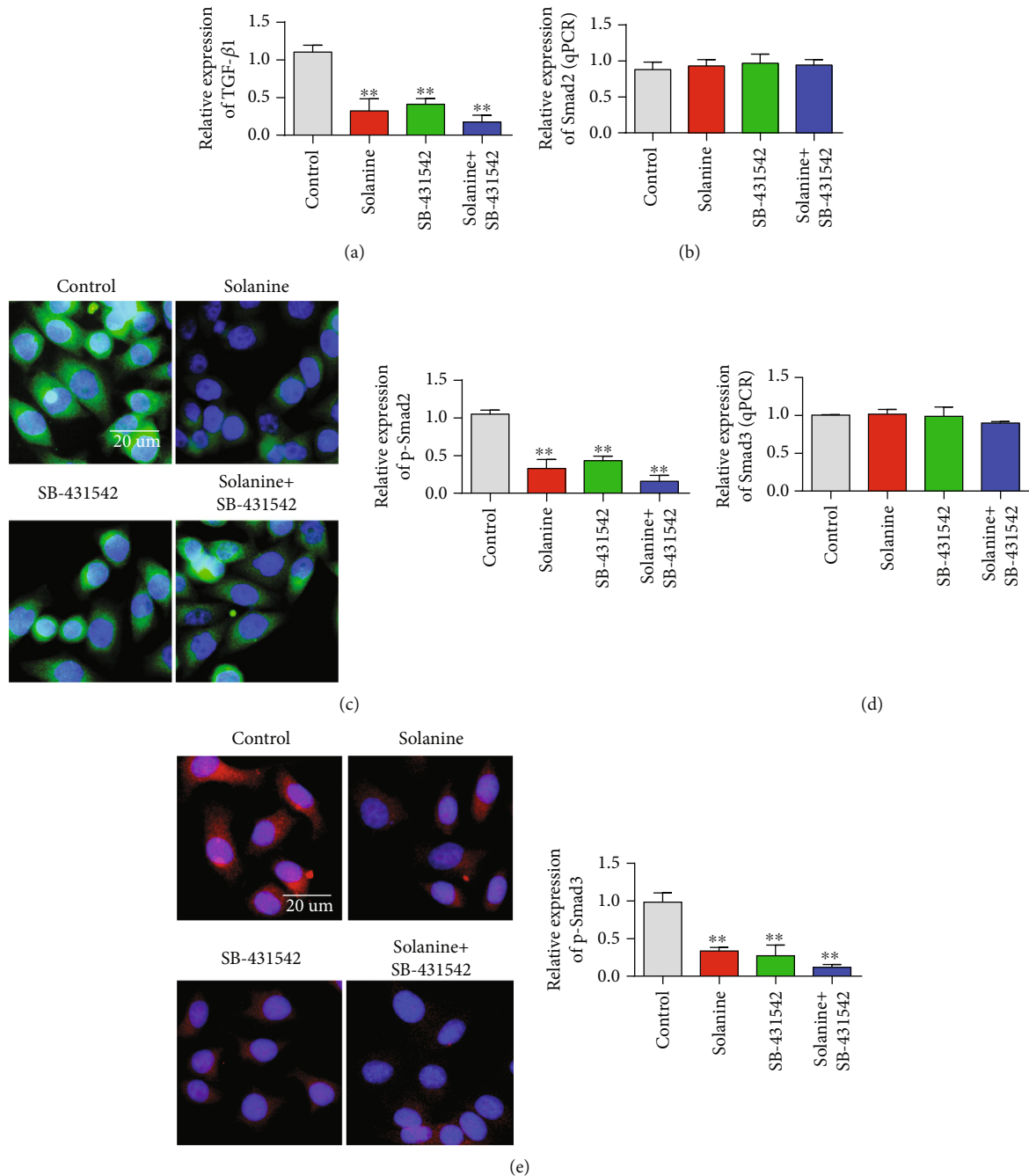


FIGURE 7: Detected the TGF β /Smad3 signaling pathway. (a) Detection of TGF β content in sorted cells. (b) Detection of Smad2 content in sorted cells. (c) Detection of p-Smad2 content in sorted cells. (d) Detection of Smad3 content in sorted cells. (e) Detection of p-Smad3 content in sorted cells. * $P < 0.05$ and ** $P < 0.01$.

5. Conclusion

The results of this study showed that solanine inhibited the growth of transplanted hepatoma in mice and reduced the proportion of CD4⁺CD25⁺Foxp3⁺ Treg and the expression levels of Foxp3 and TGF β mRNA. Mechanism studies have shown that solanine reduces the proportion of local CD4⁺CD25⁺Foxp3⁺ Treg in mice and enhances the body's antitumor immune response by inhibiting the TGF β /Smad

signaling pathway. Therefore, as a potential drug for effective treatment of liver cancer, solanine has potential application prospect. This study provides experimental basis for the future clinical application of solanine.

Data Availability

All the experimental data can be made public by the requirements of qualified researchers.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by the Medical Health Science and Technology Project of Zhejiang Provincial Health Commission (2017KY519) and the Scientific Research Project of Zhejiang Chinese Medical University (2017ZZ06).

References

- [1] C. Yi, G. Yang, M. Sun, K. Chang, R. Long, and Z. Jiang, "The Expression and Clinical Significance of Treg Cells in Chronic Myelocytic Leukemia," *Biomedical Research*, vol. 28, no. 21, pp. 9188–9192, 2017.
- [2] X. Wu, M. Peng, B. Huang et al., "Immune microenvironment profiles of tumor immune equilibrium and immune escape states of mouse sarcoma," *Cancer Letters*, vol. 340, no. 1, pp. 124–133, 2013.
- [3] T. L. Whiteside, M. Mandapathil, M. Szczepanski, and M. Szajnik, "Mechanisms of tumor escape from the immune system: adenosine-producing Treg, exosomes and tumor-associated TLRs," *Bulletin du Cancer*, vol. 98, no. 2, pp. E25–E31, 2011.
- [4] J. Meng, X. Ai, Y. Lei et al., "USP5 promotes epithelial-mesenchymal transition by stabilizing SLUG in hepatocellular carcinoma," *Theranostics*, vol. 9, no. 2, pp. 573–587, 2019.
- [5] J. Meng, S. Chen, Y. Y. Lei et al., "Hsp90 β promotes aggressive vasculogenic mimicry via epithelial-mesenchymal transition in hepatocellular carcinoma," *Oncogene*, vol. 38, no. 2, pp. 228–243, 2019.
- [6] Y. Shiono, T. Kanaseki, S. Miyamoto et al., "Loss of tapasin in human lung and colon cancer cells and escape from tumor-associated antigen-specific CTL recognition," *Oncoimmunology*, vol. 6, no. 2, p. e1274476, 2017.
- [7] S. M. Toor and E. Elkord, "Therapeutic prospects of targeting myeloid-derived suppressor cells and immune checkpoints in cancer," *Immunology & Cell Biology*, vol. 96, no. 9, pp. 888–897, 2018.
- [8] M. Terme, S. Pernot, E. Marcheteau et al., "VEGF-a-induced Treg proliferation, a novel mechanism of tumor immune escape in colorectal cancer: effects of anti-VEGF/VEGFR therapies," *Annals of Oncology*, vol. 23, article ix75, 2012.
- [9] C. Mu, G. Zhang, J. Huang, and B. Qi, "pDC induced Treg proliferation through PD-1/PD-L1 signal and promote tumor immune escape of lung cancer with MPE," *European Respiratory Journal*, vol. 14, 2014.
- [10] S. Mocellin and D. Nitti, "Therapeutics targeting tumor immune escape: towards the development of new generation anticancer vaccines," *Medicinal Research Reviews*, vol. 28, no. 3, pp. 413–444, 2008.
- [11] T. Miyazaki, K. Ikeda, W. Sato, K. Horie-Inoue, and S. Inoue, "Extracellular vesicle-mediated EBAG9 transfer from cancer cells to tumor microenvironment promotes immune escape and tumor progression," *Oncogene*, vol. 7, no. 1, p. 7, 2018.
- [12] J. A. Bluestone and A. K. Abbas, "Natural versus adaptive regulatory T cells," *Nature Reviews Immunology*, vol. 3, no. 3, pp. 253–257, 2003.
- [13] I. Apostolou and H. Von Boehmer, "In vivo instruction of suppressor commitment in naive T cells," *The Journal of Experimental Medicine*, vol. 199, no. 10, pp. 1401–1408, 2004.
- [14] M. A. Curotto de Lafaille, A. C. Lino, N. Kutchukhidze, and J. J. Lafaille, "CD25⁻ T cells generate CD25⁺ Foxp3⁺ regulatory T cells by peripheral expansion," *The Journal of Immunology*, vol. 173, no. 12, pp. 7259–7268, 2004.
- [15] A. Facciabene, G. T. Motz, and G. Coukos, "T-regulatory cells: key players in tumor immune escape and angiogenesis," *Cancer Research*, vol. 72, no. 9, pp. 2162–2171, 2012.
- [16] A. M. Wolf, D. Wolf, M. Steurer, G. Gastl, E. Gunsilius, and B. Grubeck-Loebenstern, "Increase of regulatory T cells in the peripheral blood of cancer patients," *Clinical Cancer Research*, vol. 9, no. 2, pp. 606–612, 2003.
- [17] M. Miyara and S. Sakaguchi, "Human FoxP3⁺ CD4⁺ regulatory T cells: their knowns and unknowns," *Immunology & Cell Biology*, vol. 89, no. 3, pp. 346–351, 2011.
- [18] M. Kleinewietfeld, M. Starke, D. Di Mitri et al., "CD49d provides access to "untouched" human Foxp3⁺ Treg free of contaminating effector cells," *Blood*, vol. 113, no. 4, pp. 827–836, 2009.
- [19] L. Lu, J. Ma, X. Wang et al., "Synergistic effect of TGF-beta superfamily members on the induction of Foxp3⁺ Treg," *European Journal of Immunology*, vol. 40, no. 1, pp. 142–152, 2010.
- [20] C. Lv, H. Kong, G. Dong et al., "Antitumor efficacy of α -solanine against pancreatic cancer in vitro and in vivo," *PLoS One*, vol. 9, no. 2, article e87868, 2014.
- [21] M. Mohsenikia, A. M. Alizadeh, S. Khodayari et al., "The protective and therapeutic effects of alpha-solanine on mice breast cancer," *European Journal of Pharmacology*, vol. 718, no. 1–3, pp. 1–9, 2013.
- [22] W. F. Zhong, S. P. Liu, B. Pan, Z. F. Tang, J. G. Zhong, and F. J. Zhou, "Solanine inhibits prostate cancer Du145 xenograft growth in nude mice by inducing cell cycle arrest in G1/S phase," *Nan Fang Yi Ke Da Xue Xue Bao*, vol. 36, no. 5, pp. 665–670, 2016.
- [23] J. Zhang and G. W. Shi, "Inhibitory effect of solanine on prostate cancer cell line PC-3 in vitro," *National Journal of Andrology*, vol. 17, no. 3, pp. 284–287, 2011.
- [24] H. Wang, W. Zhong, J. Zhao et al., "Oleanolic acid inhibits epithelial-mesenchymal transition of hepatocellular carcinoma by promoting iNOS dimerization," *Molecular Cancer Therapeutics*, vol. 18, no. 1, pp. 62–74, 2019.
- [25] T. Xiao, W. Zhong, J. Zhao et al., "Polyphyllin I suppresses the formation of vasculogenic mimicry via Twist1/VE-cadherin pathway," *Cell Death & Disease*, vol. 9, no. 9, p. 906, 2018.
- [26] W. Zhong, W. Yang, Y. Qin et al., "6-Gingerol stabilized the p-VEGFR2/VE-cadherin/ β -catenin/actin complex promotes microvessel normalization and suppresses tumor progression," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, p. 285, 2019.
- [27] M. Hahne, "Melanoma cells express Fas ligand: implications for tumor immune escape," *Immunology Letters*, vol. 56, no. 1–3, p. 343, 1997.
- [28] F. H. Igney and P. H. Krammer, "Immune escape of tumors: apoptosis resistance and tumor counterattack," *Journal of Leukocyte Biology*, vol. 71, no. 6, 2002.
- [29] M. Hahne, D. Rimoldi, M. Schroter et al., "Melanoma cell expression of Fas(Apo-1/CD95) ligand: implications for tumor immune escape," *Science*, vol. 274, no. 5291, pp. 1363–1366, 1996.

- [30] M. Campoli, C. C. Chang, and S. Ferrone, "HLA class I antigen loss, tumor immune escape and immune selection," *Vaccine*, vol. 20, Suppl 4, pp. A40–A45, 2003.
- [31] J. R. B. Engel and A. Honig, "Mechanisms of tumor immune escape in triple-negative breast cancers (TNBC) with and without mutated BRCA 1," *Archives of Gynecology & Obstetrics*, vol. 289, no. 1, pp. 141–147, 2014.
- [32] A. Buzyn, F. Petit, M. Ostankovitch et al., "Membrane-bound Fas (Apo-1/CD95) ligand on leukemic cells: a mechanism of tumor immune escape in leukemia patients," *Blood*, vol. 94, no. 9, pp. 3135–3140, 1999.
- [33] A. Garcia-Lora, I. Algarra, and F. Garrido, "MHC class I antigens, immune surveillance, and tumor immune escape," *Journal of Cellular Physiology*, vol. 195, no. 3, pp. 346–355, 2003.
- [34] X. Xi, N. Liu, Q. Wang et al., "ACT001, a novel PAI-1 inhibitor, exerts synergistic effects in combination with cisplatin by inhibiting PI3K/AKT pathway in glioma," *Cell Death & Disease*, vol. 10, no. 10, p. 757, 2019.
- [35] M.-L. Chen, M. J. Pittet, L. Gorelik et al., "Regulatory T cells suppress tumor-specific CD8 T cell cytotoxicity through TGF- β signals in vivo," *Proceedings of the National Academy of Sciences*, vol. 102, no. 2, pp. 419–424, 2005.
- [36] D. A. Horwitz, S. G. Zheng, J. Wang, and J. D. Gray, "Critical role of IL-2 and TGF- β in generation, function and stabilization of Foxp3+CD4+ Treg," *European Journal of Immunology*, vol. 38, no. 4, pp. 912–915, 2008.
- [37] T. Duhon, R. Duhon, A. Lanzavecchia, F. Sallusto, and D. J. Campbell, "Functionally distinct subsets of human FOXP3+ Treg cells that phenotypically mirror effector Th cells," *Blood*, vol. 119, no. 19, pp. 4430–4440, 2012.
- [38] O. Palomares, B. Rückert, T. Jartti et al., "Induction and maintenance of allergen-specific FOXP3+ Treg cells in human tonsils as potential first-line organs of oral tolerance," *Journal of Allergy and Clinical Immunology*, vol. 129, no. 2, pp. 510–520.e9, 2012.
- [39] A. Schneider, M. Rieck, S. Sanda, C. Pihoker, C. Greenbaum, and J. H. Buckner, "The effector T cells of diabetic subjects are resistant to regulation via CD4+FOXP3+regulatory T cells," *Journal of Immunology*, vol. 181, no. 10, pp. 7350–7355, 2008.

Research Article

Immune Infiltration Landscape in Lung Squamous Cell Carcinoma Implications

Jungang Zhao,¹ Wenming Bao,¹ and Weiyang Cai² 

¹Department of Hepatobiliary Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

²Department of Gastroenterology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

Correspondence should be addressed to Weiyang Cai; wish0611@126.com

Received 10 June 2020; Revised 3 September 2020; Accepted 18 September 2020; Published 10 October 2020

Academic Editor: Tao Huang

Copyright © 2020 Jungang 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.

Intrinsic cancer cells and the tumor-infiltrating immune cells (TIICs) recruited to the immune microenvironment define the malignant phenotype of lung squamous cell carcinoma (LUSC). Understanding more about the immune microenvironment of LUSC enables the selection of high-risk patients who would derive benefit from immunotherapy. Based on large public LUSC cohorts obtained from TCGA and GEO datasets, 22 types of infiltrating immune cell subgroups were evaluated by CIBERSORT. Meta-analysis, principal component analysis (PCA), single-sample gene set enrichment analysis (ssGSEA), and hierarchical clustering analysis were used to evaluate specific immune responses of LUSC. The distribution of TIICs of LUSC was entirely different from normal. TIIC subpopulations were also found to be closely associated with clinical features and molecular subtypes. Unsupervised clustering analysis revealed that three distinct TIIC subgroups existed with different survival patterns. TIICs are extensively implicated in the pathogenesis and development of LUSC. Characterizing the composition of TIICs influences the metabolism, pathological stage, and survival of tumor patients. It is hoped that this immune landscape could provide a more accurate understanding of the development and immunotherapy of LUSC.

1. Introduction

Lung cancer is currently the leading cause of cancer-related death in the world, with high morbidity and mortality, in which non-small-cell lung cancer (NSCLC) accounts for 85% [1]. The main modalities of treatment in use for LUSC currently include surgery, chemotherapy, radiation therapy, molecular targeted therapy, and immunotherapy [2–4]. Notably, despite these options, the overall five-year survival rate of lung cancer was 19%, one of the cancers with poor prognosis [5]. Therefore, it is crucial to uncover effective biomarkers for diagnosis and accelerate development of new classes of antitumor targeted drug.

Immunotherapy, a strategy using immune checkpoint inhibitors to inhibit or activate different immune cell subtypes, can improve the antitumor immune response and get better clinical outcomes. The prognosis of patients receiving immune checkpoint blockers can be significantly improved [6]. Thus, the distribution of TIICs in tumors has been a vital subject for researches. TIICs, consisting of B cells, T cells,

macrophages, and other immune cells, are one of the essential components of the tumor microenvironment and play an important role in monitoring antitumor immune responses [7–9]. Drugs targeting specific TIICs have also been found to be associated with better clinical outcomes [10, 11]. Researchers have also revealed the dual role of TIICs in improving body immunity and promoting tumorigenesis, which means that there are complex and undiscovered interactions between TIICs and tumor cells [10, 12].

Newman et al. developed a deconvolution algorithm called CIBERSORT based on gene expression for electronic quantification of various types of immunocytes in heterogeneous samples [13]. CIBERSORT can properly determine the diversity and pattern of TIICs. 22 immune cell types can be enumerated at a time by using CIBERSORT. Because of these, this analysis has received increasing attention in the research of cell heterogeneity [9]. In this research, we used CIBERSORT to assess the proportions of 22 immune cell types in tumor and adjacent samples and analyze their relationship with molecular subpopulations and overall survival.

This study is aimed at exploring the intricate relationship between intratumoral immune cell heterogeneity, tumor molecular subtypes, and disease progression in LUSC.

2. Materials and Methods

2.1. Study Design and Participants. The TIIC composition and clinical information used for analysis were obtained from TCGA database up until October 2019. All gene expression data of primary LUSC patients were thought to be qualified, without applying specific criteria for exclusion. Patients with complete information on clinical characteristic, such as gender, age, survival time, and TNM stage, were further analyzed while the rest of the patients were excluded. A total of 424 LUSC expression profiles were included for further study. Table S1 summarizes LUSC patient demographic information. We also screened eligible LUSC microarray in the GEO dataset (<http://www.ncbi.nlm.nih.gov/gds/>) up until October 2019. All chips with detailed data of gene expression (containing at least 15 samples) from LUSC patients were considered useful, without applying specific criteria for exclusion. A total of 8 chips were included in the further verified study (Table S2): GSE10929, GSE19804, GSE21933, GSE33356, GSE33479, GSE33532, GSE51855, and GSE67061. The flowchart in Figure 1 details the study design.

2.2. Evaluation of Tumor-Infiltrating Immune Cells. Taking $P < 0.05$ and $[\log FC] > 1$ as the cutoff criteria, we normalized gene expression data and then identified the differently expressed genes (DEGs) via the limma package. Then, the CIBERSORT algorithm was used to explode the normalized data, which inferred the relative ratio of 22 types of infiltrating immune cell. The CIBERSORT algorithm basing on parts of reference gene expression values derives the proportions of various immune cell types from tumor samples mixed with kinds of multiple cells [12]. The CIBERSORT P value represents a measure of confidence in the results using Monte Carlo sampling. Immune cell profiles from TCGA and GEO databases were analyzed using CIBERSORT, and the number of permutations was set to 100. The geometric mean of GZMA and PRF1 was calculated to represent immune cytolytic activity [14].

2.3. Meta-Analysis. Meta-analysis was manipulated by Review Manager to infer each infiltrating immune cell in LUSC. Continuous outcomes were reported as a standard mean difference (SMD) with a 95% confidence interval (CI). The DerSimonian and Laird method was characterized as the standard method, which offered an average impact estimate of the heterogeneity of effects across a series of chips. We took $P < 0.05$ as statistically significant.

2.4. Principal Component Analysis (PCA). PCA involves a mathematical process in which several potentially related variables are converted into a smaller number of unrelated variables called principal components. The eigenvector associated with the largest eigenvalue has the same path as the first principal component. As a result, the group bias and

the individual difference errors have been examined by PCA, thus proving the credibility of the in-depth results.

2.5. Single-Sample Gene Set Enrichment Analysis (ssGSEA) and Gene Set Enrichment Analysis (GSEA). The enrichment scores of the PD-1 immunotherapeutic were calculated by ssGSEA by the gsva package [15, 16]. The PD-1 score was defined as the average of the standardized values of IDO2, TIM-3, IDO2, PDL-1, CTLA4, LAG3, and TIGIT [17]. The correlation between the composition of the TIICs and the immune score was calculated using the Pearson correlation. GSEA based on gene expression profiles was constructed by GSEA.4.0.1. The KEGG pathways significantly altered were identified using cutoff FDR 0.1.

2.6. Evaluation of Immunoscore. TMB was defined as the total number of somatic mutations (except silent mutations) and plotted by maftools. According to gene expression profiles, the ESTIMATE algorithm quantifies the immune activity for each tumor sample.

2.7. Single Nucleotide Polymorphism (SNP) Analysis. Affymetrix SNP 6.0 arrays were gained by using the VarScan method to analyze the germ cell/somatic cell mutation site data obtained from the second-generation sequencing data of LUSC tissues in TCGA database (<https://cancergenome.nih.gov/>). We assess the influence of mutant genes on the prognosis of patients with LUSC using the Kaplan-Meier plot and screened out the most frequent mutated genes for further analysis.

2.8. Hierarchical Clustering Analysis. Samples were grouped using hierarchical agglomerative clustering, which functioned in assigning a similar sample into a data frame and finding the closest cluster pair. The distribution of samples was shown in the consensus matrix heat map.

2.9. Statistical Analyses. The Mann-Whitney U test was used for continuous variables. Meta-analysis was conducted by Review Manager 5.3 to verify the proportion of immune cell infiltration, and results were remarked as standard mean difference (SMD) with a 95% confidence interval (CI). The clinical endpoints of this investigation included overall survival (OS) and disease-free survival (PFS). Patients with a CIBERSORT P value of ≤ 0.05 were included in the survival analysis. Log-rank survival analysis was employed to determine the effect of various immune cell infiltrations on patient OS and PFS. Statistical analysis was conducted with SPSS, R language (3.4.4), and GraphPad Prism. Differences among variables were determined to be statistically significant when the P value was ≤ 0.05 (one/two tails).

3. Results

3.1. Performance of CIBERSORT for Characterizing TIIC Composition of LUSC. CIBERSORT was firstly implemented to calculate the different TIIC subpopulations of LUSC tissue. As shown in Figure 2(a) and sFig 1, the proportion of immune cells in LUSC tissues is quite different from that in normal lung tissues. Meanwhile, the proportions of TIIC

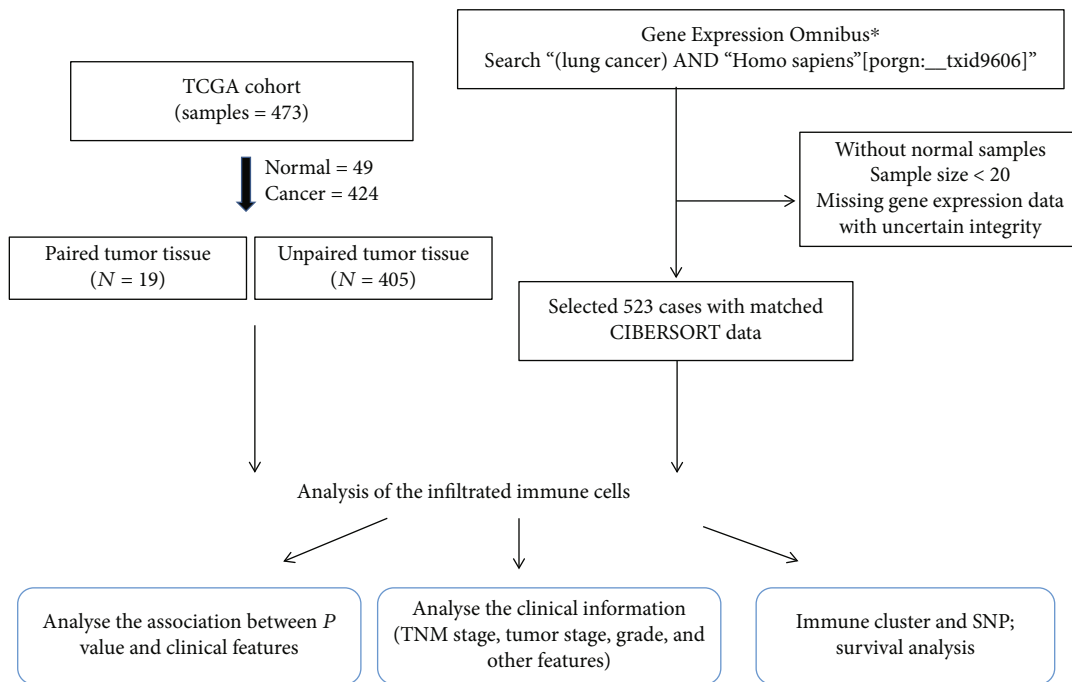


FIGURE 1: Flowchart detailing the procedure of sample collection and analysis.

composition showed a distinct group-bias clustering and individual difference (Figure 2(b)). We also particularly investigated the TIIC composition between matched cancer and adjacent tissues from 19 patients. As shown in the sFig 2, the majority proportions of immune cells were similar within intergroup. Compared to adjacent tissue, the proportions of plasma, M0 macrophage, and M1 macrophage were higher in LUSC tissues, whereas the fraction of resting CD4 memory T cells and M2 macrophage was relatively lower (Figures 2(c) and 2(d) and sFig 1, $P < 0.05$). On the other hand, we also checked the CIBERSORT outcome in which barcode genes were randomly removed in increments of 10% [18]. As expected, the P value was highly sensitive to the reduction of expressed barcode genes (Fig S3). Convincingly, these data did not show cohort bias, which speaks volume for the high reliability of our results.

3.2. Meta-Analysis of the Proportions of TIICs in LUSC Tissue. To confirm the accuracy of the above outcome, we further verified its accuracy in other independent LUSC chips both containing tumor and adjacent normal specimens. As shown in Table S2, we downloaded all of the eligible LUSC chips from GEO datasets. Overall, there were 289 LUSC cases and 234 normal tissue samples enrolled in the following exploration. Fig S4 represents the summary of each TIIC composition of the included studies. We merged different platform datasets and eliminated binding batch effect. As shown in Figure 2(e), the proportions of each chip's TIIC subpopulations showed no evident cohort bias to TCGA both in normal and cancer tissues.

As is universally acknowledged, meta-analysis is an effective standard way to sum up the results of studies rather than subjective judgment. So, we thoroughly conducted a meta-

analysis for all significantly different TIIC compositions. As shown in Figure 3, plasma cells (95% CI, 3.30-4.84; $P < 0.01$), Tfh (95% CI, 0.55 to 1.38; $P < 0.01$), Tregs (95% CI, 0.83 to 1.36; $P < 0.01$), and M1 macrophages (95% CI, 5.73 to 8.13; $P < 0.01$) exhibited an increasing tendency, whereas monocytes (95% CI, -2.10 to -1.02; $P < 0.01$), M2 macrophages (95% CI, -6.46 to -4.80; $P < 0.01$), and resting mast cells (95% CI, -2.41 to -1.11; $P < 0.01$) exhibited downtrends in LUSC tissues in agreement with the previous conclusions (except for resting CD4 memory). Thus, the CIBERSORT results were powerful enough to discriminate TIIC subpopulations in LUSC. Collectively, it highlighted the role of specific TIICs in LUSC, which could provide valuable candidates as diagnostic markers and potential therapy targets for patients.

3.3. Proportional Distribution of TIIC Subpopulation and Clinical Characteristics. We further combined clinical characteristics with TIIC composition to investigate whether the TIIC subpopulations were statistically associated with LUSC development. As shown in Figs S5A-S5D, the proportion of activated mast cells, resting mast cells, neutrophils, and T follicular helper cells were correlated with advanced T stage. And the proportions of M1 macrophage and CD8 T cell subpopulation had a strong connection with lymph node metastasis (Figs S5E and S5F). These TIIC subpopulations were also evaluated by univariate Cox regression analysis. As Figs S5G and S5H showed in the forest plot, M2 macrophages (OR: HR = 1.09, $P = 0.0093$), regulatory T cells (OR: HR = 1.53, $P = 0.029$), T follicular helper cells (OR: HR = 1.39, $P = 0.00862$), activated CD4 memory T cells (OR: HR = 1.30, $P = 0.0011$), M0 macrophages (OR: HR = 1.29, $P = 0.045$), and plasma cells (OR: HR = 1.58, $P = 0.043$) were

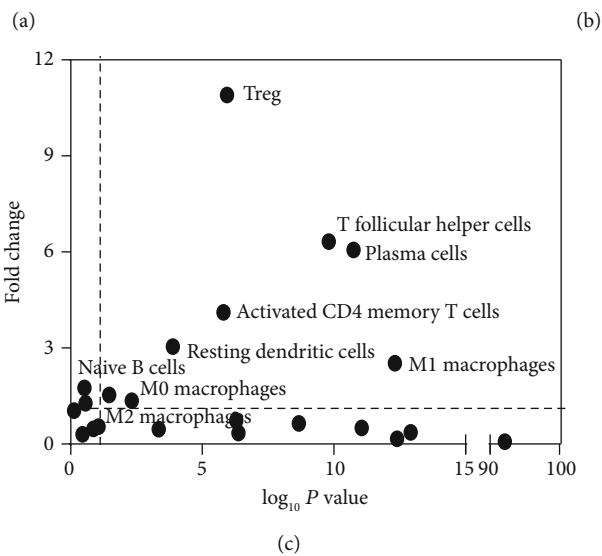
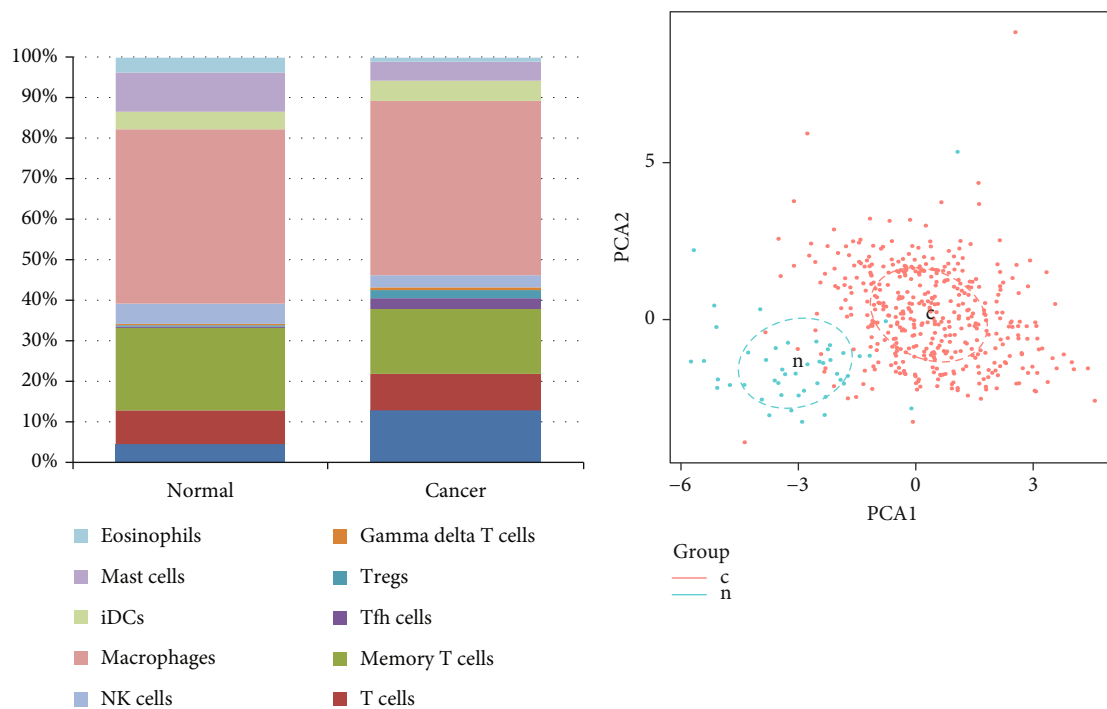


FIGURE 2: Continued.

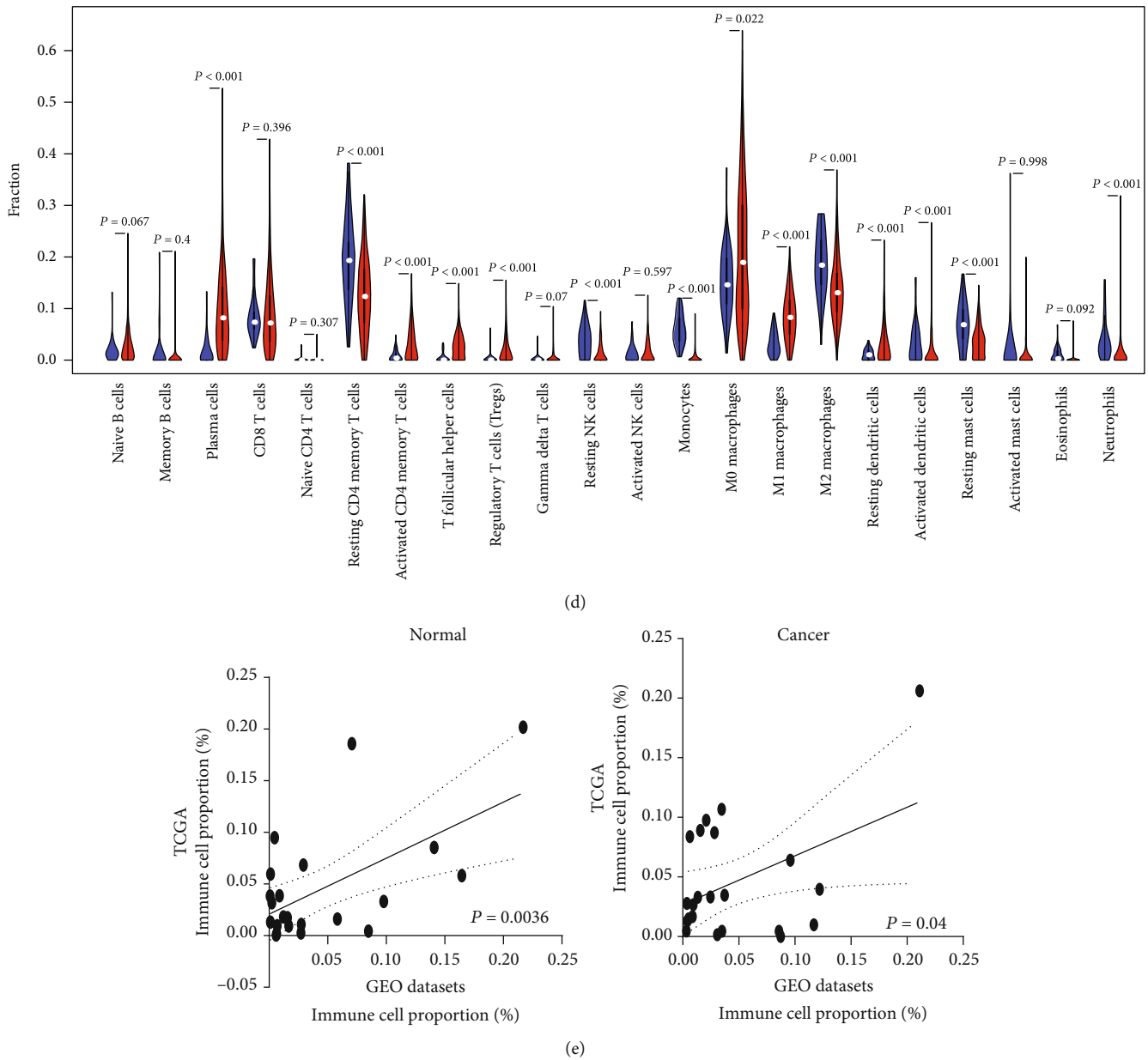
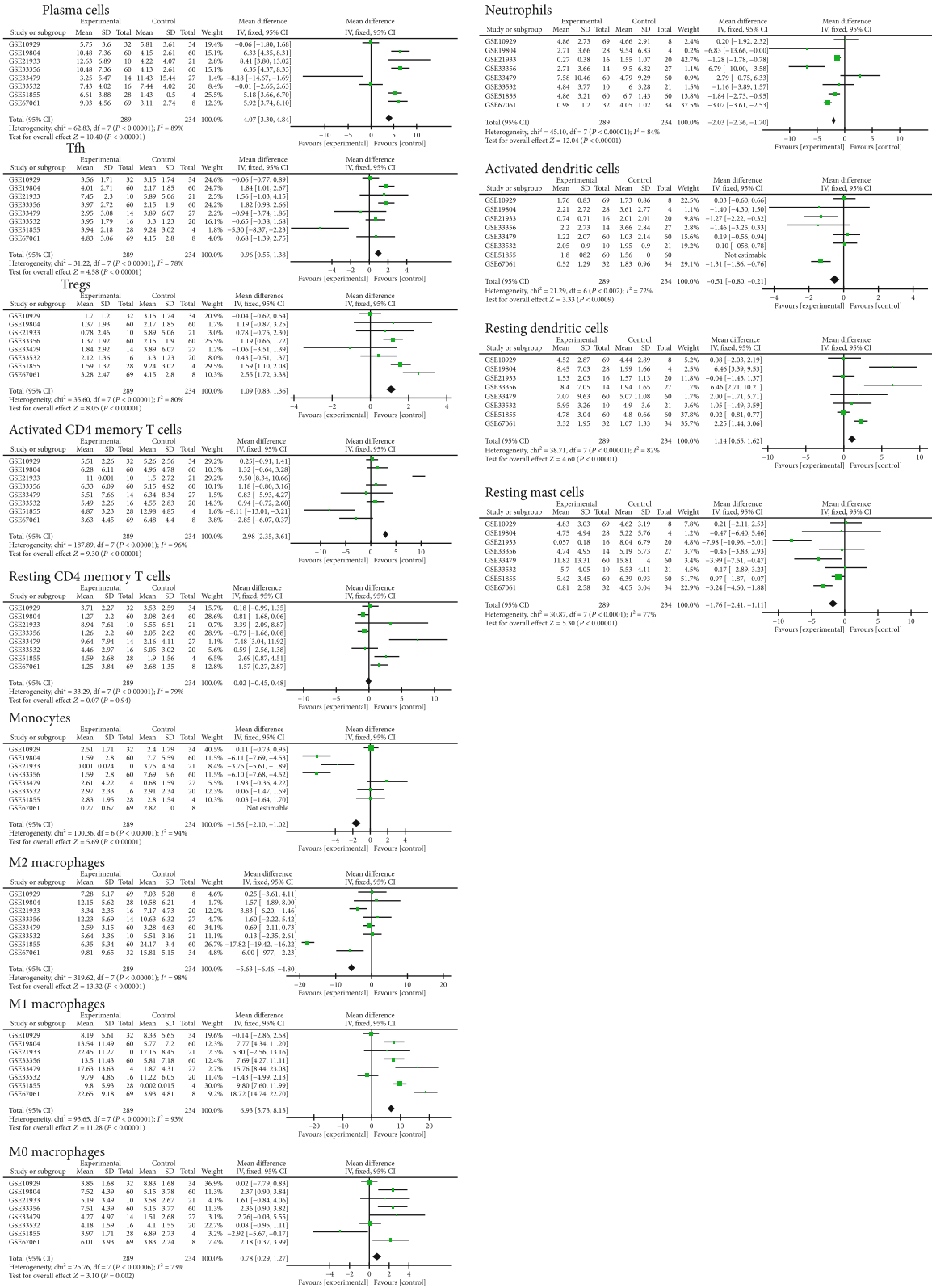


FIGURE 2: Landscape of microenvironment TIIC composition in LUSC. (a) The composition of TIICs of normal and LUSC tissues. (b) The proportions of TIIC composition from normal and cancer tissues displayed distinct group-bias clustering and individual differences. (c) Volcano plot visualizing the differentially TIIC composition. The point outside the dotted line represents the differential subpopulations with statistical significance ($P < 0.05$) between cancer and normal samples. (d) Violin plot of the proportions of TIIC subpopulation (blue represents normal tissue, red represents LUSC). (e) Relative proportions of 22 TIIC subpopulations are compared between two independent datasets (TCGA and GEO cohort).

significantly related with poor overall survival, whereas resting mast cells (HR = 0.74, $P = 0.00049$) were correlated with improved OS. Resting mast cells (OR: HR = 0.62, $P = 0.035$) and T follicular helper cells (OR: HR = 1.61, $P = 0.013$) were closely associated with PFS.

3.4. Identification of Immune Cluster in LUSC. The CIBERSORT P value reflected the ratio of immune cells and nonimmune cells, and a greater proportion of immune cells would produce a corresponding smaller P value. It is well acknowl-

edged that inflammation cytolytic activity could be described by the geometric mean of GZMA and PRF1 expression [14]. Strong relationship existed in P value and inflammation cytolytic activity in both the GEO and TCGA datasets (Figures 4(a) and 4(b), Fig S6). Meanwhile, cytolytic activity of inflammation was most strongly related with the proportion of CD8 T cells ($R^2 = 0.3274$, $P < 0.0001$) and activated memory T cells ($R^2 = 0.26$, $P < 0.0001$). Based on the ESTIMATE algorithm, we calculated each patient's immune score distributing between 0.68 and 1.14 and divided them into



(a) (b)
 FIGURE 3: Meta-analysis verified different expression TIIC composition in LUSC.

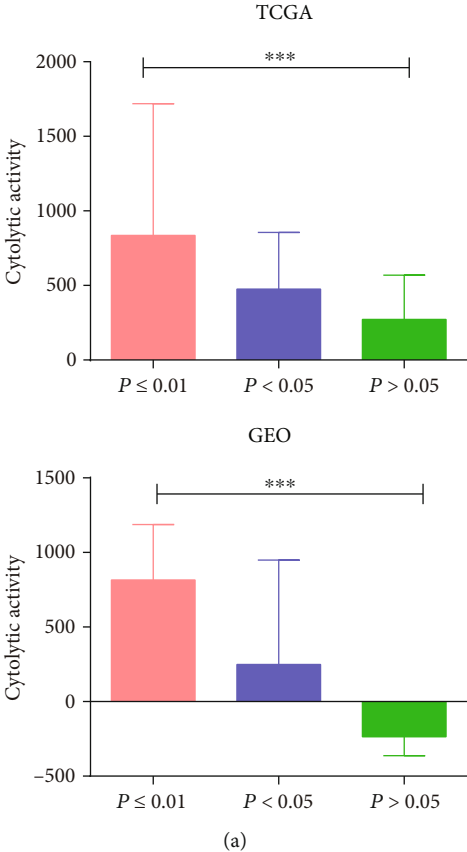
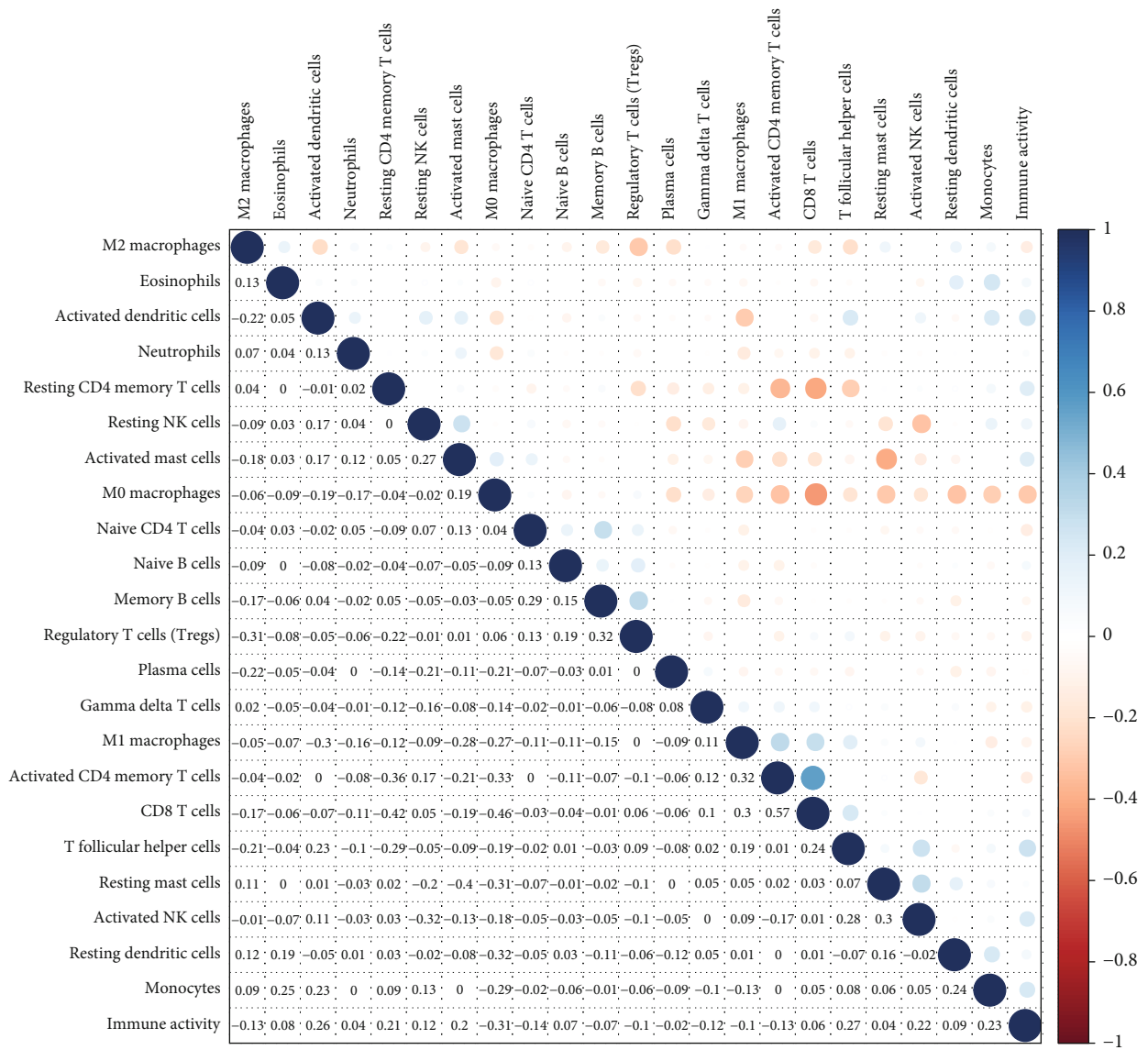
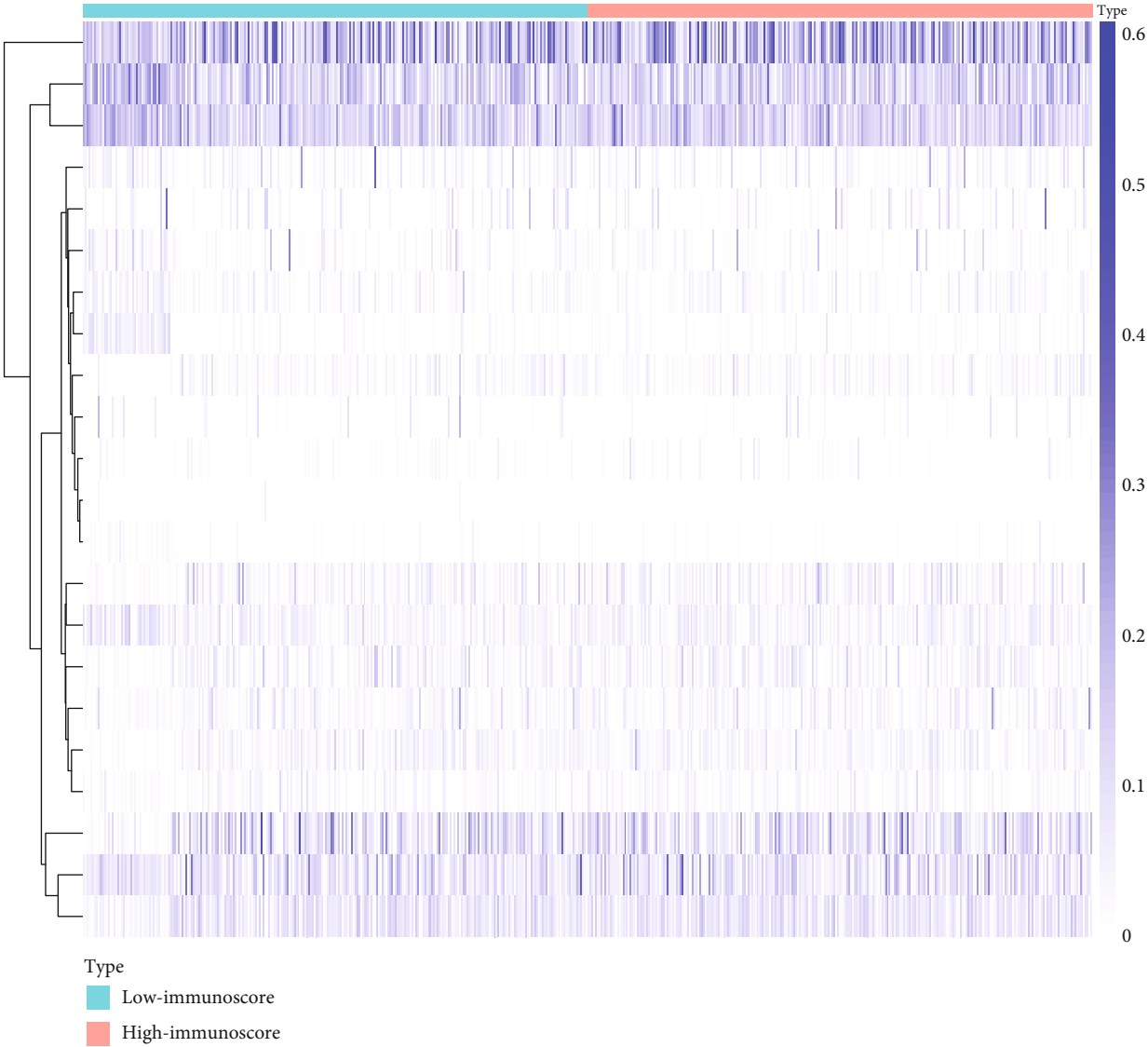


FIGURE 4: Continued.



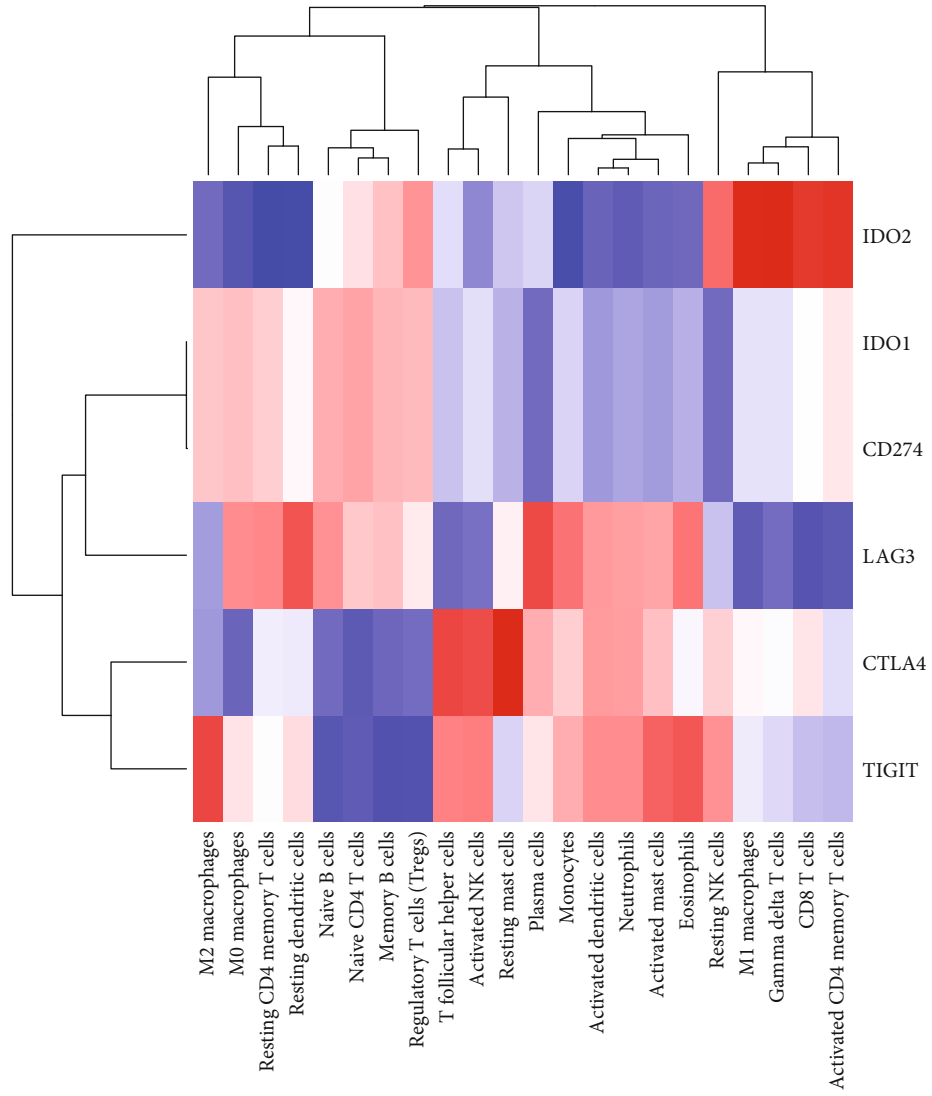
(b)

FIGURE 4: Continued.



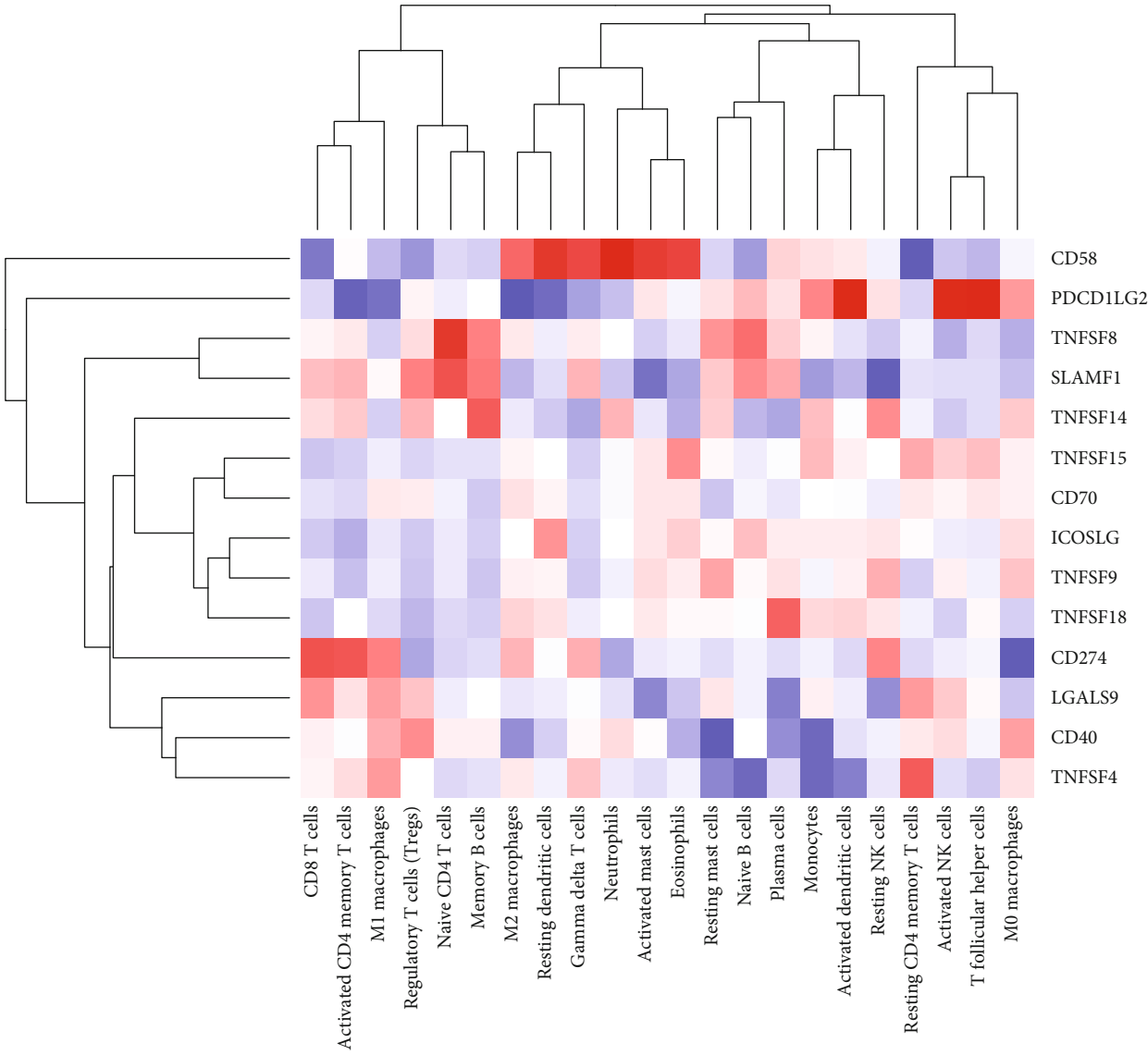
(c)

FIGURE 4: Continued.



(d)

FIGURE 4: Continued.



(e)
FIGURE 4: Continued.

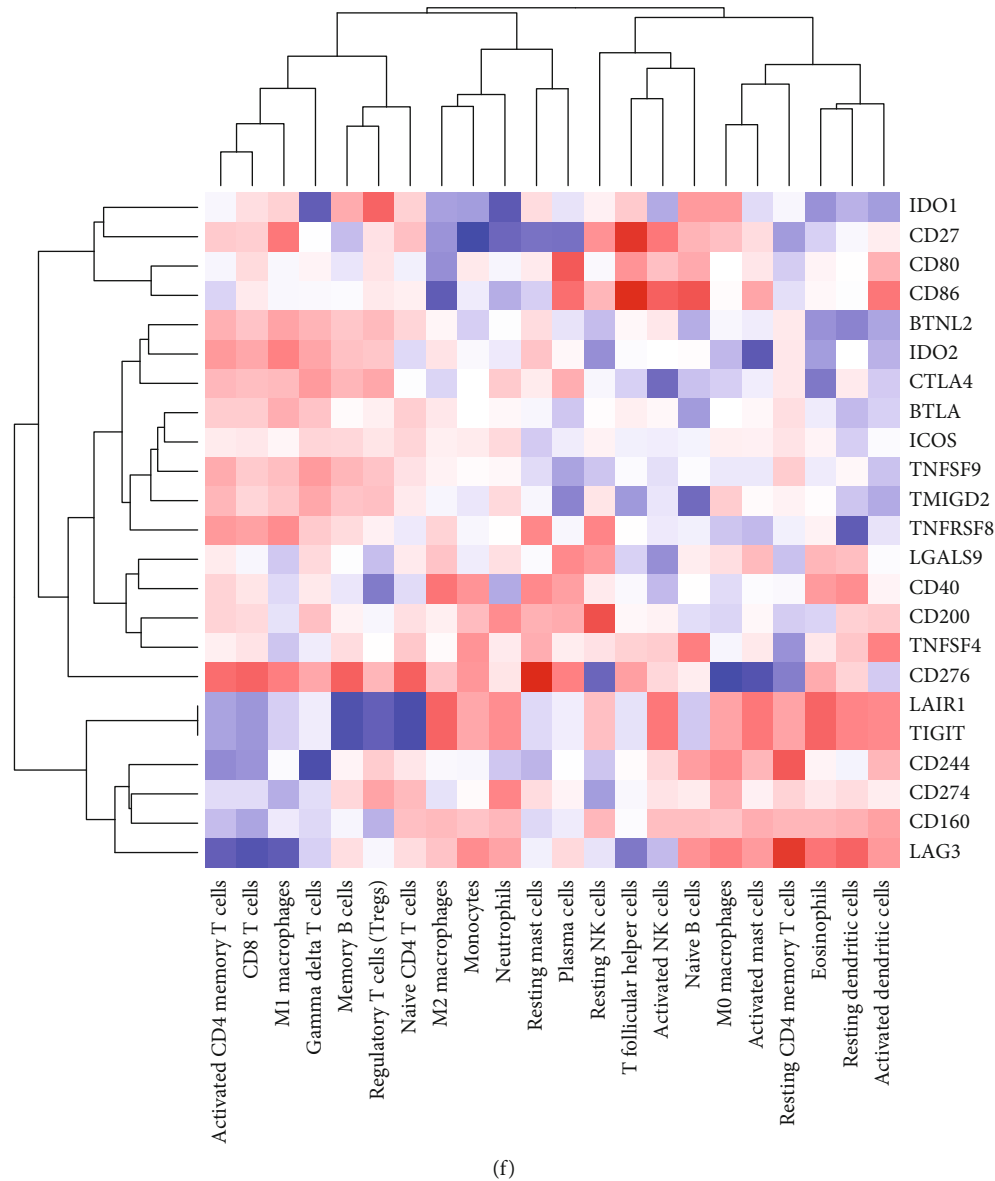
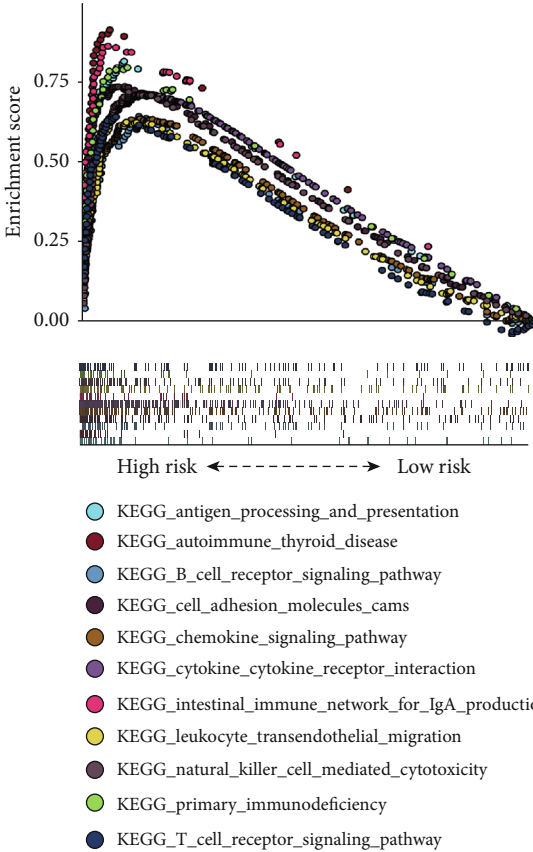


FIGURE 4: TIIC subpopulation was closely correlated with immune score. (a) CIBERSORT P values reflect the overall proportion of immune cells. (b) Spearman correlation matrix of all 22 TIIC compositions. (c) Heat map of the 22 TIIC compositions in high and low immunoscore subgroups. The horizontal axis shows the clustering information of samples which were divided into high and low major clusters. (d) The expression levels of TIIC composition are associated with PD-L1. (e) The expression levels of TIIC composition are associated with APC. (f) The expression levels of TIIC composition are associated with checkpoint.

high vs. low immune score groups in Figure 4(c). Then, we particularly studied the relationship between immune cells and specific immune signatures. Figures 4(d)–4(f), respectively, depict the relationship between immune cells and PD-L1, APC, and checkpoint. It was worth mentioning that the high immune fractions were notably enriched in the T cell receptor signaling pathway, B cell receptor signaling pathway, cytokine interaction, and chemokine signaling pathway (Figure 5(a)).

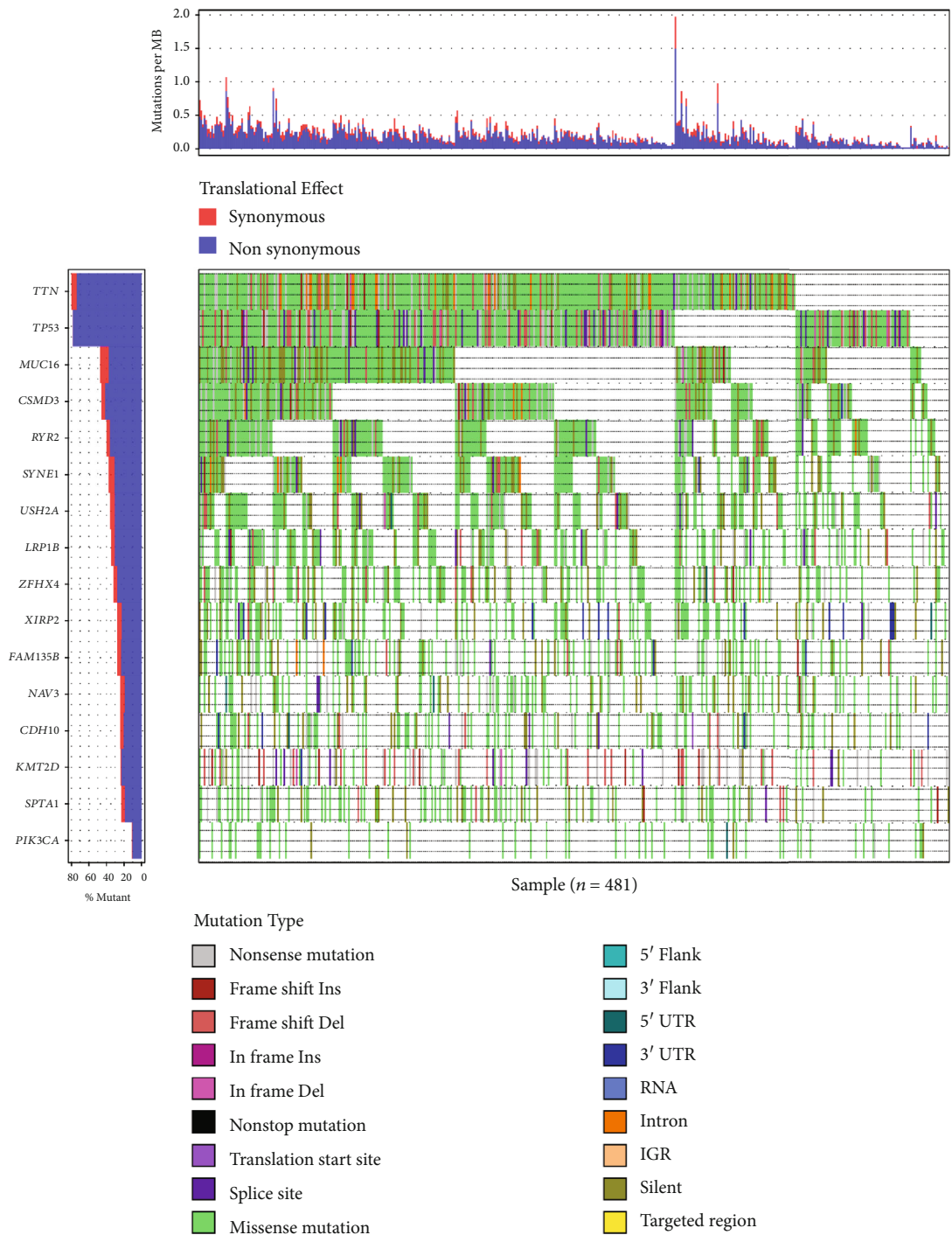
3.5. Immune Cells Associated with Prognosis and Molecular Subtypes. Single nucleotide polymorphisms (SNPs) most frequently occurred in genomic mutations. Then, we firstly analyzed the most frequent SNP mutation in LUSC. As shown in

Figure 5(b), TTN and TP53 (respectively, accounting for 81.08% and 80.87%) were the most characteristic SNPs in LUSC. TTN and TP53 mutant distinctly influenced the overall survival of LUSC patients (Figures 5(c) and 5(d)). There was a significant difference of TIIC proportion between the mutant and wild-type subgroups. Surprisingly, both TP53 and TTN mutant subgroups were reversely enriched in the T cell receptor signaling pathway, B cell receptor signaling pathway, cytokine interaction, and chemokine signaling pathway, which was opposite to immunoscore (Figures 5(a), 5(c), and 5(d)). Therefore, we made a detailed analysis of TIIC prognostic effect in TP53 and TTN molecular subtypes. As shown in Figures 5(e) and 5(f), resting dendritic cells played an important role both in TP53 (HR = 1.37, $P = 0.04$) and



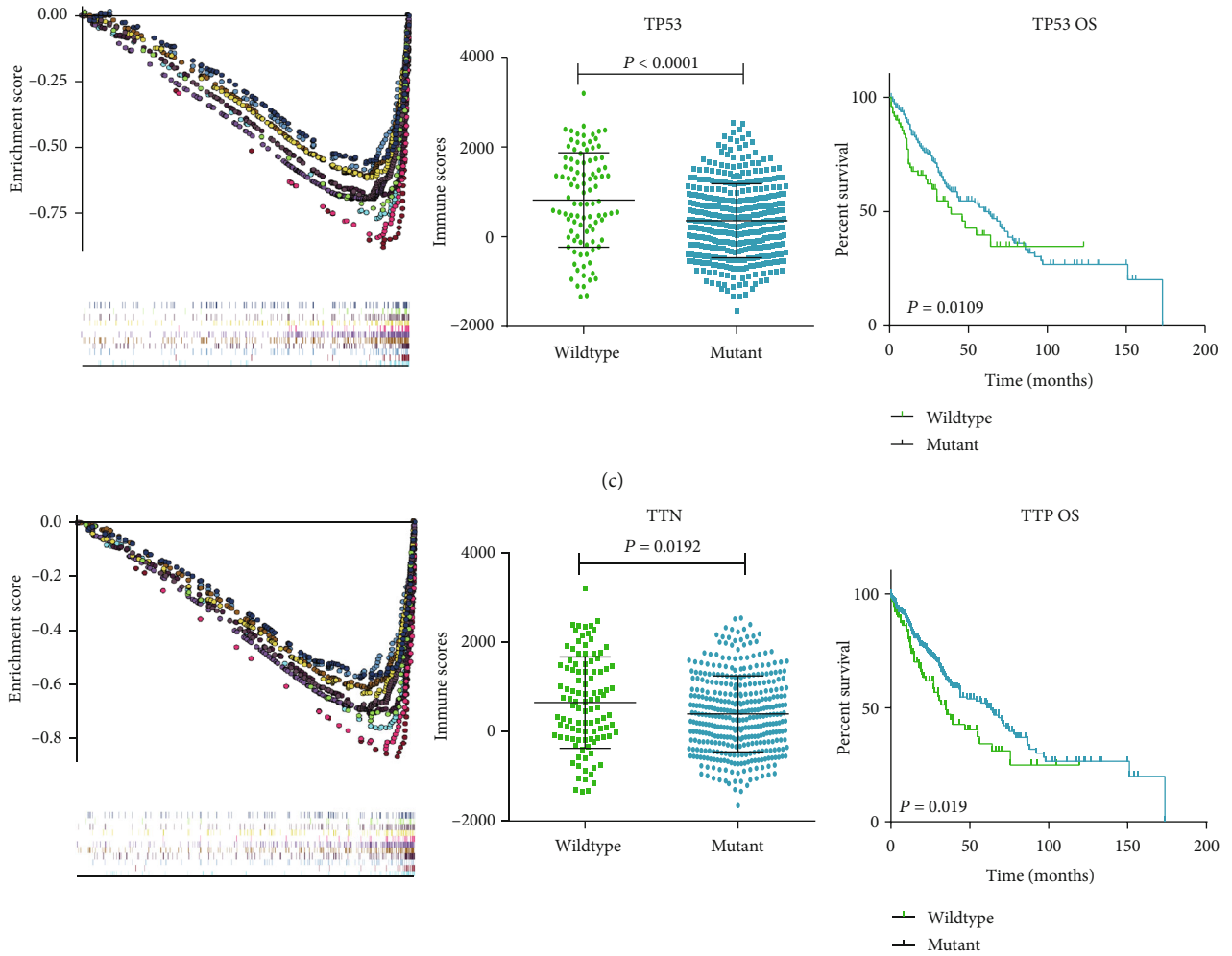
(a)

FIGURE 5: Continued.

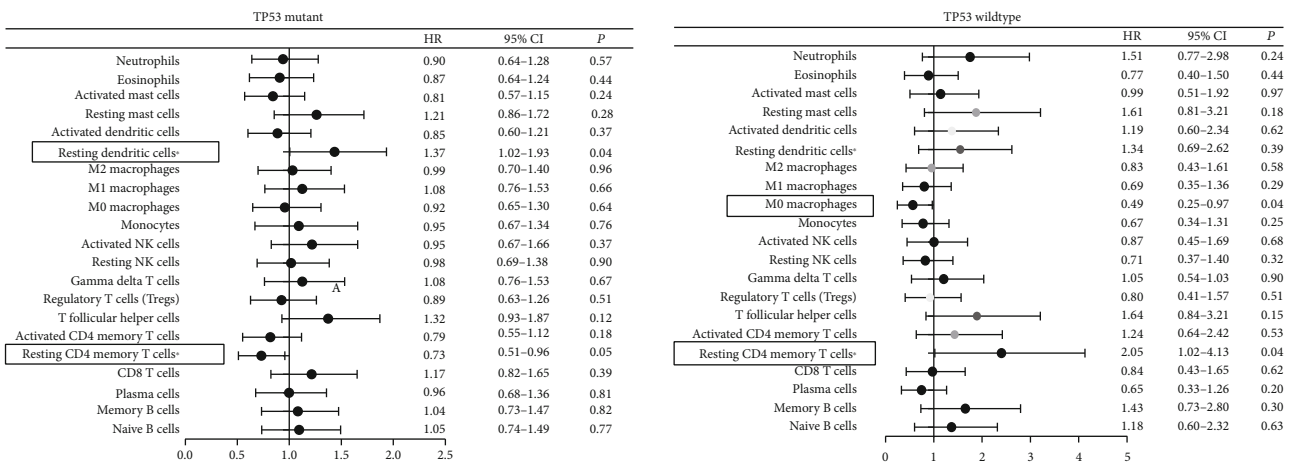


(b)

FIGURE 5: Continued.

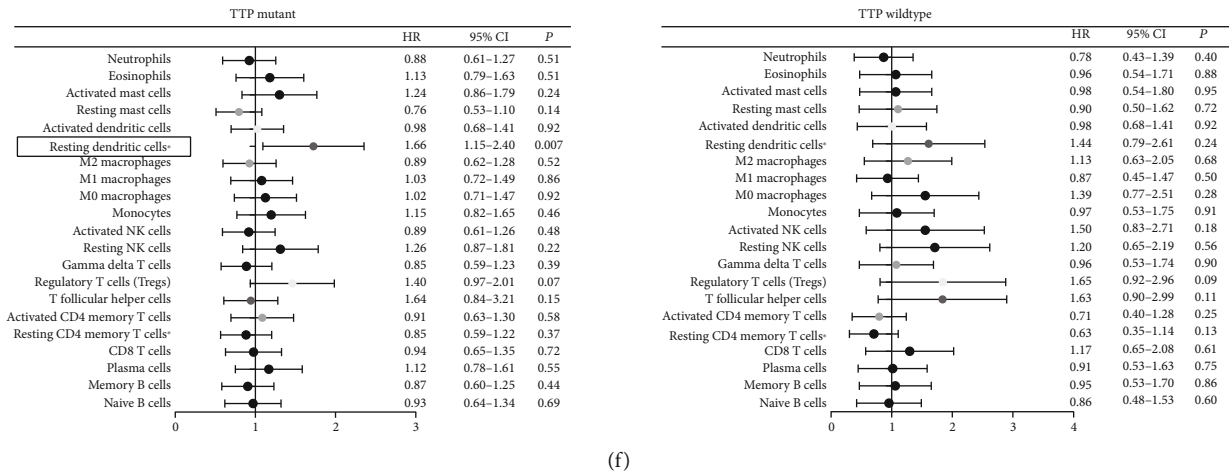


(c)



(d)

FIGURE 5: Continued.



(f)

FIGURE 5: Prognostic associations of subsets of immune cells and molecular subtypes. (a) GSEA differentiates the molecular mechanism of diverse immune score group. (b) Waterfall map depicts the top15 SNPs in TCGA-LUSC. (c, d) GSEA differentiates the molecular mechanism of top 2 SNP wild-type and mutant subgroups. Scattergram of the distribution of immune score between wild-type and mutant subgroups. Survival plots of SNP wild-type and mutant subgroups. (e) Subgroup overall survival analyses of TP53 mutant. The prognostic effect of 22 immune cell subsets by TP53 mutant and wild type. (f) Subgroup overall survival analyses of TTP mutant. The prognostic effect of 22 immune cell subsets by TTP mutant and wild type.

TTP (HR = 1.66, $P = 0.007$) mutant subgroups. Resting CD4 memory T cells were associated with a favorable outcome in the TP53 mutant subgroup (HR = 0.73, $P = 0.05$), whereas they were associated with an unfavorable outcome in the wild-type subgroup (HR = 2.05, $P = 0.04$).

3.6. Identification of Immune Cluster in LUSC. On the basis of our above findings, the change of TIIC subsets could significantly influence tumor progression and affect prognosis. In order to explore whether distinct patterns of immune infiltration can be distinguished, we performed hierarchical clustering analysis of 22 TIIC subpopulations and selected the optimal number of clusters by the elbow method. As shown in Figures 6(a) and 6(b), 3 clusters were identified as individualized clusters. Clusters were correlated with distinct immune patterns, and the survival curve is depicted in Figure 6(c). Cluster 1 was characterized by high levels of CD8 T cells, resting dendritic cells, and M0 macrophages. Cluster 2 was enriched with M1 macrophages and resting dendritic cells; cluster 3 was abundant with Tregs and activated dendritic cells. Collectively, characteristic immune clusters could influence clinical outcome (Figure 6(c)).

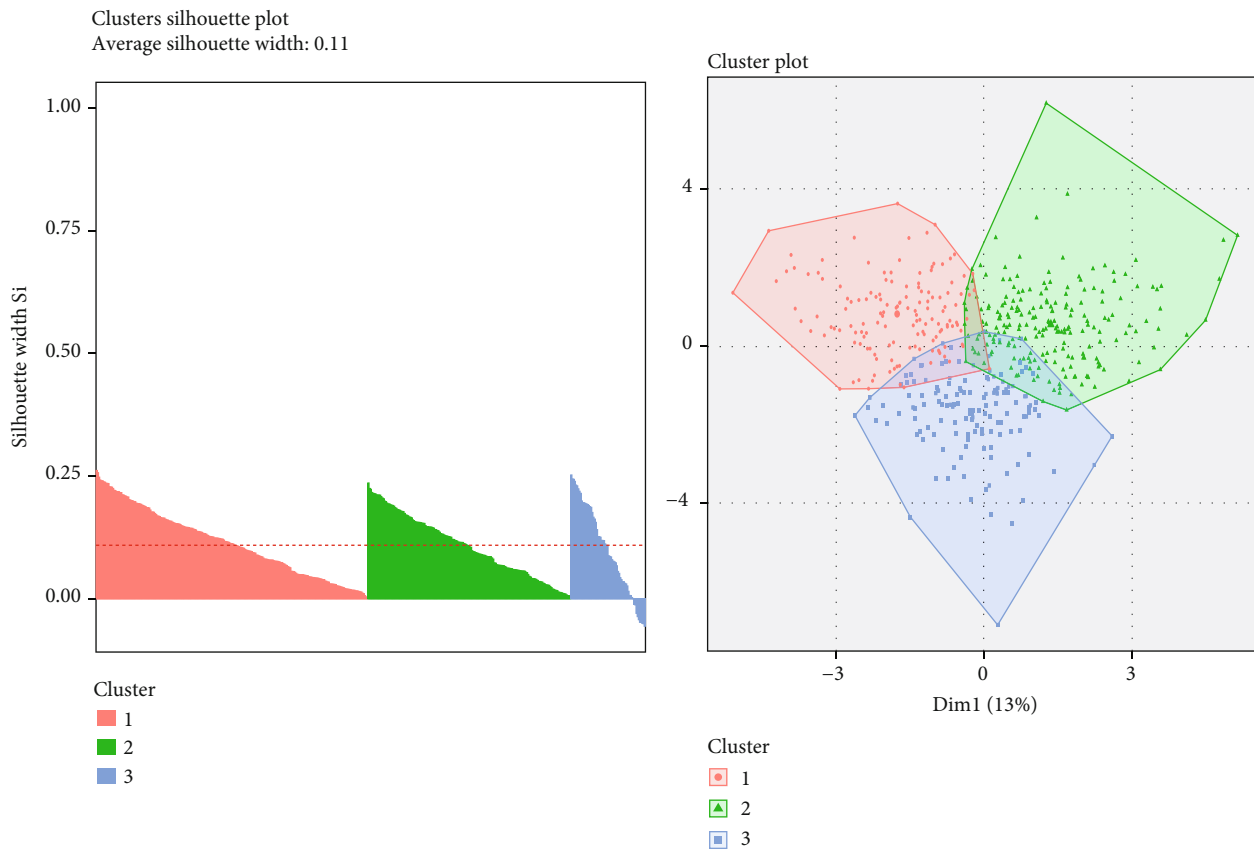
4. Discussion

Lung cancer is currently the leading cause of cancer-related deaths worldwide with a low 5-year survival rate [5]. The main treatments for pulmonary carcinoma are surgery, molecular targeted therapies, and immunotherapy which gradually emerge in recently years [4]. Researchers now focused on the complexity of the tumor microenvironment for its important role in tumorigenesis and suggested that the types and proportions of TIICs might be associated with cancer prognosis [19]. Thus, it is hope that exploring the underlying mechanisms of the relationship between the

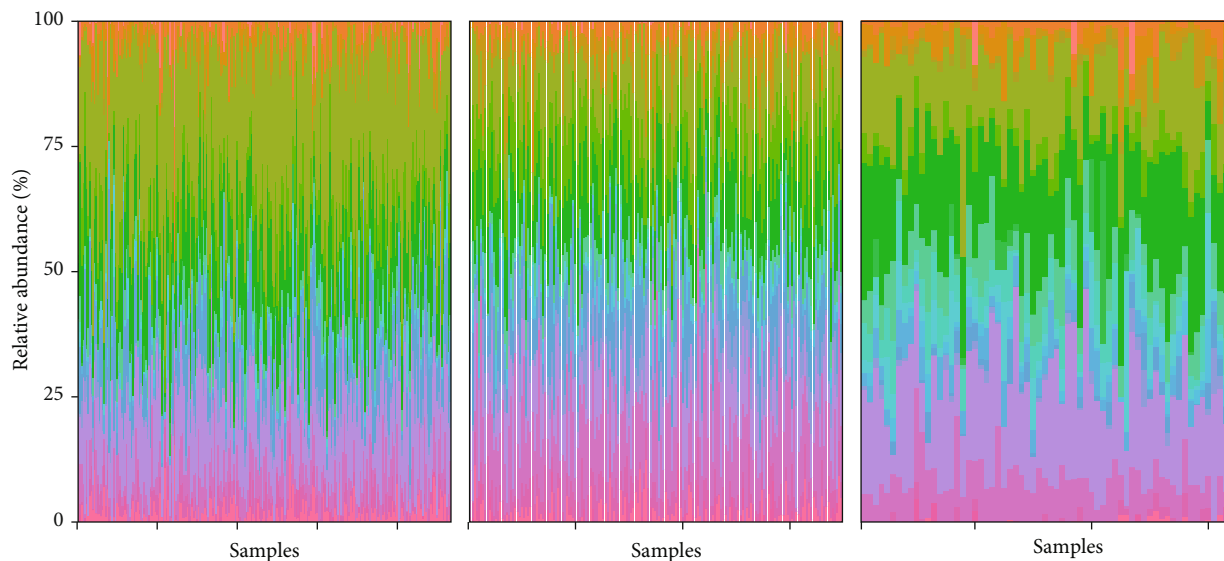
immune infiltrating cells and prognosis and diagnosis of lung cancer can contribute to discover more effective treatments.

In this study, CIBERSORT analysis, based on the gene expression profiles and deconvolution, was performed to obtain the proportions of 22 immune infiltrating cells in LUSC tissues and paracancerous tissues rather than the analysis of immunohistochemistry which relied on a single marker to distinguish TIIC subsets. Resting CD4 memory T cells, monocytes, M2 macrophage, resting mast cells, resting NK cells, and neutrophils accounted for higher proportions in LUSC tissues than adjacent tissues, opposite to the proportions of plasma cells, resting dendritic cells, M1 macrophage, regulatory T cells, T follicular helper cells, activated CD4 memory T cells, and activated NK cells. Macrophages are one of the main ingredients in tumor microenvironment, which exist in the center and margin of the tumor [19]. A series of experiments have confirmed that M1 macrophages are involved in antitumorigenesis and inflammatory response, while M2 macrophages have an opposite effect of M1 [20]. The effect on tumorigenesis in opposite directions of M1 macrophages and M2 macrophages in lung cancer was also confirmed in other studies [21]. Analysis with PCA plot also identified significant group-bias clustering and individual differences in the ratio of immune infiltrating cells. To further verify the reliability of the results, a meta-analysis was performed on 289 LUSC tissues and 234 adjacent tissues. Regulatory T cells, T follicular helper cells, activated CD4 memory T cells, and M1 macrophages were not conducive to tumorigenesis, while monocytes, M2 macrophages, and resting mast cells had protumor effects.

T follicular helper cells were at an advantage in its relation with favorable OS and PFS. Resting mast cells were bound up with poor OS and PFS. Mast cells are important regulators of the immune response [22]. However, the generalizability of the findings to mast cell inactivation in cancer is still unknown. Oxidized natural polyamines, a kind of



(a)



(b)

FIGURE 6: Continued.

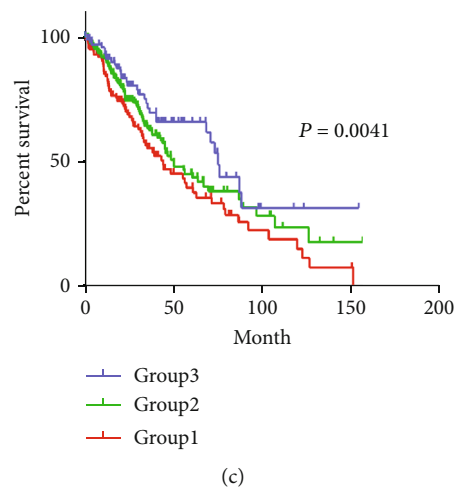


FIGURE 6: Immune hierarchical clusters associated with LUSC prognosis. (a) Consensus matrix heat map defining 3 clusters of samples. (b) Stacked bar charts of samples ordered by hierarchical cluster. (c) Kaplan-Meier curves for immune score cluster.

tumor-derived secretions, might be the reason leading to the inhibition of mast cells. Natural spermidine and spermine, oxidized by polyamine oxidase, can prevent IgE from living in vitro [25] and the level of polyamine was high in malignant cells [23, 24]. T follicular helper cells can express Bcl-6, IL-21, and CD40L signals to facilitate the proliferation and differentiation of B cells [25]. Research findings highlight the importance of immune checkpoint therapy to induce T follicular helper cell to activate B cells and inhibit tumor development [26]. Moreover, Tfh may be useful to develop or support ELs, which can recruit CD8+ T cells, NK cells, and macrophages to participate in antitumor immune responses [27]. Therefore, our findings raise the interesting possibility that T follicular helper cells and resting mast cells might be potential as biological markers in survival prognosis and diagnosis in LUSC patients.

Obvious enrichment of TTN and TP53 mutations in the T cell receptor signaling pathway, B cell receptor signaling pathway, cytokine interaction, and chemokine signaling pathway was observed in LUSC by analyzing mutation points of 481 tissue samples, which was opposite of immunoscore. Resting dendritic cells were positively associated with TP53 mutant and TTP mutant. The TTN mutant and TP53 mutant were noted to have better OS and PFS in contrast to the wild-type TTN and TP53. Resting CD4 memory T cells were negatively associated with TP53 mutation, which in turn had a positive correlation with wild-type TP53. As we know, TP53, a tumor suppressor gene, is the most common mutant gene in many malignancies and the mutation of TP53 is closely relevant to cancer progression, which can be found in around 50% of human cancers [28, 29]. In addition, TP53 mutations also had several strong links to poor prognosis across several cancers such as breast and colorectal [30]. Mutant TP53 has been verified for its ability to promote tumorigenesis, growth, migration, and invasion [31–33]. Most mutations in the coding region are missense mutations (87.9%), while missense mutations account for only about 40% outside of this region and most mutations are nonsense or frameshift mutations [34]. Because mutant TP53 protein

accumulates at higher levels in tumors, targeting mutant TP53, including renewing wild-type TP53 activity and depleting mutant TP53, at present shows to be a new therapeutic strategy [35]. Reactivating the resting CD4 memory T cells existing in the lung tumor microenvironment can induce brisker proliferative capacity and secretion of IFN- γ to eliminate tumor cells [36]. Blocking CTLA-4 can cause a dramatic expansion of the CTL response to TP53 and the expansion of memory T cells which was closely related to helper T cells [37]. Combined with our conclusions, there may be a corresponding connection between CD4 memory T cells and TP53 in the process of tumorigenesis and development.

In this study, we observed a strong association between the proportions of some TIIC subpopulations and immune cytolytic activity by analyzing the correlation matrix of each proportion and immune cytolytic activity between LUSC tissues and adjacent tissues. To ensure the reliability of the data, the data from TCGA and GEO databases were analyzed, respectively, and the trends of the results are roughly consistent. Moreover, in this study, T follicular helper cells and plasma cells were highly associated with TIM-3 receptors as well as existing high degrees of correlation between regulatory T cells and PDL-1. Programmed cell death protein-1 (PD-1), an important immune checkpoint receptor on the surface of immune cells, plays a pivotal role in regulating immune response; the dislocation and lack of PD-1 can cause autoimmune diseases [38]. PD-L1, one of the ligands of PD-1, has the ability to help cancer cells to evade the immune system and inhibit the antitumor immune response [39]. Inhibitors against PD-1/PD-L1 have proved to be effective in antitumor response in lung cancer and other tumors [40]. Compared with chemotherapy, pembrolizumab, nivolumab, and atezolizumab, inhibitors of PD-1/PD-L1 immune checkpoint can significantly improve overall survival of NSCLC patients [41–43]. High levels of Tregs are associated with increased tumor infiltration [44]. And the combination of PD-L1 and PD-1 can inhibit T cell receptor-mediated lymphocyte proliferation and cytokine secretion [45]. As

described above, it is worth to further explore the relationship between Tregs and PD-1 which might have potential therapeutic value in clinic.

After these, we divided LUSC cancer patients into 3 clusters according to the relative content of 22 TIIC subtypes. The majority of these samples were classified as cluster 1. Moreover, there were classification overlaps among these three clusters. Unlike cluster 1 and cluster 2, cluster 3 conferred better prognosis to patients. Cluster 1 was defined by high levels of CD8 T cells, resting dendritic cells, and M0 macrophages. Cluster 2 was enriched with M1 macrophages and resting dendritic cells, which was consistent with previous associations with survival outcomes. Each cluster has corresponding characteristic function enrichment term. These findings indicate that the character of immune infiltration across lung cancer has considerable variability, which can affect the clinical results.

Nevertheless, this retrospective work still has some limitations. Firstly, the data used for the CIBERSORT analysis was on the basis of TCGA and GEO databases, which lack basic information of patients and contain the unpaired samples. Secondly, though cohort bias had been eliminated by using statistical methods, potential heterogeneity in these data still exists. Thirdly, CIBERSORT can only estimate the relative abundance of immune cells, which means some cell types maybe overestimated or underestimated. Thus, if possible, further experiments would be performed in vivo and in vitro to verify these results and overcome some of these limitations.

In conclusion, we analyzed the proportions of the 22 immune cells in LUSC tissues and adjacent tissues, associated with tumorigenesis. And some specific immune infiltrating cells have the potential for diagnosis and prognosis of LUSC. Moreover, our findings of mutation points are also promising to also contribute to the implementation of immunotherapy and provide the possibility for the development of new immunotherapeutic drugs.

Abbreviations

| | |
|---------|--|
| LUSC: | Lung squamous cell carcinoma |
| TIICs: | Tumor-infiltrating immune cells |
| PCA: | Principal component analysis |
| ssGSEA: | Single-sample gene set enrichment analysis |
| NSCLC: | Non-small-cell lung cancer |
| DEGs: | Differently expressed genes |
| SMD: | Standard mean difference |
| CI: | Confidence interval |
| GSEA: | Gene set enrichment analysis |
| OS: | Overall survival |
| PFS: | Progression-free survival |
| SNP: | Single nucleotide polymorphism |
| Tregs: | Regulatory T cells. |

Data Availability

The data are available in TCGA and GEO datasets.

Conflicts of Interest

The authors have declared that no conflict of interest exists.

Authors' Contributions

WeiYang Cai and Jungang Zhao conceived and designed the experiments. Wenming Bao performed the data collection. WeiYang Cai analyzed the data. Jungang Zhao and Wenming Bao wrote the manuscript.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant No. 81703310) and the Scientific Research Foundation of the First Affiliated Hospital of Wenzhou Medical University (Grant No. FHY2019002).

Supplementary Materials

Supplementary Fig 1: specific TIIC proportion expression in normal and LUCS tissues. Supplementary Fig 2: the composition of TIICs of paired cancer and paracancerous tissue. Supplementary Fig 3: box plot of the distribution of CIBERSORT *P* value and average Pearson's correlation using datasets with progressively fewer (10% increments) barcode genes from the LUSC cohort. The *P* value was highly sensitive to the diminishing representation of the barcode genes. Supplementary Fig 4: landscape of GEO TIIC composition in LUSC. (A) Bar charts summarize the proportion of chips with different *P* value thresholds. (B) Bar charts summarize GEO chip concrete immune cell subset proportions of RCC tissues. Supplementary Fig 5: association between the TIICs and clinicopathological features in LUSC. (A–D) Infiltrating immune cells functioned in distinguishing the T stage. (E, F) Infiltrating immune cells functioned in distinguishing the N stage. (G, H) Forest plots showed the association with each immune cell subset and overall survival and PFS. Supplementary Fig 6: correlation matrix of TIIC proportion expression and inflammation activity. Table S1: the clinical information of TCGA LUSC patients Table S2: the detailed information of GEO chips. (*Supplementary Materials*)

References

- [1] K. C. Arbour and G. J. Riely, "Systemic therapy for locally advanced and metastatic non-small cell lung cancer: a review," *JAMA*, vol. 322, no. 8, pp. 764–774, 2019.
- [2] J. Vansteenkiste, L. Crinó, C. Dooms et al., "2nd ESMO Consensus Conference on Lung Cancer: early-stage non-small-cell lung cancer consensus on diagnosis, treatment and follow-up," *Annals of Oncology*, vol. 25, no. 8, pp. 1462–1474, 2014.
- [3] R. Timmerman, R. Paulus, J. Galvin et al., "Stereotactic body radiation therapy for inoperable early stage lung cancer," *JAMA*, vol. 303, no. 11, pp. 1070–1076, 2010.
- [4] F. R. Hirsch, G. V. Scagliotti, J. L. Mulshine et al., "Lung cancer: current therapies and new targeted treatments," *Lancet*, vol. 389, no. 10066, pp. 299–311, 2017.

- [5] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.
- [6] J. R. Brahmer, S. S. Tykodi, L. Q. Chow et al., "Safety and activity of anti-PD-L1 antibody in patients with advanced cancer," *The New England Journal of Medicine*, vol. 366, no. 26, pp. 2455–2465, 2012.
- [7] R. D. Bense, C. Sotiriou, M. J. Piccart-Gebhart et al., "Relevance of tumor-infiltrating immune cell composition and functionality for disease outcome in breast cancer," *Journal of the National Cancer Institute*, vol. 109, no. 1, article djw192, 2016.
- [8] B. Mlecnik, G. Bindea, H. K. Angell et al., "Integrative analyses of colorectal cancer show immunoscore is a stronger predictor of patient survival than microsatellite instability," *Immunity*, vol. 44, no. 3, pp. 698–711, 2016.
- [9] R. Zhou, J. Zhang, D. Zeng et al., "Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer," *Cancer Immunology, Immunotherapy*, vol. 68, no. 3, pp. 433–442, 2019.
- [10] H. S. Kim, J. H. Lee, S. J. Nam et al., "Association of PD-L1 expression with tumor-infiltrating immune cells and mutation burden in high-grade neuroendocrine carcinoma of the lung," *Journal of Thoracic Oncology*, vol. 13, no. 5, pp. 636–648, 2018.
- [11] A. R. Dadabayev, M. H. Sandel, A. G. Menon et al., "Dendritic cells in colorectal cancer correlate with other tumor-infiltrating immune cells," *Cancer Immunology, Immunotherapy*, vol. 53, no. 11, pp. 978–986, 2004.
- [12] S. Zhang, E. Zhang, J. Long et al., "Immune infiltration in renal cell carcinoma," *Cancer Science*, vol. 110, no. 5, pp. 1564–1572, 2019.
- [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] 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.
- [15] M. Foroutan, D. D. Bhuva, R. Lyu, K. Horan, J. Cursons, and M. J. Davis, "Single sample scoring of molecular phenotypes," *BMC Bioinformatics*, vol. 19, no. 1, p. 404, 2018.
- [16] M. M. Croken, W. Qiu, M. W. White, and K. Kim, "Gene Set Enrichment Analysis (GSEA) of *Toxoplasma gondii* expression datasets links cell cycle progression and the bradyzoite developmental program," *BMC Genomics*, vol. 15, no. 1, p. 515, 2014.
- [17] Z. Jiang, Z. Liu, M. Li, C. Chen, and X. Wang, "Increased glycolysis correlates with elevated immune activity in tumor immune microenvironment," *eBioMedicine*, vol. 42, pp. 431–442, 2019.
- [18] Y. Xiong, K. Wang, H. Zhou, L. Peng, W. You, and Z. Fu, "Profiles of immune infiltration in colorectal cancer and their clinical significance: a gene expression-based study," *Cancer Medicine*, vol. 7, no. 9, pp. 4496–4508, 2018.
- [19] S. Gnjatic, V. Bronte, L. R. Brunet et al., "Identifying baseline immune-related biomarkers to predict clinical outcome of immunotherapy," *Journal for Immunotherapy of Cancer*, vol. 5, no. 1, p. 44, 2017.
- [20] R. Noy and J. W. Pollard, "Tumor-associated macrophages: from mechanisms to therapy," *Immunity*, vol. 41, no. 1, pp. 49–61, 2014.
- [21] A. Yuan, Y. J. Hsiao, H. Y. Chen et al., "Opposite effects of M1 and M2 macrophage subtypes on lung cancer progression," *Scientific Reports*, vol. 5, no. 1, article 14273, 2015.
- [22] T. Nakano, C. Y. Lai, S. Goto et al., "Immunological and regenerative aspects of hepatic mast cells in liver allograft rejection and tolerance," *PLoS One*, vol. 7, no. 5, article e37202, 2012.
- [23] R. A. Casero Jr. and L. J. Marton, "Targeting polyamine metabolism and function in cancer and other hyperproliferative diseases," *Nature Reviews Drug Discovery*, vol. 6, no. 5, pp. 373–390, 2007.
- [24] N. Rohr-Udilova, F. Klinglmuller, R. Schulte-Hermann et al., "Deviations of the immune cell landscape between healthy liver and hepatocellular carcinoma," *Scientific Reports*, vol. 8, no. 1, article 6220, 2018.
- [25] S. K. Lee, R. J. Rigby, D. Zotos et al., "B cell priming for extra-follicular antibody responses requires Bcl-6 expression by T cells," *The Journal of Experimental Medicine*, vol. 208, no. 7, pp. 1377–1388, 2011.
- [26] D. P. Hollern, N. Xu, A. Thennavan et al., "B cells and T follicular helper cells mediate response to checkpoint inhibitors in high mutation burden mouse models of breast cancer," *Cell*, vol. 179, no. 5, pp. 1191–1206.e21, 2019.
- [27] S. Crotty, "T follicular helper cell biology: a decade of discovery and diseases," *Immunity*, vol. 50, no. 5, pp. 1132–1148, 2019.
- [28] S. Surget, M. P. Khoury, and J. C. Bourdon, "Uncovering the role of p 53 splice variants in human malignancy: a clinical perspective," *Oncotargets and Therapy*, vol. 7, pp. 57–68, 2013.
- [29] P. A. J. Muller and K. H. Vousden, "Mutant p53 in cancer: new functions and therapeutic opportunities," *Cancer Cell*, vol. 25, no. 3, pp. 304–317, 2014.
- [30] A. Petitjean, M. I. W. Achatz, A. L. Borresen-Dale, P. Hainaut, and M. Olivier, "TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes," *Oncogene*, vol. 26, no. 15, pp. 2157–2165, 2007.
- [31] P. A. J. Muller and K. H. Vousden, "p53 mutations in cancer," *Nature Cell Biology*, vol. 15, no. 1, pp. 2–8, 2013.
- [32] R. Brosh and V. Rotter, "When mutants gain new powers: news from the mutant p53 field," *Nature Reviews Cancer*, vol. 9, no. 10, pp. 701–713, 2009.
- [33] D. Dittmer, S. Pati, G. Zambetti et al., "Gain of function mutations in p 53," *Nature Genetics*, vol. 4, no. 1, pp. 42–46, 1993.
- [34] M. Olivier, M. Hollstein, and P. Hainaut, "TP53 mutations in human cancers: origins, consequences, and clinical use," *Cold Spring Harbor Perspectives in Biology*, vol. 2, no. 1, article a001008, 2010.
- [35] X. Yue, Y. Zhao, Y. Xu, M. Zheng, Z. Feng, and W. Hu, "Mutant p53 in cancer: accumulation, gain-of-function, and therapy," *Journal of Molecular Biology*, vol. 429, no. 11, pp. 1595–1606, 2017.
- [36] L. Broderick, S. J. Yokota, J. Reineke et al., "Human CD4+Effector memory T cells persisting in the microenvironment of lung cancer xenografts are activated by local delivery of IL-12 to proliferate, produce IFN- γ , and eradicate tumor cells," *Journal of Immunology*, vol. 174, no. 2, pp. 898–906, 2005.
- [37] J. Hernandez, A. Ko, and L. A. Sherman, "CTLA-4 blockade enhances the CTL responses to the p 53 self-tumor antigen," *Journal of Immunology*, vol. 166, no. 6, pp. 3908–3914, 2001.
- [38] Y. Iwai, M. Ishida, Y. Tanaka, T. Okazaki, T. Honjo, and N. Minato, "Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade," *Proceedings of the National Academy of Sciences*, vol. 102, no. 12, pp. 4912–4917, 2005.

- Sciences of the United States of America*, vol. 99, no. 19, pp. 12293–12297, 2002.
- [39] M. K. Callahan, M. A. Postow, and J. D. Wolchok, “Targeting T cell co-receptors for cancer therapy,” *Immunity*, vol. 44, no. 5, pp. 1069–1078, 2016.
- [40] J. P. Allison, “Immune checkpoint blockade in cancer therapy: the 2015 Lasker-DeBakey Clinical Medical Research Award,” *JAMA*, vol. 314, no. 11, pp. 1113–1114, 2015.
- [41] R. S. Herbst, P. Baas, D. W. Kim et al., “Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial,” *Lancet*, vol. 387, no. 10027, pp. 1540–1550, 2016.
- [42] J. Brahmer, K. L. Reckamp, P. Baas et al., “Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer,” *The New England Journal of Medicine*, vol. 373, no. 2, pp. 123–135, 2015.
- [43] H. Borghaei, L. Paz-Ares, L. Horn et al., “Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer,” *The New England Journal of Medicine*, vol. 373, no. 17, pp. 1627–1639, 2015.
- [44] Q. Gao, S. J. Qiu, J. Fan et al., “Intratumoral balance of regulatory and cytotoxic T cells is associated with prognosis of hepatocellular carcinoma after resection,” *Journal of Clinical Oncology*, vol. 25, no. 18, pp. 2586–2593, 2007.
- [45] G. J. Freeman, A. J. Long, Y. Iwai et al., “Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation,” *The Journal of Experimental Medicine*, vol. 192, no. 7, pp. 1027–1034, 2000.

Research Article

The Regulation of circRNA RNF13/miRNA-1224-5p Axis Promotes the Malignant Evolution in Acute Myeloid Leukemia

Rong Zhang, Yingchun Li, Hongtao Wang, Ke Zhu, and Guojun Zhang 

Department of Hematology, Shengjing Hospital of China Medical University, Shenyang, Liaoning, China

Correspondence should be addressed to Guojun Zhang; guazhui570980395@163.com

Received 24 July 2020; Revised 30 August 2020; Accepted 22 September 2020; Published 6 October 2020

Academic Editor: Tao Huang

Copyright © 2020 Rong 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.

Objective. To study the biological function of circular RNA RNF13 (circRNF13) in acute myeloid leukemia (AML) and its relationship with prognosis. **Methods.** We constructed stable AML cell lines with downregulated expression of circRNF13, and then, we explored the effect of downregulation of circRNF13 expression on the proliferation, migration, and invasion through qRT-PCR, MTT curve, colony formation, transwell migration and invasion experiment, cell cycle, apoptosis, Caspase 3/7 assay, and other experiments. We also studied the expression of C-myc and Tenascin-C by qRT-PCR to explore the role of circRNF13. **Results.** When the expression of circRNF13 was downregulated, the proliferation rate of AML cells decreased significantly, the cell cycle was blocked to G1 phase, and apoptosis rate increased significantly. C-myc related to cell proliferation decreased significantly at RNA level. Furthermore, when the expression of circRNF13 was downregulated, the migration and invasion ability of AML cells was significantly reduced, and the expression of Tenascin-C related to migration and invasion also decreased significantly. The luciferase reporter assay system confirmed that miRNA-1224-5p was the direct target of circRNF13. **Conclusion.** CircRNF13 inhibited the proliferation, migration, and invasion of AML cells by regulating the expression of miRNA-1224-5p. This study provides some clues for the diagnosis and treatment of AML.

1. Introduction

Acute myeloid leukemia (AML) is a kind of hematopoiesis system malignant disease with strong heterogeneity, often accompanied by a variety of genetic and genetic abnormalities [1–3]. AML is mainly characterized by uncontrolled proliferation, differentiation, and apoptosis of leukemia cells. FAB (French-American-British classification) is divided into M0-M7 type according to its morphological and histochemical characteristics [4]. At present, the treatment of AML is mainly chemotherapy. Most of the first diagnosed patients can get complete remission (CR) after combined chemotherapy, but the recurrence rate is as high as 70% [5]. The drug resistance, recurrence, and side effects during chemotherapy keep the mortality at a high peak. With the development of molecular genetic technology, various biological indicators of AML have been found gradually, so that we can deeply understand the pathogenesis and progress of AML from the molecular level. So far, a variety of cytogenetic and molecular indicators have been widely recognized in clinical practice

and become a powerful tool to assist diagnosis, guide treatment, and prognosis stratification. In recent years, a large number of researches focusing on DNA mutation, microRNA, and lncRNA are widely carried out, aiming to find more biological indicators of AML and improve the understanding of the characteristics of leukemia, so as to help monitor minimal residual lesions, improve the ability of early warning, and develop new targeted therapeutic drugs [6–8].

Circular RNAs are a kind of single-strand noncoding RNAs that widely exist in organisms [9]. They were first found in plant viruses in 1970s and then successively found in hepatitis D virus and yeast mitochondria [10]. In the early studies, circRNA was thought to be formed from the wrong splicing of exon transcripts. As a random product, it does not have a biological function, so the relevant reports are relatively rare. More and more evidences show that circRNA is not produced by chance. It has high abundance, stable structure, and special expression in time and space [11–13]. CircRNA is more stable than homologous linear RNA because it has a closed-loop structure and is not easy to be degraded.

According to the construction sequence, circRNA can be divided into three categories: intron circRNA, exon circRNA, and exon-intron circRNA. All kinds of circRNA are produced in different ways, with different sequences and structures, so their biological functions are different [14, 15]. At present, its physiological function and mechanism are still unclear, and what role it plays in various diseases is rarely reported. According to the results of previous literature, some circRNAs can be used as miRNA “sponge,” which can be used as translation templates to encode proteins or participate in the transcription of regulatory genes. CIRS-7 (CDT1as) is the first circular RNA to regulate miRNA and play a competitive adsorption mechanism to participate in tumor progression [16], which is upregulated in human malignant solid tumor cells. Further analysis showed that there were nearly 70 miRNA binding sites on CIRS-7 [17]. CIRS-7 can play the role of miRNA adsorbing sponge, competitively binding miR-7, which leads to the increase of target gene expression level downstream of miR-7 and promotes the growth and proliferation of malignant solid tumor cells [18]. Based on this, the researchers further pointed out that the discovery of CIRS-7 changed the concept of miRNA regulatory mechanism, made the mechanism of noncoding RNA regulatory miRNA more complex, and also provided a direction for further research of tumor molecular network and the development of therapeutic targets. The above phenomena suggest that the mechanism of microRNA adsorbing sponge may be a common biological phenomenon in cells and also indicate the possibility of the above mechanism in AML research. However, its expression and function in the process of AML are not clear.

Circular RNA hsa_circ_0001346 is produced from the RING finger protein 13 (RNF13) at chromosome 3q25.1. Some results showed that circRNF13 was downregulated nearly 2.98 times in lung adenocarcinoma [19]. However, whether circRNF13 can regulate the development of AML through the mechanism of ceRNA has not been reported. Therefore, in this study, we conducted a series of experiments to study the expression, function, and molecular mechanism of circRNF13 in AML.

2. Methods

2.1. Tissue Samples. The blood of ten cases of AML and ten healthy volunteers was taken from the patients admitted to our hospital, and the specimens were immediately stored in liquid nitrogen. Our study was approved by the Ethical Committee and Institutional Review Board of our Hospital. All participants provided written informed consent.

2.2. Cell Culture. HL60 and Kasumi-1 cells (ATCC, USA) were all cultured in DMEM (KeyGen, China) high glucose medium containing 1% penicillin and 10% FBS (Life Technologies, Australia). They were subcultured in 5% CO₂, 95% relative humidity, and 37°C constant temperature closed incubator.

2.3. Cell Transfection. At 24 h before transfection, the same number of cells was inoculated in each well of the 6-well

plate, and the cell saturation was 80%~90% at the time of transfection. The required volume of each plasmid is calculated according to the concentration of plasmid 2 g per plasmid. Four microliters of Lipofectamine 2000 transfection reagent was added to each tube. DMEM medium without FBS was added and diluted to 50 μ l, respectively. Plasmid solution was mixed with Lipofectamine 2000 solution (total volume of 100 μ l). After 48 hours of transfection, cells were collected to extract total RNA and protein. qRT-PCR was used to detect the silencing efficiency of RNF13. si-RNF13#1: sense sequence: 5'-CCACAUGAACGCCAGAGAAU-3', antisense sequence: 5'-AAUCUCUGGGCGUUAUGUGG-3'. si-RNF13#2: sense sequence: 5'-GUAUCCAGCGAAUCUGGA-3', antisense sequence: 5'-UCCAGAUUCGCGUGGAUAC-3'. The blank control siRNA (negative control siRNA, siRNA-NC): Sense sequence: 5'-UUCUCCGAACGUGUCACGUTT-3', antisense sequence: 5'-ACGUGACACGUUCGGAGAATT-3'.

2.4. RNA Extraction and qRT-PCR Detection. The total RNA was extracted using TRIzol Reagent (Invitrogen) and the cDNA was obtained using the PrimerScript™ Reagent Kit (Takara, Dalian, China). Fluorescence quantitative PCR detection was carried out according to the instructions. The reaction system was 10 μ l, and the reaction procedure was 94°C 3 min, 94°C 30 s, 60°C 20 s, 72°C 1 min, 40 cycles. The sequence was as follows: HMGB1F: 5'-ACATAAATTCAGAAAGGTGAT-3', R: 5'-ATATGCTAAAATGTCTGCTTC-3'. The primer sequences of β -actin are F: 5'-CGCTCTCTGCTCCTCTGTTTC-3', R: 5'-ATCCGTTGACTCCGACCTTCAC-3'; the primer sequences of 1224-5p are F: 5'-GGAGCAGCATTGTACAGG-3', R: 5'-CAGTGC GTGTCGTGGA-3'. The U6 primer sequences are F: 5'-GCTTCGGCAGCACATATACTAAAAT-3', R: 5'-CGCTTCACGAATTTGCGTGTGCAT-3'. Levels of gene expression were calculated by the 2^{- $\Delta\Delta$ Ct} method.

2.5. CCK8 Assay for Cell Proliferation. After the cells were digested by trypsin, they were suspended in DMEM medium and then counted by cell counting gun. The 96-well plate was inoculated with 100 μ l (3000 cells) per well. Each group was cultured for 7 days. The number of cells needed was calculated 4-5 times. The cell suspension was prepared with a DMEM medium containing 10% FBS. After repeated blowing and mixing, the cells were inoculated on a 96-well plate. 12-24 h after the plate laying, when the cell adhered to the wall completely, suck out the complete medium in the first hole measured and add 100 μ l DMEM medium containing 10% CCK8 into each hole. After incubation at 37°C for 1 h, remove 90 μ l from each hole to a new 96-well plate, and the optical density (wavelength at 450 nm) was measured by enzyme scale. After that, the data were measured and recorded at a fixed time every day [20]. During the experiment, the fluid was changed every two days. After six days, the data were statistically analyzed.

2.6. Colony Forming Experiment. After cell count, cells were diluted with DMEM medium containing 10% FBS, and cell

suspension was inoculated in a 6 cm cell culture dish. After shaking evenly, culture in a 37°C 5% CO₂ incubator and change the complete culture medium every 3-4 days. When the visible colony was formed in the cell culture dish, absorbed the culture medium, rinsed PBS twice, fixed it with 4% paraformaldehyde for 30min, then dyed it with methanol solution containing 0.5% crystal violet for 30min, rinsed it with clear water and dried it.

2.7. Dual-Luciferase Reporter Assay. Luciferase assay was divided into four groups: mutant group, NC group, miRNA-1224-5p group, and miRNA-1224-5p+circRNA RNF13 group. On the next day, when about 70% of the cells were fused, the luciferase plasmid containing miRNA-1224-5p mimic or inhibitor was cotransfected with lipofectamine 2000. After 48 h of conventional culture, it was measured and analyzed using the Dual-Luciferase Reporter Assay System (Promega). All experiments were repeated in independent triplicate.

2.8. Detection of Cell Cycle by Flow Cytometry. The cells were fixed overnight with glacial ethanol at 4°C. After washing the cells with phosphate buffer, 0.5 ml phosphate buffer and 50 µl ethidium bromide were added to each sample; then, 100 µg/ml RNaseA and 0.2% Triton X-100 were added. After incubated at 4°C for 30 min, flow analysis was conducted by flow cytometer (FACSCalibur; BD Biosciences). All experiments were repeated in independent triplicate.

2.9. Detection of Caspase-3/7 Activation Form. According to the instructions of Caspase-Glo 3/7 Assay kit, the logarithmic growth phase cells were selected and inoculated. The blank reaction group, negative control group, si-NC group, and si-RNF13 group were set in each experiment: the blank reaction group was cultured with a medium without cells; the si-NC group and si-RNF13 group were transfected with si-NC and si-RNF13, respectively. Three multiple holes were set in each group. After 48 hours of transfection, the 96-well plates of the cells in the si-NC group and si-RNF13 group were taken out from the 37°C incubator and balanced to room temperature. Caspase-Glo 3/7 reagent of the same volume was added with the culture medium into each pore, shook for 30 s, mixed well, and incubated for 1 hour in the dark at room temperature. Then, transfer the liquid from each hole to the opaque 96-well white plate and read the fluorescence value of each hole with Bertold Centro LB 960 microplate light detector.

2.10. Transwell Experiment. After cell counting, cell suspension was diluted with DMEM medium, and cell concentration was controlled at 4×10^3 cells/ml. 800 µl DMEM medium containing 20% FBS was added to the 24-well cell culture plate, and the cell was gently placed in the 24-well plate. Slowly add the cell suspension to the chamber, incubate at room temperature for 15 min, and then transfer it to a CO₂ incubator for 40 h. After 30 min of fixation and staining, rinse and gently wipe off the extra cells in the chamber with cotton swabs. Then, took photos with a microscope and count the number of cells passing through the chamber.

2.11. Statistical Analysis. The SPSS16.0 software was used for data analysis. *T*-test was used for the normal distribution of data between the two groups, and ANOVA was used for the comparison of the two groups in case of more groups. The difference was statistically significant as $p < 0.05$.

3. Results

3.1. Expression of circRNF13 in the Blood of AML Patients. We first measured the expression of circRNF13 in the blood of AML patients (10 cases) and healthy people (10 cases). It was found that in AML patients' blood samples, the expression of circRNF13 was at least 2 times higher than the average value of the expression in the blood of healthy people, which was significantly different from that in the blood of healthy people (Figure 1(a)). Therefore, we believed that circRNF13 may exist as an oncogene in the blood of AML patients.

3.2. Effect of Downregulation of circRNF13 on the Proliferation of AML Cells. Subsequently, we selected the cell lines with high expression of circRNF13 as experimental materials for the RNA interference experiment. Endogenous circRNF13 in the existing AML cell lines Kasumi-1 and HL60 was detected. As seen from the figure (Figure 1(b)), the expression of circRNF13 in HL60 cells is significantly higher than that in Kasumi-1 cells. Therefore, we chose HL60 cells for RNA interference to construct low expression cell lines. As seen from Figure 1(c), the expression of shRNA-1 circRNF13 in the low expression cell line decreased by about 35%, and the expression of circRNF13 in shRNA-2 decreased by more than 50%. Therefore, it can be concluded that the stable cell line with low expression of circRNF13 was constructed successfully, and the inhibition effect of the expression of circRNF13 was significant (Figure 1(c)).

In order to study whether the cell lines with low expression of circRNF13 have an effect on the proliferation of AML cells, we used the obtained control cell lines shRNA-NC and the low expression cell lines shRNA-1 and shRNA-2 to conduct MTT proliferation experiments. The proliferation rate of shRNA-1 as well as shRNA-2 decreased significantly 48 h (Figure 1(d)). Therefore, we can conclude that the downregulation of circRNF13 can inhibit the proliferation of AML cells. Furthermore, the number of clones in the low expression cell lines shRNA-1 and shRNA-2 was significantly reduced, and the number of colonies formed was significantly smaller (Figure 1(e)). Again, it showed that the downregulation of circRNF13 can inhibit cell proliferation, which was consistent with the results of the MTT proliferation experiment.

3.3. Effect of Downregulation of circRNF13 Expression on Cell Cycle and Apoptosis of HL60 Cells. Later, the cell cycle of AML cells after the downregulation of circRNF13 expression was detected. It can be seen from the figure that the cells in the G₀/G₁ phase enhanced evidently, while in the S phase did not change significantly and that in the G₂/M phase suppressed (Figure 2(a)). It can be concluded that when the expression of circRNF13 was

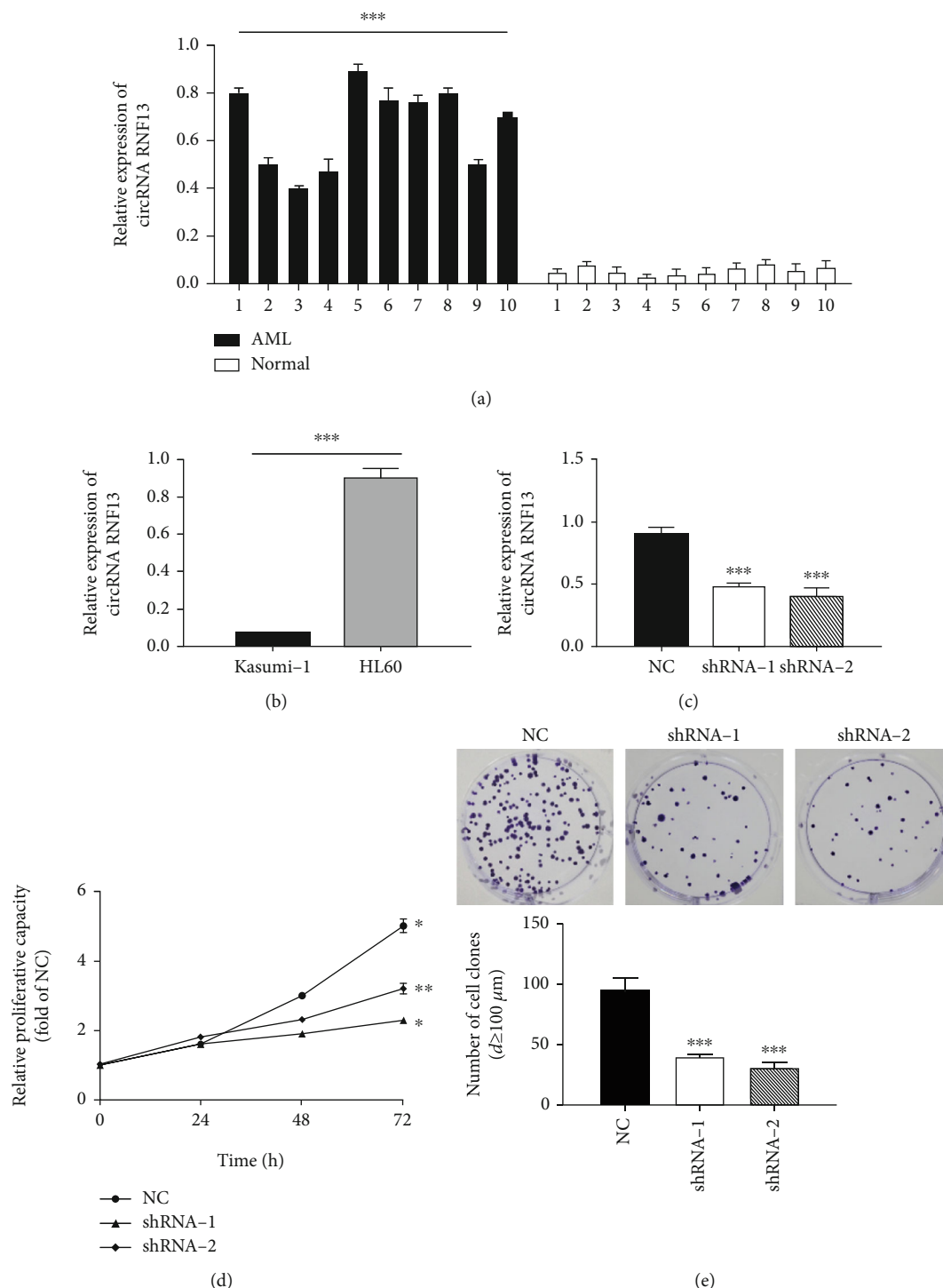


FIGURE 1: Effect of downregulation of circRNF13 on proliferation of AML cells. (a) Expression of circRNA RNF13 in AML and healthy blood. (b) Expression of circRNA RNF13 in AML cell line. (c) Detection of the expression of circRNA RNF13 in HL60 low expression stable cell line. (d) Effect of low expression of circRNA RNF13 on the proliferation of HL-60 cells. (e) Colony formation and statistical analysis of HL-60 cells after low expression of circRNA RNF13. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

downregulated, the cell cycle was blocked at G0/G1 and the cell proliferation rate was reduced.

In the previous experiment, we have detected that the expression of circRNF13 in AML increased significantly. In order to verify its effect on apoptosis, we carried out Annexin V/PI double staining experiment. It can be seen from the

result chart that the number of early apoptosis cells in the experimental group was 3-4 times that in the control group (Figure 2(b)). So we concluded that the downregulation of circRNF13 can promote the early apoptosis of AML cells. It was found that the relative activity of Caspase 3/7 in the experimental group was significantly higher

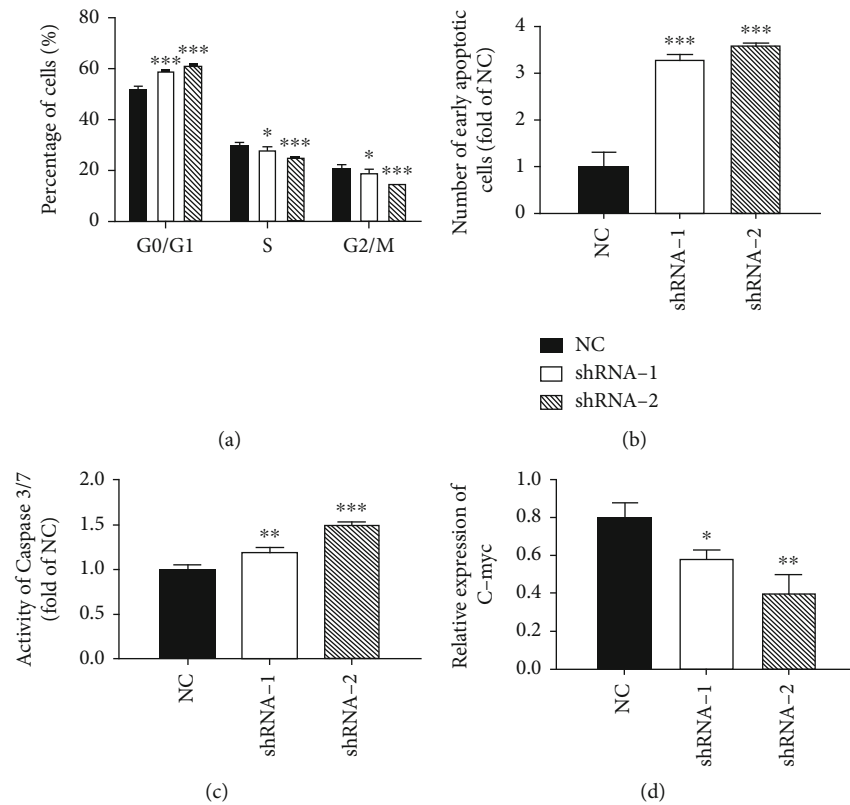


FIGURE 2: Effect of downregulation of circRNF13 expression on cell cycle and apoptosis of HL60 cells. (a) Effect of low expression of circRNA RNF13 on cell cycle of HL-60 cells. (b) Effect of low expression of circRNA RNF13 on early apoptosis of HL-60 cells. (c) Detection of caspase 3/7 in the low expression cell line of circRNA RNF13. (d) Expression of c-myc in the low expression cell line of circRNA RNF13. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

(Figure 2(c)). So it can be concluded that the downregulation of circRNF13 expression may promote the apoptosis by activating Caspase 3/7.

C- myc is an important regulator of cell proliferation. Therefore, we detected the expression of C-myc at the molecular level. As shown in Figure 2(d), the expression of shRNA-1 and shRNA-2 in C-myc low expression cell lines was lower. Therefore, we suggested that circRNF13 downregulated the expression of C-myc at the mRNA level.

3.4. Effect of Downregulation of circRNF13 Expression on the Migration and Invasion of AML Cells. In order to investigate the effect of low expression of circRNF13 on the migration ability of AML cells, we used Transwell cell for migration experiments. As shown in Figure 3(a), compared with the control cell line shRNA-NC, the number of cells passing through the basement membrane of the chamber in the low expression cell lines shRNA-1 and shRNA-2 decreased significantly. Therefore, the low expression of circRNF13 can inhibit the migration of AML cells. Compared with the control cell lines shRNA-NC, the number of cells with low expression shRNA-1 and shRNA-2 passing through Transwell cells decreased significantly (Figure 3(b)). Therefore, we believed that the low expression of circRNF13 can also

inhibit the invasion of AML cells, which was consistent with the migration results.

In order to test Tenascin-C, an important signal molecule affecting cell migration and invasion, we used qRT-PCR to measure Tenascin-C mRNA. As shown in Figure 3(c), the downregulation of circRNF13 expression significantly reduced Tenascin-C mRNA. Therefore, we suggested that the downregulation of circRNF13 may attenuate the migration and invasion of AML cells by inhibiting the expression of Tenascin-C.

3.5. Confirmation of circRNF13 as Direct Target of miRNA-1224-5p. In order to predict the miRNA interacting with circRNF13, we used bioinformatics to analyze it. When circRNF13 was the direct site of miRNA interaction, it can be used to identify the 3'UTR region or other sites for complete or incomplete complementary pairing, then further affecting the target gene. After preliminary screening, it was considered that miRNA 1224-5p was the interaction miRNA of circRNF13.

In order to confirm the correctness of the bioinformatics prediction results, we first transfected the stable cell line with the overexpression of circRNF13 into NC and miRNA-1224-5p, then carried out qRT-PCR 48 h later, and obtained the results as shown in Figure 4(a). In the stable cell line with overexpression of circRNF13, the expression of circRNF13

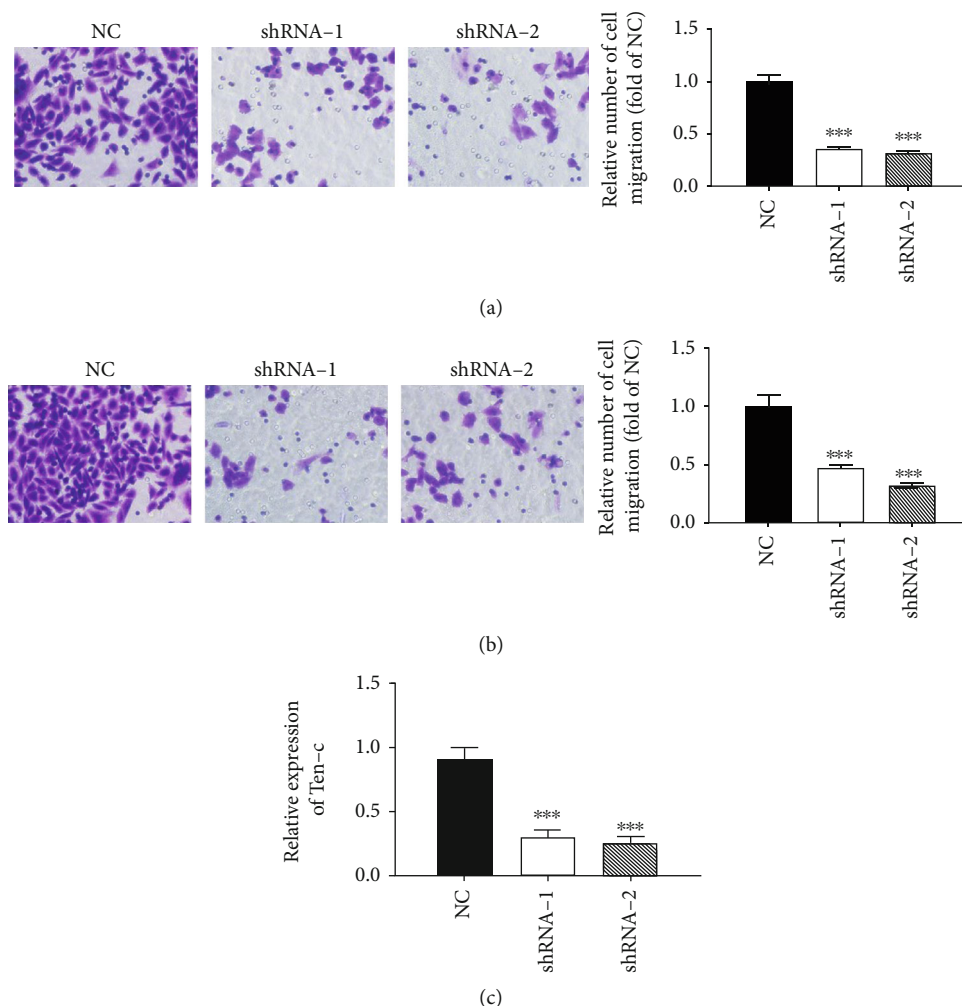


FIGURE 3: Effect of downregulation of circRNF13 expression on the migration and invasion of AML cells. (a) Migration and statistical analysis of HL-60 cells after low expression of circRNA RNF13 under a microscope. (b) Microscopically, the invasion of HL-60 cells after low expression of circRNA RNF13 and the results of statistical analysis. (c) Expression of tenascin-C in the low expression cell line of circRNA RNF13. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Magnification: 200x.

was significantly reduced at mRNA level after transfection of miRNA-1224-5p compared with the control group (Figure 4(b)). Therefore, we can preliminarily think that miRNA-1224-5p interacted with circRNF13, further inhibiting its expression.

As shown in Figure 4(c), the relative activity of firefly luciferase in the experimental group transfected with circRNF13-1, circRNF13-2, and circRNF13-3 plasmid recombinants was reduced. Therefore, we believed that the two predicted binding sites in circRNF13 can interact with miRNA-1224-5p and performed their functions. Therefore, we concluded that circRNF13 was the direct target of miRNA-1224-5p. The relative activity of firefly luciferase did not change significantly in the experimental group (Figure 4(c)). Therefore, we further proved that miRNA-1224-5p regulated the function of circRNF13 by recognizing specific sequences.

3.6. Effect of Overexpression of miRNA-1224-5p on the Proliferation of AML Cells. In order to investigate the effect

of overexpression of miRNA-1224-5p on the proliferation of AML cells, we first carried out the MTT proliferation curve experiment. We found that overexpression of miRNA-1224-5p can significantly reduce the cell proliferation ability through the MTT proliferation experiment (Figures 5(a) and 5(b)). Then, in order to further verify the inhibitory effect of miRNA-1224-5p on cell proliferation, we also carried out cell colony formation experiments. The same number of cells was inoculated in the six-well culture plate. After 10-14 days of culture, the colonies were counted and observed, and then, the colonies with a diameter of more than 100 μm were statistically analyzed with an inverted microscope. Overexpressed miRNA-1224-5p decreased significantly, and the number of formed colonies was smaller (Figure 5(c)). The results of the clonogenesis experiment were consistent with the results of the MTT proliferation curve, so we thought that the overexpression of miRNA-1224-5p can attenuate the proliferation. In addition, the cell cycle was blocked in G0/G1 after the overexpression of miRNA-1224-5p, which led to the slowdown of cell

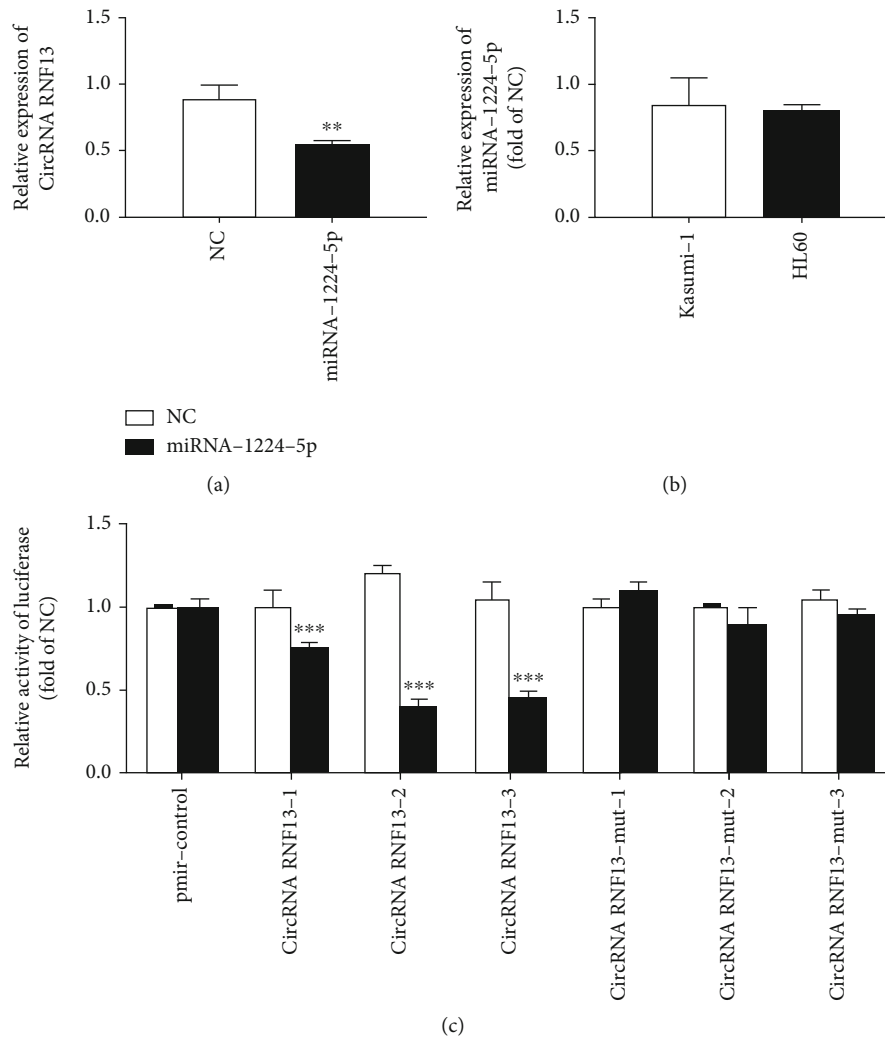


FIGURE 4: Confirmation of circRNF13 as a direct target of miRNA-1224-5p. (a) Effect of overexpression of miRNA-1224-5p on the expression of circRNA RNF13 in Kasumi-1 cells. (b) Expression of miRNA-1224-5p in AML cell line. (c) Luciferase double report system test results. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

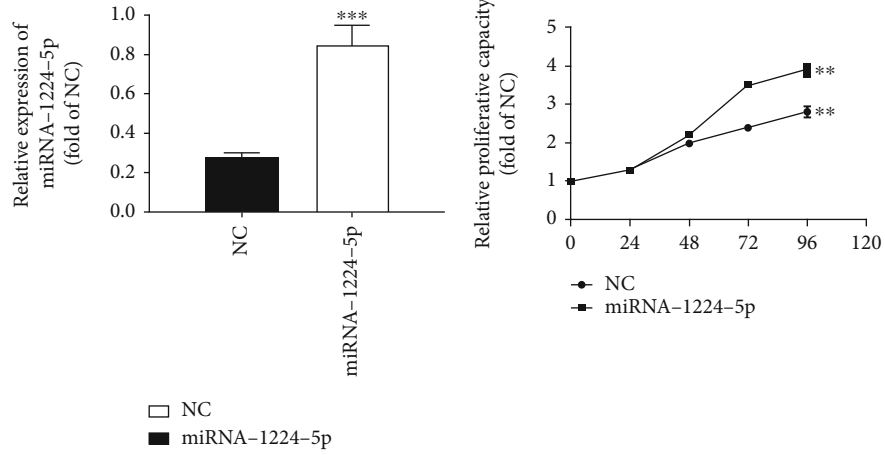
proliferation (Figure 5(d)). Overexpression of miRNA-1224-5p can promote the early apoptosis of AML cells (Figure 5(e)). After overexpression of miRNA-1224-5p, Caspase-3/7 was activated to promote the apoptosis of AML cells, which was consistent with the results of flow cytometry (Figures 5(f) and 5(g)).

3.7. Effect of Overexpression of miRNA-1224-5p on the Migration and Invasion of AML Cells. In order to further study the effect of overexpression of miRNA-1224-5p on the metastasis and infiltration of AML cells, we used the Transwell cell to detect its migration, cultured in the cell for 24 h, stained with crystal violet, observed, and photographed under the microscope. As shown in Figure 6(a), compared with NC, the number of cells passing through Transwell's ependyma after overexpression of miRNA-1224-5p decreased significantly, which showed that the overexpression of miRNA-1224-5p can inhibit the migration of AML cells. Overexpression of miRNA-1224-5p also inhibited the invasion of AML cells, which was consistent with the migra-

tion results (Figure 6(b)). In order to test Tenascin-C, an important signal molecule affecting cell migration and invasion, we used qRT-PCR to measure Tenascin-C mRNA. As shown in Figure 6(c), the overexpression of miRNA-1224-5p decreased the expression of Tenascin-C at mRNA level. It was suggested that overexpression of miRNA-1224-5p may inhibit the migration and invasion of AML cells by inhibiting the expression of Tenascin-C.

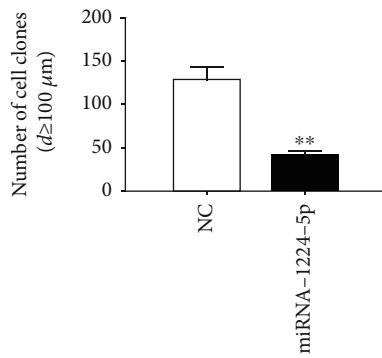
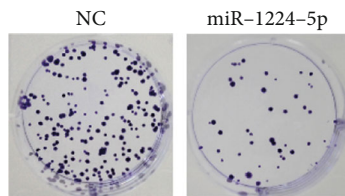
4. Discussion

CircRNA is closely related to human diseases, especially in tumors. CircRNA can be secreted into body fluids, such as saliva, blood, and exosomes [21, 22]. In clinical standard blood samples, hundreds of circRNA are more abundant than the corresponding linear mRNAs, so circRNA can be used as a new tumor marker in clinical detection [23]. In recent years, the high expression level of circRNAs in AML, including hsa_circ_0004277, hsa_circ_00750, has been screened by a gene chip [24, 25]. It is found that the

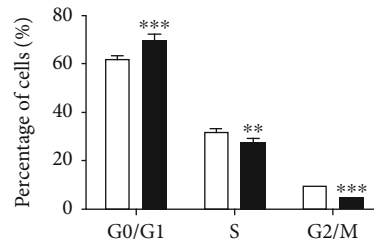


(a)

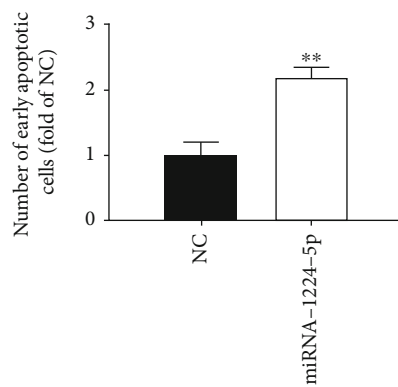
(b)



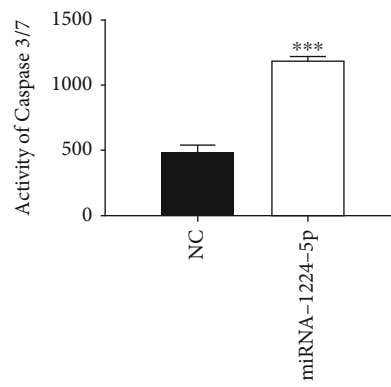
(c)



(d)



(e)



(f)

FIGURE 5: Continued.

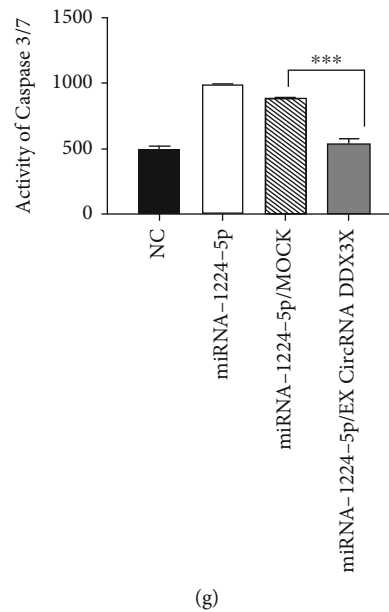


FIGURE 5: Effect of overexpression of miRNA-1224-5p on the proliferation of AML cells. (a) Detection of transfection efficiency of Kasumi-1 cells transfected with miRNA-1224-5p. (b) Effect of overexpression of miRNA-1224-5p on the proliferation of Kasumi-1 cells. (c) Microscopically, the colony formation and statistical analysis of Kasumi-1 cells after overexpression of miRNA-1224-5p. (d) Effect of overexpression of miRNA-1224-5p on the cell cycle of Kasumi-1. (e) Effect of overexpression of miRNA-1224-5p on the early apoptosis of Kasumi-1 cells. (f) Detection of caspase 3/7 in Kasumi-1 cells after the overexpression of miRNA-1224-5p. (g) The recovery effect of miRNA1224-5p-induced apoptosis after the overexpression of circRNA RNF13. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

expression levels of these circRNAs in AML present corresponding dynamic changes with the evolution of AML. At the biological level, they have competitive or cleavage effects with homologous linear RNAs and have high potential as prognostic indicators of AML. It is predicted that this may be related to the absorption of miRNA as a “sponge.” Therefore, the function of these circRNAs may be worth exploring in our future research projects.

hsa_circ_RNF13, located in chr3:149563797-149639014, is derived from the ring finger protein 13 (RNF13) gene. It was found that the expression of circRNF13 in cancer tissue was 2.98 times lower than that in the surrounding normal lung tissue. In vitro, circRNF13 can inhibit the invasion and metastasis of the lung adenocarcinoma cell line. Bioinformatics analysis and RIP experiments showed that circRNF13 can interact with RNA binding protein Ago2 and act as a sponge of miR-93-5p, which provided a new way to further study the molecular function of circRNF13 [19]. The above data indicated that circRNF13 was a new potential LAC biomarker and therapeutic target.

In our study, we found that the downregulation of circRNF13 expression suppressed the proliferation, migration, and invasion of AML cells and studied its mechanism. Through literature review, it was found that miRNA and 3'UTR region or other parts of mRNA were complementary pairing and that miRNA and lncRNA could also be a completely complementary pairing or incomplete complementary pairing. Therefore, we used bioinformatics to predict the miRNA of the direct effect of circRNF13. After the prediction and literature search, we believed that miRNA-1224-5p regulated tumor cells with circRNF13 as the target.

At present, it has been found that miRNA-1224-5p is abnormally expressed in a variety of tumors, and it can be used as an oncogene or as an antioncogene and can be regulated at the protein expression levels of Caspase-3, Bcl-2, and Bax [26]. It was found that the expression of miRNA-1224-5p in lung cancer changed significantly and inhibited the formation of keloid fibroblasts [27, 28].

Our study of circRNF13 is the first time to be studied in AML. With its involvement in the mechanism of action of AML gradually explored, circRNF13 may become a new target of targeted therapy. At the same time, this subject still needs further research, including the mechanism of a possible specific function, the molecules of interaction, the pathway involved, whether it is verified in animal experiments, and whether it can be easily detected in CSF and blood. Meanwhile, with the discovery of more and more circRNA, its function has been paid more and more attention by researchers. The function and mechanism of circRNA are diverse. circRNA can affect the life process, and its role in tumors is also concerned. However, at present, most researches focus on its formation mechanism, but the understanding of its molecular mechanism in the process of disease occurrence and development is still limited. More and more circRNA will be found and explained, and the mystery of its role in AML and other tumor diseases will be gradually revealed.

5. Conclusion

In conclusion, we concluded that circRNF13 was highly expressed in AML compared with normal brain tissue, the

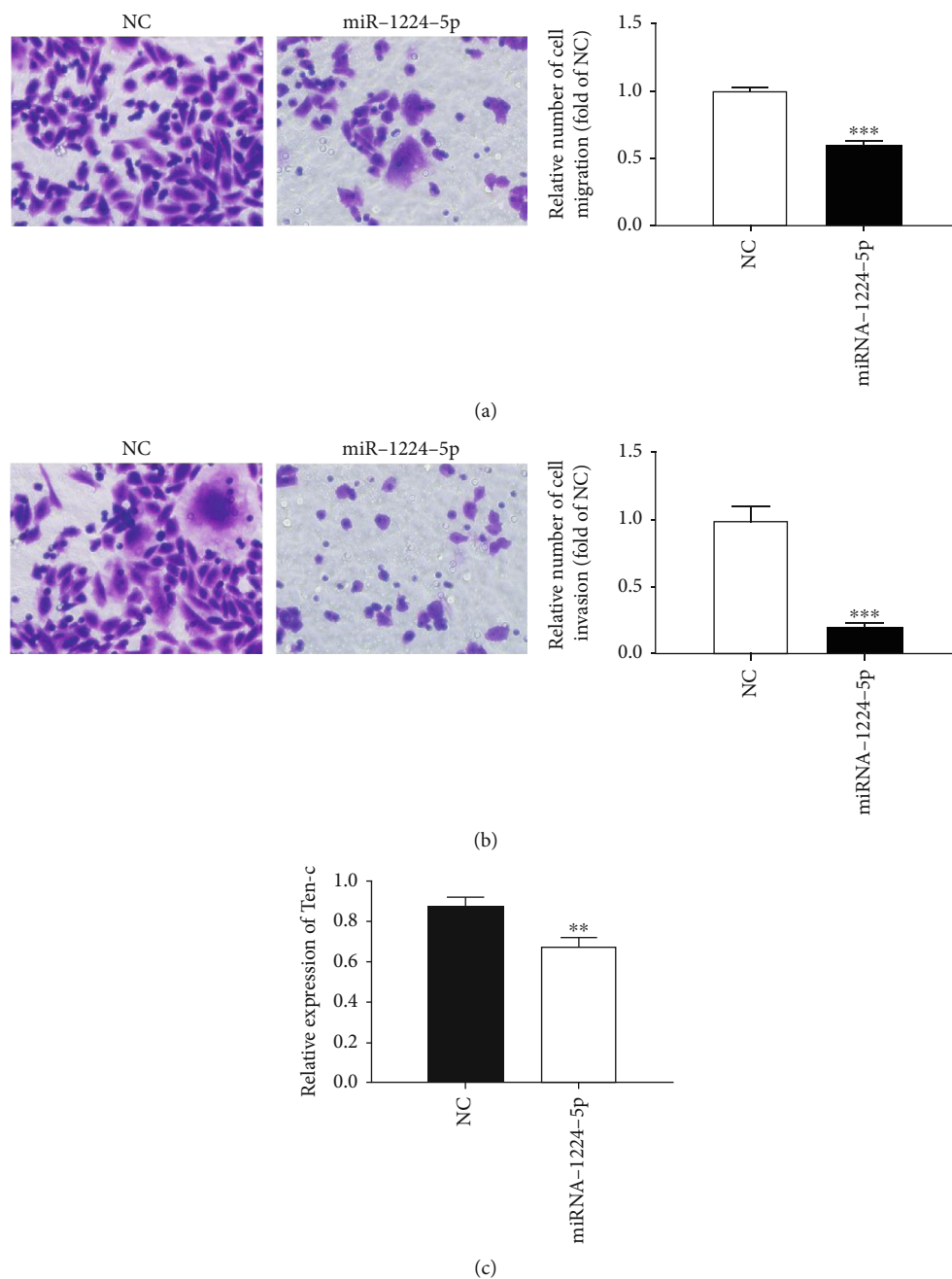


FIGURE 6: Effect of overexpression of miRNA-1224-5p on migration and invasion of AML cells. (a) Microscopically, the migration and statistical analysis of Kasumi-1 cells after the overexpression of miRNA-1224-5p. (b) Microscopically, the invasion and statistical analysis of Kasumi-1 cells after the overexpression of miRNA-1224-5p. (c) Tenascin-C expression after overexpression of miRNA-1224-5p. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Magnification: 200x.

proliferation of AML cells was inhibited by cell cycle block and apoptosis induction after the downregulation of circRNF13 expression, and the migration and invasion of AML cells were inhibited significantly after the downregulation of circRNF13 expression. miRNA-1224-5p inhibits the proliferation, migration, and invasion of AML cells by regulating the expression of circRNF13. This study may provide some clues for the diagnosis and treatment of AML.

Data Availability

We can provide our data when others need it.

Conflicts of Interest

There is no conflict of interest.

References

- [1] F. R. Appelbaum, J. M. Rowe, J. Radich, and J. E. Dick, "Acute myeloid leukemia," *Hematology*, vol. 2001, no. 1, pp. 62–86, 2001.
- [2] R. F. Schlenk, K. Döhner, J. Krauter et al., "Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia," *New England Journal of Medicine*, vol. 358, no. 18, pp. 1909–1918, 2008.
- [3] S. Fröhling, R. F. Schlenk, J. Breitruck et al., "Prognostic significance of activating FLT3 mutations in younger adults (16 to 60 years) with acute myeloid leukemia and normal cytogenetics: a study of the AML Study Group Ulm," *Blood*, vol. 100, no. 13, pp. 4372–4380, 2002.
- [4] G. Cimino, M. C. Rapanotti, L. Elia et al., "*ALL-1* gene rearrangements in acute myeloid leukemia: association with M4-M5 French-American-British classification subtypes and young age," vol. 55, no. 8, p. 1625, 1995.
- [5] P. H. Wiernik, "Optimal therapy for adult patients with acute myeloid leukemia in first complete remission," *Current Treatment Options in Oncology*, vol. 15, no. 2, pp. 171–186, 2014.
- [6] J. Zheng, Y. Song, Z. Li, A. Tang, Y. Fei, and W. He, "The implication of lncRNA expression pattern and potential function of lncRNA RP4-576H24.2 in acute myeloid leukemia," *Cancer Medicine*, vol. 8, no. 17, pp. 7143–7160, 2019.
- [7] B. Zhang, Y.-F. Sun, X.-M. Zhang, N. Jiang, and Q. Chen, "TUG1 weakens the sensitivity of acute myeloid leukemia cells to cytarabine by regulating miR-655-3p/CCND1 axis," vol. 24, no. 9, pp. 4940–4953, 2020.
- [8] S. Feng, N. Liu, X. Chen, Y. Liu, and J. An, "Long non-coding RNA NEAT1/miR-338-3p axis impedes the progression of acute myeloid leukemia via regulating CREBFR," *Cancer Cell International*, vol. 20, no. 1, 2020.
- [9] R. Kishore, V. N. S. Garikipati, and C. Gonzalez, "Role of circular RNAs in cardiovascular disease," *Journal of Cardiovascular Pharmacology*, vol. 76, no. 2, pp. 128–137, 2020.
- [10] A. Murayama, N. Yamada, Y. Osaki et al., "N-terminal PreS1 sequence regulates efficient infection of cell culture-generated hepatitis B virus," *Hepatology*, 2020.
- [11] Z. Wang, Y. Zhang, K. Dai et al., "circEgg regulates histone H3K9me3 by sponging bmo-miR-3391-5p and encoding circEgg-P122 protein in the silkworm, *Bombyx mori*," *Insect Biochemistry and Molecular Biology*, vol. 124, article 103430, 2020.
- [12] F. Khanipouyani, H. Akrami, and M. R. Fattahi, "Circular RNAs as important players in human gastric cancer," *Clinical and Translational Oncology*, 2020.
- [13] J. Wu, W. Ren, Z. Zheng et al., "Mmu_circ_003795 regulates osteoblast differentiation and mineralization in MC3T3-E1 and MDPC23 by targeting COL15A1," *Molecular Medicine Reports*, vol. 22, no. 3, pp. 1737–1746, 2020.
- [14] J. R. Welden, A. Pawluchin, J. van Doorn, and S. Stamm, "Use of *Alu* element containing minigenes to analyze circular RNAs," *Journal of Visualized Experiments*, no. 157, 2020.
- [15] X. Zhang, H. Chu, L. Wen et al., "Competing endogenous RNA network profiling reveals novel host dependency factors required for MERS-CoV propagation," *Emerging Microbes & Infections*, vol. 9, no. 1, pp. 733–746, 2020.
- [16] L. Peng, X. Qing, Y. Guan, and C. Li, "The emerging landscape of circular RNA ciRS-7 in cancer (Review)," *Oncology Reports*, vol. 33, no. 6, pp. 2669–2674, 2015.
- [17] S. Memczak, M. Jens, A. Elefsinioti et al., "Circular RNAs are a large class of animal RNAs with regulatory potency," *Nature*, vol. 495, no. 7441, pp. 333–338, 2013.
- [18] T. B. Hansen, T. I. Jensen, B. H. Clausen et al., "Natural RNA circles function as efficient microRNA sponges," *Nature*, vol. 495, no. 7441, pp. 384–388, 2013.
- [19] L. Wang, S. Liu, Y. Mao et al., "CircRNF13 regulates the invasion and metastasis in lung adenocarcinoma by targeting miR-93-5p," *Gene*, vol. 671, pp. 170–177, 2018.
- [20] C. Gu, Z. Huang, X. Chen et al., "TEAD4 promotes tumor development in patients with lung adenocarcinoma via ERK signaling pathway," *Biochimica et Biophysica Acta - Molecular Basis of Disease*, vol. 1866, no. 12, p. 165921, 2020.
- [21] M. Yang and W. Huang, "Circular RNAs in nasopharyngeal carcinoma," *Clinica Chimica Acta*, vol. 508, pp. 240–248, 2020.
- [22] A. Francavilla, S. Turoczi, S. Tarallo, P. Vodicka, B. Pardini, and A. Naccarati, "Exosomal microRNAs and other non-coding RNAs as colorectal cancer biomarkers: a review," *Mutagenesis*, vol. 35, 2020.
- [23] L. V. Cappelli, M. Meggendorfer, F. Dicker et al., "DNMT3A mutations are over-represented in young adults with NPM1 mutated AML and prompt a distinct co-mutational pattern," *Leukemia*, vol. 33, no. 11, pp. 2741–2746, 2019.
- [24] Y. Ding, Y. Dong, H. Lu et al., "Circular RNA profile of acute myeloid leukaemia indicates circular RNA annexin A2 as a potential biomarker and therapeutic target for acute myeloid leukaemia," *American Journal of Translational Research*, vol. 12, no. 5, pp. 1683–1699, 2020.
- [25] D. Papaioannou, S. Volinia, D. Nicolet et al., "Clinical and functional significance of circular RNAs in cytogenetically normal AML," *Blood Advances*, vol. 4, no. 2, pp. 239–251, 2020.
- [26] Z. Pan, G.-F. Li, M.-L. Sun et al., "MicroRNA-1224 splicing circularRNA-Filip1l in an Ago2-dependent manner regulates chronic inflammatory pain via targeting Ubr5," *The Journal of Neuroscience*, vol. 39, no. 11, pp. 2125–2143, 2019.
- [27] Q. Singleton, K. Vaibhav, M. Braun et al., "Bone marrow derived extracellular vesicles activate osteoclast differentiation in traumatic brain injury induced bone loss," *Cells*, vol. 8, no. 1, p. 63, 2019.
- [28] B. Li, P. Wu, W. Fu et al., "The role and mechanism of miRNA-1224 in thePolygonatum sibiricumPolysaccharide regulation of bone marrow-derived macrophages to osteoclast differentiation," *Rejuvenation Research*, vol. 22, no. 5, pp. 420–430, 2019.

Research Article

JAK3 and TYK2 Serve as Prognostic Biomarkers and Are Associated with Immune Infiltration in Stomach Adenocarcinoma

Lingkai Meng,¹ Ling Ding,² Yue Yu,³ and Wang Li¹

¹NHC Key Laboratory of Hormones and Development, Tianjin Key Laboratory of Metabolic Diseases, Chu Hsien-I Memorial Hospital & Tianjin Institute of Endocrinology, Tianjin Medical University, Tianjin 300134, China

²Department of Anesthesiology, Tianjin Union Medical Center, Tianjin 300191, China

³University of Alberta, Edmonton, Alberta, Canada

Correspondence should be addressed to Wang Li; lwbaker@126.com

Received 1 July 2020; Revised 19 August 2020; Accepted 3 September 2020; Published 6 October 2020

Academic Editor: Tao Huang

Copyright © 2020 Lingkai Meng 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. Stomach adenocarcinoma (STAD) is one of the most common malignant tumors. The Janus kinases (JAKs) play a significant part in cellular biological process, inflammation, and immunity. The roles of JAKs in STAD are still not systematically described. **Methods.** A series of bioinformatics tools were used to clarify the role of JAKs in STAD. **Results.** JAK3/TYK2 levels were significantly increased in STAD during subgroup analyses based on gender, tumor grade, cancer stages, and nodal metastasis status. STAD patients with high levels of JAK3/TYK2 had poor overall survival, postprogression survival, and first progression. Immune infiltration revealed a significant correlation between JAK3/TYK2 expression and the abundance of immune cells as well as immune biomarker expression in STAD. JAK3/TYK2 was associated with the adaptive immune response, chemokine signaling pathway, and JAK-STAT signaling pathway. **Conclusions.** JAK3 and TYK2 serve as prognostic biomarkers and are associated with immune infiltration in STAD.

1. Introduction

Gastric cancer (GC) is one of the most common malignant tumors, with the fifth largest incidence and third largest mortality rate among all malignant tumors [1]. Stomach adenocarcinoma (STAD) is the most common subtype of GC, accounting for over 95% of all GC cases.

Although the identification of *Helicobacter pylori* has reduced the incidence of gastric cancer, it is estimated that 1,033,701 patients would be initially diagnosed with GC worldwide in 2018 [2]. Moreover, the molecular mechanisms concerning the tumorigenesis and progress of GC is far from clarified and the therapeutic measures for GC are limited, resulting in a poor patient prognosis. Furthermore, the overall survival of patients with advanced or metastatic GC is only approximately 1 year [3]. These sobering data illustrate a critical need for novel prognostic biomarkers and therapeutic targets for STAD.

Janus kinases (JAKs) are major activators of signal transducers and play a significant role in cellular biological processes, inflammation, and immunity [4–7]. JAK/STAT signaling is a key regulator of gene expression, transcriptional programs, and immune response. In all, four members have been identified in the JAK family: JAK1/2/3 and TYK2. Genetic alterations of JAKs are involved in tumor cell proliferation, migration, apoptosis, and metastasis in certain types of cancers [8]. Increasing evidence demonstrates JAKs as a prognostic biomarker and therapeutic target for many cancers or other diseases, such as JAK3 for renal cell carcinoma, JAK2 for acute lymphoblastic leukemia [9], JAK2 for skin cutaneous melanoma [10], and TYK2 for hepatocellular carcinoma [11]. However, specific functions of JAKs in STAD remain to be systematically described.

Therefore, we aimed to explore the expression of JAKs and prognostic value of the association between immune infiltration and JAKs in STAD. We further evaluated the cor-

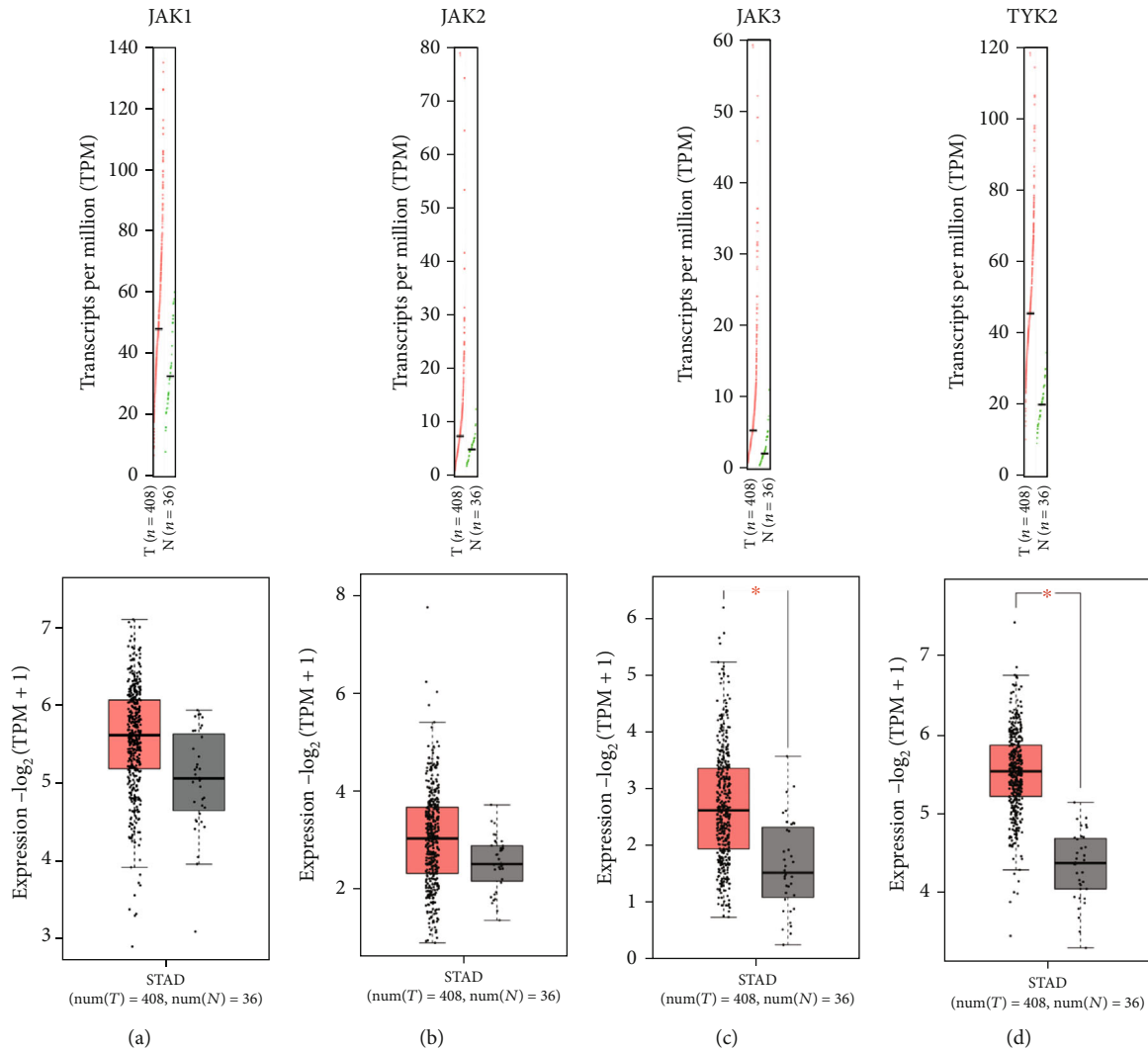


FIGURE 1: The expression of JAKs in STAD (GEPIA). The expression of JAK3 and TYK2 were significantly elevated in STAD tissues at mRNA level. STAD: stomach adenocarcinoma; * $P < 0.05$; T: tumor tissues; N: normal tissues.

relation between JAK expression and the clinicopathological parameters of patients as well as immune infiltration in STAD. Our results may provide additional evidence about the prognostic biomarkers and therapeutic targets for STAD.

2. Materials and Methods

2.1. GEPIA. GEPIA is a novel bioinformatics web server for analyzing RNA sequencing expression data across The Cancer Genome Atlas Program (TCGA) cancers [12]. TCGA is a landmark cancer genomics program that has molecularly characterized more than 20,000 primary cancers and matched normal samples spanning 33 cancer types. Tumor/normal differential expression analysis of JAKs in STAD was explored using the TCGA STAD dataset ($n = 415$) in GEPIA with analysis of variance (ANOVA). A P value less than 0.05 indicated statistical significance.

2.2. UALCAN. UALCAN is designed for gene expression analysis, prognosis analysis, and methylation analysis based on the data from TCGA and Clinical Proteomic Tumor Analysis Consortium (CPTAC) [13]. In the current study, the correlation between JAK3 and TYK2 expression and the clinicopathological parameters of STAD patients, including the race, gender, age, *H. pylori* infection status, histological subtype, tumor grade, cancer stage, and nodal metastasis status of patients, were analyzed using the TCGA STAD dataset ($n = 415$). A P value less than 0.05 indicated statistical significance.

2.3. The Kaplan–Meier Plotter (KM Plotter). The KM plotter is designed for the prognostic analysis of 54k genes (mRNA, miRNA, and protein) in certain types of cancers including breast, lung, and gastric cancer [14]. Here, the significance of JAK3 and TYK2 in determining the overall survival (OS), postprogression survival (PPS), and first progression (FP) of STAD was analyzed using the Kaplan–Meier curve.

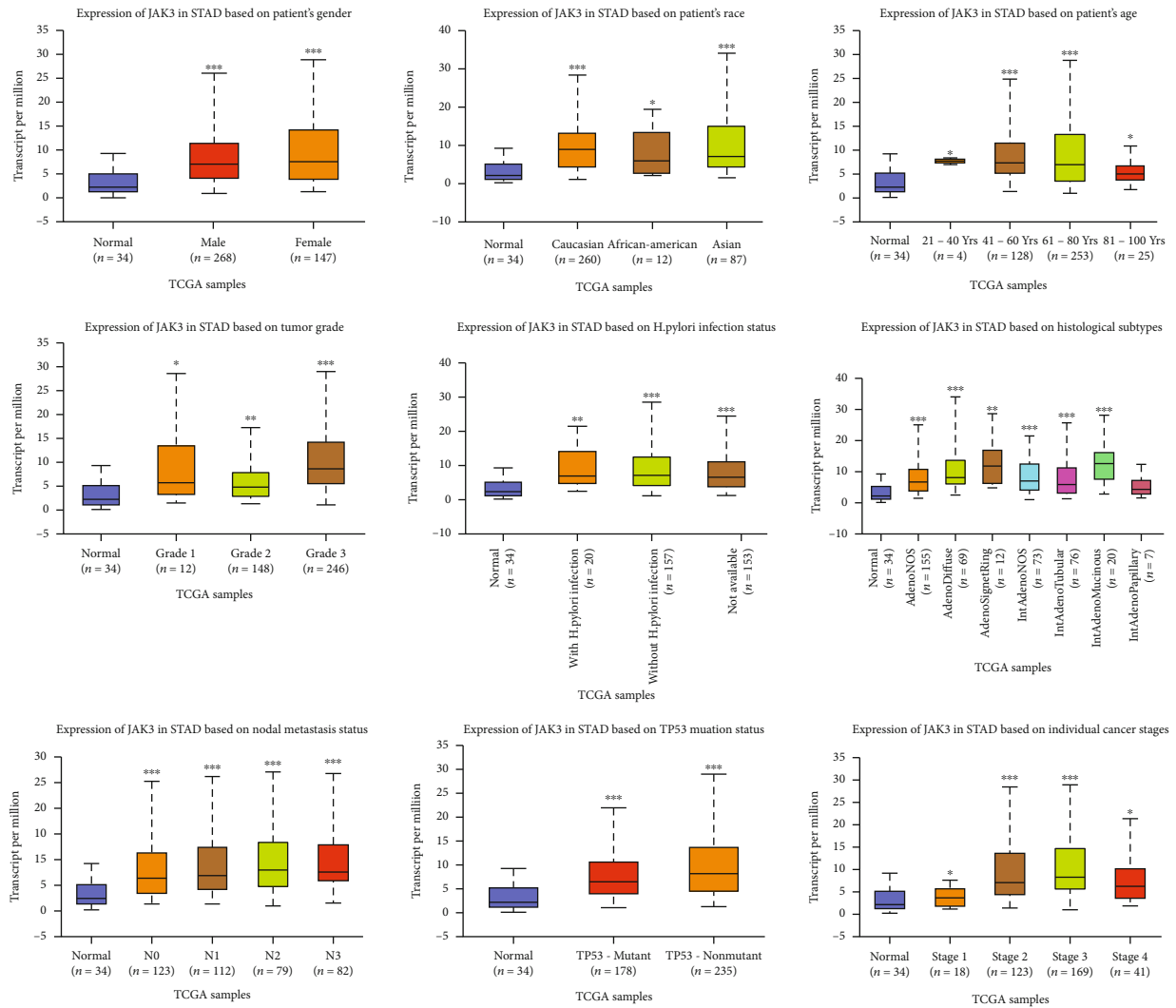


FIGURE 2: The expression of JAK3 in STAD in subgroup analyses (UALCAN). Subgroup analyses were performed based on patients’ race, patients’ gender, patients’ age, H. pylori infection status, histological subtypes, tumor grade, individual cancer stages, and nodal metastasis status. STAD: stomach adenocarcinoma; * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

The medium value of the JAK3 and TYK2 expressions was used to split patients into high-/low-expression groups.

2.4. cBioPortal. cBioPortal is a cancer genomics portal designed for exploring multidimensional cancer genomics data using the TCGA dataset [15]. We used cBioPortal to explore, visualize, and analyze the genetic alterations and mutations of JAK3 and TYK2 in STAD using the TCGA STAD dataset ($n = 415$). Furthermore, mRNA expression z scores (RNA Seq V2 RSEM) were obtained (z score threshold, ± 2.0). Protein expression z scores (RPPA) were also obtained (z score threshold, ± 2.0).

2.5. LinkedOmics. LinkedOmics is a bioinformatics web portal designed for accessing, analyzing, and comparing cancer multiomics data of various cancer types [16]. Complete data of 415 TCGA STAD patients were used to explore JAK3- and TYK2-associated genes via the Spearman correlation analysis. Moreover, Gene Set Enrich-

ment Analysis (GSEA) was performed to explore JAK3- and TYK2-associated functions (GO analysis and KEGG pathway analysis) in STAD, with the minimum number of genes being three and the P value threshold being 0.05. The transcription factor targets of JAK3 and TYK2 were also analyzed via GSEA.

2.6. TIMER. TIMER is a comprehensive resource for the systematic analysis of immune infiltrates across diverse cancer types [17]. In the current study, the Spearman correlation analysis was used to explore the correlation between the expression levels of JAK3/TYK2 and the abundance of immune cell infiltrates and the expression of gene biomarkers of immune cells [18–20]. The two-sided Wilcoxon rank-sum test was used to evaluate the effect of somatic copy number alterations (SCNAs) of JAK3/TYK2 on immune cell infiltrates. A P value less than 0.05 indicated statistical significance.

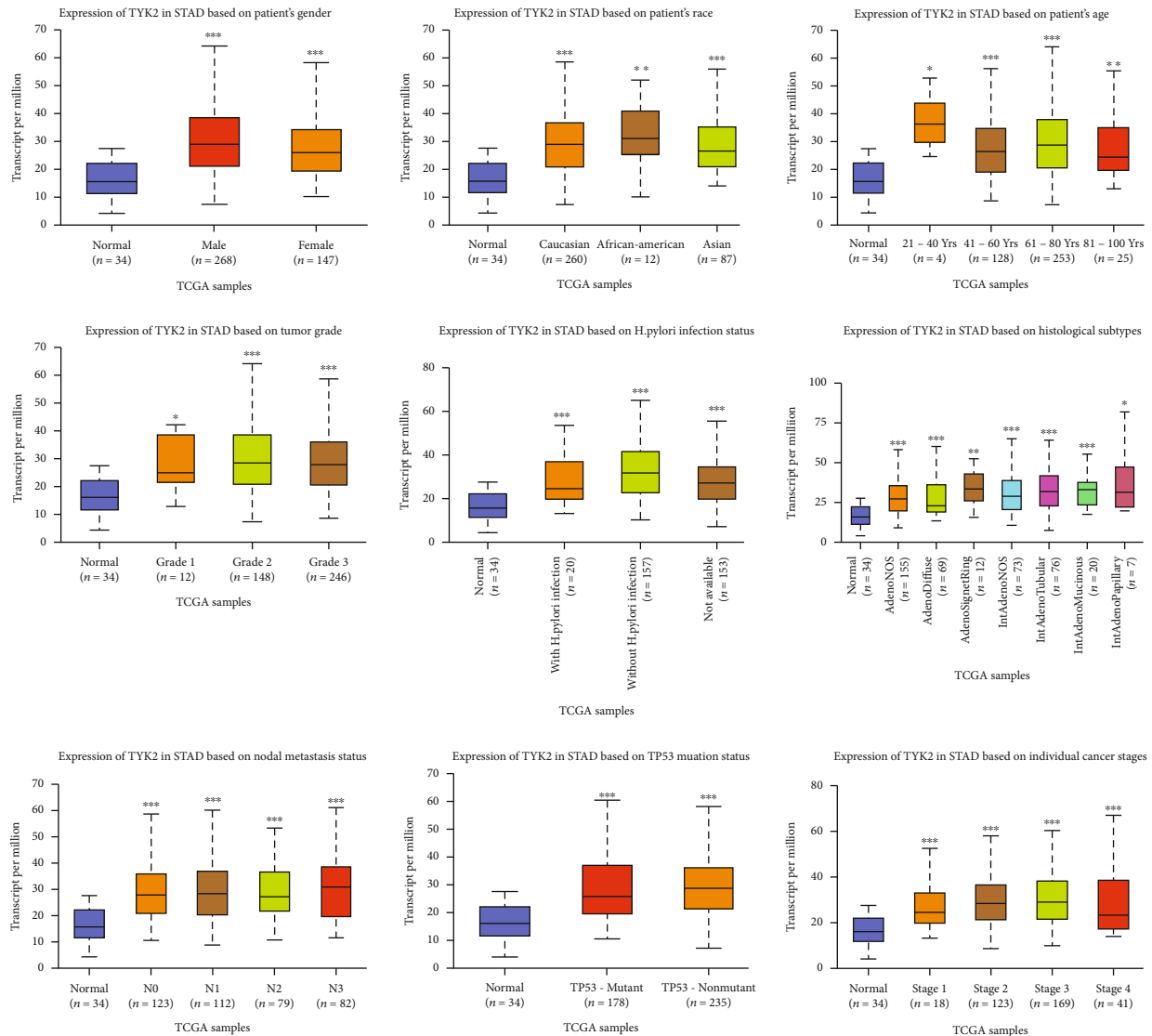


FIGURE 3: The expression of TYK2 in STAD in subgroup analyses (UALCAN). Subgroup analyses were performed based on patients' race, patients' gender, patients' age, *H. pylori* infection status, histological subtypes, tumor grade, individual cancer stages, and nodal metastasis status. STAD: stomach adenocarcinoma; * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

3. Results

3.1. JAK Expression in STAD. The level of JAKs in primary STAD was first determined using GEPIA. As shown in Figure 1, the expression levels of JAK3 (Figure 1(c), $P < 0.05$) and TYK2 (Figure 1(d), $P < 0.05$) were significantly elevated in STAD tissues compared with normal tissues. However, there was no difference in the expression levels of JAK1 (Figure 1(a)) and JAK2 (Figure 1(b)) between STAD tissues and normal tissues. We then analyzed the correlation between the expression levels of JAK3/TYK2 and the clinicopathological parameters of STAD patients. As expected, the mRNA levels of JAK3 were significantly increased in STAD during subgroup analyses based on the race, gender, age, *H. pylori* infection status, histological sub-

type, tumor grade, cancer stage, and nodal metastasis status of patients (Figure 2). The same results were obtained for TYK2, and the mRNA levels of JAK3 were significantly increased in STAD during subgroup analyses based on the race, gender, age, *H. pylori* infection status, histological subtype, tumor grade, individual cancer stage, and nodal metastasis status of patients (Figure 3). Therefore, JAK3 and TYK2 may play a significant role in the tumorigenesis, progression, and aggressiveness of STAD.

3.2. JAK3/TYK2 As a Prognostic Biomarker in STAD. The prognostic value of JAK3/TYK2 in STAD was evaluated using the KM plotter. We found that STAD patients with high JAK3 levels had poor OS (HR = 1.45 (1.22-1.71), $P = 2 \times 10^{-5}$), FP (HR = 1.41 (1.15-1.72), $P = 0.00076$), and PPS

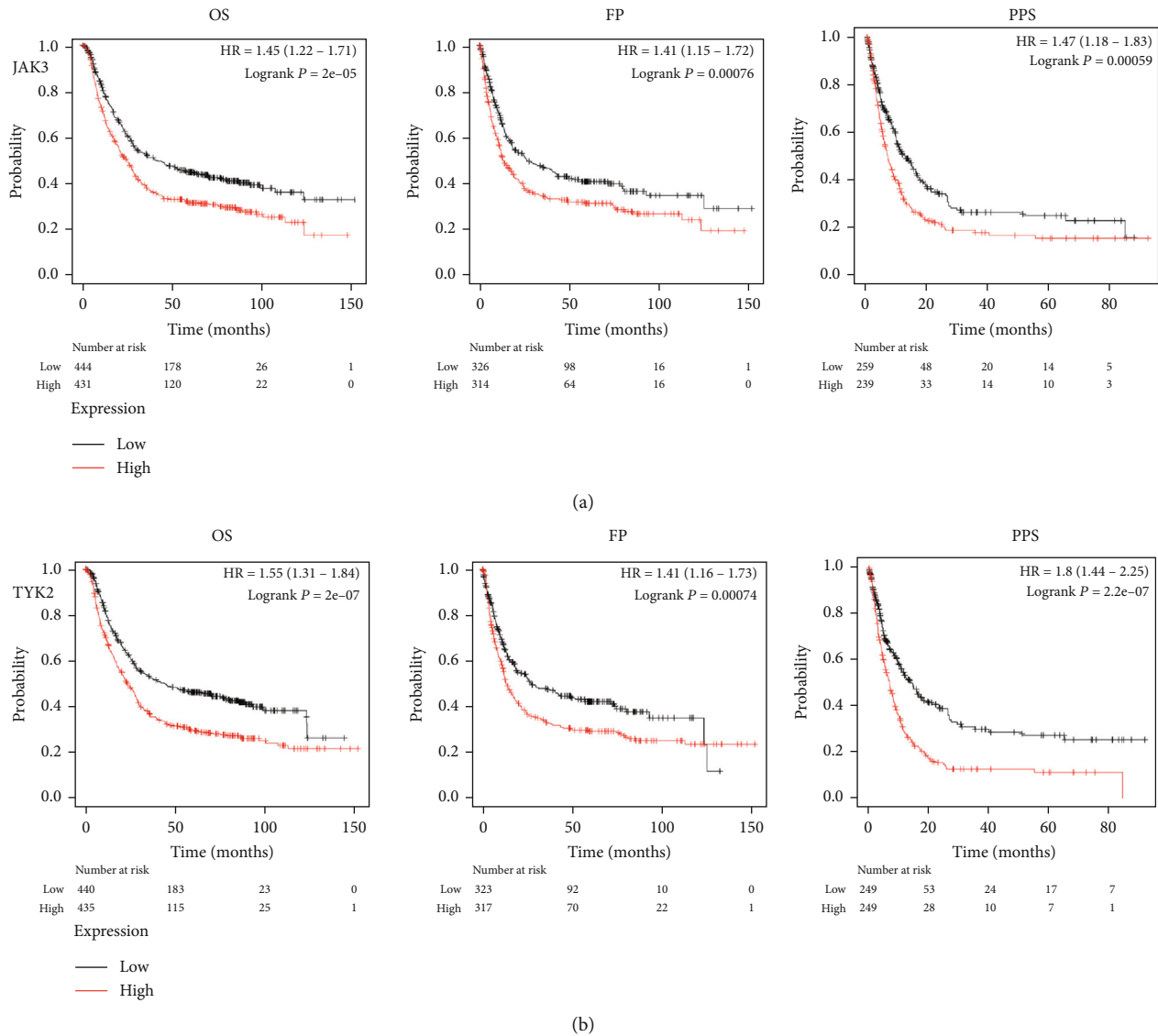


FIGURE 4: The prognostic value of JAK3/TYK2 in STAD (KM plotter). (a) STAD patients with high mRNA level of JAK3 had worse OS, PF, and PPS. (b) STAD patients with high mRNA level of TYK2 had worse OS, PF, and PPS. All the analyses were performed with Kaplan–Meier analysis. HR: hazard ratio; OS: overall survival; PPS: postprogression survival; FP: first progression.

(HR = 1.47 (1.18-1.83), $P = 0.00059$) (Figure 4(a)). Moreover, even STAD patients with high TYK2 levels had poor OS (HR = 1.55 (1.31-1.84), $P = 4e^{-7}$), FP (HR = 1.41 (1.16-1.73), $P = 0.00074$), and PPS (HR = 1.8 (1.44-2.25), $P = 2.2e^{-7}$). Thus, JAK3/TYK2 served as a prognostic biomarker in STAD (Figure 4(b)).

To better understand how the expression levels of JAK3 and TYK2 impact the prognosis of STAD patients, we also analyzed the correlation between the expression of JAK3 and TYK2 and clinical characteristics of TCGA STAD patients using the KM plotter. JAK3 and TYK2 overexpression was associated with worse OS (Table 1) and PFS (Table 2) in male and female patients as well as in patients with intestinal and diffuse type Lauren classification ($P < 0.05$). Further, the overexpression of JAK3 and TYK2 was associated

with worse FP (Table 3) in male and female patients ($P < 0.05$). STAD patients with poor differentiation and high JAK3 levels had worse OS (Table 1) and PFS (Table 2), although the P value in OS analysis was 0.059. We further found that the overexpression of JAK3 and TYK2 was associated with worse OS (Table 1) and PFS (Table 2) in patients with stage 2 and 3 disease ($P < 0.05$). STAD patients with regional lymph node metastasis (N stage 1 or 1+2+3) and high JAK3 expression had significantly worse OS (Table 1), PFS (Table 2), and PF (Table 3). Similarly, STAD patients with regional lymph node metastasis (N stage 1, 2, or 1+2+3) and high TYK2 expression levels had worse OS (Table 1), PFS (Table 2), and PF (Table 3). Therefore, JAK3/TYK2 level can impact the prognosis of STAD patients with lymph node metastasis.

TABLE 1: Correlation of JAK3/TYK2 mRNA expression and overall survival in STAD with different clinicopathological factors (Kaplan–Meier plotter).

| Pathological parameters | N | Overall survival | | | | |
|-------------------------|-----|----------------------|---------|-----|----------------------|-----------|
| | | JAK3 Hazard radio | P value | N | TYK2 Hazard radio | P value |
| Sex | | | | | | |
| Female | 236 | 1.93 (1.36-2.74) | 0.00017 | 236 | 1.5 (1.06-2.14) | 0.023 |
| Male | 544 | 1.59 (1.23-2.04) | 0.00028 | 544 | 1.86 (1.5-2.3) | $1e^{-8}$ |
| Stage | | | | | | |
| 1 | 67 | 3.61 (1.35-9.65) | 0.0062 | 67 | 2.02 (0.75-5.44) | 0.16 |
| 2 | 140 | 2.39 (1.3-4.38) | 0.0037 | 140 | 1.92 (1.03-3.56) | 0.036 |
| 3 | 305 | 1.56 (0.17-2.09) | 0.0023 | 305 | 1.66 (1.22-2.26) | 0.001 |
| 4 | 148 | 0.76 (0.49-1.17) | 0.21 | 148 | 0.71 (0.48-1.06) | 0.091 |
| Stage T | | | | | | |
| 2 | 241 | 1.47 (0.93-2.3) | 0.094 | 241 | 1.28 (0.82-2) | 0.27 |
| 3 | 204 | 0.78 (0.54-1.13) | 0.19 | 204 | 1.47 (0.97-2.22) | 0.065 |
| 4 | 38 | 0.33 (0.1-1.11) | 0.059 | 38 | 0.67 (0.28-1.59) | 0.36 |
| Stage N | | | | | | |
| 0 | 74 | 2.03 (0.6-6.85) | 0.24 | 74 | 1.75 (0.75-4.07) | 0.19 |
| 1 | 225 | 2.63 (1.49-4.66) | 0.00054 | 225 | 1.63 (1.07-2.48) | 0.02 |
| 2 | 121 | 0.72 (0.45-1.15) | 0.17 | 121 | 0.64 (0.41-1) | 0.048 |
| 3 | 76 | 0.7 (0.37-1.31) | 0.26 | 76 | 0.64 (0.38-1.1) | 0.11 |
| 1+2+3 | 422 | 1.35 (1.01-1.8) | 0.043 | 422 | 1.38 (1.3-1.54) | 0.03 |
| Stage M | | | | | | |
| 0 | 444 | 1.3 (0.97-1.75) | 0.082 | 444 | 1.32 (1-1.76) | 0.053 |
| 1 | 56 | 0.77 (0.4-1.5) | 0.44 | 56 | 0.37 (0.16-0.82) | 0.011 |
| Lauren classification | | | | | | |
| Intestinal | 320 | 3.32 (1.5-3.59) | 0.0001 | 320 | 1.74 (1.27-2.39) | 0.00047 |
| Diffuse | 241 | 1.38 (0.97-1.97) | 0.07 | 241 | 1.15 (0.81-1.62) | 0.44 |
| Differentiation | | | | | | |
| Poor | 165 | 1.5 (0.98-2.3) | 0.059 | 165 | 0.83 (0.55-1.23) | 0.34 |
| Moderate | 67 | 0.57 (0.29-1.11) | 0.096 | 67 | 0.56 (0.29-1.07) | 0.075 |

3.3. *Genetic Alterations of JAK3/TYK2 in STAD.* cBioPortal was used to determine the genetic alterations of JAK3/TYK2 in STAD. We found that JAK3 and TYK2 were altered in 6% and 8% of all TCGA STAD cases, respectively (Figure 5(a)).

Genetic alterations of JAK3 and TYK2 in STAD comprised missense mutation, truncating mutation, amplification, deep deletion, high mRNA levels, and low mRNA levels. Thus, mutation is the most common type of JAK3/TYK2 genetic alteration. The mutation sites of JAK3/TYK2 in STAD are shown in Figures 5(b) and 5(c).

3.4. *JAK3/TYK2 Correlated with Immune Infiltration in STAD.* An increasing number of studies have suggested an interaction between immune response and pathophysiological processes [21, 22]. Moreover, JAKs play a critical role in immune regulation by invoking intracellular signaling pathways in cancers [23]. Therefore, we next evaluated the correlation between JAK3/TYK2 and immune infiltration in

STAD. As shown in Figure 6, JAK3 levels showed a positive correlation with the abundance of CD8+ T cells ($Cor = 0.521, P = 3.87e^{-27}$), CD4+ T cells ($Cor = 0.509, P = 1.52e^{-25}$), macrophages ($Cor = 0.332, P = 5.33e^{-12}$), neutrophils ($Cor = 0.497, P = 1.62e^{-24}$), and dendritic cells ($Cor = 0.588, P = 6.21e^{-36}$) (Figure 6(a)). We also found a positive correlation between TYK2 levels and the abundance of CD8+ T cells ($Cor = 0.103, P = 0.0468$), CD4+ T cells ($Cor = 0.249, P = 1.44e^{-06}$), neutrophils ($Cor = 0.129, P = 0.0127$), and dendritic cells ($Cor = 0.148, P = 0.00428$) (Figure 6(b)). Interestingly, SCNA of JAK3/TYK2 could partially inhibit immune infiltration in STAD (Figures 6(c) and 6(d)).

We also evaluated the correlation between JAK3/TYK2 and immune biomarkers in STAD. Previous studies have reported these biomarkers of immune cells [18–20]. As expected, the expression levels of JAK3/TYK2 were positively correlated with the expression levels of immune biomarkers in STAD (Tables 4 and 5). We found that the expression

TABLE 2: Correlation of JAK3/TYK2 mRNA expression and postprogression survival in STAD with different clinicopathological factors (Kaplan–Meier plotter).

| Pathological parameters | Post progression survival | | | | | |
|-------------------------|---------------------------|----------------------|----------|-----|----------------------|---------------------|
| | N | JAK3 Hazard radio | P value | N | TYK2 Hazard radio | P value |
| Sex | | | | | | |
| Female | 149 | 1.84 (1.19-2.85) | 0.0053 | 149 | 2.36 (1.49-3.73) | 0.00015 |
| Male | 348 | 1.71 (1.31-2.22) | 0.000055 | 348 | 2.32 (1.79-3.02) | 1.1e ⁻¹⁰ |
| Stage | | | | | | |
| 1 | 31 | 2.86 (0.5-16.46) | 0.22 | 31 | 1.66 (0.37-7.43) | 0.51 |
| 2 | 105 | 2.5 (1.28-4.87) | 0.0053 | 105 | 2.39 (1.23-4.64) | 0.0081 |
| 3 | 142 | 1.45 (0.92-2.28) | 0.1 | 142 | 2.47 (1.6-3.83) | 2.7e ⁻⁵ |
| 4 | 104 | 1.66 (1.01-2.72) | 0.045 | 104 | 0.68 (0.42-1.11) | 0.12 |
| Stage T | | | | | | |
| 2 | 196 | 1.66 (1.05-2.64) | 0.029 | 196 | 1.72 (1.09-2.73) | 0.019 |
| 3 | 150 | 1.21 (0.8-1.82) | 0.36 | 150 | 1.91 (1.17-3.12) | 0.0083 |
| 4 | 29 | 0.45 (0.15-1.38) | 0.15 | 29 | 0.61 (0.22-1.64) | 0.32 |
| Stage N | | | | | | |
| 0 | 41 | 2.59 (0.72-9.35) | 0.13 | 41 | 2.57 (0.77-8.57) | 0.11 |
| 1 | 169 | 2.52 (1.6-3.98) | 0.000039 | 169 | 2.48 (1.58-3.91) | 4.8e ⁻⁵ |
| 2 | 105 | 0.7 (0.43-1.15) | 0.16 | 105 | 1.67 (0.99-2.81) | 0.054 |
| 3 | 63 | 1.65 (0.86-3.18) | 0.13 | 63 | 0.53 (0.29-0.95) | 0.032 |
| 1+2+3 | 337 | 1.44 (1.06-1.95) | 0.02 | 337 | 1.62 (1.21-2.17) | 0.0011 |
| Stage M | | | | | | |
| 0 | 342 | 1.38 (1-1.9) | 0.051 | 342 | 1.91 (1.4-2.6) | 2.7e ⁻⁵ |
| 1 | 36 | 2.35 (1.08-5.13) | 0.028 | 36 | 0.73 (0.33-1.63) | 0.44 |
| Lauren classification | | | | | | |
| Intestinal | 192 | 1.69 (1.08-2.66) | 0.02 | 192 | 1.86 (1.22-2.84) | 0.0037 |
| Diffuse | 176 | 1.5 (0.99-2.26) | 0.053 | 176 | 1.55 (1.05-2.29) | 0.028 |
| Differentiation | | | | | | |
| Poor | 49 | 3.3 (1.59-6.88) | 0.00076 | 49 | 1.68 (0.82-3.41) | 0.15 |
| Moderate | 24 | 2.06 (0.75-5.63) | 0.15 | 24 | 0.63 (0.26-1.56) | 0.32 |

levels of biomarkers of CD8+ T cells (CD8A and CD8B), T cells (CD3D, CD3E, and CD2), B cells (CD19 and CD79A), monocytes (CD86 and CD115), and TAMs (CD68 and IL10) positively correlated with the expression levels of JAK3 and TYK2 in STAD. The expression levels of INOS, IRF5, CD163, VSIG4, MS4A4A, CD11b, and CCR7 were positively correlated with JAK3/TYK2 levels in STAD. All biomarkers of natural killer cells (KIR2DL1, KIR2DL3, KIR2DL4, KIR3DL1, KIR3DL2, KIR3DL3, and KIR2DS4) showed positive correlation with JAK3 expression. Similarly, all biomarkers of dendritic cells (KIR2DL1, KIR2DL3, KIR2DL4, KIR3DL1, KIR3DL2, KIR3DL3, and KIR2DS4), Th1 cells (TBX21, STAT4, STAT1, IIFNG, and TNF), Th2 cells (GATA3, STAT6, STAT5A, and IL13), and Tfh cells (BCL6 and IL21) showed a positive correlation with the JAK3 and TYK2 expressions. Moreover, levels of immune biomarkers of Treg cells (FOXP3, CCR8, and STAT5B) and T cell exhaustion (PD-1, CTLA4, LAG3, TIM-3, and GZMB) were positively associ-

ated with JAK3 and TYK2 levels. These results indicate that JAK3 and TYK2 played a vital role in immune escape in the STAD microenvironment.

3.5. Enrichment Analysis of JAK3/TYK2 in STAD. The function module of LinkedOmics was used to performed enrichment analysis of JAK3/TYK2 in STAD. In all, 7855 genes (dark red dots) were positively correlated with JAK3, whereas 4687 genes (dark green dots) were negatively correlated with JAK3 in STAD (Supplementary Figure 2A, $P < 0.05$). Further, 50 significant gene sets that positively and negatively correlated with JAK3 in STAD are presented in Supplementary Figure 2A-2C, respectively. Enrichment analysis performed via GSEA suggested that JAK3 is associated with adaptive immune response, protein transmembrane transport, DNA damage response, DNA damage detection, preribosomal structure, respiratory chain, cytokine binding, translation factor activity, RNA binding, snoRNA binding, and tRNA

TABLE 3: Correlation of JAK3/TYK2 mRNA expression and first progression in STAD with different clinicopathological factors (Kaplan–Meier plotter).

| Pathological parameters | N | First progression | | | | |
|-------------------------|-----|----------------------|---------|-----|----------------------|--------------------|
| | | JAK3 Hazard radio | P value | N | TYK2 Hazard radio | P value |
| Sex | | | | | | |
| Female | 201 | 2.02 (1.38-2.95) | 0.00021 | 201 | 1.37 (0.94-2) | 0.097 |
| Male | 437 | 1.42 (1.1-1.84) | 0.0076 | 437 | 2.07 (1.6-2.68) | 1.4e ⁻⁸ |
| Stage | | | | | | |
| 1 | 60 | 2.37 (0.79-7.1) | 0.11 | 60 | 0.54 (0.18-1.68) | 0.28 |
| 2 | 131 | 1.5 (0.79-2.84) | 0.21 | 131 | 1.42 (0.77-2.61) | 0.25 |
| 3 | 186 | 1.52 (1.04-2.24) | 0.031 | 186 | 1.36 (0.94-1.97) | 0.1 |
| 4 | 141 | 0.64 (0.41-1.02) | 0.057 | 141 | 0.73 (0.48-1.11) | 0.14 |
| Stage T | | | | | | |
| 2 | 239 | 1.51 (0.92-2.45) | 0.097 | 239 | 1.21 (0.8-1.83) | 0.36 |
| 3 | 204 | 0.7 (0.49-1.01) | 0.054 | 204 | 1.35 (0.9-2.02) | 0.15 |
| 4 | 39 | 0.46 (0.2-1.09) | 0.072 | 39 | 0.77 (0.36-1.66) | 0.5 |
| Stage N | | | | | | |
| 0 | 72 | 2.22(0.66-7.49) | 0.19 | 72 | 1.56 (0.66-3.65) | 0.3 |
| 1 | 222 | 2.27 (1.34-3.82) | 0.0016 | 222 | 1.45 (0.98-2.15) | 0.059 |
| 2 | 125 | 0.84 (0.53-1.32) | 0.44 | 125 | 0.59 (0.38-0.91) | 0.015 |
| 3 | 76 | 1.31 (0.72-2.37) | 0.37 | 76 | 0.73 (0.41-1.33) | 0.3 |
| 1+2+3 | 423 | 1.32 (1-1.74) | 0.049 | 423 | 0.89 (0.67-1.17) | 0.4 |
| Stage M | | | | | | |
| 0 | 443 | 1.35 (0.98-1.85) | 0.066 | 443 | 1.21 (0.92-1.59) | 0.17 |
| 1 | 56 | 0.6 (0.33-1.11) | 0.099 | 56 | 0.41 (0.19-0.92) | 0.026 |
| Lauren classification | | | | | | |
| Intestinal | 263 | 1.74 (1.15-2.62) | 0.0078 | 263 | 1.28 (0.88-1.88) | 0.2 |
| Diffuse | 231 | 1.25(0.87-1.79) | 0.22 | 231 | 0.85 (0.59-1.24) | 0.4 |
| Differentiation | | | | | | |
| Poor | 121 | 1.32 (0.84-2.09) | 0.23 | 121 | 0.67 (0.41-1.08) | 0.095 |
| Moderate | 67 | 0.63 (0.33-1.19) | 0.15 | 67 | 0.6 (0.32-1.13) | 0.11 |

binding during GO analysis (Supplementary Figure 2D-2F). Moreover, KEGG analysis revealed that JAK3 was involved in cytokine-cytokine receptor interactions, chemokine signaling pathway, NF-kappa B signaling pathway, Th17 cell differentiation, and T cell receptor signaling pathway and that JAK3 was associated with cell adhesion molecules (CAMs) (Supplementary Figure 2G and Supplementary Figure 3).

The results of enrichment analysis of TYK2 in STAD are shown Supplementary Figure 4. We found that 5756 genes (dark red dots) were positively correlated with TYK2, whereas 3993 genes (dark green dots) were negatively correlated with TYK2 in STAD (Supplementary Figure 4A, $P < 0.05$). Further, 50 significant gene sets that positively and negatively correlated with TYK2 in STAD are presented in Supplementary Figure 4B and 4C, respectively. Enrichment analysis performed by GSEA suggested that TYK2 was associated with the regulation of leukocyte activation, adaptive immune responses, translational initiation, mitochondrial matrix, ribosomal structure, translation factor activity, cytokine receptor activity, rRNA

binding, and protein transporter activity during GO analysis (Supplementary Figure 4D-4F). Furthermore, KEGG analysis revealed that JAK3 was associated with ribosomal structure, cytokine-cytokine receptor interaction, JAK-STAT signaling pathway, RNA transport, CAMs, and Th1 and Th2 cell differentiation (Supplementary Figure 4G and Supplementary Figure 5).

4. Discussion

Increasing evidence has revealed that JAKs play an important role in the regulation of cytokine signaling, thus affecting basic cellular mechanisms, such as cell invasion, proliferation, apoptosis, and cellular immunity [5, 24]. Moreover, JAK-associated signaling pathways are associated with tumorigenesis and progression of cancers, including lung cancer, renal cell carcinoma, and lung cancer [25–27]. However, specific functions of the JAK family in STAD remain to be systematically described. Therefore, our study was conducted to clarify the role of JAKs in STAD.

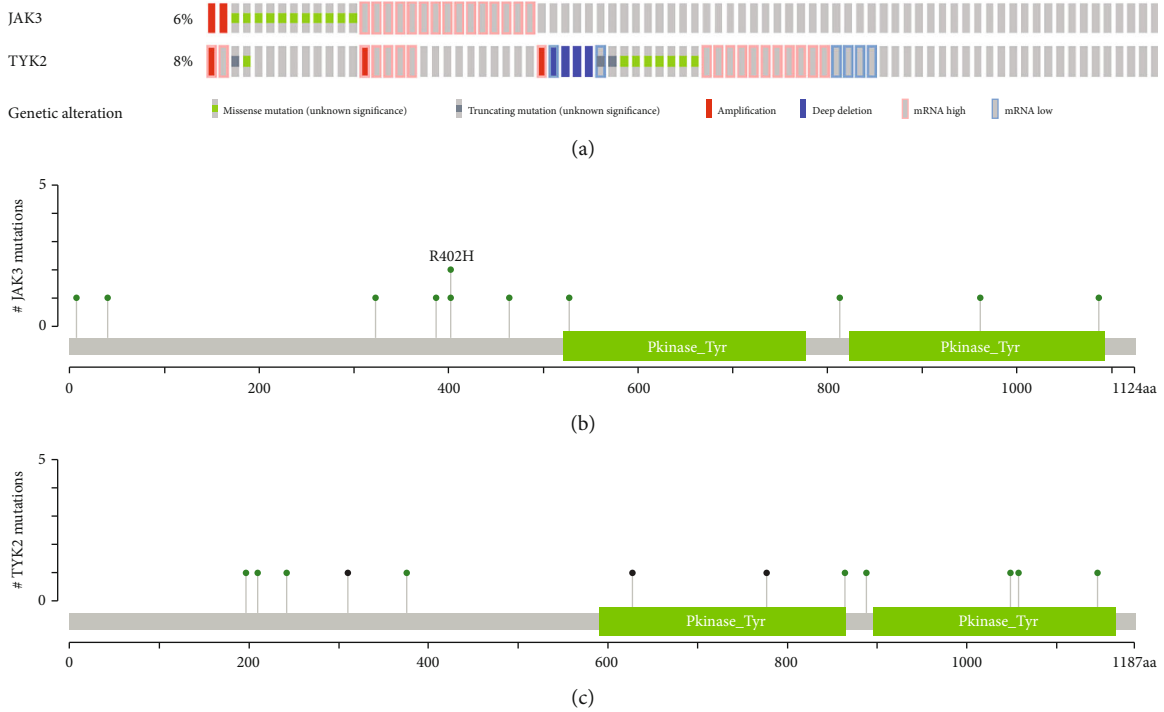


FIGURE 5: Genetic alteration of JAK3/TYK2 in STAD (cBioPortal). (a) OncoPrint of JAK3/TYK2 alterations in STAD. (b, c) Mutation sites of JAK3/TYK2 in STAD.

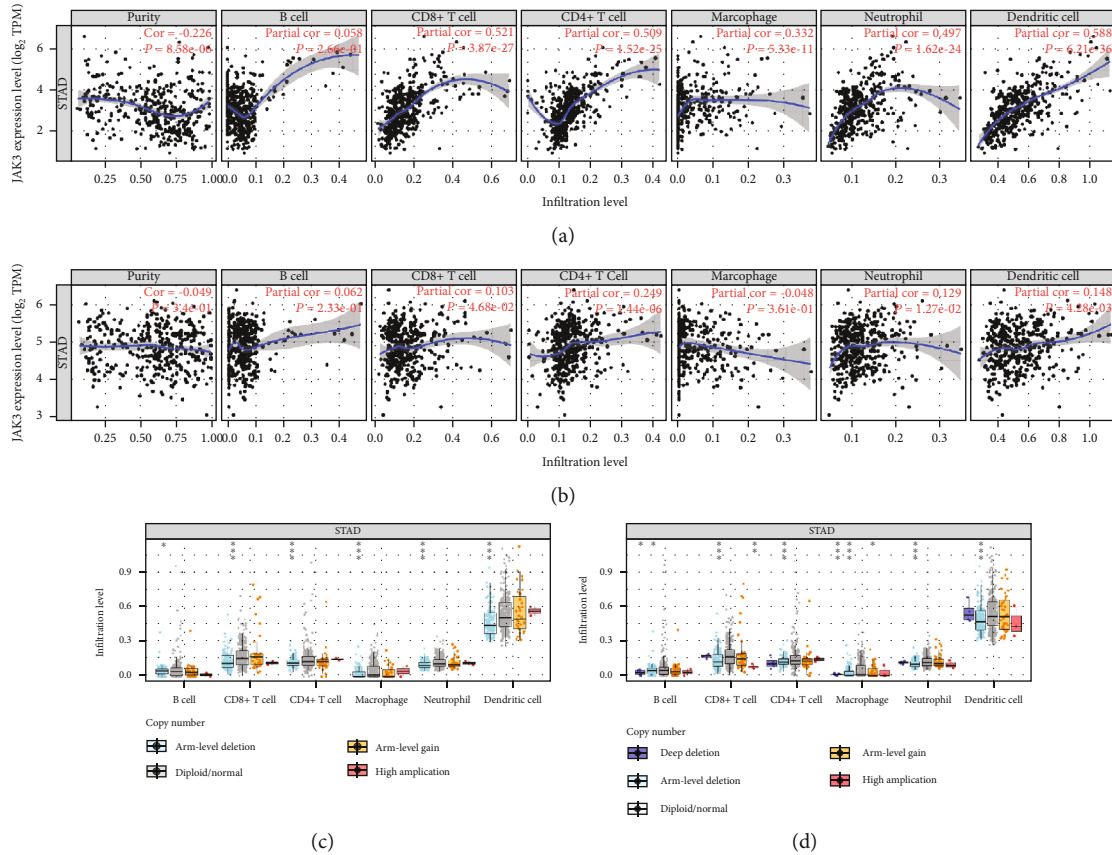


FIGURE 6: The correlation between JAK3/TYK2 and immune infiltration (TIMER). (Aa, b) The correlation between JAK3/TYK2 expression and the abundance of CD8+ T cells, CD4+ T cells, macrophage, neutrophils, and dendritic cells. (c, d) The correlation between SCNA of JAK3/TYK2 and immune cell infiltration. SCNA: somatic copy number alterations; * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

TABLE 4: Correlation analysis between JAK3 and gene biomarkers of immune cells in STAD (TIMER).

| Description | Biomarkers | STAD | | | |
|---------------------|-----------------|------|-------|--------|---------|
| | | None | Cor | Purity | P value |
| CD8+ T cell | CD8A | | 0.7 | | *** |
| | CD8B | | 0.553 | | *** |
| T cell (general) | CD3D | | 0.711 | | *** |
| | CD3E | | 0.735 | | *** |
| | CD2 | | 0.701 | | *** |
| B cell | CD19 | | 0.658 | | *** |
| | CD79A | | 0.629 | | *** |
| Monocyte | CD86 | | 0.562 | | *** |
| | CD115(CSF1R) | | 0.541 | | *** |
| TAM | CCL2 | | 0.441 | | *** |
| | CD68 | | 0.318 | | *** |
| | IL10 | | 0.482 | | *** |
| M1 macrophage | INOS (NOS2) | | 0.135 | | ** |
| | IRF5 | | 0.401 | | *** |
| | COX2(PTGS2) | | 0.036 | | 0.465 |
| M2 macrophage | CD163 | | 0.482 | | *** |
| | VSIG4 | | 0.389 | | *** |
| | MS4A4A | | 0.474 | | *** |
| Neutrophils | CD66b (CEACAM8) | | 0.054 | | 0.269 |
| | CD11b (ITGAM) | | 0.563 | | *** |
| | CCR7 | | 0.725 | | *** |
| Natural killer cell | KIR2DL1 | | 0.283 | | ** |
| | KIR2DL3 | | 0.264 | | *** |
| | KIR2DL4 | | 0.301 | | *** |
| | KIR3DL1 | | 0.29 | | *** |
| | KIR3DL2 | | 0.433 | | *** |
| | KIR3DL3 | | 0.103 | | * |
| | KIR2DS4 | | 0.283 | | ** |
| Dendritic cell | HLA-DPB1 | | 0.581 | | *** |
| | HLA-DQB1 | | 0.466 | | *** |
| | HLA-DRA | | 0.496 | | *** |
| | HLA-DPA1 | | 0.495 | | *** |
| | BDCA-1(CD1C) | | 0.52 | | *** |
| | BDCA-4(NRP1) | | 0.472 | | *** |
| | CD11c (ITGAX) | | 0.64 | | *** |
| Th1 | T-bet (TBX21) | | 0.753 | | *** |
| | STAT4 | | 0.759 | | *** |
| | STAT1 | | 0.466 | | *** |
| | IFN-g (IFNG) | | 0.438 | | *** |
| | TNF-a (TNF) | | 0.361 | | *** |

TABLE 4: Continued.

| Description | Biomarkers | STAD | | | |
|-------------------|----------------|-------|---------|--------|---------|
| | | None | STAD | Purity | |
| | | Cor | P value | Cor | P value |
| Th2 | GATA3 | 0.633 | *** | 0.625 | *** |
| | STAT6 | 0.296 | *** | 0.321 | *** |
| | STAT5A | 0.581 | *** | 0.576 | *** |
| | IL13 | 0.214 | *** | 0.218 | *** |
| Tfh | BCL6 | 0.426 | *** | 0.413 | *** |
| | IL21 | 0.384 | *** | 0.362 | *** |
| Th17 | STAT3 | 0.455 | *** | 0.467 | *** |
| | IL17A | 0.075 | 0.12 | 0.087 | 0.0919 |
| Treg | FOXP3 | 0.711 | *** | 0.679 | *** |
| | CCR8 | 0.657 | *** | 0.644 | *** |
| | STAT5B | 0.534 | *** | 0.545 | *** |
| | TGFb (TGFB1) | 0.536 | *** | 0.52 | *** |
| T cell exhaustion | PD-1 (PDCD1) | 0.725 | *** | 0.718 | *** |
| | CTLA4 | 0.645 | *** | 0.623 | *** |
| | LAG3 | 0.608 | *** | 0.597 | *** |
| | TIM-3 (HAVCR2) | 0.567 | *** | 0.545 | *** |
| | GZMB | 0.449 | *** | 0.407 | *** |

* $P < 0.05$, ** $P < 0.001$, and *** $P < 0.001$.

In this study, we found that the expression levels of JAK3 and TYK2 were higher in tumor tissues than in normal tissues in STAD. Further analysis revealed that JAK3 and TYK2 served as prognostic biomarkers in STAD and were associated with tumorigenesis, progression, and metastasis of STAD. Previous studies have also suggested that JAKs serve as biomarkers in certain types of cancers. In clear cell renal cell carcinoma, JAK3 acted as a novel biomarker and was associated with immune infiltration [26]. Another study revealed that JAK2 was a prognostic biomarker in skin cutaneous melanoma and was involved in gene regulation [10]. Moreover, JAK2 and TYK2 were suggested to be potential biomarkers for the diagnosis of hepatocellular carcinoma.

Another significant finding of our study is that JAK3 and TYK2 were associated with the abundance of immune cells, including CD8+ T cells, CD4+ T cells, neutrophils, and dendritic cells. Moreover, the expression levels of JAK3/TYK2 were positively correlated with the expression levels of immune biomarkers in STAD, demonstrating that JAK3 and TYK2 may play a vital role in immune escape in the STAD microenvironment. Previous studies have also clarified the significant role of JAK3 and TYK2 in the tumor microenvironment and immune response. JAK3 has been reported to be involved in hematopoiesis during T cell development by mediating innate and adaptive immunity-associated signaling [28]. Another study has reported that JAK3 deficiency can inhibit the development of innate lymphoid cells [29]. In lung cancer, JAK3 variants can

promote PD-L1 induction in the tumor immune microenvironment and JAK3 activation may contribute to the long-term efficacy of PD-L1 [30]. A CTLA-4-TYK2-STAT3 axis has been reported in B cell lymphoma cells and tumor-associated B cells and is relevant to immune checkpoint therapy [31].

In this study, enrichment analysis was performed, which revealed the functions and pathways of JAK3 and TYK2 in STAD, indicating that JAK3 and TYK2 were mainly associated with adaptive immune responses, translational initiations, DNA damage responses, chemokine signaling pathway, NF-kappa B signaling pathway, ribosomal structure, and JAK-STAT signaling pathway. It is well known that NF-kappa B signaling pathway is involved in inflammation and innate immunity and plays a vital role in cancer initiation and progression [32]. Moreover, NF-kB suppression can inhibit tumor cell growth and promote cell apoptosis in cholangiocarcinoma [33]. Increasing evidence has also highlighted the significant role of JAK/STAT/NF-kB signaling pathway in the immune response, axial spondyloarthritis, type 2 diabetes, metabolic disorders, and cancers [34–38]. Thus, JAK3 and TYK2 may exert functions in STAD via JAK-STAT and NF-kB signaling pathway.

This study has some limitations. First, in our study, we performed analysis at an mRNA level; it would be better to verify our results at a protein level. Furthermore, validation of our results by performing *in vivo* and *in vitro* experiments is warranted.

TABLE 5: Correlation analysis between TYK2 and gene biomarkers of immune cells in STAD (TIMER).

| Description | Biomarkers | STAD | | | |
|---------------------|-----------------|------|-------|--------|---------|
| | | None | Cor | Purity | P value |
| CD8+ T cell | CD8A | | 0.298 | | *** |
| | CD8B | | 0.161 | | ** |
| T cell (general) | CD3D | | 0.214 | | *** |
| | CD3E | | 0.295 | | *** |
| | CD2 | | 0.257 | | *** |
| B cell | CD19 | | 0.296 | | *** |
| | CD79A | | 0.231 | | *** |
| Monocyte | CD86 | | 0.213 | | *** |
| | CD115(CSF1R) | | 0.302 | | *** |
| TAM | CCL2 | | 0.039 | | 0.432 |
| | CD68 | | 0.267 | | *** |
| | IL10 | | 0.288 | | *** |
| M1 macrophage | INOS (NOS2) | | 0.17 | | *** |
| | IRF5 | | 0.37 | | *** |
| | COX2(PTGS2) | | 0.009 | | 0.858 |
| M2 macrophage | CD163 | | 0.31 | | *** |
| | VSIG4 | | 0.159 | | ** |
| | MS4A4A | | 0.18 | | *** |
| Neutrophils | CD66b (CEACAM8) | | 0.018 | | 0.713 |
| | CD11b (ITGAM) | | 0.411 | | *** |
| | CCR7 | | 0.31 | | *** |
| Natural killer cell | KIR2DL1 | | 0.036 | | 0.463 |
| | KIR2DL3 | | 0.037 | | 0.448 |
| | KIR2DL4 | | 0.158 | | ** |
| | KIR3DL1 | | 0.114 | | 0.02 |
| | KIR3DL2 | | 0.145 | | ** |
| | KIR3DL3 | | 0.094 | | 0.0554 |
| | KIR2DS4 | | 0.09 | | 0.0673 |
| Dendritic cell | HLA-DPB1 | | 0.26 | | *** |
| | HLA-DQB1 | | 0.249 | | *** |
| | HLA-DRA | | 0.263 | | *** |
| | HLA-DPA1 | | 0.259 | | *** |
| | BDCA-1(CD1C) | | 0.175 | | *** |
| | BDCA-4(NRP1) | | 0.26 | | *** |
| | CD11c (ITGAX) | | 0.393 | | *** |
| Th1 | T-bet (TBX21) | | 0.372 | | *** |
| | STAT4 | | 0.31 | | *** |
| | STAT1 | | 0.373 | | *** |
| | IFN-g (IFNG) | | 0.24 | | *** |
| | TNF-a (TNF) | | 0.254 | | *** |

TABLE 5: Continued.

| Description | Biomarkers | STAD | | | |
|-------------------|----------------|-------|---------|--------|---------|
| | | None | STAD | Purity | |
| | | Cor | P value | Cor | P value |
| Th2 | GATA3 | 0.218 | *** | 0.235 | *** |
| | STAT6 | 0.435 | *** | 0.438 | *** |
| | STAT5A | 0.546 | *** | 0.561 | *** |
| | IL13 | 0.101 | 0.039 | 0.118 | 0.0213 |
| Tfh | BCL6 | 0.232 | *** | 0.233 | *** |
| | IL21 | 0.216 | *** | 0.211 | *** |
| Th17 | STAT3 | 0.457 | *** | 0.456 | *** |
| | IL17A | 0.057 | 0.25 | 0.068 | 0.189 |
| Treg | FOXP3 | 0.439 | *** | 0.467 | *** |
| | CCR8 | 0.372 | *** | 0.379 | *** |
| | STAT5B | 0.493 | *** | 0.489 | *** |
| | TGFb (TGFB1) | 0.288 | *** | 0.293 | *** |
| T cell exhaustion | PD-1 (PDCD1) | 0.411 | *** | 0.448 | *** |
| | CTLA4 | 0.322 | *** | 0.347 | *** |
| | LAG3 | 0.27 | *** | 0.296 | *** |
| | TIM-3 (HAVCR2) | 0.306 | *** | 0.314 | *** |
| | GZMB | 0.159 | ** | 0.174 | *** |

* $P < 0.05$, ** $P < 0.001$, and *** $P < 0.001$.

In conclusion, our results demonstrated that JAK3 and TYK2 serve as prognostic biomarkers and are associated with immune infiltration in STAD, providing additional data about biomarkers, STAD prognosis, and therapy.

Data Availability

The analyzed data sets generated during the study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Lingkai Meng and Wang Li were responsible for the design of the study and writing the manuscript. Ling Ding and Yue Yu were responsible for data analysis work. All authors read and approved the final manuscript.

Acknowledgments

This study was funded by the Scientific Research Funding of Tianjin Medical University Chu Hsien-I Memorial Hospital (account number 2018ZDKF03, 2018ZDKF04).

Supplementary Materials

Supplementary Figure 1 The expression of JAKs in STAD (UALCAN). The expressions of JAK1, JAK2, JAK3, and TYK2 were significantly elevated in STAD tissues at mRNA level. STAD: stomach adenocarcinoma; *** $P < 0.001$. Supplementary Figure 2. The enrichment analysis of JAK3 in STAD (LinkedOmics). (A) A Pearson test was used to analyze correlations between JAK3 and genes differentially expressed in STAD. (B, C) Heat maps showing genes positively and negatively correlated with JAK3 in STAD (Top 50). Red indicates positively correlated genes, and green indicates negatively correlated genes. (D–F) Heat map of GO enrichment in CC terms, BP terms, and MF terms. (G) KEGG pathways analysis. GO and KEGG were performed by Gene Set Enrichment Analysis. GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; BP: biological process; CC: molecular function; MF: molecular functions. Supplementary Figure 3. KEGG pathway annotations of the cytokine-cytokine receptor interaction. GO and KEGG were performed by Gene Set Enrichment Analysis. Supplementary Figure 4. The enrichment analysis of TYK2 in STAD (LinkedOmics). (A) A Pearson test was used to analyze correlations between TYK2 and genes differentially expressed in STAD. (B, C) Heat maps showing genes positively and negatively correlated with TYK2 in STAD (Top 50). Red indicates positively correlated genes, and green indicates negatively correlated genes. (D–F) Heat map of GO enrichment in CC terms, BP terms, and MF terms. (G) KEGG pathways

analysis. GO and KEGG were performed by Gene Set Enrichment Analysis. GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; BP: biological process; CC: molecular function; MF: molecular functions. Supplementary Figure 5. KEGG pathway annotations of the ribosome. KEGG: Kyoto Encyclopedia of Genes and Genomes. (*Supplementary Materials*)

References

- [1] M. Venerito, A. C. Ford, T. Rokkas, and P. Malfertheiner, "Prevention and management of gastric cancer," *Helicobacter*, vol. 25, p. e12740, 2020.
- [2] 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.
- [3] A. Digkila and A. D. Wagner, "Advanced gastric cancer: current treatment landscape and future perspectives," *World Journal of Gastroenterology*, vol. 22, no. 8, pp. 2403–2414, 2016.
- [4] B. Groner and V. von Manstein, "Jak Stat signaling and cancer: opportunities, benefits and side effects of targeted inhibition," *Molecular and Cellular Endocrinology*, vol. 451, pp. 1–14, 2017.
- [5] J. Pencik, H. T. T. Pham, J. Schmoellerl et al., "JAK-STAT signaling in cancer: from cytokines to non-coding genome," *Cytokine*, vol. 87, pp. 26–36, 2016.
- [6] G. Colon, T. Saccon, A. Schneider et al., "The enigmatic role of growth hormone in age-related diseases, cognition, and longevity," *Geroscience*, vol. 41, no. 6, pp. 759–774, 2019.
- [7] S. Gonzalo and N. Coll-Bonfill, "Genomic instability and innate immune responses to self-DNA in progeria," *Geroscience*, vol. 41, no. 3, pp. 255–266, 2019.
- [8] D. E. Johnson, R. A. O'Keefe, and J. R. Grandis, "Targeting the IL-6/JAK/STAT3 signalling axis in cancer," *Nature Reviews. Clinical Oncology*, vol. 15, no. 4, pp. 234–248, 2018.
- [9] M. Forero-Castro, C. Robledo, R. Benito et al., "Mutations in TP53 and JAK2 are independent prognostic biomarkers in B-cell precursor acute lymphoblastic leukaemia," *British Journal of Cancer*, vol. 117, no. 2, pp. 256–265, 2017.
- [10] W. Zhang, H. Zhao, J. Chen et al., "Mining database for the expression and gene regulation network of JAK2 in skin cutaneous melanoma," *Life Sciences*, vol. 253, p. 117600, 2020.
- [11] X. Wang, X. Liao, T. Yu et al., "Analysis of clinical significance and prospective molecular mechanism of main elements of the JAK/STAT pathway in hepatocellular carcinoma," *International Journal of Oncology*, vol. 55, no. 4, pp. 805–822, 2019.
- [12] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, "GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses," *Nucleic Acids Research*, vol. 45, no. W1, pp. W98–w102, 2017.
- [13] 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.
- [14] A. M. Szász, A. Lániczky, Á. Nagy et al., "Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients," *Oncotarget*, vol. 7, no. 31, pp. 49322–49333, 2016.
- [15] J. Gao, B. A. Aksoy, U. Dogrusoz et al., "Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal," *Science Signaling*, vol. 6, no. 269, p. p11, 2013.
- [16] S. V. Vasaikar, P. Straub, J. Wang, and B. Zhang, "LinkedOmics: analyzing multi-omics data within and across 32 cancer types," *Nucleic Acids Research*, vol. 46, no. D1, pp. D956–D963, 2018.
- [17] T. Li, J. Fan, B. Wang et al., "TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells," *Cancer Research*, vol. 77, no. 21, pp. e108–e110, 2017.
- [18] N. O. Siemers, J. L. Holloway, H. Chang et al., "Genome-wide association analysis identifies genetic correlates of immune infiltrates in solid tumors," *PLoS One*, vol. 12, no. 7, article e0179726, 2017.
- [19] P. Danaher, S. Warren, L. Dennis et al., "Gene expression markers of tumor infiltrating leukocytes," *Journal for Immunotherapy of Cancer*, vol. 5, no. 1, p. 18, 2017.
- [20] S. Sousa and J. Maatta, "The role of tumour-associated macrophages in bone metastasis," *Journal of Bone Oncology*, vol. 5, no. 3, pp. 135–138, 2016.
- [21] G. H. Royce, H. M. Brown-Borg, and S. S. Deepa, "The potential role of necroptosis in inflammaging and aging," *Geroscience*, vol. 41, no. 6, pp. 795–811, 2019.
- [22] R. Santín-Márquez, A. Alarcón-Aguilar, N. E. López-Diazguerrero, N. Chondrogianni, and M. Königsberg, "Sulforaphane - role in aging and neurodegeneration," *Geroscience*, vol. 41, no. 5, pp. 655–670, 2019.
- [23] G. Giordano, P. Parcesepe, M. R. D'Andrea et al., "JAK/Stat5-mediated subtype-specific lymphocyte antigen 6 complex, locus G6D (LY6G6D) expression drives mismatch repair proficient colorectal cancer," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, p. 28, 2019.
- [24] T. A. Waldmann and J. Chen, "Disorders of the JAK/STAT pathway in T cell lymphoma pathogenesis: implications for immunotherapy," *Annual Review of Immunology*, vol. 35, no. 1, pp. 533–550, 2017.
- [25] M. Shen, Z. Xu, W. Xu et al., "Inhibition of ATM reverses EMT and decreases metastatic potential of cisplatin-resistant lung cancer cells through JAK/STAT3/PD-L1 pathway," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, p. 149, 2019.
- [26] F. Liang, H. Liang, Z. Li, and P. Huang, "JAK3 is a potential biomarker and associated with immune infiltration in kidney renal clear cell carcinoma," *International Immunopharmacology*, vol. 86, p. 106706, 2020.
- [27] H. Fan, C. Li, Q. Xiang et al., "PIK3CA mutations and their response to neoadjuvant treatment in early breast cancer: a systematic review and meta-analysis," *Thoracic Cancer*, vol. 9, no. 5, pp. 571–579, 2018.
- [28] C. H. Chuang, P. G. Greenside, Z. N. Rogers et al., "Molecular definition of a metastatic lung cancer state reveals a targetable CD109-Janus kinase-Stat axis," *Nature Medicine*, vol. 23, no. 3, pp. 291–300, 2017.
- [29] M. L. Robinette, M. Cella, J. B. Telliez et al., "Jak3 deficiency blocks innate lymphoid cell development," *Mucosal Immunology*, vol. 11, no. 1, pp. 50–60, 2018.
- [30] E. M. Van Allen, H. G. Golay, Y. Liu et al., "Long-term benefit of PD-L1 blockade in lung cancer associated with JAK3 activation," *Cancer Immunology Research*, vol. 3, no. 8, pp. 855–863, 2015.

- [31] A. Karjalainen, S. Shoebridge, M. Krunić et al., “TYK2 in tumor immunosurveillance,” *Cancers*, vol. 12, no. 1, p. 150, 2020.
- [32] B. Hoesel and J. A. Schmid, “The complexity of NF- κ B signaling in inflammation and cancer,” *Molecular Cancer*, vol. 12, no. 1, p. 86, 2013.
- [33] W. Seubwai, C. Wongkham, A. Puapairoj et al., “Aberrant expression of NF- κ B in liver fluke associated cholangiocarcinoma: implications for targeted therapy,” *PLoS One*, vol. 9, no. 8, article e106056, 2014.
- [34] S. Tan, J. Xu, A. Lai et al., “Curculigoside exerts significant anti-arthritic effects in vivo and in vitro via regulation of the JAK/STAT/NF- κ B signaling pathway,” *Molecular Medicine Reports*, vol. 19, no. 3, pp. 2057–2064, 2019.
- [35] J. Świerkot, R. Sokolik, A. Czarny et al., “Activity of JAK/STAT and NF- κ B in patients with axial spondyloarthritis,” *Postępy Higieny i Medycyny Doświadczalnej (Online)*, vol. 69, pp. 1291–1298, 2015.
- [36] A. Kauppinen, T. Suuronen, J. Ojala, K. Kaarniranta, and A. Salminen, “Antagonistic crosstalk between NF- κ B and SIRT1 in the regulation of inflammation and metabolic disorders,” *Cellular Signalling*, vol. 25, no. 10, pp. 1939–1948, 2013.
- [37] H. Y. Bako, M. A. Ibrahim, M. S. Isah, and S. Ibrahim, “Inhibition of JAK-STAT and NF- κ B signalling systems could be a novel therapeutic target against insulin resistance and type 2 diabetes,” *Life Sciences*, vol. 239, p. 117045, 2019.
- [38] L. von Hoff, E. Kärger, V. Franke et al., “Autocrine LTA signaling drives NF- κ B and JAK-STAT activity and myeloid gene expression in Hodgkin lymphoma,” *Blood*, vol. 133, no. 13, pp. 1489–1494, 2019.

Retraction

Retracted: The Relationship between Urinary Stones and Gut Microbiomeby 16S Sequencing

BioMed Research International

Received 12 March 2024; Accepted 12 March 2024; Published 20 March 2024

Copyright © 2024 BioMed Research International. 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.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] C. Zhou, K. Li, L. Zhao et al., "The Relationship between Urinary Stones and Gut Microbiomeby 16S Sequencing," *BioMed Research International*, vol. 2020, Article ID 1582187, 7 pages, 2020.

Research Article

The Relationship between Urinary Stones and Gut Microbiome by 16S Sequencing

Chenhao Zhou, Kai Li, Lun Zhao, Wei Li, Zongbao Guo, Jingyao Xu, Xiaofei Qi, and Hexing Yuan 

Department of Urology Surgery, The First Affiliated Hospital of Soochow University, Soochow, China

Correspondence should be addressed to Hexing Yuan; hezong3815008241@163.com

Received 24 July 2020; Revised 16 September 2020; Accepted 27 September 2020; Published 5 October 2020

Academic Editor: Tao Huang

Copyright © 2020 Chenhao 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.

Objective. To understand the relationship between urinary stones and the gut microbiome and to screen for microbial species that may be involved in stone formation. **Methods.** Stool samples were collected from patients with urolithiasis and healthy patients between March and December 2017. The samples were analyzed by 16S sequencing to determine differences in the microbiome profiles between the two groups. The mouse model was established and was divided into two groups. Fecal samples were collected from the mice before gavage and three weeks postgavage for microbiome analysis. The microbial population of each group was analyzed to screen for microbial species that may affect the formation of urinary stones. Differences in the number of crystals in the renal tubules of the mice were examined by necropsy. **Results.** The microbial composition was different between urolithiasis patients and healthy controls. The urolithiasis patients had significantly reduced microbial abundance; however, increased proportions of Bacteroidetes and Actinobacteria were detected compared to healthy controls. Furthermore, the abundance of *Alistipes indistinctus* and *Odoribacter splanchnicus* was significantly increased in the urolithiasis patients compared to the healthy controls. In addition, the incidence of urolithiasis was much higher in the experimental mouse group (stone solution + urolithiasis patient stool) than in the control mouse group. However, the microbial abundance before gavage was not significantly different from that seen three weeks postgavage. **Conclusion.** Urolithiasis patients in this study had a different gut microbiome when compared with that of healthy individuals. The altered microbiome increased the rate of crystal formation in renal tubules and accelerated urinary stone formation in the mouse model of urolithiasis.

1. Introduction

Urolithiasis is a common disease among inpatients in the division of urology. Epidemiological studies from Europe and the U.S. have shown that 5%-10% of individuals develop urinary stones at least once in their lifetime. In recent years, the incidence of urolithiasis has been increasing in China and the country now has one of the three highest incidences of urolithiasis in the world. As our understanding of the etiology of urolithiasis deepens, metabolic risk factors for urinary stones have gained increased attention from urologists.

The scientific community has demonstrated great interest in the study of the gut microbiome in recent years. Studies have shown that the gut microbiome is associated with

numerous human diseases but its role in the pathophysiology of urolithiasis is not known. A recent study found that the composition of the gut microbiome was significantly different between urolithiasis patients and nonurolithiasis patients [1]. However, it is unclear whether the difference in microbial abundance between urolithiasis patients and controls was the cause of stone formation or the consequence of other factors, such as antibiotic exposure and diet [2]. In 1985, Allison et al. reported that *Oxalobacterium formigenes* (Oxf) was involved in the formation of calcium oxalate stones [3]. However, the potential impact of other microbial species on urolithiasis has been rarely investigated. Therefore, this study is aimed at examining the effect of the gut microbiome and its mechanism of action in urinary stone formation in order to provide a basis for further research in this area.

2. Materials and Methods

2.1. Materials

2.1.1. Experimental Animals. Twenty clean grade healthy adult mice were purchased from Shanghai SLAC Laboratory Animal Co., Ltd. Experimental animal welfare and ethics review were conducted, and the study was approved by the Animal Ethics Committee of Soochow University.

2.1.2. Reagents and Instrument. An Olympus optical microscope was used in this study. Ethylene glycol and ammonium chloride were obtained from Sigma-Aldrich.

2.1.3. Preparation of Main Solutions

(1) *Urolithiasis-Inducing Solution (Stone Solution).* 5 ml ethylene glycol and 5 g ammonium chloride were dissolved in 500 ml phosphate-buffered saline (PBS).

(2) *Stone Solution + Urolithiasis Patient Stool.* Stool samples were collected from urolithiasis patients and homogenized in distilled water. After particulates in the solution settled, the supernatant was collected and mixed with the stone solution.

(3) *Stone Solution + Healthy Control Stool.* Stool samples were collected from patients in the Division of Urology without a history of urolithiasis and homogenized in distilled water. After the particulates in the solution settled, the supernatant was collected and mixed with the stone solution.

(4) *Stone Solution + Mouse Feces.* Mouse feces were homogenized in equal volumes of distilled water. After particulates in the solution settled, the supernatant was collected and mixed with stone solution. This solution was used as a control.

(5) *Stone Solution + PBS.* PBS and an equal volume of supernatant were mixed with stone solution.

(6) *Urolithiasis Patient Stool Solution.* Stool samples were collected from urolithiasis patients and homogenized in distilled water. After particulates in the solution settled, the supernatant was collected and numbered to identify them.

(7) *Healthy Control Stool Solution.* Stool samples were collected from patients in the Division of Urology without a history of urolithiasis and homogenized in distilled water. After particulates in the solution settled, the supernatant was collected and numbered to identify them.

(8) *Mouse Fecal Solution.* Mouse feces were homogenized in an equal volume of distilled water. After particulates in the solution settled, the supernatant was collected and numbered.

2.2. Methods

2.3. Sample Collection. Stool samples were collected from patients diagnosed with urolithiasis in our hospital between March and December 2017 (experimental group) and from

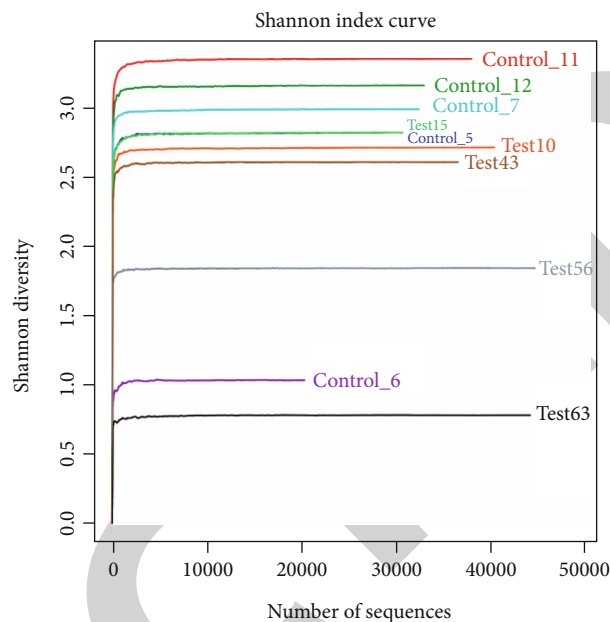


FIGURE 1: The Shannon index distribution among the experimental and control groups.

patients without urolithiasis treated in the Division of Urology during the same period (control group). Five patients from the experimental group were paired with five patients in the control group, and the urolithiasis and non-urolithiasis patient pairs were gender-matched and within two years of age. The selected patients in the experimental group were test10, test15, test43, test56, and test63. The selected patients in the control group were control_5, control_6, control_7, control_11, and control_12. The fecal microbiomes of the patients were analyzed by 16S sequencing.

Twenty healthy adult mice were randomly divided into four groups of five mice per group. The mice in each group were fed different solutions ((1) stone solution + urolithiasis patient stool, (2) stone solution + mouse feces, (3) stone solution + healthy control stool, and (4) stone solution + PBS) for four weeks. The mice were euthanized by cervical dislocation, placed in a supine position, and the abdomen was sterilized with 75% alcohol. The skin and muscle layers were incised using surgical scissors and forceps to expose the organs. The abdominal fat was gently pulled toward the head of the mouse by forceps to expose the kidneys, which were then harvested, fixed in 10% formalin, and numbered.

2.3.1. Diagnostic Criteria for Urolithiasis. Patients who met any one of the following diagnostic criteria was definitively diagnosed with urolithiasis: (1) patients with stone shadows by X-ray, (2) the indication of stones by ultrasound, or (3) surgical removal of stones or excretion of stones from the urethra following extracorporeal lithotripsy. Exclusion criteria include kidney failure, urinary tract malformations, urinary tract infections, and hyperthyroidism.

2.3.2. Inclusion Criteria for the Control Group. Patients in the Division of Urology who were previously healthy were included in the control group.

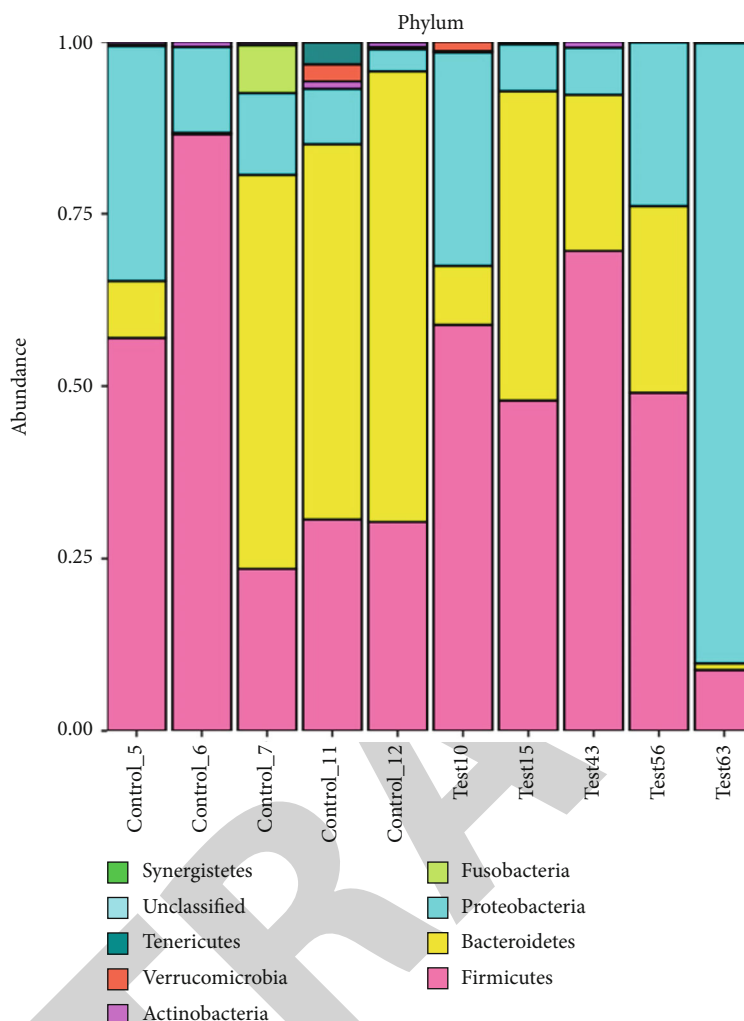


FIGURE 2: A comparison of the microbial composition at the phylum level between the experimental and control groups.

2.3.3. Exclusion criteria. Patients with a history of systemic disease, urinary stones, or other urinary tract diseases were excluded from the control group.

2.4. Statistical Analysis. All data were statistically analyzed using SPSS software (SPSS Inc., Chicago, IL, United States). Counted data are expressed as the number of cases or percentage and were compared between groups using the χ^2 test. $P < 0.05$ was considered statistically significant.

3. Results

3.1. Differences in the Microbiome. The Shannon index curves of the microbial profiles of the 10 experimental and control group patients are shown in Figure 1.

As shown in Figure 1, the Shannon index distribution was significantly different among the experimental and control groups, except for control 6; control 6 was a patient with an open fracture. The patient underwent surgery and received antibiotics upon admission, and the stool sample was collected after surgery. Therefore, the deviation in this patient may be attributable to the effect of antibiotics on the gut microbiome.

The proportions of Bacteroidetes and Firmicutes were significantly different between the healthy controls and urolithiasis patients (Figure 2).

3.1.1. Effect of the Microbiome on Stone Formation. Mice were fed four different stone solution mixtures for four weeks. Upon euthanization by cervical dislocation, the spleens were collected and examined. The number of renal tubules containing crystals and the ratio of crystal-containing tubules to the total number of tubules were calculated and compared among the samples.

As shown in Figure 3, mice fed stone solution + urolithiasis patient stool had significantly higher crystal-containing tubule to total tubule ratios than the other three groups. In addition, the crystal-containing tubule to total tubule ratio varied among all four groups, indicating that the urolithiasis patient stools accelerated the rate of crystal formation and hence stone formation in the renal tubules of mice.

3.1.2. Analysis of Microbial Species. We identified 400 microbial species belonging to the Bacteroides, Heliobacillus, Clostridium, and Fusobacterium genera.

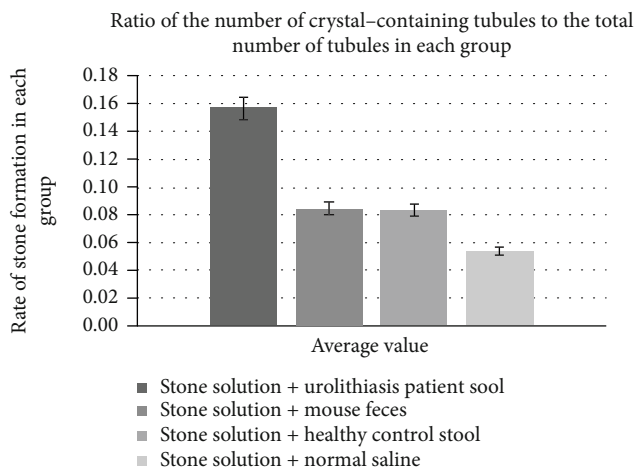


FIGURE 3: Rate of stone formation in each group.

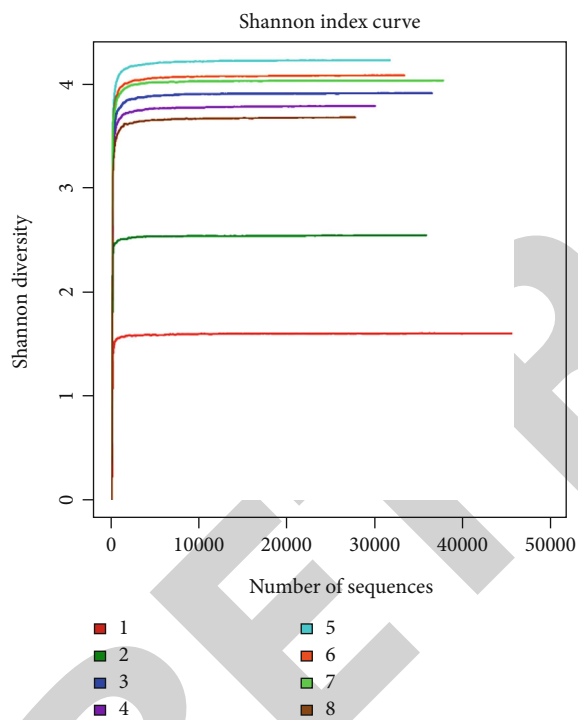


FIGURE 4: Shannon index. (1) Healthy control stool. (2) Urolithiasis patient stool. (3) Mouse feces before PBS gavage. (4) Mouse feces after PBS gavage. (5) Mouse feces before urolithiasis patient stool gavage. (6) Mouse feces after urolithiasis patient stool gavage. (7) Mouse feces before healthy control stool gavage. (8) Mouse feces after healthy control stool gavage.

Figures 4 and 5 show differences in the Shannon index and phylum between the eight groups. Group 2 (urolithiasis patient stool) and group 6 (mouse feces after urolithiasis patient stool gavage) had significantly higher Shannon indices than the other groups. At the phylum level, group 2 (urolithiasis patient stool) had significantly increased proportions of Proteobacteria and Fusobacteria than group 1 (healthy control stool), a difference also seen between group 5 (mouse feces before urolithiasis patient stool gavage) and group 6

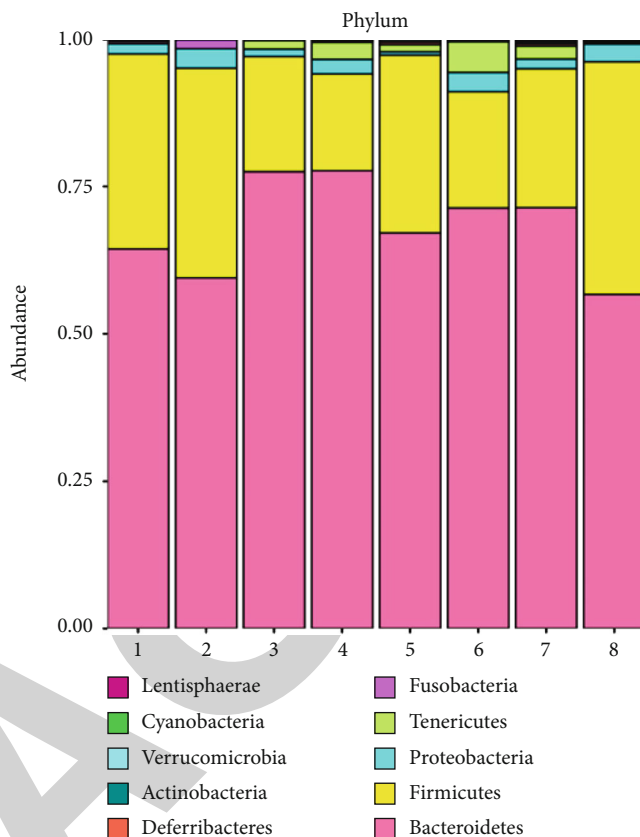


FIGURE 5: Phylum. (1) Healthy control stool. (2) Urolithiasis patient stool. (3) Mouse feces before PBS gavage. (4) Mouse feces after PBS gavage. (5) Mouse feces before urolithiasis patient stool gavage. (6) Mouse feces after urolithiasis patient stool gavage. (7) Mouse feces before healthy control stool gavage. (8) Mouse feces after healthy control stool gavage.

(mouse feces after urolithiasis patient stool gavage). These findings further support our earlier conclusion.

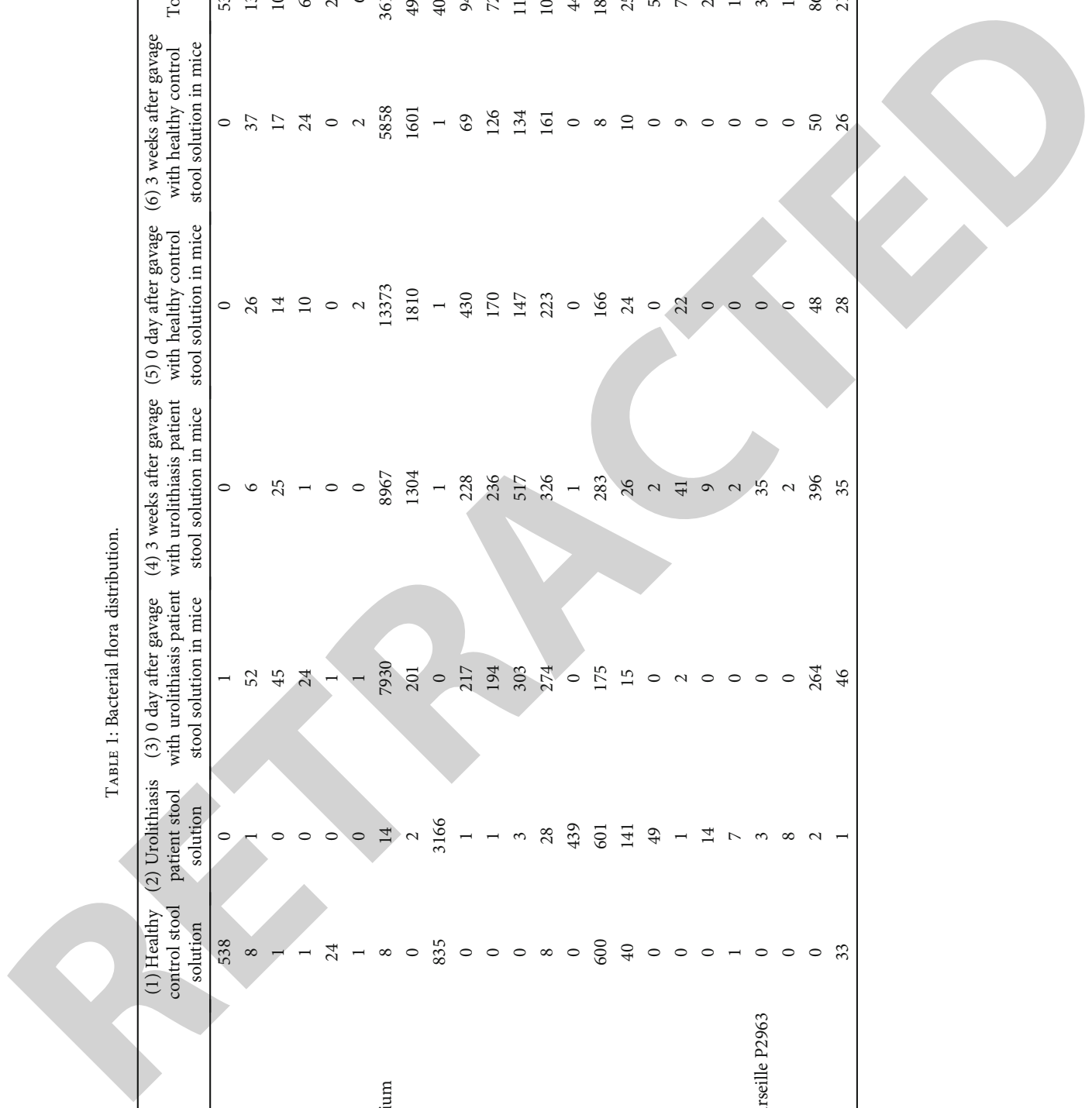
3.1.3. Screening of Target Species. We compared the fecal microbial profiles of urolithiasis patients, healthy controls, mice before and after gavage with urolithiasis patient stool, and mice before and after gavage with healthy control stool. By examining the synchronous changes in the abundance of microbial species before and after gavage, we identified 24 candidate species that may influence stone formation (Table 1).

After removing some species that have not yet been studied, we selected two species with the greatest change in abundance after fecal gavage, *Aslitipes indistinctus* and *Odoribacter splanchnicus*.

A. indistinctus is a Gram-negative bacterium that metabolizes glucose into succinic acid and acetic acid in PYG broth medium base. This bacterium is spherical or rod shaped, strictly anaerobic, and about $0.5-0.7 \times 1.0-3.8 \mu\text{m}$ in size. Gray, slightly opaque round colonies measuring 0.1-0.5 mm in diameter can be observed after four days of anaerobic culture in the modified GAM agar medium. The major end products of glucose fermentation by *A. indistinctus* in the PYG broth medium base are succinic acid and acetic acid.

TABLE 1: Bacterial flora distribution.

| Name of species | (1) Healthy control stool solution | (2) Urolithiasis patient stool solution | (3) 0 day after gavage with urolithiasis patient stool solution in mice | (4) 3 weeks after gavage with urolithiasis patient stool solution in mice | (5) 0 day after gavage with healthy control stool solution in mice | (6) 3 weeks after gavage with healthy control stool solution in mice | Total |
|---|------------------------------------|---|---|---|--|--|-------|
| Bacteroides_coprocola | 538 | 0 | 1 | 0 | 0 | 0 | 539 |
| Unidentified OUT 112 | 8 | 1 | 52 | 6 | 26 | 37 | 130 |
| Uncultured bacterium OUT 148 | 1 | 0 | 45 | 25 | 14 | 17 | 102 |
| OUT 215 | 1 | 0 | 24 | 1 | 10 | 24 | 60 |
| Dorea formicigenans | 24 | 0 | 1 | 0 | 0 | 0 | 25 |
| Unidentified OUT 416 | 1 | 0 | 1 | 0 | 2 | 2 | 6 |
| Uncultured Bacteroidales bacterium | 8 | 14 | 7930 | 8967 | 13373 | 5858 | 36150 |
| Uncultured organism OUT 11 | 0 | 2 | 201 | 1304 | 1810 | 1601 | 4918 |
| OUT 13 | 835 | 3166 | 0 | 1 | 1 | 1 | 4004 |
| Uncultured bacterium OUT 32 | 0 | 1 | 217 | 228 | 430 | 69 | 945 |
| Uncultured bacterium OUT 39 | 0 | 1 | 194 | 236 | 170 | 126 | 727 |
| Uncultured bacterium OUT 46 | 0 | 3 | 303 | 517 | 147 | 134 | 1104 |
| Firmicutes bacterium ASF 500 | 8 | 28 | 274 | 326 | 223 | 161 | 1020 |
| Odoribacter splanchnicus | 0 | 439 | 0 | 1 | 0 | 0 | 440 |
| Gut metagenome | 600 | 601 | 175 | 283 | 166 | 8 | 1833 |
| (Clostridium) lactatifermentans | 40 | 141 | 15 | 26 | 24 | 10 | 256 |
| OUT 193 | 0 | 49 | 0 | 2 | 0 | 0 | 51 |
| Oscillibacter ruminantium | 0 | 1 | 2 | 41 | 22 | 9 | 75 |
| Alistipes indistinctus | 0 | 14 | 0 | 9 | 0 | 0 | 23 |
| Uncultured bacterium OUT 359 | 1 | 7 | 0 | 2 | 0 | 0 | 10 |
| Ruminococcaceae bacterium Marseille P2963 | 0 | 3 | 0 | 35 | 0 | 0 | 38 |
| Uncultured bacterium OUT 432 | 0 | 8 | 0 | 2 | 0 | 0 | 10 |
| Dorea sp. 5-2 | 0 | 2 | 264 | 396 | 48 | 50 | 862 |
| OUT 128 | 33 | 1 | 46 | 35 | 28 | 26 | 218 |



A. indistinctus is intolerant to 20% bile; does not hydrolyze heptaphylline or gelatin; does not reduce nitrate; does not produce indole, oxidase, or urease; and produces catalase. Furthermore, *A. indistinctus* metabolizes cellobiose, glucose, lactose, maltose, D-mannose, melezitose, raffinose, 1-rhamnose, salicin, sucrose, trehalose, and D-xylose to synthesize organic acids [4].

Odoribacter splanchnicus is a Gram-negative, anaerobic bacterium of the *Odoribacter* genus. *O. splanchnicus* metabolizes various sugars, including glucose, galactose, arabinose, lactose, and mannose, producing metabolites such as acetic acid, propionic acid, succinic acid, butyric acid, isovaleric acid, and isobutyric acid. *O. splanchnicus* does not reduce nitrate and contains highly active enzymes involved in the pentose metabolism pathway. This bacterium is not sensitive to aminoglycosides and polymyxins but is sensitive to tetracycline, lincomycin, clindamycin, erythromycin, and rifampicin [5].

4. Discussion

The human body is inhabited by a vast number of microbes and the interaction between these microorganisms, and their host may be critical to the health and disease of the host. However, it was not until the emergence of new molecular techniques that the diversity of this microbiome began to gain the attention of researchers. A recent study demonstrated a complex bidirectional relationship between the gut microbiome and the host and reported that this relationship may be critical to human health and may be associated with the pathogenesis of disease. The gut microbiome is not only involved in food digestion and nutrient extraction but can also alter the host immune response, prevent infection, metabolize drugs, and participate and regulate metabolism of the host [6].

Studies demonstrated that kidney stones are mineral tuberculosis in the calyces and pelvis, which are found to be free or attached to the renal papilla. When urine is oversaturated with minerals, stones (called kidney stones or urinary stones) form in the urethra, causing crystals to form, grow, accumulate, and stay in the kidneys. Globally, approximately 80% of kidney stones are made of a mixture of calcium oxalate (CaOx) and calcium phosphate (CaP). Stones composed of uric acid, magnesium phosphate, and cystine are also common, accounting for about 9%, 10%, and 1% of the total number of stones, respectively. Stones may also be oversaturated with some relatively insoluble drugs or their metabolites, which may cause the urine components to crystallize in the renal collecting duct (iatrogenic stones) [7]. About 75% of kidney stones are mainly composed of calcium oxalate, and urinary oxalate is considered a risk factor.

According to the composition of stones, urinary stones are mainly divided into five categories: calcium oxalate, calcium phosphate, magnesium ammonium phosphate, uric acid, and cystine. Among them, oxalate occupies the vast majority, exceeding 85%. The microbiota refers to the entire microbial population that colonizes in a specific location and includes not only bacteria but also other microorganisms such as fungi, archaea, viruses, and protozoa [8]. The scientific community has shown great interest in the gut microbiota in recent years; the gut microbiota is associated with

many human diseases, such as intestinal diseases such as inflammatory bowel disease (IBD) [9] and irritable bowel syndrome (IBS) [10], metabolic diseases such as obesity and diabetes [11], and allergic diseases [12] to neurodevelopmental diseases, although the strength of some evidence is not sound in many of them. It has long been believed that the gut microbiota has a significant functional role in maintaining the healthy intestines of normal individuals and humans as a whole. There is now more and more evidence that studies on humans and sterile mice support these speculations. The United States Human Microbiome Project (HMP) [13], European Human Gut Microbial Genomics (MetaHIT) [14], and several other studies have proven that normal intestinal flora is beneficial to health at the genetic level. These effects make it possible for the gut microbiota to also affect the absorption and secretion of solutes related to kidney stones.

To date, little is known about the general role of the gut microbiota in the pathophysiology of kidney stones. A recent study found that the gut microbiome of patients with kidney stones is significantly different from that of patients without stones. Whether these differences in bacterial abundance seen by stone formers and controls are the cause of stone formation or secondary to other variables such as antibiotic exposure or diet is uncertain. Allison and colleagues discovered the oxalobacteriumformigenes (*Oxf*) in 1985, which has attracted considerable attention due to its involvement in calcium oxalate stone disease. But for other bacteria that may affect the formation of urinary calculi, there are few reports.

In this study, we found that the gut microbiome of healthy controls and urolithiasis patients was significantly different and we determined the correlation between the gut microbiome and urolithiasis using fecal transplants in mice. *A. indistinctus* and *O. splanchnicus* are both Gram-negative anaerobes involved in food hydrolysis and metabolism in the intestine and may contribute to crystal formation in the kidneys. However, the factors influencing urolithiasis and the mechanisms by which these bacteria participate in metabolism need further investigation.

In summary, the urolithiasis patients in this study had a different gut microbiome when compared with that of healthy individuals. The altered microbiome increased the rate of crystal formation in renal tubules and accelerated urinary stone formation in the mouse model of urolithiasis.

Data Availability

Our data would be available if checking process required.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

- [1] I. E. Mulder, B. Schmidt, C. R. Stokes et al., "Environmentally-acquired bacteria influence microbial diversity and natural

Research Article

Identification of Latent Oncogenes with a Network Embedding Method and Random Forest

Ran Zhao,¹ Bin Hu,² Lei Chen ,¹ and Bo Zhou ³

¹College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China

²State Key Laboratory of Livestock and Poultry Breeding, Guangdong Public Laboratory of Animal Breeding and Nutrition, Guangdong Provincial Key Laboratory of Animal Breeding and Nutrition, Institute of Animal Science, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China

³Shanghai University of Medicine & Health Sciences, Shanghai 201318, China

Correspondence should be addressed to Lei Chen; chen_lei1@163.com and Bo Zhou; zhoubo@sumhs.edu.cn

Received 20 July 2020; Revised 9 September 2020; Accepted 14 September 2020; Published 23 September 2020

Academic Editor: Shijia Zhu

Copyright © 2020 Ran 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.

Oncogene is a special type of genes, which can promote the tumor initiation. Good study on oncogenes is helpful for understanding the cause of cancers. Experimental techniques in early time are quite popular in detecting oncogenes. However, their defects become more and more evident in recent years, such as high cost and long time. The newly proposed computational methods provide an alternative way to study oncogenes, which can provide useful clues for further investigations on candidate genes. Considering the limitations of some previous computational methods, such as lack of learning procedures and terming genes as individual subjects, a novel computational method was proposed in this study. The method adopted the features derived from multiple protein networks, viewing proteins in a system level. A classic machine learning algorithm, random forest, was applied on these features to capture the essential characteristic of oncogenes, thereby building the prediction model. All genes except validated oncogenes were ranked with a measurement yielded by the prediction model. Top genes were quite different from potential oncogenes discovered by previous methods, and they can be confirmed to become novel oncogenes. It was indicated that the newly identified genes can be essential supplements for previous results.

1. Introduction

Cancer is the second cause of human deaths in the world, following the cardiovascular disease. Lots of people directly died from cancer per year [1]. Although several efforts have been made in recent years, the mechanism of cancers has not been fully uncovered, which makes difficulties in designing effective treatments. Genetic background and environmental factors are widely accepted to be major causes of cancers [2]. Investigation on the mechanism of cancers with related genes is an essential way to understand the tumor initiation and development.

Oncogene is an important type of cancer-related genes, which can promote the tumor initiation. Thus, it is essential to identify latent oncogenes as much as possible, promoting the understanding of cancers. In early time, experimental techniques performed on typical cell lines or animal models

are the main way for detecting oncogenes. However, this way is time-consuming and with high cost. In recent years, with the development of computer science, this procedure can be improved aided by designing computational methods. The computational methods can give a deep insight into a large-scale data and learn hidden associations between cancers and genes, thereby making useful clues and providing latent candidates. Experimenters can do targeted tests to confirm the results. Two pioneer studies have been proposed in this regard recently. The first study proposed a network method for inferring novel oncogenes based on validated oncogenes reported in some online databases [3]. The method applied the shortest path (SP) algorithm on a protein-protein interaction (PPI) network to extract the shortest paths connecting any two proteins of oncogenes. Proteins lying on these paths were picked up and screened by three measurements. 37 possible oncogenes were obtained

by this method. The second study investigated oncogenes in a quite different way [4]. It tried to uncover the functions, including Gene Ontology (GO) terms and biological pathways, of oncogenes with machine learning algorithms. They first extracted essential GOs and pathways that can indicate the differences of oncogenes and other general genes and made prediction with them. More than 800 genes were predicted to be novel oncogenes. All of the above two studies proposed some putative oncogenes; some of which were extensively discussed. However, the limitations also exist. For the network method proposed in the first study [3], it did not contain a learning procedure, indicated that it cannot capture the essential features of oncogenes, inducing several false positive oncogenes. Although the second method [4] contained a learning procedure, it did not include the protein association information. As proteins with strong associations always share common functions, the protein association information is powerful materials for discovering novel oncogenes.

In view of the limitations of the above two studies, this study proposed a new computational method. The protein networks, derived from protein associations, were constructed, from which informative features were extracted to represent genes. The classic machine learning algorithm, random forest (RF) [5], was adopted to capture essential features of oncogenes and build the model. The latent oncogenes were ranked by a measurement yielded by the proposed method. Top latent oncogenes were quite different from those reported in previous two studies. We also analyzed some top latent oncogenes to confirm their likelihood of being oncogenes.

2. Materials and Methods

2.1. Materials. Validated oncogenes were directly downloaded from a previous study [3], which were collected from HUGO Gene Nomenclature Committee (HGNC, <https://www.genenames.org/>) [6] and Gene Set Enrichment Analysis Molecular Signatures Database (GSEA MSigDB, https://www.broadinstitute.org/gsea/msigdb/gene_families.jsp) [7, 8]. From HGNC, 251 oncogenes were collected and 330 oncogenes were retrieved from GSEA MSigDB. 543 oncogenes were obtained after combining above two sets of oncogenes. Because we used protein-protein interaction (PPI) networks to identify latent oncogenes, where proteins were represented by Ensembl IDs, proteins encoded by these 543 oncogenes were extracted and they were further mapped onto their Ensembl IDs. After excluding Ensembl IDs that were not in the PPI networks, we finally accessed 481 Ensembl IDs. With these IDs, we designed a computational method to identify new possible IDs. These new IDs suggested latent oncogenes.

2.2. Protein-Protein Interaction and Network Construction. In recent years, it is quite popular to adopt networks for investigating various diseases [3, 9–14]. Networks can organize data and information in a system level. It is beneficial to study different problems in a more complete view. Thus, we employed PPI networks to identify novel oncogenes in present study.

In this study, we adopted the PPI data reported in STRING (<https://string-db.org/>, Version 10.0) [15, 16], a well-known public database collecting known and predicted PPIs. Currently, 9,643,763 proteins from 2,031 organisms comprise huge numbers of PPIs, which were collected from a variety of sources, such as genomic background prediction, high-throughput laboratory experiments, (conservative) coexpression, automated textualization, and prior knowledge in databases. Thus, each interaction contains the physical and functional associations of two proteins and can widely measure the linkage between proteins. Compared with the PPIs reported in other databases, such as DIP (Database of Interaction Proteins) database [17] and BioGRID [18] which only include experimentally validated PPIs, PPIs in STRING contain more information and are more helpful for building models with a complete view. For human, 4,274,001 PPIs are reported in STRING, covering 19,247 human proteins. Each PPI consists of two proteins, encoded by Ensembl IDs. Further, STRING evaluates the strength of each PPI from eight different aspects and assigns eight scores to each PPI, titled by “Neighborhood”, “Fusion”, “Cooccurrence”, “Coexpression”, “Experiment”, “Database”, “Textmining”, and “Combined_score”. The last score combines other seven scores in a naive Bayesian fashion [16]. It was not adopted in the present study because it may produce redundancy with other seven scores. For each of the other scores, one PPI network was constructed in the following manner. We took the “Neighborhood” score as an example. First, 19,247 proteins were picked up as nodes. Second, two nodes were adjacent if and only if the “Neighborhood” score between corresponding proteins was larger than zero. Finally, such “Neighborhood” score was assigned to the corresponding edge as its weight. Accordingly, seven PPI networks were constructed, which were denoted by N_N , N_F , N_{CO} , N_{CE} , N_E , N_D , and N_T , respectively. The sizes (numbers of edges) of seven networks were quite different. N_T had most edges (3816497), followed by N_E (1736931 edges), N_{CE} (768962 edges), N_D (212430 edges), N_N (76214 edges), N_{CO} (23739 edges), and N_F (2060 edges).

2.3. Feature Engineering. Network is excellent to organize the associations of the proteins, which can view a specific protein in a system level. However, there is a gap between it and the traditional machine learning algorithms because these algorithms always need numerical vectors as input. Fortunately, some network embedding algorithms, such as Mashup [19], Node2vec [20], and Deepwalk [21], were proposed in recent years, which can abstract relationship in the network and output a feature vector for each node in the network. The occurrence of these algorithms connects the network and the traditional machine learning algorithm. Considering the fact that seven networks were involved in this study, Mashup [19] was adopted. It can tackle multiple networks, which is the greatest merit compared with other network embedding algorithms. Its brief descriptions were as follows.

The procedures of Mashup encoding each node consist of two stages. In the first stage, it applies the random walk with the restart (RWR) algorithm [22–26] on each network to construct a raw feature vector for each node in this network.

In detail, for a network N_j ($j \in \{N, F, CO, CE, E, D, T\}$) constructed in Section 2.2, each node p_i in this network was picked up as the node seed of the RWR algorithm one by one. When the RWR algorithm stopped, each node in this network was assigned a probability, indicating its associations to the seed node. A raw feature vector of p_i was built by collecting all these probabilities, which was denoted by V_j^i . Because some nodes can occur in multiple networks, several feature vectors were generated for these nodes. It is necessary to fuse them into one feature vector in a rigorous way. On the other hand, a dimensionality reduction procedure is also necessary due to the large dimension of the raw feature vectors. All these are the purpose of the second stage of Mashup. Let X^i be the final vector of protein p_i and W_j^i be a context feature vector of p_i in the network N_j . It is clear that Mashup tries to determine the optimal components in X^i and W_j^i , which retain the essential information in V_j^i as much as possible. Based on X^i and W_j^i , a vector, denoted by \tilde{V}_j^i , can be constructed. Its components were defined as follows:

$$\tilde{V}_{jk}^i = \frac{\exp\left((X^i)^T W_j^k\right)}{\sum_k \exp\left((X^i)^T W_j^k\right)}, k = 1, 2, \dots, n \quad (1)$$

where n is the total number of different nodes (proteins) in seven networks. The following problem is to find out the optimal components in X^i and W_j^i , which can generate \tilde{V}_j^i approximating V_j^i as much as possible. An optimization problem is set up to determine the optimal components, which is formulated as below:

$$\text{minimize}_{X^i, W_j^i} \frac{1}{n} \sum_{j=1}^m \sum_{i=1}^n D_{KL}\left(V_j^i \parallel \tilde{V}_j^i\right), \quad (2)$$

where m stands for the total number of networks and $D_{KL}(\bullet)$ stands for the function of KL-divergence (relative entropy).

The present study used the Mashup program obtained from <http://cb.csail.mit.edu/cb/mashup/>. The dimension of the output vector is a main parameter of the Mashup. Several dimensions, varying from 100 to 1000, were tried in this study because we did not know which dimension was best.

2.4. Random Forest. The network embedding algorithm, Mashup, connects the networks and traditional machine learning algorithms. With the feature vectors extracted from seven protein networks via Mashup, a specific machine learning algorithm can deeply study the characteristic of vectors of oncogenes, thereby building a prediction model. This study selected the classic machine learning algorithm, RF [5], due to its wide applications in bioinformatics and medical informatics [27–34].

RF is an ensemble classification algorithm, which consists of several decision trees. Given a dataset with n samples and m features, RF constructs each decision tree in the following manner. Randomly selected n samples, with replacement, from the original dataset. A decision tree is constructed based

on selected samples. When the tree is grown at a node, randomly select m' features, where m' is much smaller than m , and the optimal splitting way is determined based on these m' features. For an input sample, each decision tree first makes its prediction. RF integrates these predictions by majority voting. Although, decision tree is a weak classifier, RF is deemed to be much strong and competitive compared with other advanced classification algorithms [35].

In this study, a tool “RandomForest” in Weka [36] was directly employed, which implements the above-mentioned RF. Default parameters were used, where the number of decision trees was set to ten.

2.5. The Proposed Method for Inferring Latent Oncogenes. Among the 19,247 proteins occurring in seven protein networks, 481 are encoded by validated oncogenes, whereas the rest 18,766 proteins have not been labelled. It is clear that some proteins may be encoded by latent oncogenes. The proposed method can discover novel latent oncogenes with the feature vectors obtained Section 2.3 and RF described in Section 2.4.

For 18,766 unlabelled proteins, the proposed method evaluated their likelihood of being oncogenes in the following manner. For technique reasons, these 18,766 unlabelled proteins were termed as negative samples, whereas 481 proteins of oncogenes were deemed as positive samples. Evidently, negative samples were much more than positive samples (about 39 times). Thus, we randomly divided the negative samples into 39 parts. Negative samples in each part were combined with positive samples to construct a dataset, thereby yielding 39 datasets. On each dataset, a prediction model was built with RF as the classification algorithm. Accordingly, 39 RF models were produced. For each unlabelled protein p , it was fed into 38 RF models, except the RF model containing it. Each model assigned a probability to p , suggesting its likelihood to be an oncogene. The mean of all its probabilities was finally calculated to fully measure the likelihood of it being an oncogene. For an easy description, such mean value was called level value.

After all 18,766 unlabelled proteins had been evaluated with the above procedures, we ranked them in a list with the decreasing order of their level values. Evidently, those with high level values were more likely to be oncogenes.

The entire procedures of the method are illustrated in Figure 1.

2.6. Performance Evaluation. To evaluate the utility of the proposed method, a procedure similar to the jackknife test [37–39] was adopted. Each of 481 proteins encoded by oncogenes was singled out one by one as unlabelled proteins. For a specific singled out protein, we want to know whether the rest 480 proteins of oncogenes can identify it. According to the procedures of the abovementioned method, 18,767 unlabelled proteins (one singled out protein and 18,766 actual unlabelled proteins) were also randomly divided into 39 parts. Each part and the positive sample set were combined to generate a dataset, thereby producing 39 datasets. Then, the same procedures of the methods followed to yield the level value of the singled out protein. After all proteins

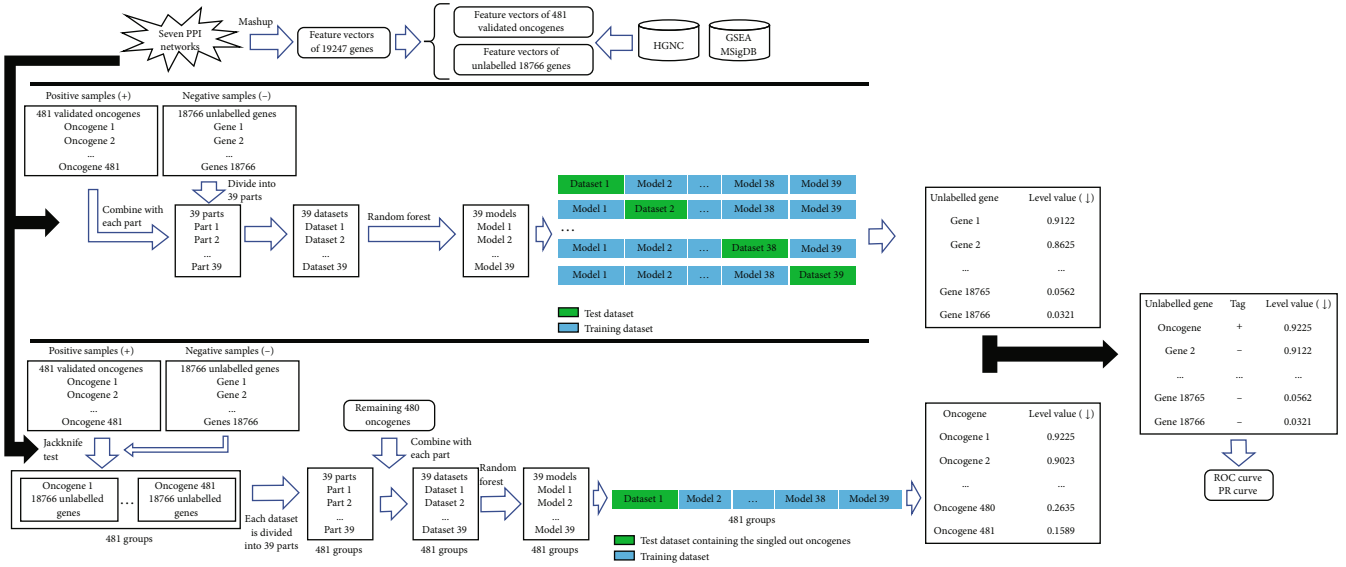


FIGURE 1: Flow chart to show the procedures of the method for inferring latent oncogenes. Seven protein-protein interaction (PPI) networks are constructed using the PPI information reported in STRING, from which feature vectors of oncogenes and unlabelled genes are extracted via Mashup. Unlabelled genes are divided into 39 parts; each of which combines the oncogenes to comprise a dataset. Each dataset induces a random forest (RF) model. A level value is computed for each unlabelled gene, which is the average of the probabilities yielded by 38 RF models. Unlabelled genes are ranked by the decreasing order of their level values. The same procedures are done for each oncogene, thereby ranking oncogenes. Finally, all genes are sorted in a list and ROC and PR curves are plotted to evaluate the performance of the method.

encoded by oncogenes had been tested, they were all assigned a level value. A protein list was created by ranking all 19,247 proteins, including proteins of oncogenes and unlabelled proteins, with the decreasing order of their level values. Some measurements can be calculated to evaluate such list, thereby indicating the utility of the method.

Given a protein list, which sorted proteins with decreasing order of their level values, whether a protein was predicted to be an oncogene (positive sample) was determined after a threshold of level value was set; that is, proteins with level values larger than the threshold were predicted to be oncogenes (positive samples); otherwise, they were predicted to be nononcogenes (negative samples). Accordingly, four values, true positive (TP), false negative (FN), false positive (FP) and true negative (TN), can be counted. Then, the sensitivity (SN) (same as recall), specificity (SP), and precision can be computed by

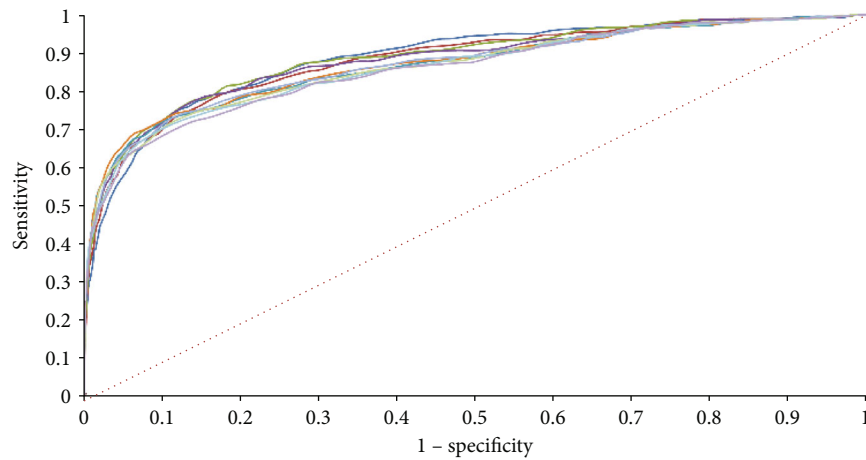
$$\begin{aligned} \text{SN}(\text{recall}) &= \frac{TP}{TP + FN}, \\ \text{SP} &= \frac{TN}{TN + FP}, \\ \text{Precision} &= \frac{TP}{TP + FP}. \end{aligned} \quad (3)$$

After setting several thresholds, we can obtain a number of SNs, SPs, and precisions. A receiver operating characteristic (ROC) [40] curve and a precision-recall (PR) curve were plotted, where the ROC curve sets SN as y -axis and $1 - \text{SP}$ as x -axis, whereas PR curve adopts precision as y -axis and recall as x -axis. Furthermore, the area under each of above two

curves can be calculated, which were called AUROC and AUPRC, respectively, in this study. Evidently, the higher the area was, the better the method was.

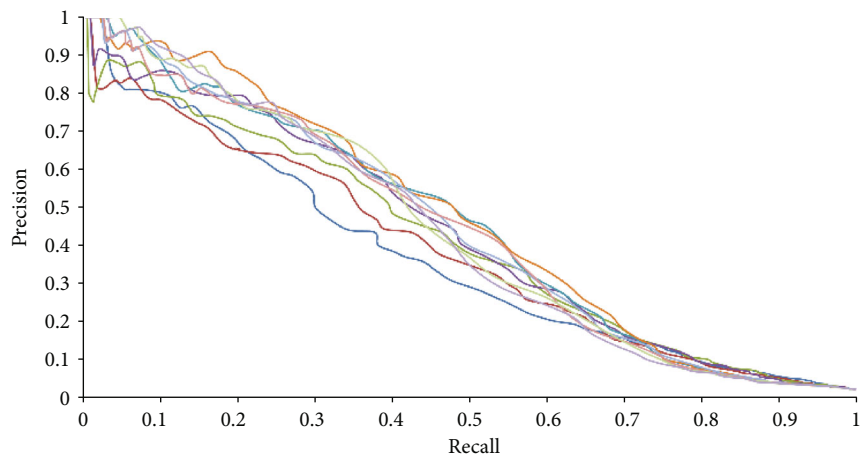
3. Results and Discussion

3.1. Performance of the Method with Different Dimensions. In this study, we used features derived from seven protein networks. Several dimensions were tried to select the best one. For each dimension, the proposed method was evaluated in the way described in Section 2.6. A ROC curve and a PR curve was plotted, as shown in Figure 2. From Figure 2(a), the method with dimension 300 yielded the highest AUROC of 0.8845, whereas the method with dimension 600 gave the highest AUPRC of 0.4676 from Figure 2(b). In general, the PR curve is a more accurate measurement than the ROC curve if the dataset is greatly imbalanced. In our study, the negative samples were about 39 times as many as positive samples. Thus, we selected the method with dimension 600 as the proposed method. To further elaborate that this selection is reasonable, we calculated the average of AUROC and AUPRC for each dimension and plotted a scatter diagram to show these averages, as illustrated in Figure 3. Evidently, the dimension 600 gave the highest average of 0.6680, supporting the above selection. The trend of average on dimension shown in Figure 3 proved the reliability of the results. Before 600, the average showed an increasing trend, while it generally descended after 600. It is reasonable because when the dimension was small, several essential information cannot be included, where the dimension was large, lots of noisy was included. All these results supported the method that with dimension 600 was the best choice because it can



| Dimension | AUROC |
|-----------|--------|
| 100 | 0.8840 |
| 200 | 0.8791 |
| 300 | 0.8845 |
| 400 | 0.8789 |
| 500 | 0.8662 |
| 600 | 0.8685 |
| 700 | 0.8672 |
| 800 | 0.8624 |
| 900 | 0.8598 |
| 1000 | 0.8560 |

(a)



| Dimension | AUROC |
|-----------|--------|
| 100 | 0.3696 |
| 200 | 0.3854 |
| 300 | 0.4136 |
| 400 | 0.4340 |
| 500 | 0.4488 |
| 600 | 0.4676 |
| 700 | 0.4431 |
| 800 | 0.4392 |
| 900 | 0.4388 |
| 1000 | 0.4320 |

(b)

FIGURE 2: ROC and PR curves to show the performance of the method with different dimensions varying from 100 to 1000. (a) ROC curves, the dimension 300 yields the highest AUROC of 0.8845; (b) PR curves, the dimension 600 yields the highest AUPRC of 0.4676.

recover actual oncogenes (positive samples) as much as possible. The unlabelled proteins predicted to be positive samples by this method were more reliable.

3.2. Inferred Oncogenes Obtained by the Proposed Method. As mentioned in Section 3.1, we selected the method with dimension 600 as the proposed method. For each unlabelled protein, it was assigned a level value by the method to indicate its likelihood of being oncogenes. The level values of all 18,766 unlabelled proteins are provided in Supplementary Material S1. Figure 4 shows the distribution of all level values.

It can be observed that one unlabelled protein was assigned the level value larger than 0.9, seven proteins were with level values between 0.8 and 0.9. These proteins are more likely to be encoded by latent oncogenes. On the other hand, majority proteins (about 96.39%) received the level values smaller than 0.6.

3.3. Comparison of Previous Studies. Two previous computational methods have been proposed for identifying possible oncogenes. The one method adopted SP algorithm to search novel oncogenes in a PPI network; thus, this method was

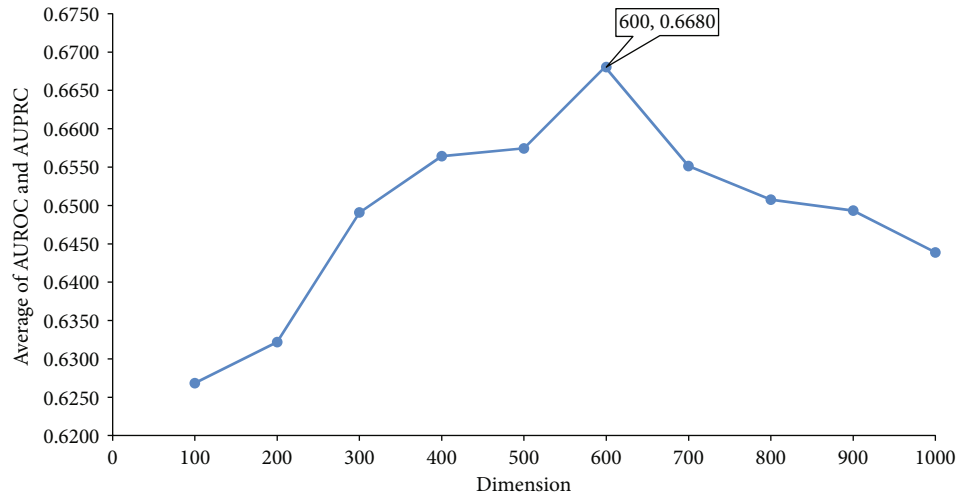


FIGURE 3: Scatter diagram to show the performance of the method with different dimensions varying from 100 to 1000. The x-axis represents dimension and the y-axis indicates the average of AUROC and AUPRC. The method with dimension 600 produces best performance.

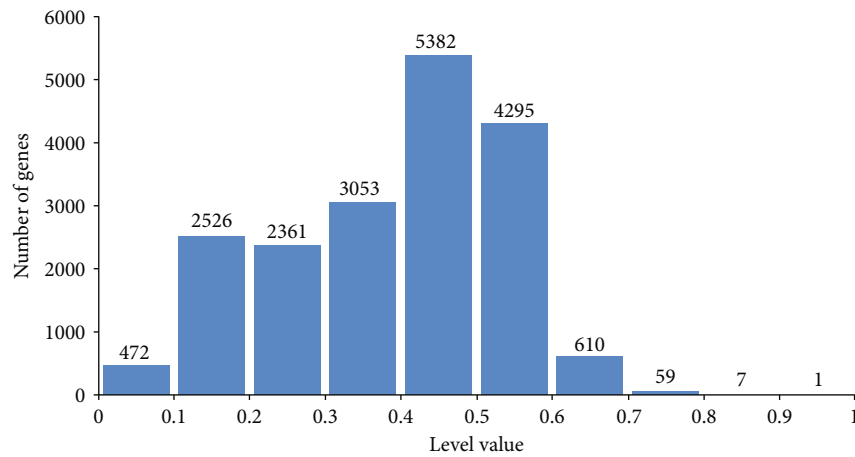


FIGURE 4: Distribution of level values on 18,766 unlabelled proteins. Only 677 unlabelled proteins (~3.61%) are assigned the level values higher than 0.6. These proteins are more likely to be encoded by latent oncogenes than others.

called the SP-based method. The other method investigated oncogenes from the point view of their functions; it was termed as function-based method in this study. These previous methods all yielded some latent oncogenes. A comparison was performed in this section.

As our method only ranks the candidates with their level values, we set some thresholds to select inferred genes to make comparisons. The thresholds included 0.8, 0.7, and 0.6, yielding eight, 67 and 677 inferred oncogenes, respectively. The intersection of these inferred oncogene sets and two oncogene sets yielded by previous methods is illustrated in Figure 5. When the threshold was set to 0.8, only one gene (HOXA10) was also identified by the SP-based method. 25 inferred oncogenes were shared by either SP- or function-based methods when the threshold was 0.7, where two genes (HOXA10, AR) were inferred by all three methods. For the threshold 0.6, this number was 246, where eight genes (HOXA10, AR, ESR1, NOTCH3, PTPN6, MYO5A, KIAA0100, and MAP2K1) were shared by all methods. The

exclusive oncogenes yielded by the proposed method occupied 87.5% when 0.8 was set as the threshold. Such percent was 62.69% and 63.66% for the thresholds 0.7 and 0.6, respectively. These results indicate that majority top inferred oncogenes of our method were not discovered by previous methods, indicating that our method can discover novel latent oncogenes that cannot be identified by previous methods.

3.4. Analysis of Top Inferred Oncogenes. In this study, some latent oncogenes were inferred by the proposed computational method. Each gene was assigned a level value to indicate its likelihood of being oncogenes. Table 1 lists the top fifteen inferred oncogenes. This section selected four of them for detailed analysis.

3.4.1. RAB31. Such gene is the top identification with level value 0.9105. RAB31 (Ras-related protein in brain 31), a member of the RAB family, encodes a protein belonging to

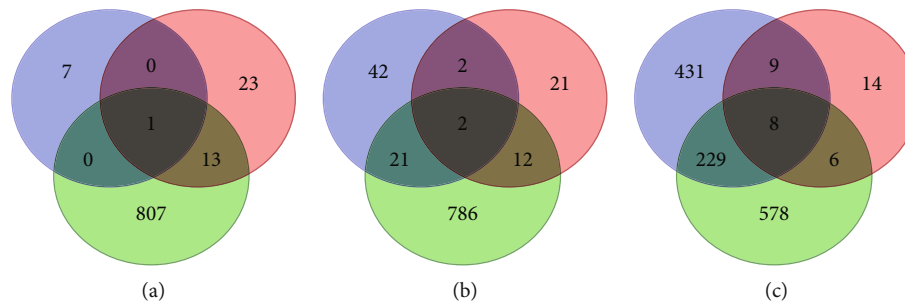


FIGURE 5: Venn diagrams to show the intersection of inferred oncogenes yielded by three methods. Blue, red, and green circles represent inferred oncogenes obtained by the proposed method, SP-based method, and function-based method, respectively. (a) Inferred oncogenes with level values larger than 0.8 are considered. (b) Inferred oncogenes with level values larger than 0.7 are considered. (c) Inferred oncogenes with level values larger than 0.6 are considered.

the Ras superfamily of small GTPases. Because it was a significant homology with RAB22 (71% sequence identity), RAB31 was also named RAB22b. Similar to other members of the RAB family, it functions as molecular switches and plays critical roles in cell adhesion molecules and membrane trafficking of growth factor receptors [41]. Therefore, it is also conceivable that RAB31-mediated dysregulation in endocytosis or recirculation may result in failure to control cell proliferation, adhesion, and migration. As expected, its promotive effect on tumor progression has been reported in several types of cancers [42]. In breast cancer, it was confirmed to be overexpressed in patients with estrogen receptor (ER) positive breast cancer [43]. It is reported that high expression of RAB31 mRNA has a significant correlation with the poor prognosis of lymph node-negative breast cancer patients [44]. Further *in vivo* and *in vitro* experiments confirmed that the overexpression of RAB31 promoted cell proliferation of breast cancer cells [45]. Immunohistochemical staining revealed that the expression of RAB31 in liver cancer tissue was significantly higher than that in adjacent liver tissue. Overexpression of RAB31 in liver cancer tissue after hepatectomy is considered to be related to a poor prognosis [46]. In addition, it was found that RAB31 is associated with the survival of glioblastoma [47]. A recent study also confirmed that overexpression of RAB31 in gastric cancer tissues was significantly related to specific clinicopathological features and shorter survival time, strongly suggesting that RAB31 can be a new oncogene for gastric cancer [48].

3.4.2. ACSL5. This gene received the level value of 0.8184. It encodes a specific transcription factor of the long-chain acyl-CoA synthetase (ACSL) family. In fatty acid metabolism, the first and essential step is the activation of fatty acids. ACSLs, responsible for activation of the most abundant long-chain fatty acids (12-20 carbons) in the diet into acyl-CoA thioesters, are generally deregulated in cancer. Such deregulation is also related to poor survival in patients with cancer [49]. The role of ACSL5 in cancer is quite complex. ACSL5 was reported to be downregulated in colorectal carcinomas [50, 51], breast cancer [52, 53], bladder cancer [53], and pancreas cancer [54]. Furthermore, ACSL5 lower regulation predicted a worse prognosis in breast cancer [52]. However, opposing results were also reported in studies on glioma [55] and gastric cancer [56], where ACSL5 was upregulated.

In addition, fibroblast growth factor receptor 2 (FGFR2)-ACSL5 chimera RNA caused clinical gastric cancer cells to be resistant to the treatment with FGFR inhibitors [57]. Evidences showed that high levels of ACSL5, as a potential downstream target of the transcription factor ONECUT2 (OC2), together with OC2, may cooperatively promote intestinal metaplasia and gastric cancer progression [56]. These contradicting results indicated that the roles of ACSL5 were different among the different cancer types. Lastly, exon 20 skipped variant of ACSL5 protein (splice, Spl) was identified. Results showed that the growth inhibitory effect produced by the Spl protein was opposed to the growth-promoting activity of the nonsplice (NSpl) isoform [58]. Therefore, due to both isoforms, ACSL5 may act either as a tumor suppressor gene or an oncogene.

3.4.3. WNT7B. This gene was assigned a level value of 0.8053. WNT7B is an extracellular matrix protein of Wnt family protein [59]. The Wnt (Wingless-INT) was derived from the wingless gene related to visual mutations in *Drosophila* and the *Int1* gene related to mouse breast cancer. Wnt signaling is a well-conserved pathway via canonical (β -catenin) and noncanonical (planar cell polarity and calcium) signaling [60]. WNT7B, as an activator of canonical Wnt/ β -catenin signaling [61], plays a critical role in normal development and tumorigenesis [62]. Because Wnt protein was first isolated from mouse breast cancer, the role of WNT7B on breast cancer has also been increasingly reported. Huguet et al. [63] explored differential expression of human Wnt Genes 2, 3, 4, and 7B in human breast cell lines and normal and disease states of human breast tissue. They further found that in 10% of tumors WNT7B expression was 30-fold higher than in normal or benign breast tissues. In addition to confirming results consistent with them, Ojalvo et al. [64] and Chen et al. [65] further validated that WNT7B expression makes connections with markers of poor prognosis. Yeo et al. [59] built a *Csf1r-icre* mouse model using a WNT7B deletion, which also illustrated a critical role of myeloid WNT7B in breast cancer progression, including the levels of angiogenesis, invasion, and metastasis. In addition to breast cancer, abnormal expression of the WNT7b leads to the pathogenesis of many other cancers. Arensman et al. [66] confirmed that WNT7B expression was increased with high activity levels of autocrine Wnt/ β -catenin signaling in pancreatic

TABLE 1: Top fifteen inferred oncogenes.

| Rank | Ensembl ID | Gene symbol | Description | Level value |
|------|-----------------|-----------------|---|-------------|
| 1 | ENSP00000304565 | RAB31 | RAB31, member RAS oncogene family | 0.9105 |
| 2 | ENSP00000421799 | ENSG00000257184 | — | 0.8868 |
| 3 | ENSP00000469872 | RAB4B-EGLN2 | RAB4B-EGLN2 Readthrough (NMD candidate) | 0.8500 |
| 4 | ENSP00000283921 | HOXA10 | Homeobox A10 | 0.8289 |
| 5 | ENSP00000385586 | HOXD12 | Homeobox D12 | 0.8184 |
| 6 | ENSP00000348429 | ACSL5 | Acyl-CoA synthetase long chain family member 5 | 0.8184 |
| 7 | ENSP00000256953 | RERG | RAS-like estrogen regulated growth inhibitor | 0.8105 |
| 8 | ENSP00000341032 | WNT7B | Wnt family member 7B | 0.8053 |
| 9 | ENSP00000321805 | RIT2 | Ras like without CAAX 2 | 0.7763 |
| 10 | ENSP00000285735 | RHOC | Ras homolog family member C | 0.7763 |
| 11 | ENSP00000282397 | FLT1 | Fms related receptor tyrosine kinase 1 | 0.7763 |
| 12 | ENSP00000264711 | DNAJC27 | DnaJ heat shock protein family (Hsp40) member C27 | 0.7737 |
| 13 | ENSP00000339787 | ACSL4 | Acyl-CoA synthetase long chain family member 4 | 0.7737 |
| 14 | ENSP00000357306 | RIT1 | Ras like without CAAX 1 | 0.7684 |
| 15 | ENSP00000301068 | RHEBL1 | RHEB like 1 | 0.7684 |

adenocarcinoma. Zheng et al. [67] found that expression of WNT7B is essential for the growth of prostate cancer cells and this effect is enhanced under androgen-deprived conditions. Their further analyses revealed that WNT7B promotes androgen-independent growth of CRPC cells likely via the activation of protein kinase C isozymes. Their results further showed that prostate cancer-produced WNT7B made osteoblast differentiation in vitro and in vivo. As for osteosarcoma (OS) [68], WNT7B expression is dramatically upregulated in OS tissue samples and cells, especially in metastatic OS cell lines. Liu et al. [68] also found that WNT7B silence within OS cells remarkably inhibited the viability and invasion and enhanced the apoptosis of OS cells, suggesting that knocking down WNT7B could inhibit the OS cell growth. Therefore, we presume that WNT7B may function as an oncogene in carcinoma tissue types.

3.4.4. FLT1. This gene was assigned a level value of 0.7763. Fms-related tyrosine kinase 1 (FLT1, also known as VEGFR-1) is a gene that encodes for a member of the VEGFR family, which presents a critical point in angiogenesis and subsequent cancer progression [69]. The expression of FLT1 is not limited to vascular endothelial cells. It is also found in cells of the hematopoietic lineage (i.e., monocytes and macrophages), dendritic cells, osteoclasts, pericytes, liver cells, placental trophoblasts [70], and smooth muscle cells [71], where it has a regulatory function. Therefore, with respect to carcinogenesis, the role of FLT1 may be more complex [72]. Recent reports also indicate that FLT1 is directly expressed on tumor cells from breast, colon, and skin origin. It is an important oncogenic driver in above cells, boosting survival, cell proliferation, and invasion in an independent manner [73–76]. In head and neck squamous cell carcinoma (HNSC), FLT1 was selectively overexpressed in tumor tissue. FLT1 was further identified as an important oncogenic driver of HNSC survival and resistance to radiotherapy

with a shRNAmir-based dropout screening setup [72]. Qian et al. [77] found that FLT1 labels a subset of macrophages in human breast cancers which are significantly enriched in metastatic sites. Furthermore, using several genetic models, they elaborated that macrophage FLT1 is important for tumor cell seeding and persistent growth during the distal metastasis. Jiang et al. [78] identified that FLT1 promotes invasion and migration of glioblastoma cells through the modulation of sonic hedgehog (SHH) signaling pathway. A further study has indicated that FLT1 knockdown can prevent the spread of glioblastoma cells in vivo. FLT1 may be a novel oncogene, and therefore, inhibition of FLT1 may serve as a potential target for the development of therapies against metastatic events.

4. Conclusions

This study proposed a computational method for the identification of latent oncogenes. From seven protein networks, informative features of proteins were extracted via a powerful network embedding algorithm. Obtained features were learned by random forest, thereby setting up the prediction model. The principle of our method was quite different from previous methods and provided some novel latent oncogenes. Some inferred genes can be confirmed to be novel oncogenes, suggesting that the newly identified oncogenes can be essential supplements for previous studies. It is hoped that the new findings reported in this study can promote the research process of cancers.

Data Availability

The validated oncogenes were collected from HUGO Gene Nomenclature Committee and Gene Set Enrichment Analysis Molecular Signatures Database.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

Ran Zhao and Bin Hu contributed equally to this work.

Acknowledgments

This study was supported by the Natural Science Foundation of Shanghai (17ZR1412500), the National Natural Science Foundation of China (61701298), the Key-Area Research and Development Program of Guangdong Province (2018B020203003), Guangzhou science and technology planning project (201707020007), Science and Technology Planning Project of Guangdong Province (2017A010405039).

Supplementary Materials

Supplementary material S1 Level values of candidate oncogenes. (*Supplementary Materials*)

References

- [1] S. McGuire, "World Cancer Report 2014. Geneva, Switzerland: World Health Organization, International Agency for Research on Cancer, WHO press, 2015," *Advances in Nutrition*, vol. 7, no. 2, pp. 418-419, 2016.
- [2] F. P. Perera, "Environment and cancer: who are susceptible?," *Science*, vol. 278, no. 5340, pp. 1068-1073, 1997.
- [3] L. Chen, B. Wang, S. P. Wang et al., "OPMSP: a computational method integrating protein interaction and sequence information for the identification of novel putative oncogenes," *Protein and Peptide Letters*, vol. 23, no. 12, pp. 1081-1094, 2016.
- [4] Z. Xing, C. Chu, L. Chen, and X. Kong, "The use of Gene Ontology terms and KEGG pathways for analysis and prediction of oncogenes," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1860, no. 11, pp. 2725-2734, 2016.
- [5] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
- [6] K. A. Gray, B. Yates, R. L. Seal, M. W. Wright, and E. A. Bruford, "Genenames.org: The HGNC resources in 2015," *Nucleic Acids Research*, vol. 43, no. D1, pp. D1079-D1085, 2015.
- [7] 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.
- [8] V. K. Mootha, C. M. Lindgren, K. F. Eriksson et al., "PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes," *Nature Genetics*, vol. 34, no. 3, pp. 267-273, 2003.
- [9] A. L. Barabasi, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease," *Nature Reviews. Genetics*, vol. 12, no. 1, pp. 56-68, 2011.
- [10] L. Chen, Z. Xing, T. Huang, Y. Shu, G. H. Huang, and H. P. Li, "Application of the shortest path algorithm for the discovery of breast cancer related genes," *Current Bioinformatics*, vol. 11, no. 1, pp. 51-58, 2016.
- [11] J. Zhang, J. Yang, T. Huang, Y. Shu, and L. Chen, "Identification of novel proliferative diabetic retinopathy related genes on protein-protein interaction network," *Neurocomputing*, vol. 217, pp. 63-72, 2016.
- [12] J. Zhang, Y. Suo, M. Liu, and X. Xu, "Identification of genes related to proliferative diabetic retinopathy through RWR algorithm based on protein-protein interaction network," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1864, no. 6, pp. 2369-2375, 2018.
- [13] L. Li, Y. S. Wang, L. An, X. Y. Kong, and T. Huang, "A network-based method using a random walk with restart algorithm and screening tests to identify novel genes associated with Menière's disease," *PLoS One*, vol. 12, no. 8, article e0182592, 2017.
- [14] F. Yuan and W. C. Lu, "Prediction of potential drivers connecting different dysfunctional levels in lung adenocarcinoma via a protein-protein interaction network," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1864, no. 6, Part B, pp. 2284-2293, 2018.
- [15] C. von 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.
- [16] C. Von Mering, L. J. Jensen, B. Snel et al., "STRING: known and predicted protein-protein associations, integrated and transferred across organisms," *Nucleic Acids Research*, vol. 33, no. Database issue, pp. D433-D437, 2005.
- [17] I. Xenarios, D. W. Rice, L. Salwinski, M. K. Baron, E. M. Marcotte, and D. Eisenberg, "DIP: the database of interacting proteins," *Nucleic Acids Research*, vol. 28, no. 1, pp. 289-291, 2000.
- [18] C. Stark, B. J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, and M. Tyers, "BioGRID: a general repository for interaction datasets," *Nucleic Acids Research*, vol. 34, no. 90001, pp. D535-D539, 2006.
- [19] H. Cho, B. Berger, and J. Peng, "Compact integration of multi-network topology for functional analysis of genes," *Cell Systems*, vol. 3, no. 6, pp. 540-548.e5, 2016.
- [20] A. Grover and J. Leskovec, "node2vec: scalable feature learning for networks," in *KDD'16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 855-864, ACM: San Francisco, CA, USA, 2016.
- [21] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: online learning of social representations," in *KDD '14: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA, 2014.
- [22] H. Tong, C. Faloutsos, and J. Pan, "Fast random walk with restart and its applications," in *Sixth International Conference on Data Mining (ICDM'06)*, Hong Kong, China, 2006.
- [23] S. Köhler, S. Bauer, D. Horn, and P. N. Robinson, "Walking the interactome for prioritization of candidate disease genes," *The American Journal of Human Genetics*, vol. 82, no. 4, pp. 949-958, 2008.
- [24] H. Liang, L. Chen, X. Zhao, and X. Zhang, "Prediction of drug side effects with a refined negative sample selection strategy," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 1573543, 16 pages, 2020.
- [25] L. Chen, T. Liu, and X. Zhao, "Inferring anatomical therapeutic chemical (ATC) class of drugs using shortest path and random walk with restart algorithms," *Biochimica et Biophysica*

- Acta (BBA) - Molecular Basis of Disease*, vol. 1864, no. 6, pp. 2228–2240, 2018.
- [26] L. Chen, Y. H. Zhang, Z. Zhang, T. Huang, and Y. D. Cai, “Inferring novel tumor suppressor genes with a protein-protein interaction network and network diffusion algorithms,” *Molecular Therapy - Methods & Clinical Development*, vol. 10, pp. 57–67, 2018.
- [27] X. Zhao, L. Chen, and J. Lu, “A similarity-based method for prediction of drug side effects with heterogeneous information,” *Mathematical Biosciences*, vol. 306, pp. 136–144, 2018.
- [28] X. Zhao, L. Chen, Z. H. Guo, and T. Liu, “Predicting drug side effects with compact integration of heterogeneous networks,” *Current Bioinformatics*, vol. 14, no. 8, pp. 709–720, 2019.
- [29] J. R. Li, L. Lu, Y.-H. Zhang et al., “Identification of synthetic lethality based on a functional network by using machine learning algorithms,” *Journal of Cellular Biochemistry*, vol. 120, no. 1, pp. 405–416, 2019.
- [30] S. Wang, D. Wang, J. R. Li, T. Huang, and Y. D. Cai, “Identification and analysis of the cleavage site in a signal peptide using SMOTE, dagging, and feature selection methods,” *Molecular Omics*, vol. 14, no. 1, pp. 64–73, 2018.
- [31] E. S. Sankari and D. Manimegalai, “Predicting membrane protein types by incorporating a novel feature set into Chou’s general PseAAC,” *Journal of Theoretical Biology*, vol. 455, pp. 319–328, 2018.
- [32] L. Wei, P. Xing, J. Tang, and Q. Zou, “PhosPred-RF: a novel sequence-based predictor for phosphorylation sites using sequential information only,” *IEEE Transactions on Nanobioscience*, vol. 16, no. 4, pp. 240–247, 2017.
- [33] L. Wei, P. W. Xing, R. Su, G. Shi, Z. S. Ma, and Q. Zou, “CPPred-RF: a sequence-based predictor for identifying cell-penetrating peptides and their uptake efficiency,” *Journal of Proteome Research*, vol. 16, no. 5, pp. 2044–2053, 2017.
- [34] Y. B. Marques, A. de Paiva Oliveira, A. T. Ribeiro Vasconcelos, and F. R. Cerqueira, “Miracle: machine learning with SMOTE and random forest for improving selectivity in pre-miRNA ab initio prediction,” *BMC Bioinformatics*, vol. 17, Supplement 18, p. 474, 2016.
- [35] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, “Do we need hundreds of classifiers to solve real world classification problems?,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3133–3181, 2014.
- [36] I. H. Witten and E. Frank, *Data mining: practical machine learning tools and techniques*, Elsevier, San Francisco, 2nd edition, 2005.
- [37] K. C. Chou and C. T. Zhang, “Prediction of protein structural classes,” *Critical Reviews in Biochemistry and Molecular Biology*, vol. 30, no. 4, pp. 275–349, 1995.
- [38] J.-P. Zhou, L. Chen, and Z.-H. Guo, “iATC-NRAKEL: an efficient multi-label classifier for recognizing anatomical therapeutic chemical classes of drugs,” *Bioinformatics*, vol. 36, no. 5, pp. 1391–1396, 2020.
- [39] L. Chen, W. M. Zeng, Y. D. Cai, K. Y. Feng, and K. C. Chou, “Predicting anatomical therapeutic chemical (ATC) classification of drugs by integrating chemical-chemical interactions and similarities,” *PLoS One*, vol. 7, no. 4, article e35254, 2012.
- [40] J. Egan, *Signal Detection Theory and ROC Analysis*, Academic Press, New York, 1975.
- [41] J. Colicelli, “Human RAS superfamily proteins and related GTPases,” *Science’s STKE*, vol. 2004, no. 250, article Re13, p. re13, 2004.
- [42] C. E. Chua and B. L. Tang, “The role of the small GTPase Rab31 in cancer,” *Journal of Cellular and Molecular Medicine*, vol. 19, no. 1, pp. 1–10, 2015.
- [43] M. C. Abba, Y. Hu, H. Sun et al., “Gene expression signature of estrogen receptor α status in breast cancer,” *BMC Genomics*, vol. 6, no. 1, p. 37, 2005.
- [44] M. Kotsch, A. M. Sieuwerts, M. Grosser et al., “Urokinase receptor splice variant uPAR-del4/5-associated gene expression in breast cancer: identification of rab31 as an independent prognostic factor,” *Breast Cancer Research and Treatment*, vol. 111, no. 2, pp. 229–240, 2008.
- [45] B. Grismayer, S. Sölch, B. Seubert et al., “Rab31 expression levels modulate tumor-relevant characteristics of breast cancer cells,” *Molecular Cancer*, vol. 11, no. 1, p. 62, 2012.
- [46] Y. Sui, X. Zheng, and D. Zhao, “Rab31 promoted hepatocellular carcinoma (HCC) progression via inhibition of cell apoptosis induced by PI3K/AKT/Bcl-2/BAX pathway,” *Tumour Biology*, vol. 36, no. 11, pp. 8661–8670, 2015.
- [47] N. V. Serão, K. R. Delfino, B. R. Southey, J. E. Beever, and S. L. Rodriguez-Zas, “Cell cycle and aging, morphogenesis, and response to stimuli genes are individualized biomarkers of glioblastoma progression and survival,” *BMC Medical Genomics*, vol. 4, no. 1, p. 49, 2011.
- [48] C. T. Tang, Q. Liang, L. Yang et al., “RAB31 targeted by MiR-30c-2-3p regulates the GII1 signaling pathway, Affecting Gastric Cancer Cell Proliferation and Apoptosis,” *Frontiers in Oncology*, vol. 8, p. 554, 2018.
- [49] Y. Tang, J. Zhou, S. Hooi, Y.-M. Jiang, and G.-D. Lu, “Fatty acid activation in carcinogenesis and cancer development: essential roles of long-chain acyl-CoA synthetases (Review),” *Oncology Letters*, vol. 16, no. 2, pp. 1390–1396, 2018.
- [50] N. Gassler, A. Schneider, J. Kopitz et al., “Impaired expression of acyl-CoA-synthetase 5 in epithelial tumors of the small intestine,” *Human Pathology*, vol. 34, no. 10, pp. 1048–1052, 2003.
- [51] F. Hartmann, D. Sparla, E. Tute et al., “Low acyl-CoA synthetase 5 expression in colorectal carcinomas is prognostic for early tumour recurrence,” *Pathology, Research and Practice*, vol. 213, no. 3, pp. 261–266, 2017.
- [52] M. C. Yen, J. Y. Kan, C. J. Hsieh, P. L. Kuo, M. F. Hou, and Y. L. Hsu, “Association of long-chain acyl-coenzyme A synthetase 5 expression in human breast cancer by estrogen receptor status and its clinical significance,” *Oncology Reports*, vol. 37, no. 6, pp. 3253–3260, 2017.
- [53] N. T. Gaisa, A. Reinartz, U. Schneider et al., “Levels of acyl-coenzyme A synthetase 5 in urothelial cells and corresponding neoplasias reflect cellular differentiation,” *Histology and Histopathology*, vol. 28, no. 3, pp. 353–364, 2013.
- [54] H. Li, X. Wang, Y. Fang et al., “Integrated expression profiles analysis reveals novel predictive biomarker in pancreatic ductal adenocarcinoma,” *Oncotarget*, vol. 8, no. 32, pp. 52571–52583, 2017.
- [55] Y. Yamashita, T. Kumabe, Y. Y. Cho et al., “Fatty acid induced glioma cell growth is mediated by the acyl-CoA synthetase 5 gene located on chromosome 10q25.1-q25.2, a region frequently deleted in malignant gliomas,” *Oncogene*, vol. 19, no. 51, pp. 5919–5925, 2000.
- [56] E. H. Seo, H. J. Kim, J. H. Kim et al., “ONECUT2 upregulation is associated with CpG hypomethylation at promoter-proximal DNA in gastric cancer and triggers ACSL5,” *International Journal of Cancer*, vol. 146, no. 12, pp. 3354–3368, 2020.

- [57] S. Y. Kim, T. Ahn, H. Bang et al., “Acquired resistance to LY2874455 in *FGFR2*-amplified gastric cancer through an emergence of novel *FGFR2-ACSL5* fusion,” *Oncotarget*, vol. 8, no. 9, pp. 15014–15022, 2017.
- [58] I. Pérez-Núñez, M. Karaky, M. Fedetz et al., “Splice-site variant in *ACSL5*: a marker promoting opposing effect on cell viability and protein expression,” *European Journal of Human Genetics*, vol. 27, no. 12, pp. 1836–1844, 2019.
- [59] E. J. Yeo, L. Cassetta, B. Z. Qian et al., “Myeloid WNT7b mediates the angiogenic switch and metastasis in breast cancer,” *Cancer Research*, vol. 74, no. 11, pp. 2962–2973, 2014.
- [60] R. Nusse and H. Clevers, “Wnt/ β -catenin signaling, disease, and emerging therapeutic modalities,” *Cell*, vol. 169, no. 6, pp. 985–999, 2017.
- [61] Z. Wang, W. Shu, M. M. Lu, and E. E. Morrisey, “Wnt7b activates canonical signaling in epithelial and vascular smooth muscle cells through interactions with Fzd1, Fzd10, and LRP5,” *Molecular and Cellular Biology*, vol. 25, no. 12, pp. 5022–5030, 2005.
- [62] M. Noda, M. Vallon, and C. J. Kuo, “The Wnt7’s tale: a story of an orphan who finds her tie to a famous family,” *Cancer Science*, vol. 107, no. 5, pp. 576–582, 2016.
- [63] E. L. Huguet, J. McMahon, A. McMahon, R. Bicknell, and A. L. Harris, “Differential expression of human Wnt genes 2, 3, 4, and 7B in human breast cell lines and normal and disease states of human breast tissue,” *Cancer Research*, vol. 54, no. 10, pp. 2615–2621, 1994.
- [64] L. S. Ojalvo, C. A. Whittaker, J. S. Condeelis, and J. W. Pollard, “Gene expression analysis of macrophages that facilitate tumor invasion supports a role for Wnt-signaling in mediating their activity in primary mammary tumors,” *Journal of Immunology*, vol. 184, no. 2, pp. 702–712, 2010.
- [65] J. Chen, T. Y. Liu, H. T. Peng et al., “Up-regulation of Wnt7b rather than Wnt1, Wnt7a, and Wnt9a indicates poor prognosis in breast cancer,” *International Journal of Clinical and Experimental Pathology*, vol. 11, no. 9, pp. 4552–4561, 2018.
- [66] M. D. Arensman, A. N. Kovichich, R. M. Kulikauskas et al., “WNT7B mediates autocrine WNT/ β -catenin signaling and anchorage-independent growth in pancreatic adenocarcinoma,” *Oncogene*, vol. 33, no. 7, pp. 899–908, 2014.
- [67] D. Zheng, K. F. Decker, T. Zhou et al., “Role of WNT7B-induced noncanonical pathway in advanced prostate cancer,” *Molecular Cancer Research*, vol. 11, no. 5, pp. 482–493, 2013.
- [68] Q. Liu, Z. Wang, X. Zhou et al., “miR-342-5p inhibits osteosarcoma cell growth, migration, invasion, and sensitivity to doxorubicin through targeting Wnt7b,” *Cell Cycle*, vol. 18, no. 23, pp. 3325–3336, 2019.
- [69] N. Ferrara, H. P. Gerber, and J. LeCouter, “The biology of VEGF and its receptors,” *Nature Medicine*, vol. 9, no. 6, pp. 669–676, 2003.
- [70] M. Shibuya and L. Claessonwelsch, “Signal transduction by VEGF receptors in regulation of angiogenesis and lymphangiogenesis,” *Experimental Cell Research*, vol. 312, no. 5, pp. 549–560, 2006.
- [71] Y. Wu, A. T. Hooper, Z. Zhong et al., “The vascular endothelial growth factor receptor (VEGFR-1) supports growth and survival of human breast carcinoma,” *International Journal of Cancer*, vol. 119, no. 7, pp. 1519–1529, 2006.
- [72] E. J. Van Limbergen, P. Zabrocki, M. Porcu, E. Hauben, J. Cools, and S. Nuyts, “FLT1 kinase is a mediator of radioresistance and survival in head and neck squamous cell carcinoma,” *Acta Oncologica*, vol. 53, no. 5, pp. 637–645, 2014.
- [73] F. Fan, J. S. Wey, M. F. McCarty et al., “Expression and function of vascular endothelial growth factor receptor-1 on human colorectal cancer cells,” *Oncogene*, vol. 24, no. 16, pp. 2647–2653, 2005.
- [74] T. H. Lee, S. Seng, M. Sekine et al., “Vascular endothelial growth factor mediates intracrine survival in human breast carcinoma cells through internally expressed VEGFR1/FLT1,” *PLoS Medicine*, vol. 4, no. 6, article e186, 2007.
- [75] B. M. Lichtenberger, P. K. Tan, H. Niederleithner, N. Ferrara, P. Petzelbauer, and M. Sibilica, “Autocrine VEGF signaling synergizes with EGFR in tumor cells to promote epithelial cancer development,” *Cell*, vol. 140, no. 2, pp. 268–279, 2010.
- [76] S. Naik, R. S. Dothager, J. Marasa, C. L. Lewis, and D. Piwnicka-Worms, “Vascular endothelial growth factor receptor-1 is synthetic lethal to aberrant β -catenin activation in colon cancer,” *Clinical Cancer Research*, vol. 15, no. 24, pp. 7529–7537, 2009.
- [77] B. Z. Qian, H. Zhang, J. Li et al., “FLT1 signaling in metastasis-associated macrophages activates an inflammatory signature that promotes breast cancer metastasis,” *The Journal of Experimental Medicine*, vol. 212, no. 9, pp. 1433–1448, 2015.
- [78] K. Jiang, Y. P. Wang, X. D. Wang et al., “Fms related tyrosine kinase 1 (Flt1) functions as an oncogene and regulates glioblastoma cell metastasis by regulating sonic hedgehog signaling,” *American Journal of Cancer Research*, vol. 7, no. 5, pp. 1164–1176, 2017.

Research Article

Inhibition Mir-92a Alleviates Oxidative Stress and Apoptosis of Alveolar Epithelial Cells Induced by Lipopolysaccharide Exposure through TLR2/AP-1 Pathway

Jian Cui,¹ Huanhuan Ding,² Yongyuan Yao,¹ and Wei Liu¹ 

¹Department of Intensive Care Unit (ICU), People's Hospital of Rizhao, Shandong Province, China

²Department of Cardiology, People's Hospital of Rizhao, Shandong Province, China

Correspondence should be addressed to Wei Liu; dishangliaoy@163.com

Received 25 July 2020; Revised 27 August 2020; Accepted 3 September 2020; Published 16 September 2020

Academic Editor: Tao Huang

Copyright © 2020 Jian Cui 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.

Objective. To probe into the role of miR-92a in alleviating oxidative stress and apoptosis of alveolar epithelial cell (AEC) injury induced by lipopolysaccharide (LPS) exposure through the Toll-like receptor (TLR) 2/activator protein-1 (AP-1) pathway. **Methods.** Acute lung injury (ALI) rat model and ALI alveolar epithelial cell model were constructed to inhibit the expression of miR-92a/TLR2/AP-1 in rat and alveolar epithelial cells (AECs), to detect the changes of oxidative stress, inflammatory response, and cell apoptosis in rat lung tissues and AECs, and to measure the changes of wet-dry weight (W/D) ratio in rat lung tissues. **Results.** Both inhibition of miR-92a expression and knockout of TLR2 and AP-1 gene could reduce LPS-induced rat ALI, alleviate pulmonary edema, inhibit oxidative stress and inflammatory response, and reduce apoptosis of lung tissue cells. In addition, the TLR2 and AP-1 levels in the lung tissues of ALI rats were noticed to be suppressed when inhibiting the expression of miR-92a, and the AP-1 level was also decreased after the knockout of TLR2 gene. Further, we verified this relationship in AECs and found that inhibition of miR-92a/TLR2/AP-1 also alleviated LPS-induced AEC injury, reduced cell apoptosis, and inhibited oxidative stress and inflammatory response. What is more, like that in rat lung tissue, the phenomenon also existed in AECs, that is, when the expression of miR-92a was inhibited, the expression of TLR2 and AP-1 was inhibited, and silencing TLR2 can reduce the expression level of AP-1. **Conclusion.** MiR-92a/TLR2/AP-1 is highly expressed in ALI, and its inhibition can improve oxidative stress and inflammatory response and reduce apoptosis of AECs.

1. Introduction

Acute lung injury (ALI) is an acute respiratory distress syndrome (ARDS) with clinical characteristics of acute hypoxic respiratory failure and bilateral pulmonary infiltration caused by multiple factors in and out of the lung. Due to the lack of specific treatment, the mortality rate can reach up to 40% [1–3]. Alveolar epithelial cells (AECs), as the main sites for gas exchange, are also one of the main components of respiratory barrier [4]. During ALI, AECs are always found damaged more, the generation of surfactant reduced, lung compliance decreased, and gas exchange blocked [5, 6].

Short noncoding RNA (miRNAs) are a class of small RNA with a length of about 18–25 nucleotides. By binding to messenger RNA (3'-UTR) and inducing RNA silencing

or degradation through miRNA-induced silencing complexes, they negatively regulate protein-coding genes' expression and participate in the regulation of various cellular processes, including inflammatory responses [7, 8]. In recent years, multiple literature has reported that miRNAs have a part to play in ALI as markers of acute lung injury and diffuse alveolar injury [9]. A case is that in the study of Song et al. [10], miR-34a could target FoxO3 to inhibit autophagy of alveolar type II epithelial cells in ALI and reduce the damage of lipopolysaccharides (LPS)-induced ALI. MiR-92a has also been found to act on development of inflammatory responses in ALI rats, and its inhibition can reduce the secretion of pro-inflammatory factors and improve inflammatory responses [11]. MiRNAs are important regulators of Toll-like receptor (TLR) signaling, among which TLR2 has been reported as

one of the targets of miR-92a, which can alleviate liver fibrosis caused by *Schistosoma japonicum* [12]. In addition, Lai et al. [13] demonstrated that TLR-mediated decrease in miR-92a expression can promote the production of inflammatory cytokines in TLR-induced macrophages. What is more, Fei et al. [14] revealed that in ALI, glycyrrhizic acid can block the TLR-2 signal cascade to inhibit the inflammatory response induced by ischemia-reperfusion lung injury. Activator Protein-1 (AP-1) is an upstream regulator of interleukin-4 (IL-4), which also participates in the ALI process. Khan et al. [15] showed that Anomalin could inhibit AP-1 to relieve mechanical pain and inhibit leukocyte infiltration in ALI rats. And AP-1 is regulated by TLR2 [16, 17].

Based on preceding research, we hypothesized whether there is a signaling axis such as miR-92a/TLR2/AP-1 that is effective in the occurrence and progression of ALI. Hence, in this study, we analyzed the effect of miR-92a/TLR2/AP-1 on ALI AECs.

2. Materials and Methods

2.1. Rat Source. A total of 100 healthy Sprague-Dawley rats, aged 10 weeks and weighed 250-300 g, were obtained from the Experimental Animal Center of Harbin Medical University. The rats were kept at room temperature of 20-25°C, relative humidity of 40%-70%, normal 12 h circadian rhythm, and they were free to eat and drink. All animal experiments in present study were approved by the Animal Care and Use Committee of our hospital and followed the guidelines of the Council for International Organizations of Medical Sciences (CIOMS).

2.2. ALI Rat Model Construction and Observation Index. All the rats were grouped into a control group (CG), a model group (MG), a miR-92a inhibitor group, a TLR2(-) group, and an AP-1(-) group at random, with 20 in each group. TLR2 and AP-1 gene knockout were performed on the rats of TLR2(-) group and AP-1 group by Sigma company, while the CG was left untreated. The MG, miR-92a inhibitor group, TLR2(-) group, and AP-1(-) group were intratracheally infused with 5 mg/kg LPS. Meanwhile, miR-92a inhibitor group was injected via tail vein with a concentration of 100 mg/kg pretreated with liposome 2000 miR-92a inhibitor carrier, and miR-92a inhibitor vector was designed and synthesized by Sigma Company. Twenty-four hours later, pentobarbital sodium (45 mg/kg, Sigma) was intraperitoneally injected into the rats to anaesthetize and kill according to the guidelines for the Care and Use of Laboratory Animals, and the lung tissues were isolated. Then, the W/D ratio, apoptosis rate, oxidative stress, and inflammatory factors in the lung tissues of rats in the five groups were detected.

2.3. Detection Methods. **W/D Ratio.** After removing both lungs from the bronchi of free rats, the left lung was taken as the object of examination. The fresh lung tissue was weighed and baked to constant weight in an oven at 80°C. The W/D ratio = dry weight/wet weight.

Apoptosis Rate. Half of the right lung tissue was cut into pieces and centrifuged at 300 × g for 1 min at 4°C in a filter

centrifuge tube with a diameter of 35 μm. After removing the supernatant, phosphate buffer solution was added to resuspend the precipitated cells, and the cell concentration was adjusted to 1 × 10⁶ ml. Then, AnnexinV-FITC and PI were added successively to incubate in the dark at room temperature for 5 min, followed by the detection conducted by FACSCalibur flow cytometry (BD Biosciences, CA, USA). The experiment was repeated 3 times to take the average. The Annexin V-FITC/PI apoptosis detection kit was purchased from Invitrogen, USA, under the article number V35113.

Detection of Oxidative Stress and Inflammatory Response Levels. The remaining half of the right lung tissue was placed in a glass homogenate tube, and the tissue was ground into a pulp. The levels of tumor necrosis factor α (TNF-α), interleukin-6 (IL-6), interleukin-10 (IL-10), malondialdehyde (MDA), and superoxide dismutase (SOD) in lung tissues were detected by ELISA with reference to the kit instructions. The MDA, TNF-α, IL-6, and IL-10 ELISA kits were purchased from Guduo Biotechnology Co., Ltd., Shanghai, China, with the article numbers of GD-BN1921, GD-DS1716, GD-DS1726, and GD-DS1731, respectively, and SOD ELISA kit was obtained from Jingkang Bioengineering Co., Ltd., Shanghai, China, with the article number of JLC2390.

2.4. Cell Experiment. AECs A549, numbered ATCC®CCL-185 cells, were acquired from American Type Culture Collection (ATCC). The cells were culture in the medium consisting of FMUI 12K medium (ATCC, 30-2004) +10% fetal bovine serum (ATCC, 30-2020) under 95% air and 5% carbon dioxide (CO₂), at a temperature of 37°C.

2.5. Cell Grouping and Treatment. Cells were divided into 5 groups, namely the blank group (BG), LPS group, inhibitor group, si-TLR2 group, and si-AP-1 group. Concerning the treatment of cells in each group, those in the BG were left untreated, while those in the inhibitor group were transfected with miR-92a inhibitor vector, those in the si-TLR2 group were transfected with si-TLR2 vector, and those in the si-AP-1 group were transfected with si-AP-1 vector. The si-TLR2 and si-AP-1 vectors were designed and synthesized by Sigma, and the corresponding vectors were transfected by means of Lipofectamine 2000 kit (Thermo Fisher China). In addition to the BG, the other 4 groups of cells were supplemented with 1 g/ml of LPS. After 24 hours of culture, the apoptosis rate and changes in oxidative stress indicators in the cells were detected, with the same detection method as above.

2.6. qRT-PCR. Trizol kit (Invitrogen) was applied to separate the Total RNA from tissues/cells. The EasyScript One-Step RT-PCR SuperMix kit was acquired from Transgen Biotechnology, Co., Ltd, Beijing, China, and the specific detection steps were referred to the kit instructions. RNA Template: 1 μg, Forward GSP (10 μM): 0.4 μL, Reverse GSP (10 μM): 0.4 μL, 2*One-Step Reaction Mix: 10 μL, EasyScript One-Step Enzyme Mix: 0.4 μL, and RNase-free Water was added to complete the reaction volume of 20 μL. The reaction

conditions were 40°C for 30 min, 94°C for 5 min, 94°C for 30s, 60°C for 30s, 72°C for 2 kb/min, and 72°C for 10 min, totaling 40 cycles. Three replicate wells were set up in the experiment, and U6 was used as the internal reference for the reaction. The results were analyzed by $2^{-\Delta\Delta Ct}$ method.

2.7. Western Blot. The protein in the cells/tissues was extracted by repeated freeze-thaw method, and the protein concentration was determined with the aid of BCA. Next, the protein was made to 4 μ g/ μ L, electrophoretically separated by 12% SDS-PAGE before it was processed under the initial voltage of 90 V, and then an increased voltage of 120 V to shift the sample to the suitable site on the separation gel. Upon the completion of electrophoresis, the membrane was transferred, with 100 V constant pressure for 100 min and sealed for 60 min at 37°C. Then, the membrane was placed in 5% skim milk for sealing before immune reaction. The membrane was subsequently cultivated overnight at 4°C after adding with primary antibody (1:1000), followed by warm washing with PBS three times the next day, 5 min each, and then incubated with secondary antibody (1:1000) at room temperature for 1 hour. After that, ECL luminescent reagent was developed and fixed. Quantity One software was employed for statistical analysis of the bands after film scanning, and the protein's relative expression level was equal to the gray value of the bands/the grays value of the internal parameters. BCA protein kit, trypsin, and ECL luminescence kit were all acquired from Thermo Scientific™, with the corresponding article number of 23250, 35055, and 90058. Rabbit anti-TLR2 monoclonal antibody, rabbit anti-AP-1 polyclonal antibody, rabbit anti-bax, goat anti-rabbit IgG secondary antibody, and bcl-2 monoclonal antibody were obtained from Abcam, USA, under the article numbers of ab209217, ab21981, ab32503, ab185002, and ab6721, respectively.

2.8. Statistical Analysis. SPSS19.0, which was purchase from Chicago, IL, USA, was employed for statistical analysis of the collected data. The measurement data was described in the form of mean \pm SD. The comparison between two groups was conducted using Student's *t*-test, while that among multiple groups was carried out by one-way ANOVA. LSD test was adopted for post hoc test. Two-tailed $P < 0.05$ was considered statistically significant. Graph-Pad Prism 8.0 (La Jolla, CA) was responsible for picture drawing.

3. Results

3.1. Effects of miR-92a/TLR2/AP-1 on ALI Rats

3.1.1. Analysis of miR-92a/TLR2/AP-1 Level in 5 Groups of Rats. After LPS stimulation, the level of miR-92a/TLR2/AP-1 in the lung tissues in the MG was noticeably higher than that in the CG ($P < 0.05$). When compared with the MG, the miR-92a in the lung tissue in the miR-92a inhibitor group was decreased ($P < 0.05$), the TLR2 in the TLR2(-) group was declined ($P < 0.05$), and the AP-1 in the AP-1(-) group was reduced ($P < 0.05$). Moreover, it was found that after inhibiting the expression of miR-92a, the expression levels of TLR2 and AP-1 in the lung tissues of the miR-92a inhibitor group

were also lower than those of the MG ($P < 0.05$), and the AP-1 in the TLR2(-) group also went down while inhibiting the expression of TLR2 ($P < 0.05$) (Figure 1).

3.1.2. Changes of W/D Ratio in ALI Rats' Lung Tissues. Compared to the CG, the W/D ratio of lung tissues in the MG increased remarkably after LPS stimulation, but decreased in the miR-92a inhibitor group, TLR2(-) group, and AP-1(-) group after inhibiting the expression of miR-92a/TLR2/AP-1. The W/D ratio of lung tissues differed little among miR-92a inhibitor group, TLR2(-) group, and AP-1(-) group ($P > 0.05$) (Figure 2).

3.1.3. Changes of Apoptosis Level in ALI Rats' Lung Tissues. Compared with the CG, LPS stimulation markedly elevated the apoptosis rate and bax and bcl-2 levels in the MG ($P < 0.05$), while inhibition of miR-92a/TLR2/AP-1 expression resulted in decreased apoptosis rate and bax levels, and increased bcl-2 levels in the miR-92a inhibitor group, TLR2(-) group, and AP-1(-) group than the MG ($P < 0.05$). The level of apoptosis in the lung tissues did not identify any significant difference among the miR-92a inhibitor group, TLR2(-) group, and AP-1(-) group ($P > 0.05$) (Figure 3).

3.1.4. Changes of Oxidative Stress Levels in ALI Rats' Lung Tissues. For the purpose of evaluating the effect of miR-92a/TLR2/AP-1 on oxidative stress levels in ALI rats' lung tissues, we observed the changes of oxidative stress in the lung tissues of ALI rats. It turned out that after LPS stimulation, the oxidative stress levels in the lung tissues of rats in the MG rose dramatically, SOD level dropped ($P < 0.05$), and MDA level boosted ($P < 0.05$). However, after inhibiting the expression of miR-92a/TLR2/AP-1, the levels of oxidative stress in ALI rats' lung tissues decreased, the SOD level elevated ($P < 0.05$), and the MDA level declined ($P < 0.05$). In particular, the inhibitory effect of miR-92a on oxidative stress was the most obvious. Among the three groups of miR-92a inhibitor group, TLR2(-) group and AP-1(-) group, the level of SOD in the lung tissues of miR-92a inhibitor group was the highest ($P < 0.05$) (Figure 4).

3.1.5. Changes of Inflammatory Response Levels in ALI Rats' Lung Tissues. Further, we assessed the effect of miR-92a/TLR2/AP-1 on inflammatory response levels in ALI rats' lung tissues. The levels of TNF- α , IL-6, and IL-10 in the lung tissues of the MG were obviously increased ($P < 0.05$). Inhibition of miR-92a/TLR2/AP-1 brought lowered TNF- α and IL-6 levels ($P < 0.05$), elevated IL-10 level ($P < 0.05$), and suppressed inflammatory response in ALI rats, and miR-92a showed the strongest anti-inflammatory effect. The TNF- α and IL-6 in the lung tissues in the miR-92a inhibitor group were lower than those in the TLR2(-) group and the AP-1(-) group, with a higher IL-10 ($P < 0.05$) (Figure 5).

3.2. Effects of miR-92a/TLR2/AP-1 on LPS Exposed AECs

3.2.1. Results of miR-92a/TLR2/AP-1 Transfection in AECs. After 48 h of transfection, miR-92a, TLR2, and AP-1 levels in cells of inhibitor group, si-TLR2 group, and si-AP-1 group all showed corresponding decreases ($P < 0.05$). In addition,

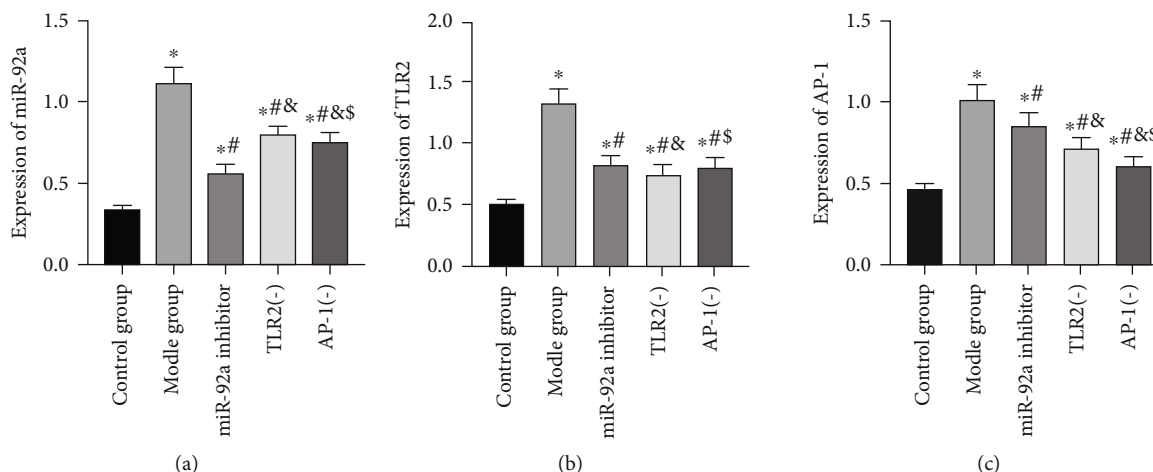


FIGURE 1: Analysis of miR-92a/TLR2/AP-1 level in 5 groups of rats. (a) Changes in miR-92a expression level in rat lung tissues. (b) Changes in TLR2 expression level in rat lung tissues. (c) Changes in AP-1 expression level in rat lung tissues. * indicates $P < 0.05$ compared with the CG, # indicates $P < 0.05$ compared with the MG, & indicates $P < 0.05$ compared with the miR-92a inhibitor group, and \$ indicates $P < 0.05$ compared with the TLR2(-) group.

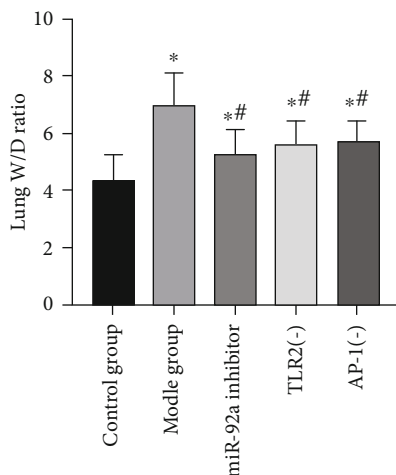


FIGURE 2: Changes of W/D ratio of lung tissues in ALI rats. * indicates $P < 0.05$ compared with the CG, # indicates $P < 0.05$ compared with the MG, & indicates $P < 0.05$ compared with the miR-92a inhibitor group, and \$ indicates $P < 0.05$ compared with the TLR2(-) group.

compared with the LPS group, the TLR2 and AP-1 levels in cells of the inhibitor group were remarkably decreased after downregulating miR-92a ($P < 0.05$), while the expression level of AP-1 was also reduced after downregulating the expression of TLR2 ($P < 0.05$) (Figure 6).

3.3. Effects of miR-92a/TLR2/AP-1 on Apoptosis Level of LPS Exposed AECs. After exposure to LPS, the apoptosis level of AECs increased dramatically, and the apoptosis rate and bax and bcl-2 levels in the LPS group were higher than those in the BG ($P < 0.05$). While inhibited miR-92a/TLR2/AP-1 led to decreased apoptosis level of AECs, and the apoptosis rate and bax and bcl-2 levels in the inhibitor group, si-TLR2 group, and si-AP-1 group were lower than those in LPS group, and the bcl-2 level increased ($P < 0.05$) (Figure 7).

3.4. Effects of miR-92a/TLR2/AP-1 on Oxidative Stress Levels of LPS Exposed AECs. After LPS stimulation, SOD level in AECs declined ($P < 0.05$), MDA level boosted ($P < 0.05$), and oxidative stress level enhanced. While inhibited miR-92a/TLR2/AP-1 reduced the oxidative stress stimulation of LPS on AECs, increased SOD, and decreased MDA ($P < 0.05$). (Figure 8)

3.5. Effects of miR-92a/TLR2/AP-1 on Inflammatory Response Levels of LPS Exposed AECs. The level of inflammatory response of AECs in the LPS group was higher than the BG. After being stimulated by LPS, TNF- α , IL-6, and IL-10 levels increased noticeably ($P < 0.05$). While after the inhibition of miR-92a/TLR2/AP-1 in cells, the inflammatory response inhibited, TNF- α and IL-6 decreased, and IL-10 increased ($P < 0.05$) (Figure 9).

4. Discussion

ALI is a refractory respiratory dysfunction disease whose pathogenesis is not yet fully understood, so specific prevention and treatment approaches have been lacking, leading to a high mortality rate globally [18]. MiRNAs exert marked effects on cell development, genomic imprinting, and regulation of cell functions [19]. With the exception of Fu et al. [11], little research in recent years has reported the part miR-92a plays in ALI. The present study once again verified the effect of miR-92a on ALI rats and further analyzed miR-92a's effects on the survival of AECs exposed to LPS. It was found that miR-92a attenuated the oxidative stress response and apoptosis of LPS-induced AECs through TLR2/AP-1 signaling axis.

In this study, we revealed that both inhibition of miR-92a expression and knockout of TLR2 and AP-1 genes could reduce LPS-induced rat ALI, alleviate pulmonary edema, inhibit oxidative stress and inflammatory response, and reduce apoptosis of lung tissue cells. In addition, it was noticed that in ALI rats' lung tissues, TLR2 and AP-1 were

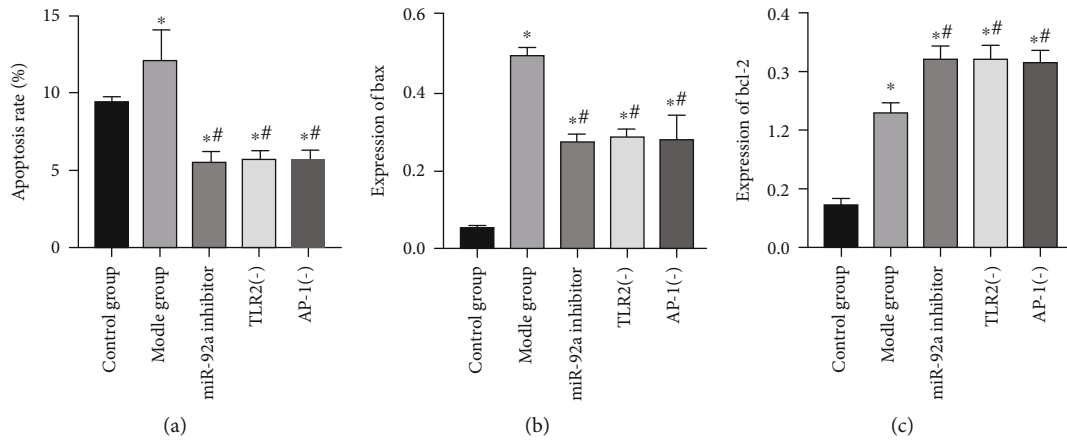


FIGURE 3: Changes of apoptosis level in lung tissues of ALI rats. (a) Results of apoptosis detection in rat lung tissues. (b) Changes in bax expression level in rat lung tissues. (c) Changes in bcl-2 expression level in rat lung tissues. * indicates $P < 0.05$ compared with the CG, and # indicates $P < 0.05$ compared with the MG.

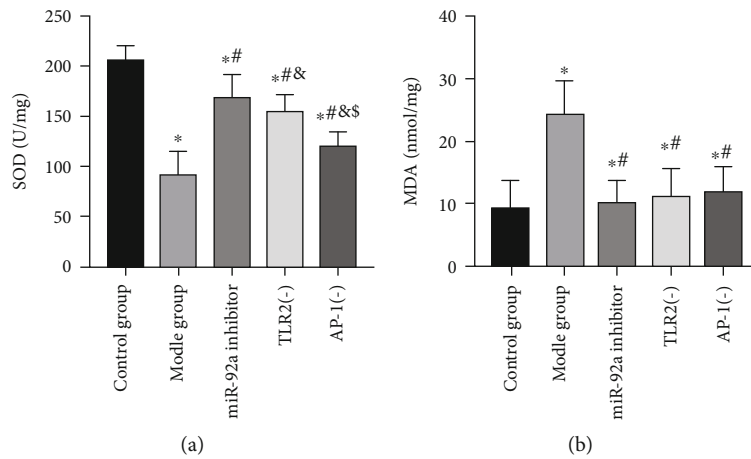


FIGURE 4: Analysis of oxidative stress levels in lung tissues of rats in 5 groups. (a) Changes in SOD expression level in rat lung tissues. (b) Changes in MDA expression level in rat lung tissues. * indicates $P < 0.05$ compared with the CG, # indicates $P < 0.05$ compared with the MG, & indicates $P < 0.05$ compared with the miR-92a inhibitor group, and \$ indicates $P < 0.05$ compared with the TLR2(-) group.

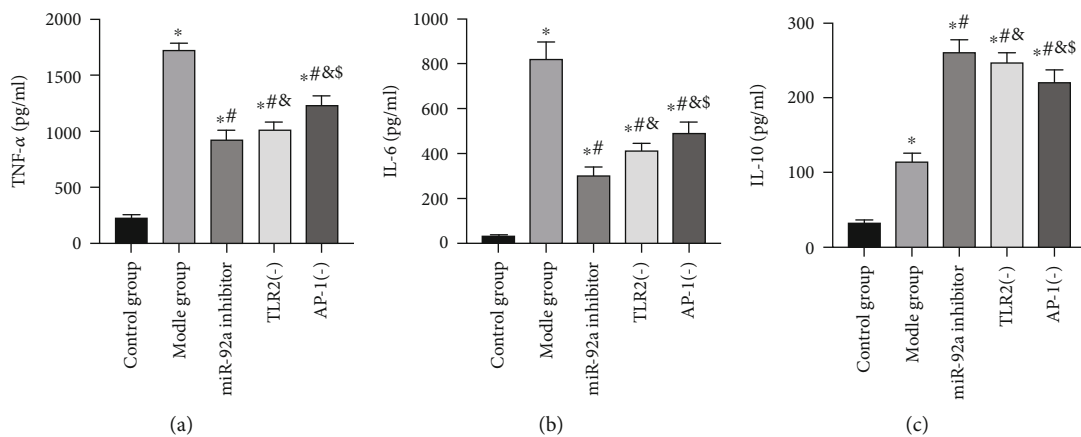


FIGURE 5: Analysis of inflammatory reaction levels in lung tissues of rats in 5 groups. (a) Changes in TNF-α expression level in rat lung tissues. (b) Changes in IL-6 expression level in rat lung tissues. (c) Changes in IL-10 expression level in rat lung tissues. * indicates $P < 0.05$ compared with the CG, # indicates $P < 0.05$ compared with the MG, & indicates $P < 0.05$ compared with the miR-92a inhibitor group, and \$ indicates $P < 0.05$ compared with the TLR2(-) group.

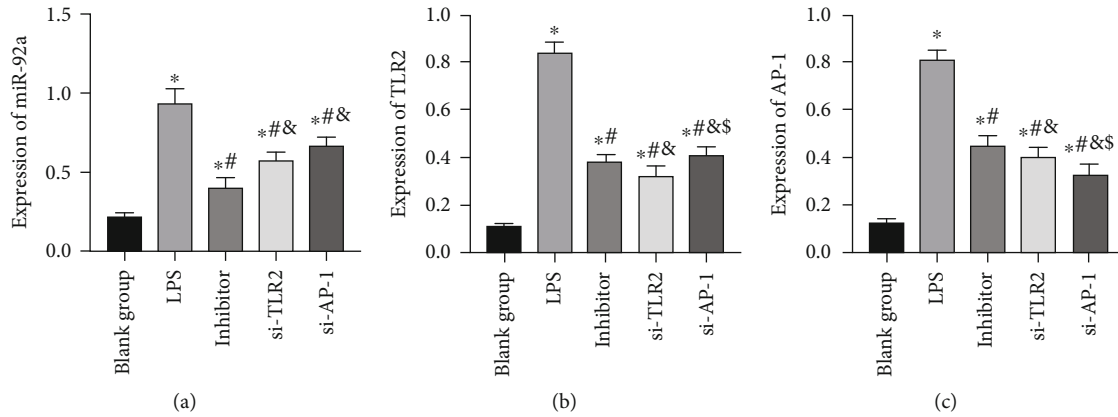


FIGURE 6: Results of miR-92a/TLR2/AP-1 transfection in AECs. (a) Changes in miR-92a expression level in AECs. (b) Changes in TLR2 expression level in AECs. (c) Changes in AP-1 expression level in AECs. * indicates $P < 0.05$ compared with the BG, # indicates $P < 0.05$ compared with the LPS group, & indicates $P < 0.05$ compared with the inhibitor group, and \$ indicates $P < 0.05$ compared with the si-TLR2 group.

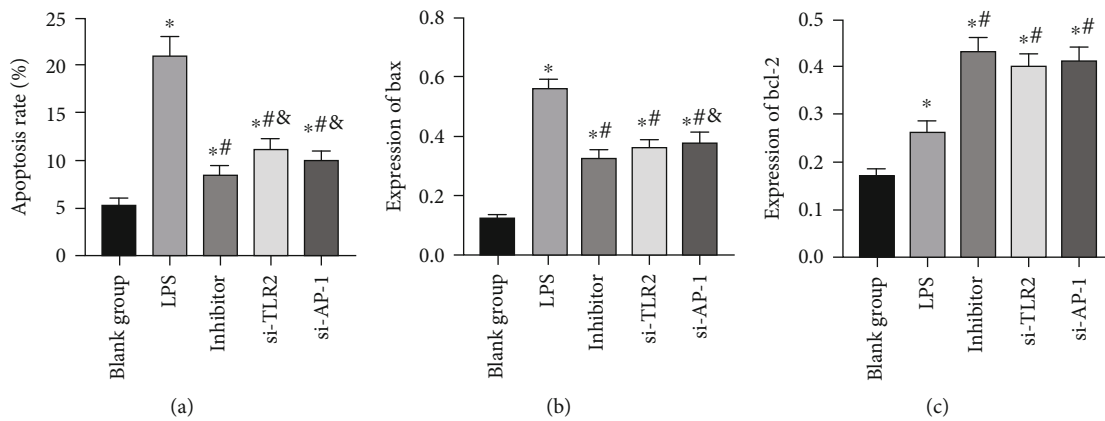


FIGURE 7: Analysis of apoptosis level of AECs. (a) Apoptosis rate of AECs (b) Changes in bax expression level in AECs. (c) Changes in bcl-2 expression level in AECs. * indicates $P < 0.05$ compared with the BG, # indicates $P < 0.05$ compared with the LPS group, and & indicates $P < 0.05$ compared with the inhibitor group.

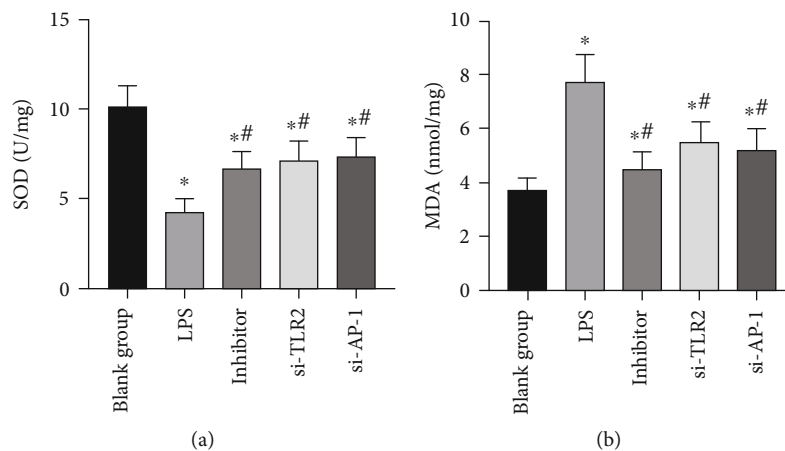


FIGURE 8: Changes in oxidative stress levels in AECs. (a) Changes in SOD expression level in AECs. (b) Changes in MDA expression level in AECs. * indicates $P < 0.05$ compared with the BG, and # indicates $P < 0.05$ compared with the LPS group.

suppressed after inhibiting the expression of miR-92a, and the expression of AP-1 was decreased after knocking out the TLR2 gene, suggesting that miR-92a/TLR2/AP-1 might

be a signaling axis that exerts marked effects on the occurrence and development of ALI. The regulatory effect of miRNAs on TLR protein has been confirmed in many studies

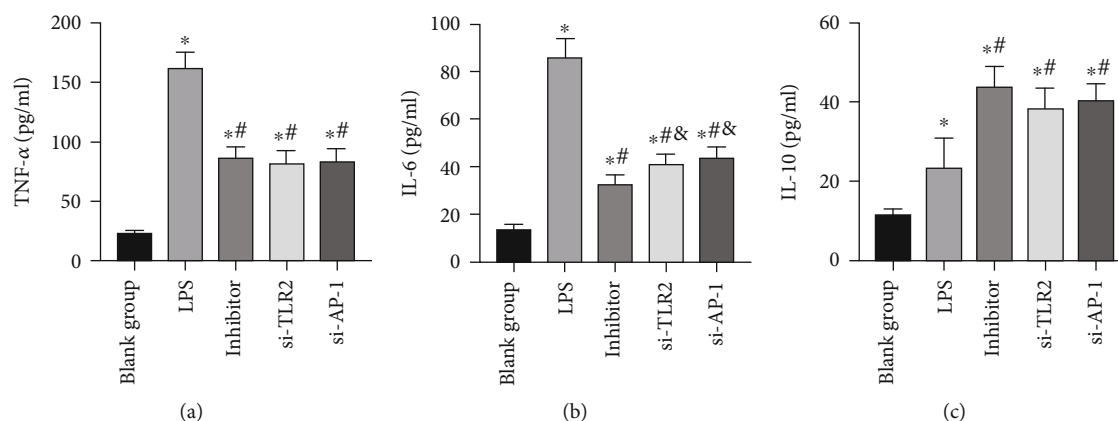


FIGURE 9: Analysis of inflammatory response levels of AECs. (a) Changes in TNF- α expression level in AECs. (b) Changes in IL-6 expression level in AECs. (c) Changes in IL-10 expression level in AECs. * indicates $P < 0.05$ compared with the BG, # indicates $P < 0.05$ compared with the LPS group, and & indicates $P < 0.05$ compared with the inhibitor group.

[20], while only Zhao et al. [12] reported that miR-92a could target the regulation of TLR2 expression. Apart from that, it is well established that AP-1 can be regulated by TLR2. As reported by Wan et al. [21], geranyl diphosphate synthase 1 alleviates ventilator-induced lung injury through the TLR2/4/AP-1 signaling axis. Moreover, the TLR2-JNK-AP-1 pathway also plays a crucial part in pulmonary fibrosis caused by *Mycobacterium tuberculosis* [22]. This also supports our speculation, and we verified this relationship in AECs. We found that inhibition of miR-92a/TLR2/AP-1 also alleviated LPS-induced AEC injury, reduced cell apoptosis, and inhibited oxidative stress and inflammatory response. What is more, like that in rat lung tissues, the phenomenon also existed in AECs, that is, when miR-92a was inhibited, the expression of TLR2 and AP-1 was also inhibited, and silencing TLR2 could reduce the expression level of AP-1. Therefore, it can be found from our study that miR-92a/TLR2/AP-1 is related to the onset and progression of ALI, and inhibition of it can effectively improve ALI symptoms and reduce apoptosis of AECs.

MiR-92a is a gene closely associated with lung disease and has been shown in many studies to be bound up with lung cancer metastasis. For example, Hsu et al. [23] reported in their study that bone marrow-derived cells could release vesicles coated with miR-92a to promote liver metastasis of lung cancer. In the study of Borzi et al. [24], miR-92a was found to promote the proliferation of lung bronchial cells and establish an ecological environment for intralung metastasis of lung cancer. However, earlier evidence [25] only revealed that inhibited miR-92a could promote the WNT1-inducible signaling pathway protein 1 (WISP1) in lung fibroblasts, a fibrogenic medium, but they did not further analyze fibroblast changes, so miR-92a may inhibit the development of pulmonary fibrosis. This suggests that miR-92a has various roles in different lung diseases, but it is unclear whether miR-92a also has a two-way effect in ALI, which needs to be verified by more studies. Although this study studied the role of miR-92a/TLR2/AP-1 in ALI from animal models in vivo and cell models in vitro, further clinical trials are still needed.

To sum up, miR-92a/TLR2/AP-1 is highly expressed in ALI, and its inhibition can improve oxidative stress and inflammatory response and reduce apoptosis of AECs.

Data Availability

The corresponding data of this manuscript can be available if any researcher required.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

WL and JC designed the project and are responsible for the overall content. JC and HHD performed statistical analysis. JC and YYY carried out all experiments. JC prepared the manuscript. WL contributed to revise the manuscript. All authors have seen and approved the final manuscript. Jian Cui contributed equally to this work.

References

- [1] V. Neudecker, K. S. Brodsky, E. T. Clambey et al., "Neutrophil transfer of miR-223 to lung epithelial cells dampens acute lung injury in mice," *Science Translational Medicine*, vol. 9, no. 408, article eaah5360, 2017.
- [2] M. Sklar, B. Patel, J. Beitler, T. Piraino, and E. Goligher, "Optimal ventilator strategies in acute respiratory distress syndrome," *Seminars in Respiratory and Critical Care Medicine*, vol. 40, no. 1, pp. 081–093, 2019.
- [3] M. A. Matthay, R. L. Zemans, G. A. Zimmerman et al., "Acute respiratory distress syndrome," *Nature Reviews Disease Primers*, vol. 5, no. 1, p. 18, 2019.
- [4] D. Chakraborty, S. Zenker, J. Rossaint et al., "Alarmin S100A8 activates alveolar epithelial cells in the context of acute lung injury in a TLR4-dependent manner," *Frontiers in Immunology*, vol. 8, p. 1493, 2017.
- [5] L. P. Kebaabetswe, A. K. Haick, and T. A. Miura, "Differentiated phenotypes of primary murine alveolar epithelial cells

- and their susceptibility to infection by respiratory viruses," *Virus Research*, vol. 175, no. 2, pp. 110–119, 2013.
- [6] H. Laycock and A. Rajah, "Acute lung injury and acute respiratory distress syndrome: a review article," *BJMP*, vol. 3, no. 2, p. 324, 2010.
- [7] M. L. Motti, S. D' Angelo, and R. Meccariello, "MicroRNAs, cancer and diet: facts and new exciting perspectives," *Current Molecular Pharmacology*, vol. 11, no. 2, pp. 90–96, 2018.
- [8] A. Ramassone, S. Pagotto, A. Veronese, and R. Visone, "Epigenetics and micro RNAs in cancer," *International Journal of Molecular Sciences*, vol. 19, no. 2, p. 459, 2018.
- [9] A. Ferruelo, Ó. Peñuelas, and J. A. Lorente, "MicroRNAs as biomarkers of acute lung injury," *Annals of Translational Medicine*, vol. 6, no. 2, p. 34, 2018.
- [10] L. Song, F. Zhou, L. Cheng et al., "MicroRNA-34a suppresses autophagy in alveolar type II epithelial cells in acute lung injury by inhibiting FoxO3 expression," *Inflammation*, vol. 40, no. 3, pp. 927–936, 2017.
- [11] L. Fu, P. Zhu, S. Qi, C. Li, and K. Zhao, "MicroRNA-92a antagonism attenuates lipopolysaccharide (LPS)-induced pulmonary inflammation and injury in mice through suppressing the PTEN/AKT/NF- κ B signaling pathway," *Biomedicine & Pharmacotherapy*, vol. 107, pp. 703–711, 2018.
- [12] Y. Zhao, Z. Dang, and S. Chong, "Mmu-miR-92a-2-5p targets TLR2 to relieve *Schistosoma japonicum*-induced liver fibrosis," *International Immunopharmacology*, vol. 69, pp. 126–135, 2019.
- [13] L. Lai, Y. Song, Y. Liu et al., "MicroRNA-92a negatively regulates Toll-like receptor (TLR)-triggered inflammatory response in macrophages by targeting MKK4 kinase," *Journal of Biological Chemistry*, vol. 288, no. 11, pp. 7956–7967, 2013.
- [14] L. Fei, F. Jifeng, W. Tiantian, H. Yi, and P. Linghui, "Glycyrrhizin ameliorate ischemia reperfusion lung injury through downregulate TLR2 signaling cascade in alveolar macrophages," *Frontiers in Pharmacology*, vol. 8, p. 389, 2017.
- [15] A. Khan, S. Khan, H. Ali et al., "Anomalin attenuates LPS-induced acute lungs injury through inhibition of AP-1 signaling," *International Immunopharmacology*, vol. 73, pp. 451–460, 2019.
- [16] X. Yang, L. Li, J. Liu, B. Lv, and F. Chen, "Extracellular histones induce tissue factor expression in vascular endothelial cells via TLR and activation of NF- κ B and AP-1," *Thrombosis Research*, vol. 137, pp. 211–218, 2016.
- [17] L. A. McKiel and L. E. Fitzpatrick, "Toll-like receptor 2-dependent NF- κ B/AP-1 activation by damage-associated molecular patterns adsorbed on polymeric surfaces," *ACS Biomaterials Science & Engineering*, vol. 4, no. 11, pp. 3792–3801, 2018.
- [18] S. Qin, M. Chen, H. Ji et al., "miR-21-5p regulates type II alveolar epithelial cell apoptosis in hyperoxic acute lung injury," *Molecular Medicine Reports*, vol. 17, no. 4, pp. 5796–5804, 2018.
- [19] L. Ke, J. Wang, X. Jiang, and C. Zhang, "A review of Micro RNAs related to the occurrence, diagnosis, and prognosis of non-small cell lung Cancer," *Clinical surgery research communications*, vol. 2, no. 3, 2018.
- [20] Z. Zhu, L. Liang, R. Zhang et al., "Whole Blood Mir-181a, Mir-92a, and Mir-424 are risk factors of acute respiratory distress syndrome," in *C32. Lung Injury, Ards, and Sepsis*, no. article A5223, 2017 American Thoracic Society, 2017.
- [21] B. Wan, W. Xu, M. Chen et al., "Geranylgeranyl diphosphate synthase 1 knockout ameliorates ventilator-induced lung injury via regulation of TLR2/4-AP-1 signaling," *Free Radical Biology & Medicine*, vol. 147, pp. 159–166, 2020.
- [22] H. S. Lee, H. S. Hua, C. H. Wang, M. C. Yu, B. C. Chen, and C. H. Lin, "Mycobacterium tuberculosis induces connective tissue growth factor expression through the TLR2-JNK-AP-1 pathway in human lung fibroblasts," *The FASEB Journal*, vol. 33, no. 11, pp. 12554–12564, 2019.
- [23] Y.-L. Hsu, M.-S. Huang, J.-Y. Hung et al., "Bone-marrow-derived cell-released extracellular vesicle miR-92a regulates hepatic pre-metastatic niche in lung cancer," *Oncogene*, vol. 39, no. 4, pp. 739–753, 2020.
- [24] C. Borzi, L. Calzolari, A. M. Ferretti et al., "c-Myc shuttled by tumour-derived extracellular vesicles promotes lung bronchial cell proliferation through miR-19b and miR-92a," *Cell Death & Disease*, vol. 10, no. 10, 2019.
- [25] B. Berschneider, D. C. Ellwanger, H. A. Baarsma et al., "miR-92a regulates TGF- β 1-induced WISP1 expression in pulmonary fibrosis," *The International Journal of Biochemistry & Cell Biology*, vol. 53, pp. 432–441, 2014.

Research Article

circHIPK3 Acts as Competing Endogenous RNA and Promotes Non-Small-Cell Lung Cancer Progression through the miR-107/BDNF Signaling Pathway

Weijun Hong,¹ Yajuan Zhang,¹ Jingyi Ding,¹ Qilian Yang,² Haixiang Xie ,³ and Xiwen Gao ¹

¹Department of Respiratory Medicine, Minhang Hospital, Fudan University, China

²Pharmacy Department, Minhang Hospital, Fudan University, China

³Department of Emergency Medicine, Minhang Hospital, Fudan University, China

Correspondence should be addressed to Haixiang Xie; 13916494378@163.com and Xiwen Gao; xiwengao@sina.com

Received 2 July 2020; Accepted 21 August 2020; Published 15 September 2020

Guest Editor: Tao Huang

Copyright © 2020 Weijun Hong 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.

Circular RNAs (circRNAs) act as a crucial part in many human diseases, particularly in cancers. circRNA HIPK3 (circHIPK3) is a special circRNA that may participate in the oncogenesis of non-small-cell lung cancer (NSCLC), even though its latent regulatory mechanism is not very clear. Here, we studied the roles of circHIPK3 in NSCLC. qRT-PCR assay was applied to study the expression of circHIPK3 in NSCLC. The influence of circHIPK3 on NSCLC was estimated by silencing circHIPK3 and miR-107 mock transfection and brain-derived neurotrophic factor (BDNF) overexpression, and the correlation between circHIPK3, miR-107, and BDNF was evaluated by dual-luciferase reporter assay. The results showed that circHIPK3 expression was upregulated in NSCLC cells. circHIPK3 knockdown inhibited the migration and proliferation of NSCLC cells by promoting the expression of miR-107. circHIPK3 could be used as a miR-107 sponge to promote BDNF cell proliferation. The dual-luciferase reporter assay proved that miR-107 was the target of circHIPK3, and miR-107 had an interaction with the 3' untranslated region of BDNF. miR-107 overexpression inhibited BDNF-mediated NSCLC cell proliferation. These results indicate that circHIPK3 promotes tumor progression through a new circHIPK3/miR-107/BDNF axis, which offers potential markers and medical treatment for NSCLC.

1. Background

Cancer-related diseases have become the biggest leading cause of death worldwide. Lung tumor ranked the first site of all cancers and accounted for 11.6 percent of newly diagnosed cancers and 18.4 percent of cancer-related deaths from the latest report in 2018 [1, 2]. Lung cancer (LC) includes two main types: small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC), which account for about 85% of all LC cases [3]. As imaging modality in medical diagnosis is widely used, the accuracy of diagnosis in NSCLC has been greatly improved [4]. Early diagnosis of patient with NSCLC has not yet been satisfactorily effective for expensive examination fees and low accuracy of clinical image interpretation. The

five-year survival rate of patient with NSCLC is often low, due to the diagnosed time in an advanced stage and the lack of tumor-specific agents [5]. Therefore, it is important to find an appropriate target of NSCLC.

circRNAs are a class of RNA modules that form a continuous cycle of covalent closures. Due to the limitations of technique and cognition, circRNAs are initially thought to be metabolic wastes of RNA missplicing. In recent years, circRNAs have been found to have many functions in the metabolic activities of cells. Unlike linear RNAs, circRNAs are relatively stable presenting in the body, because they have a unique ring-like structure [6–8]. The most common function of circRNAs is to act as miRNA sponges in the regulation of the cell signaling pathway. For example, Cao et al. found that

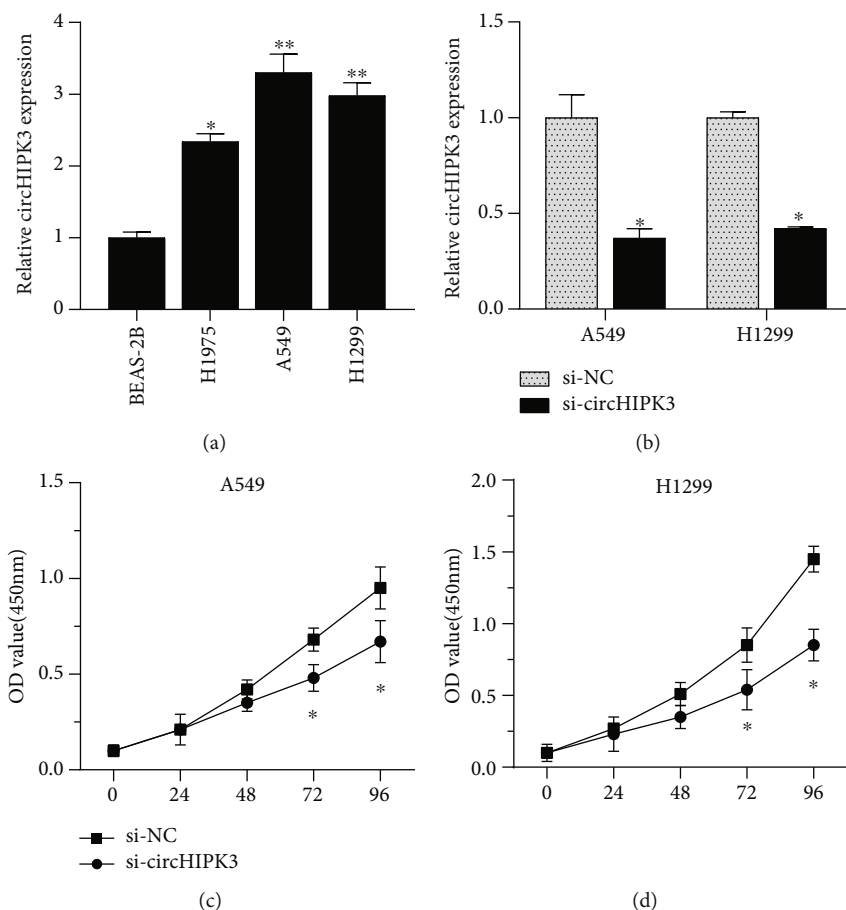


FIGURE 1: circHIPK3 downregulation inhibited NSCLC cell proliferation. (a) The circHIPK3 in the NSCLC cell lines was increased compared to that in BEAS-2B. (b) Effect of small interfering RNA (si-circHIPK3) directed against circHIPK3 in H1299 and A549 cells. (c, d) circHIPK3 silencing had an inhibitory effect on the proliferation of H1299 and A549 cells. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

hsa_circ_0037251 enhanced the progression of glioma via sponging miR-1229-3p [9]. Moreover, circRNAs can also engage in other cell activities. For instance, circRNA circ-Foxo3 can interact with RNA-binding proteins CDK2 and p21, leading to cell cycle arrest [10]. circRNAs can also strengthen the stability of mRNA by binding it, regulate parental gene transcription, and even translate into protein [8]. Previous studies have revealed the abnormal expression of circRNAs is associated with the development of cancer [11]. For example, Han et al. discovered circ-ABCB10 was remarkably upregulated in breast cancer tissue, promoting cell proliferation [12]. A recent study demonstrated that circ-ITCH was suppressed in bladder cancer samples, and upregulation of circ-ITCH inhibited bladder cell proliferation, migration, invasion, and metastasis [13]. There have been related researches about circHIPK3 in lung cancer. The result shows that the circHIPK3 can affect the related pathway through different regulatory axes [14, 15].

MicroRNA (miRNA) is also a very important class of noncoding RNA. It can regulate translation via interacting with the 3'UTR region of mRNA [2]. miRNAs play pivotal roles in the process of gene translation, acting as gene regulators, but miRNAs also are regulated by other factors, such as

lncRNA and circRNA. For instance, it is demonstrated that lncRNA APF affects ATG7 expression by regulating miR-188-3p, thereby promoting autophagy death and myocardial infarction [16]. circRNA_010567 affects the expression of the miR-141 and TGF- β 1 and promotes the resection of fibrosis-associated protein in cardiac fibroblasts [17]. Few studies have pointed out that the circRNAs serve as miRNA sponges in regulating the gene expression [18].

In this research, we found that circHIPK3 was strongly overexpressed in NSCLC tissue. Based on the finding, we conducted a series of assays to explore the roles of circHIPK3 in the progression of NSCLC. The results revealed that circHIPK3 could act as a sponge for miR-107 to promote the NSCLC cell tumorigenesis and relieve miRNA repression for downstream target gene BDNF. In short, our results showed that circHIPK3 might act as an oncogenic gene in NSCLC progression and could be a potential biomarker for screening and treatment of NSCLC.

2. Materials and Methods

2.1. Cell Culture. H1299, A549, and BEAS-2B were obtained from ATCC and cultured in DMEM (Life Technologies,

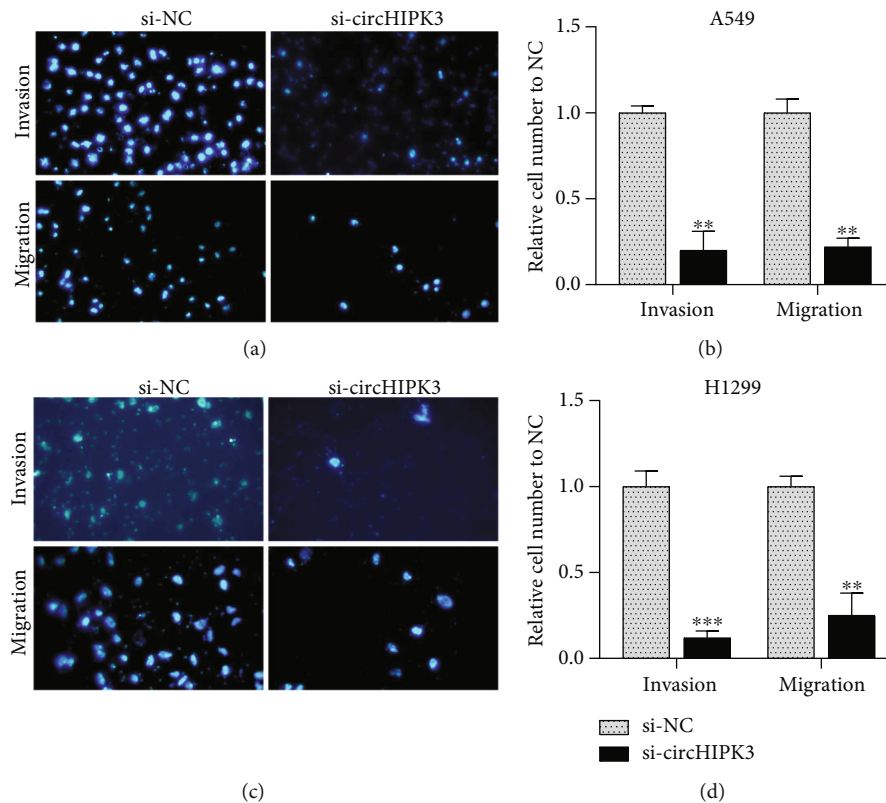


FIGURE 2: circHIPK3 downregulation inhibited NSCLC cell metastasis. (a, c) Picture of A549 (a) and H1299 (b) cells transfected with si-circHIPK3 and si-NC under a fluorescence microscope. (b, d) Knockdown of circHIPK3 suppressed cell metastasis in A549 (b) and H1299 (c). * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Carlsbad, CA) with 10% fetal bovine serum (FBS), 100 U/mL penicillin, and 100 $\mu\text{g}/\text{mL}$ streptomycin, in a 37°C incubator containing 5% CO_2 .

2.2. Plasmid Construction and Transfection Assay. To overexpress miR-107, miR-107 mimics and miR-NC mimics were obtained from the GenePharma company. We used Lipofectamine 2000 (Invitrogen, CA) to transfect the cells with 50 nm mimics. To inhibit the expression of miR-107, NSCLC cells were transfected with miR-107 specific inhibitor (Invitrogen) for 48 hours prior to other experiments. For the expression analysis of BDNF, we used the pCDNA3.1 vector and transfected the plasmids into cells with Lipofectamine 2000. To analyze the expression of circHIPK3, Lipofectamine 2000 was used to transfect small interfering RNA (siRNA) for circHIPK3 (GenePharma) into H1299 and A549 cells at 50 nM.

2.3. Total RNA Isolation and Real-Time Fluorescent Quantitative PCR (qRT-PCR). We extracted the total RNA by using TRIzol reagent (Invitrogen) and decided the concentration via the NanoDrop ND-1000. The primers for the detection of circHIPK3, miR-107, and BDNF were designed and purchased from GenePharma. qRT-PCR was applied using an AB7300 thermo-recycler (Applied Biosystems, Carlsbad, CA) with a TaqMan Universal PCR Master Mix. The relative expression of targets was determined by the $2^{-\Delta\Delta\text{Ct}}$ method.

2.4. Assay of Cell Proliferation. Cell growth ability was measured by using the CCK-8 kit depending on the instruction (Dojindo Laboratories, Japan) [19].

2.5. Transwell Assay. Transwell chambers (Corning, NY) were applied to observe the cell migration and invasion. After cultured for 2 days, the cells on the upper surface were removed. The cells on the lower surface were fixed and stained with DAPI. For the detection of invasion, the cells were seeded into the upper chamber which was precoated with a 2 mg/mL matrix gel.

2.6. Dual-Luciferase Reporter Assay. First, the mutant and wild-type circHIPK3 and BDNF1 3'UTR were cloned into pmirGLO vectors. WT/Mut-pmirGLO-circHIPK3 or WT/Mut-BDNF and miR-107 were cotransfected into NSCLC cells with Lipofectamine 2000. After transfection for 2 days, the cells were collected and detected the luciferase activity with the luciferase reporter assay system (Promega, Madison, WI). The relative activity was set as an internal control index of renal dual-luciferase.

2.7. Statistical Analysis. We used the SPSS software to analyze data. Student's t -test or the Mann-Wintney nonparametric test was used to determine the difference between two groups, and $p < 0.05$ was considered to be significant. The p value was corrected by the FDR (False Discovery Rate).

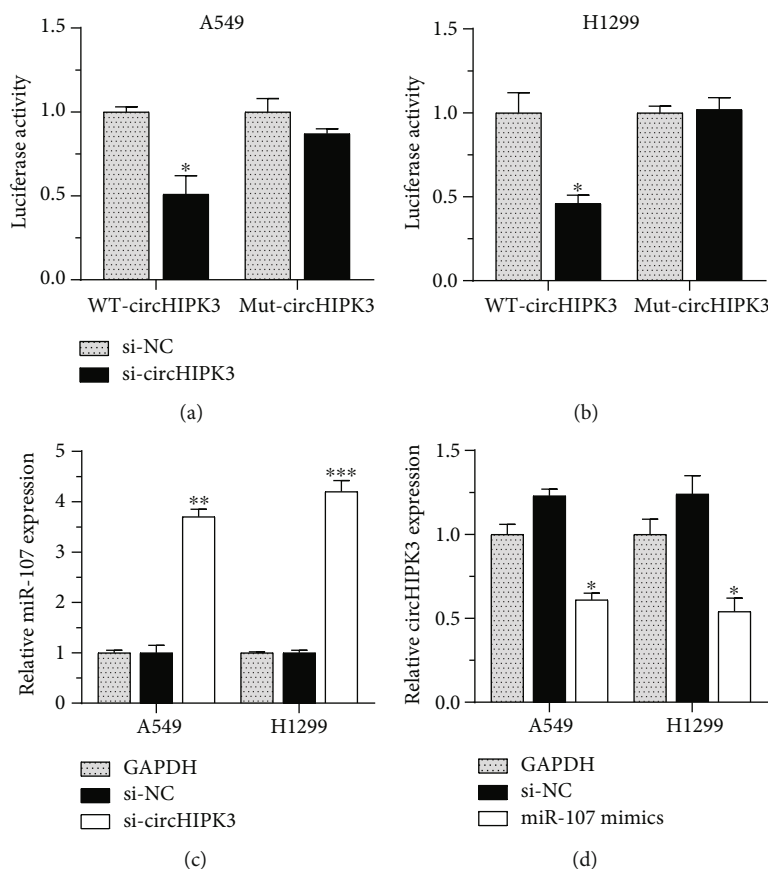


FIGURE 3: miR-107 was a target for circHIPK3 in NSCLC cells. (a, b) A549 (a) and H1299 (b) cells cotransfected with miR-107 and WT-circHIPK3 have lower relative luciferase activity. (c) The downregulation of circHIPK3 significantly upregulated the expression of miR-107. (d) Overexpression of miR-107 suppressed circHIPK3 expression. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

3. Results

3.1. circHIPK3 Downregulation Inhibited NSCLC Cell Proliferation and Invasion. The potential regulatory mechanism of circHIPK3 was still unclear. The results showed that the circHIPK3 in the H1975, A549, and H1299 cell lines was increased compared to that in BEAS-2B (Figure 1(a)). H1299 and A549 cells were used for functional exploration, because their expression levels were the highest. To detect the potential function of circHIPK3 in the invasion and proliferation of NSCLC cells, siRNA cells were transfected for 48 h and compared with the NC groups. As presented in Figure 1(b), it was remarkable that the expression of circHIPK3 was downregulated. The CCK-8 assay showed that circHIPK3 silencing inhibited the proliferation of A549 and H1299 cells depending on the time (Figures 1(c) and 1(d)). Transwell assay also showed that silencing circHIPK3 had an inhibitory effect on the migration and invasion of NSCLC (Figures 2(a)–2(d)).

3.2. miR-107 Was a Target for circHIPK3 in NSCLC Cells. The bioinformatics analyses showed miR-107 was a potential binding miRNA for circHIPK3. We predicted the downstream target genes by bioinformatics tools (RegRNA) [20], and the prediction websites were <http://regrna2.mbc.nctu.edu.tw/>. The wild-type and mutant circHIPK3 luciferase

reporter plasmids were constructed. Dual-luciferase reporter assay indicated cotransfection of miR-107, and WT-circHIPK3 reduced relative luciferase activity, while cotransfection of miR-107 and Mut-circHIPK3 did not result in the decrease of luciferase activity. These data confirmed that miR-107 was a direct target for circHIPK3 (Figures 3(a) and 3(b)). In order to explore the relationship between circHIPK3 and miR-107, we transfected miR-107 mimics or si-circHIPK3 into H1299 and A549 cells. The downregulation of circHIPK3 significantly upregulated the expression of miR-107 (Figure 3(c)), and the overexpression of miR-107 suppressed circHIPK3 expression (Figure 3(d)).

3.3. circHIPK3 Knockdown Suppressed H1299 and A549 Cell Proliferation through miR-107. To determine whether miR-107 was involved in the effects of circHIPK3 knockdown on cell proliferation, H1299 and A549 cells were transfected with a miR-107 inhibitor that suppressed the expression of miR-107 (Figure 4(a)). The CCK-8 assay showed that the downregulation of circHIPK3 inhibited the proliferation of A549 and H1299 cells (Figures 4(b) and 4(c)). The downregulation of miR-107 reversed the suppressive effect, indicating that miR-107 was the downstream of circHIPK3. Our results proved miR-107 also played a key role in regulating tumorigenesis of NSCLC cell.

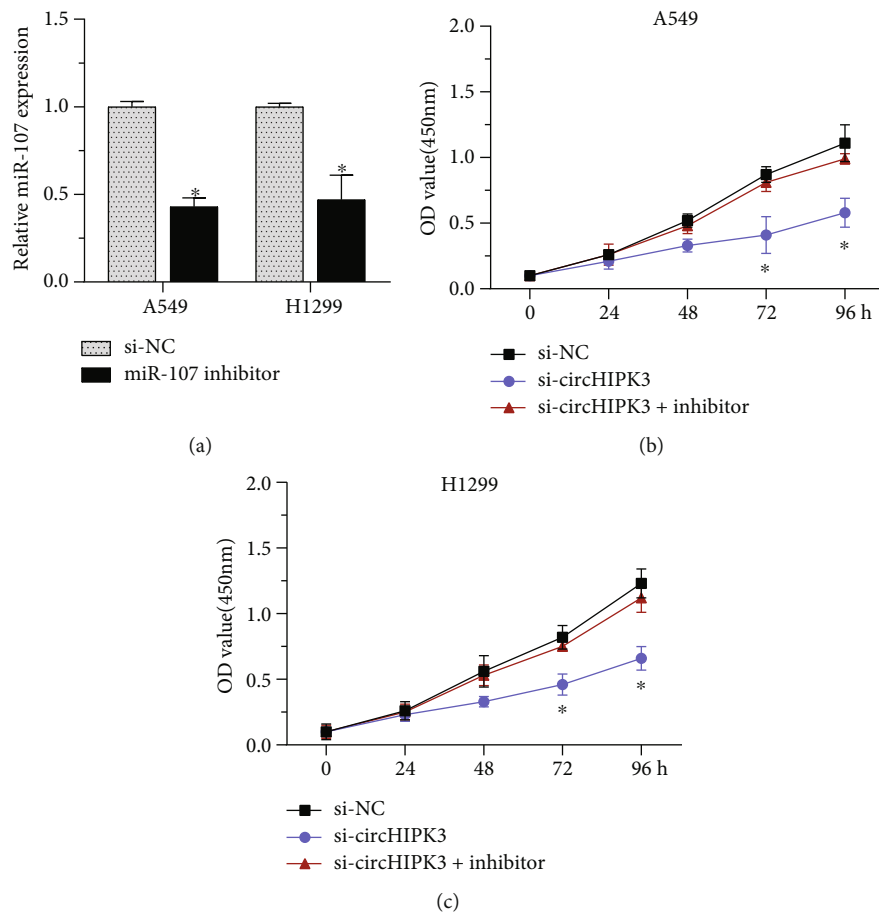


FIGURE 4: circHIPK3 knockdown suppressed H1299 and A549 cell proliferation through miR-107. (a) The miR-107 expression level of H1299 and A549 cells was significantly reduced after transfection with miR-107 inhibitor. (b, c) Downregulation of circHIPK3 suppressed the growth of A549 (b) and H1299 (c) cells. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

3.4. BDNF Was a Direct Target of miR-107. BDNF plays a vital part in elevating the ability of tumor metastasis and proliferation. We explored the roles of BDNF in circHIPK3/miR-107-mediated tumor progression in NSCLC. Dual-luciferase reporter assay discovered cotransfection of miR-107, and BDNF 3'UTR reporter reduced dual-luciferase activity, while cotransfection of miR-107 and Mut-BDNF vectors proved to have no significant impact on relative luciferase activity. Our findings indicated BDNF was a direct target of miR-107 (Figures 5(b) and 5(c)). To further confirm the interaction between miR-107 and BDNF, the levels of BDNF mRNA in H1299 and A549 cells were detected after the upregulation of miR-107. The overexpression of miR-107 remarkably suppressed the expression of BDNF mRNA (Figure 5(d)), revealing that BDNF was a downstream regulator of miR-107. To confirm that BDNF was a target of circHIPK3, we detected BDNF mRNA levels after circHIPK3 knockdown, and the results showed circHIPK3 silencing suppressed BDNF mRNA levels (Figure 5(e)), and this suppression could be reversed by miR-107 inhibitors (Figure 5(f)).

Next, we constructed a BDNF overexpression vector (Figure 6(a)) and transfected it into NSCLC cells. The overexpression of BDNF reversed the inhibition of miR-107

mimics on the proliferation of H1299 and A549 cells, as proved in the CCK-8 assay (Figure 6(c)).

4. Discussion

Previous investigations demonstrated that circRNAs acted as a key regulator in the progression and development of cancers [11, 18]. According to the findings, we summarized that circHIPK3 played an important role in the proliferation and metastasis of NSCLC cells. In this study, we reported circHIPK3 was upregulated in NSCLC cells. Knockdown of circHIPK3 obviously reduced NSCLC cell proliferation, migration, and invasion. Our results showed the levels of miR-107 and BDNF expression were also regulated by circHIPK3 knockdown. This suggested that circHIPK3 promoted NSCLC development via the miR-107/BDNF axis.

It is evident that the circRNAs have a crucial part in the regulation of gene expression. circRNAs negatively regulate the miRNA level by direct binding, thus working as miRNA sponges in gene translation regulatory networks [21]. Our results showed that circHIPK3 bound to miR-107, which was further confirmed by the dual-luciferase reporter assay. Previous studies suggested that the expression of miR-107

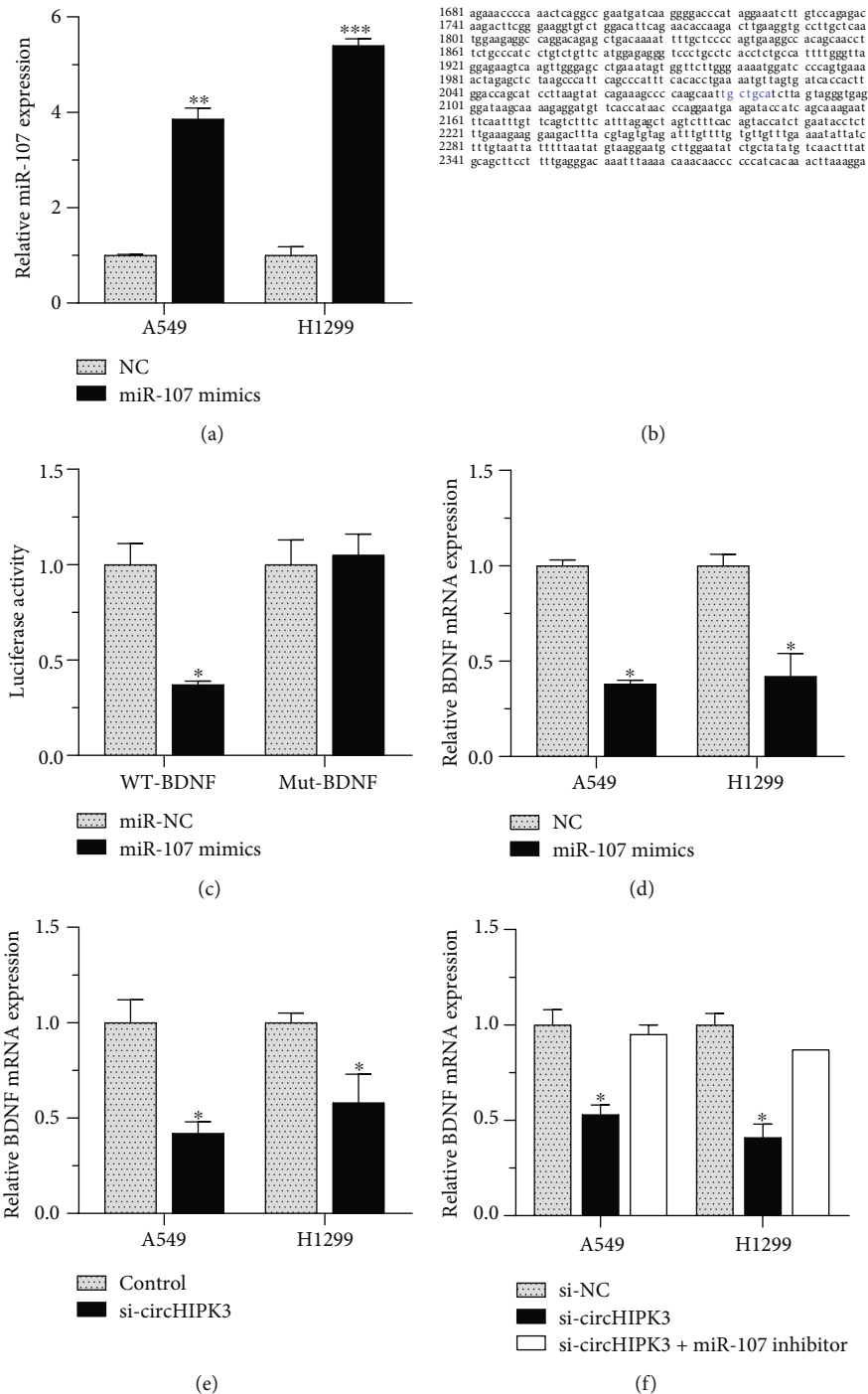


FIGURE 5: BDNF was a direct target of miR-107. (a) miR-107 expression increased after the transfection of miR-107 mimics in H1299 and A549 cells. (b, c) BDNF was a direct target of miR-107. (d) miR-107 overexpression remarkably inhibited the expression of BDNF. (e) circHIPK3 silencing suppressed BDNF mRNA levels. (f) miR-107 inhibitors had reversion function on the suppression caused by circHIPK3 silencing. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

was associated with the development of the disease. In Alzheimer's disease, miR-107 might target β -site amyloid precursor protein-cleaving enzyme 1 [22]. Chen et al. proved that miR-107 directly interacted with miRNA let-7 to negatively regulate the tumor suppressor, thereby promoting breast cancer cell growth [23]. Lee et al. found that the expression of cyclin-dependent kinase 6 in pancreatic cancer

lines MiaPACA-2 and PANC-1 was regulated by the down-regulation of miR-107 [24]. In this study, the knockdown of circHIPK3 upregulated miR-107 expression. However, the expression of downregulation of miR-107 reversed the circHIPK3-silencing-induced inhibition of the progression of NSCLC cells; accordingly, miR-107 upregulation suppressed NSCLC cell proliferation and metastasis.

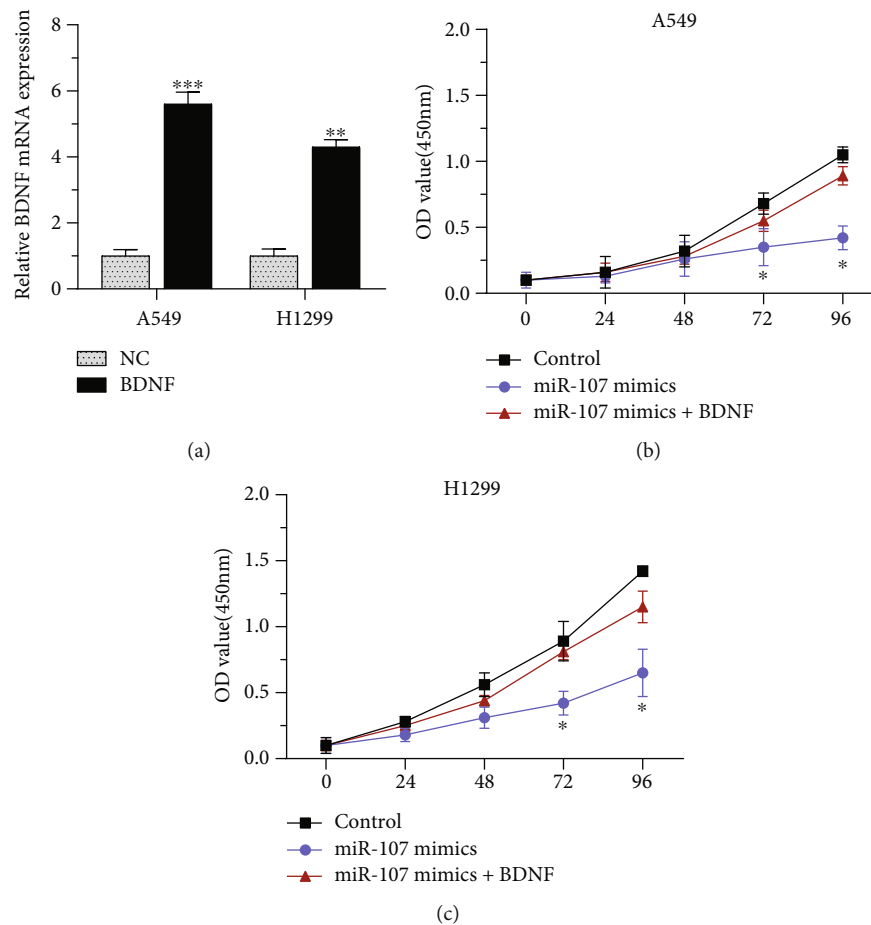


FIGURE 6: BDNF reversed the inhibition of miR-107 overexpression. (a) BDNF overexpression vector increased the BDNF mRNA expression. (b) The overexpression of BDNF had reversion function on the inhibition of miR-107 mimics on the proliferation of H1299 and A549 cells. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

BDNF exists in the adult central nervous system and human platelets [25]. At first, it was known BDNF had a crucial role in the development and function of the nervous system and cardiovascular system [26]. A series of studies indicated BDNF was associated with the development of many diseases. These findings showed that BDNF was aberrantly expressed in multiple human cancers, including bladder cancer and colorectal cancer [27, 28]. Wang et al. demonstrated that BDNF promoted thyroid cancer cell growth via regulating downstream signal pathway PI3K/AKT [29]. Xia et al. revealed miR-107 upregulation could suppress the progression of NSCLC through the downregulation of BDNF [30]. It was confirmed by our experimental results that miR-107 interacted with BDNF. In this study, mechanism assays suggested after circHIPK3 knock-down, miR-107 was released and the miR-107 downstream target genes, such as BDNF, were downregulated. Meanwhile, when we knocked down the expression of circHIPK3, the progression of NSCLC cells was suppressed, accompanied by the reduced expression of the miR-107 downstream targeted genes BDNF.

There are a number of factors, including the occurrence and development of NSCLC, including smoking, air pollution, and genetic [31, 32]. Apparently, a single study pre-

cludes a complete understanding of the etiology of NSCLC. Therefore, more basic studies are needed in the future to get more information about NSCLC.

In short, our results uncovered circHIPK3 new mechanism in regulating NSCLC proliferation and metastasis via acting as a miR-107 sponge. We highlighted circHIPK3/miR-107/BDNF as a novel tumor screening biomarker for NSCLC. Thus, we hoped that our findings would help to the screening and treatment of NSCLC. This regulatory mechanism could help us to explore the transcriptional regulation level in NSCLC and could be also explored in other diseases. This research result provided an important supplement to the regulatory study of the ceRNA regulation network.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Weijun Hong and Yajuan Zhang contributed equally to this work.

Acknowledgments

This work is supported by the Shanghai Municipal Health Commission's special clinical research projects in the health industry (Grant Number 201940102), applicant: Xiwen Gao.

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] S. Gu, L. Jin, F. Zhang, P. Sarnow, and M. A. Kay, "Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs," *Nature Structural & Molecular Biology.*, vol. 16, no. 2, pp. 144–150, 2009.
- [3] M. G. Oser, M. J. Niederst, L. V. Sequist, and J. A. Engelman, "Transformation from non-small-cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin," *The Lancet Oncology.*, vol. 16, no. 4, pp. e165–e172, 2015.
- [4] Y.-L. Lv, D.-M. Yuan, K. Wang et al., "Diagnostic performance of integrated positron emission tomography/computed tomography for mediastinal lymph node staging in non-small cell lung cancer: a bivariate systematic review and meta-analysis," *Journal of Thoracic Oncology.*, vol. 6, no. 8, pp. 1350–1358, 2011.
- [5] J. He, Y. Hu, M. Hu, and B. Li, "Development of PD-1/PD-L1 Pathway in Tumor Immune Microenvironment and Treatment for Non-Small Cell Lung Cancer," *Scientific Reports.*, vol. 5, no. 1, 2015.
- [6] L. Yang, J. Fu, and Y. Zhou, "Circular RNAs and Their Emerging Roles in Immune Regulation," *Frontiers in Immunology.*, vol. 9, 2018.
- [7] Q. Shang, Z. Yang, R. Jia, and S. Ge, "The novel roles of circRNAs in human cancer," *Molecular Cancer.*, vol. 18, no. 1, 2019.
- [8] R. Zhou, Y. Wu, W. Wang et al., "Circular RNAs (circRNAs) in cancer," *Cancer Letters.*, vol. 425, pp. 134–142, 2018.
- [9] Q. Cao, Y. Shi, X. Wang et al., "Circular METRN RNA hsa_circ_0037251 promotes glioma progression by sponging miR-1229-3p and regulating mTOR expression," *Scientific Reports.*, vol. 9, no. 1, p. 19791, 2019.
- [10] W. W. Du, W. Yang, E. Liu, Z. Yang, P. Dhaliwal, and B. B. Yang, "Foxo 3 circular RNA retards cell cycle progression via forming ternary complexes with p 21 and CDK2," *Nucleic Acids Research.*, vol. 44, no. 6, pp. 2846–2858, 2016.
- [11] H.-d. Zhang, L.-h. Jiang, D.-w. Sun, J.-c. Hou, and Z.-l. Ji, "CircRNA: a novel type of biomarker for cancer," *Breast Cancer.*, vol. 25, no. 1, pp. 1–7, 2018.
- [12] X.-T. Han, J. Q. Jiang, M. Z. Li, and Q. M. Cong, "Circular RNA circ-ABCB10 promotes the proliferation and invasion of thyroid cancer by targeting KLF6," *European Review for Medical and Pharmacological Sciences.*, vol. 24, no. 3, pp. 1271–1277, 2020.
- [13] C. Yang, W. Yuan, X. Yang et al., "Circular RNA circ-ITCH inhibits bladder cancer progression by sponging miR-17/miR-224 and regulating p21, PTEN expression," *Molecular Cancer.*, vol. 17, no. 1, p. 19, 2018.
- [14] X. Chen, R. Mao, W. Su et al., "Circular RNA circHIPK3 modulates autophagy via MIR124-3p-STAT3-PRKAA/AMPK α signaling in STK11 mutant lung cancer," *Autophagy.*, vol. 16, no. 4, pp. 659–671, 2020.
- [15] H. Lu, X. Han, J. Ren, K. Ren, Z. Li, and Z. Sun, "Circular RNA HIPK3 induces cell proliferation and inhibits apoptosis in non-small cell lung cancer through sponging miR-149," *Cancer Biology & Therapy.*, vol. 21, no. 2, pp. 113–121, 2020.
- [16] K. Wang, C.-Y. Liu, L.-Y. Zhou et al., "APF lncRNA regulates autophagy and myocardial infarction by targeting miR-188-3p," *Nature Communications.*, vol. 6, no. 1, 2015.
- [17] B. Zhou and J.-W. Yu, "A novel identified circular RNA, circRNA_010567, promotes myocardial fibrosis via suppressing miR-141 by targeting TGF- β 1," *Biochemical and Biophysical Research Communications.*, vol. 487, no. 4, pp. 769–775, 2017.
- [18] S. Meng, H. Zhou, Z. Feng et al., "CircRNA: functions and properties of a novel potential biomarker for cancer," *Molecular Cancer.*, vol. 16, no. 1, p. 94, 2017.
- [19] J. Bai, X. Zhu, J. Ma, and W. Wang, "miR-205 regulates A549 cells proliferation by targeting PTEN," *International Journal of Clinical and Experimental Pathology.*, vol. 8, no. 2, pp. 1175–1183, 2015.
- [20] T.-H. Chang, H.-Y. Huang, J. B.-K. Hsu, S.-L. Weng, J.-T. Horng, and H.-D. Huang, "An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs," *BMC Bioinformatics.*, vol. 14, no. S2, 2013.
- [21] J. Cheng, H. Zhuo, M. Xu et al., "Regulatory network of circRNA-miRNA-mRNA contributes to the histological classification and disease progression in gastric cancer," *Journal of Translational Medicine.*, vol. 16, no. 1, p. 216, 2018.
- [22] W.-X. Wang, B. W. Rajeev, A. J. Stromberg et al., "The Expression of MicroRNA miR-107 Decreases Early in Alzheimer's Disease and May Accelerate Disease Progression through Regulation of -Site Amyloid Precursor Protein-Cleaving Enzyme 1," *Journal of Neuroscience.*, vol. 28, no. 5, pp. 1213–1223, 2008.
- [23] P.-S. Chen, J.-L. Su, S.-T. Cha et al., "miR-107 promotes tumor progression by targeting the let-7 microRNA in mice and humans," *The Journal of Clinical Investigation.*, vol. 121, no. 9, pp. 3442–3455, 2011.
- [24] K. H. Lee, C. Lotterman, C. Karikari et al., "Epigenetic silencing of microRNA miR-107 regulates cyclin-dependent kinase 6 expression in pancreatic cancer," *Pancreatology.*, vol. 9, no. 3, pp. 293–301, 2009.
- [25] H. Yamamoto and M. Gurney, "Human platelets contain brain-derived neurotrophic factor," *The Journal of Neuroscience.*, vol. 10, no. 11, pp. 3469–3478, 1990.
- [26] E. Pius-Sadowska and B. Machaliński, "BDNF – a key player in cardiovascular system," *Journal of Molecular and Cellular Cardiology.*, vol. 110, pp. 54–60, 2017.
- [27] H. Zhi and J. Lian, "LncRNA BDNF-AS suppresses colorectal cancer cell proliferation and migration by epigenetically repressing GSK-3 β expression," *Cell Biochemistry and Function.*, vol. 37, no. 5, pp. 340–347, 2019.
- [28] P. C. Lai, T. H. Chiu, and Y. T. Huang, "Overexpression of BDNF and TrkB in human bladder cancer specimens," *Oncology Reports.*, vol. 24, no. 5, pp. 1265–1270, 2010.

- [29] P. Wang, X. Meng, Y. Huang et al., "MicroRNA-497 inhibits thyroid cancer tumor growth and invasion by suppressing BDNF," *Oncotarget*, vol. 8, no. 2, pp. 2825–2834, 2017.
- [30] H. Xia, Y. Li, and X. Lv, "MicroRNA-107 inhibits tumor growth and metastasis by targeting the BDNF-mediated PI3K/AKT pathway in human non-small lung cancer," *International Journal of Oncology*, vol. 49, no. 4, pp. 1325–1333, 2016.
- [31] C. Jiang, X. Fang, H. Zhang et al., "AMD3100 combined with triptolide inhibit proliferation, invasion and metastasis and induce apoptosis of human U2OS osteosarcoma cells," *Bio-medicine & Pharmacotherapy.*, vol. 86, pp. 677–685, 2017.
- [32] N. Akhtar and J. G. Bansal, "Risk factors of lung cancer in non-smoker," *Current Problems in Cancer.*, vol. 41, no. 5, pp. 328–339, 2017.

Research Article

SRCIN1 Regulated by circCCDC66/miR-211 Is Upregulated and Promotes Cell Proliferation in Non-Small-Cell Lung Cancer

Weijun Hong,¹ Suyun Yu,¹ Yaqing Zhuang,¹ Qingqing Zhang,¹ Jiqin Wang^{ID},²
and Xiwen Gao^{ID}¹

¹Department of Respiratory Medicine, Minhang Hospital, Fudan University, China

²Department of Emergency Medicine, Minhang Hospital, Fudan University, China

Correspondence should be addressed to Jiqin Wang; wjqshcn@163.com and Xiwen Gao; xiwengao@sina.com

Received 30 June 2020; Accepted 20 August 2020; Published 10 September 2020

Guest Editor: Tao Huang

Copyright © 2020 Weijun Hong 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 incidence and mortality of lung cancer were extremely high. The present study showed that SRCIN1 was an oncogene in non-small-cell lung cancer (NSCLC). Public dataset analysis showed SRCIN1 was significantly overexpressed in NSCLC samples. Also, we found that NSCLC patients with higher SRCIN1 expression had shorter OS time by analyzing TCGA, Kaplan-Meier Plotter, GSE30219, GSE50081, and GSE19188 databases. Overexpression or knockdown of SRCIN1 significantly induced or reduced A549 and H1299 cell proliferation. Furthermore, we found SRCIN1 was directly targeted by miR-211. Overexpression or knockdown of miR-211 suppressed or induced SRCIN1 levels in NSCLC. Moreover, we found that miR-211 affected NSCLC cell proliferation through SRCIN1. Previous studies demonstrated that circRNAs could act as miRNA sponges in cancer cells. In this study, we showed that knockdown of circCCDC66 induced expression of miR-211. Luciferase assay demonstrated that miR-211 suppressed the activity of luciferase reporter-contained circCCDC66 sequences. Moreover, knockdown of circCCDC66 significantly inhibited SRCIN1 levels in both A549 and H1299 cells. These results showed that circCCDC66 acted as a miRNA sponge to affect the miR-211/SRCIN1 axis. Of note, we for the first time revealed that circCCDC66 suppression reduced cell proliferation by about 65% in A549 and by about 40% in H1299 cells. We thought this study could provide novel potential biomarkers for NSCLC.

1. Background

As global cancer statistics showed in 2018, the incidence and mortality of lung cancer were extremely higher compared to other tumors [1]. Lung cancer is classified into NSCLC (non-small-cell lung cancer) and SCLC (small-cell lung cancer) [2]. Most lung cancer patients belong to NSCLC, which includes squamous cell carcinoma, adenocarcinoma, and large carcinoma [3]. Although medical imaging is a largely used strategy to screen lung cancer, an increasing number of NSCLCs can be diagnosed at the initial phase, and the case fatality rate for NSCLC is still high [4]. Due to the massive population base, the increasing number of smokers, and high health care costs in some countries, a majority of people are diagnosed with late-stage disease [5, 6]. Therefore, although

many treatment approaches have been developed for NSCLC, the therapeutic outcomes are not optimal [7]. Therefore, determining the underlying role of NSCLC and further uncovering the unknown indicator are really of importance, followed by ameliorating early diagnosis and providing promising treatment for NSCLC patients.

SRC Kinase Signaling Inhibitor 1 (SRCIN1) acts as a regulator for inhibiting cell spreading and migration and is also involved in calcium-dependent exocytosis. SRCIN1 is identified to be an inhibitor in lung cancer and breast cancer along with osteosarcoma [8]. For instance, Wang et al. showed that increased expression of SRCIN1 could result in reduced cell growth of osteosarcoma via promoting E-cadherin expression in vitro [9]. Chen et al. demonstrated that enhanced SRCIN1 repressed proliferation of human liver cancer cells

and blocked epithelial-mesenchymal transition in cell line HepG2 [10]. Although some researches have been carried out on SRCIN1, there is a very little scientific understanding of the roles of SRCIN1 in the development of tumors. It is therefore necessary to study the roles and functions of SRCIN1 in tumorigenesis and progression of NSCLC.

miRNAs are small noncoding RNAs containing less than 30 nucleotides. miRNAs are crucial posttranscriptional regulators in cells. Mature miRNAs target mRNA transcripts through complementary base-pairing to the 3' UTR, thus resulting in target mRNA degradation or translational inhibition. In this way, miRNAs act as a regulator widespread in cancer cells [11]. Therefore, changes in miRNA expression or miRNA imbalance may affect the cell cycle progression, which in turn affects the fate and behavior of tumor cells [12]. Previous research has shown that miRNAs played a carcinogenic or suppressive role in the progression of NSCLC. For example, highly expressed miR-10a would induce cell proliferation migration and invasion of NSCLC by targeting PTEN [13]. Jiang et al. revealed that miR-17, miR-20a, and miR-20b were regarded as inhibitors of TGF-beta receptor 2, thus rescuing cisplatin-resistant and retarding metastasis of NSCLC [14]. TRAIL, a TNF-related apoptosis-inducing ligand, is a novel molecular against tumors because it has the ability of selective inhibition of apoptosis in tumors in the absence of side effects on nearby normal cells [15]. The TRAIL-mediated suppressed tumor route is reduced in numerous tumors containing lung cancer [16]. Iaboni et al. demonstrated that overexpression of the tumor suppressor miR-212 could restore the effectiveness of TRAIL treatment by inhibiting PED/PEA-15 (antiapoptotic protein) in NSCLC cells [17]. The abovementioned studies showed that miRNAs played a key role in the progression and development of NSCLC. Therefore, it is important to explore molecular mechanism aspects to better understand NSCLC pathogenesis.

circRNA in eukaryotes was discovered in the 1970s by transmission electron microscopy [18]. However, we know very little about the structures and functions of circRNA before the maturity of high-throughput sequencing and bioinformatics. Numerous studies have found that circRNAs are abundant and have many functions in eukaryotic cells, including regulating parental gene level and functioning as microRNA (miRNA) sponges. Besides, some studies have demonstrated that circRNAs participated in the growth of cancer [19]. For example, Li et al. found that the hsa_circ_002059 level was obviously linked to distal metastasis, TNM stage, gender, and age in GC (gastric cancer) patients [20]. Yao et al. showed that higher expression of hsa_circ_100876 was positively related to the metastasis of the lymph node and carcinoma stage in NSCLC [21]. Zhong et al. indicated that the expression of circRNA-MYLK and circTCF25 was evidently enhanced in bladder cancer tissues [22]. All the data indicated that the abnormal expression of circRNAs could act as a newly produced indicator for tumor development.

Competitive endogenous RNA (ceRNA) regulation suggested a complex network of transcriptional RNAs including long noncoding RNAs (lncRNAs) and circRNA, which can act as natural miRNA sponges to inhibit miRNA functions and modulate mRNA expression [23, 24]. Here, we

attempted to investigate the functions and mechanisms of circCCDC66, SRCIN1, and miR-211 in NSCLC. Our results showed that decreased SRCIN1 could inhibit cell proliferation, migration, and invasion. Furthermore, luciferase assay showed that SRCIN1 was a direct target of miR-211, which was also sponged by circCCDC66. Collectively, our findings suggested several novel biomarkers for NSCLC.

2. Methods and Materials

2.1. Tissues. Twenty NSCLC tissues and twenty normal ones on average 5 cm from the indicated tumor of patients which was surgically removed were provided by Minhang Hospital, Fudan University. Among them, none were subjected to radiation and chemotherapy beforehand. All the experiments were approved by the corresponding Ethics Committee and unanimous consent by all subjects with signed informed documents. The tissues in this study were quickly put into liquid nitrogen after dissecting from patients, followed by preserving them in -80°C for long-term use.

2.2. Cells. H1299 and A549 (human NSCLC cells) were purchased from the ATCC (Manassas, USA) and then cultured in RPMI 1640 medium (Gibco, USA) containing 10% FBS (Gibco) and 1% penicillin/streptomycin under 37°C incubator with 5% CO_2 .

2.3. qRT-PCR (Quantitative Real-Time Reverse Transcription PCR). Trizol (TaKaLa, China) was used to extract total RNA from cells and tissues as described in the manual. Reverse transcription system was performed as follows: 10 μL volume included 500 ng RNA with Prime Script RT Master Mix (RiboBio, China) and RNAase free ddH_2O , followed by being subjected to qRT-PCR with indicated primers by a SYBR Master Mix (ABI, USA) on an ABI 7500 system (ABI). We finally calculated relative RNA expression by the $2^{-\Delta\Delta\text{Ct}}$ method.

2.4. Plasmid Construction and Transfection. All the oligonucleotides were synthesized by GenePharma (Shanghai, China). The designed siRNA targeting circCCDC66 (si-circCCDC66) was for the covalently closed junction. PcDNA3.1 vector expressing si-circCCDC66 ($5'$ -GAGCAU CAGGAAACAGUAC- $3'$) was constructed to ablate the circCCDC66 level. 50 pmol/mL of si-circCCDC66 and negative control (NC) inhibitor was separately transfected into indicated cells by Lipofectamine 2000 (Invitrogen) as per the manual's instructions, followed by changing the medium at 6 hours posttransfection.

2.5. CCK-8 Assay. Cell proliferation assay was conducted by a Cell Counting Kit (Dojindo, Japan) at 48 hours posttransfection. 1×10^4 cells of H1299 and A549 per well were reseeded in 96 wells and then 10 μL of CCK-8 solution was added at the indicated time for 2 hours of incubation. Cell proliferation was measured at specified days. Absorbance value in 450 nm was detected by an Infinite M200 plate reader (Tecan, Switzerland).

2.6. Transwell Assay. 8 μm pores of a transwell chamber (Costar, USA) and Matrigel film (BD Biosciences) used for coating upper chambers were successively applied to conduct an invasion assay. The treatment chamber including 100 μL of medium absence serum and the lower chamber with 600 μL of medium containing 5% FBS were seeded with 1×10^4 cells as indicated at 37°C with 5% CO_2 overnight. On the following day, nonimmigrated or noninvasive cells on the top side were eliminated by cotton swab. The insert was fixed by methanol for 20 minutes and then dyed by DAPI at concentration of 10 $\mu\text{g}/\text{mL}$ for 5 minutes. We counted migrated or invade cells to the membrane bottom and captured images by microscope from three independent experiments in triplicate.

2.7. Luciferase Assay. The circCCDC66 fragment containing mutated (mut) or wild-type (wt) seed region was inserted into the psiCHECK-2 construct (ABI). Wt or mut circCCDC66, miR-211 mimic, or mimic control was separately transfected into 1×10^5 cells per well of A549 and H1299 by Lipofectamine 2000 (Invitrogen). Relative luciferase activity was detected by a dual-luciferase reporter kit at 48-hour induction (Promega, USA). WT 3'UTR of SRCIN1 mRNA with the assumed binding site of miR-211 was ligated downstream of the firefly luciferase expression cassette in the pMIR-REPORT vector (Thermo Scientific). The positive clones were named by pMIR-SRCIN1-3'UTR (SRCIN1-wt). The pMIR-SRCIN1-3'UTR-mut (SRCIN1-mut) plasmid was generated by the QuikChange Mutagenesis kit (Stratagene, USA) which referred to the abovementioned. SRCIN1-wt or SRCIN1-mut with miR-211 mimics or NC mimics was then transfected into A549 and H1299 cells by Lipofectamine 2000.

2.8. Statistical Analysis. The representative data are shown as mean \pm SD. All the data was calculated after three independent experiments, followed by the limma package. The value of gene expression conformed to the normal distribution. The difference existing in two comparison groups or multiple groups in the indicated experiments was determined by Student's *t*-test. The differences between tumor and normal tissues were counted by paired-sample *t*-test. The linear relationship occurring in either two groups in NSCLC tissues was detected by the Pearson correlation coefficient. The obvious difference was indicated as *P* value less than 0.05.

3. Results

3.1. SRCIN1 Was Upregulated and Correlated to Shorter Survival Time in NSCLC. As presented in Figure 1(a), we observed that SRCIN1 was upregulated in both lung adenocarcinoma and lung squamous cell carcinoma samples compared to that in match normal tissues, suggesting that SRCIN1 was related to the tumorigenesis of NSCLC.

Also, we calculated the association between SRCIN1 levels and overall survival (OS) time by using TCGA data (Figure 1(b)). We found that NSCLC patients with higher SRCIN1 expression had shorter OS time. To further confirm this, we analyzed microarray data related to NSCLC. We also

observed that overexpression of SRCIN1 was related to shorter OS time in NSCLC by analyzing the Kaplan-Meier Plotter (Figure 1(c)), GSE30219 (Figure 1(d)), GSE50081 (Figure 1(e)), and GSE19188 databases (Figure 1(f)).

3.2. SRCIN1 Acted as an Oncogene in NSCLC Cells. We next validated the SRCIN1 role involved in A549 and H1299 cell metastases. Our findings revealed that upregulated SRCIN1 could result in a higher level of SRCIN1 in both A549 and H1299 cells (Figure 2(a)). Then, we aimed to assess the influences of SRCIN1 induced on cell proliferation. The CCK-8 assay was applied and showed that SRCIN1 would induce A549 and H1299 cell growth (Figures 2(b) and 2(c)). However, knockdown of SRCIN1 suppressed cell proliferation in A549 and H1299 (Figures 2(e) and 2(f)). These results showed that SRCIN1 was an oncogene in NSCLC.

3.3. SRCIN1 Was a Target of miR-211 in NSCLC Cells. Here, we wanted to predict and investigate the upstream miRNAs of SRCIN1 by bioinformatics. For the following studies, we selected miR-211 as the candidate miRNA that targeted SRCIN1. miR-211 was reported as a tumor suppressor in multiple cancers. Higher expression of miR-211 was related to longer OS time in NSCLC, especially in low mutation-burden NSCLC (Figures 3(a) and 3(b)), suggesting that miR-211 might be a tumor suppressor.

Then, we transfected miRNA mimics or inhibitors to modulate miR-211 ($P < 0.05$, Figure 3(c)). qRT-PCR showed that SRCIN1 was suppressed by the miR-211 and induced by the inhibitors (Figure 3(d)). Of note, a dual-luciferase reporter assay confirmed that SRCIN1 was a direct target of miR-211 (Figures 3(e) and 3(f)). Furthermore, we detected the effect of miR-211/SRCIN1 on cell growth in NSCLC cells (Figures 4(a) and 4(b)). Our results showed that miR-211 upregulation suppressed the A549 and H1299 proliferation rates, but miR-211 downregulation enhanced the A549 and H1299 proliferation rates. Moreover, we found that SRCIN1 overexpression rescued the suppressive effect of miR-211 overexpression on cell proliferation in both A549 and H1299.

3.4. circCCDC66 Targeted miR-211 and Affected SRCIN1 Expression. Previous studies demonstrated that circRNAs could act as miRNA sponges in cancer cells. The prediction showed that circCCDC66 targeted miR-211. The qRT-PCR assay showed that the knockdown of circCCDC66 suppressed endogenous levels of this circRNA and induced the expression of miR-211 (Figures 5(a) and 5(b)). Luciferase assay demonstrated that miR-211 suppressed the activity of luciferase reporter-contained circCCDC66 sequences (Figures 5(c) and 5(d)). Moreover, knockdown of circCCDC66 significantly inhibited SRCIN1 levels in both A549 and H1299 cells.

3.5. Knockdown circCCDC66 Suppresses Cell Proliferation in NSCLC. We moved forward to identify the functionality of circCCDC66 in NSCLC. Our data revealed that inhibited circCCDC66 reduced cell proliferation by about 65% in A549 and by about 40% in H1299 cells (Figures 5(f) and 5(g)). This finding showed that circCCDC66 acted as an oncogenetic circRNA in NSCLC.

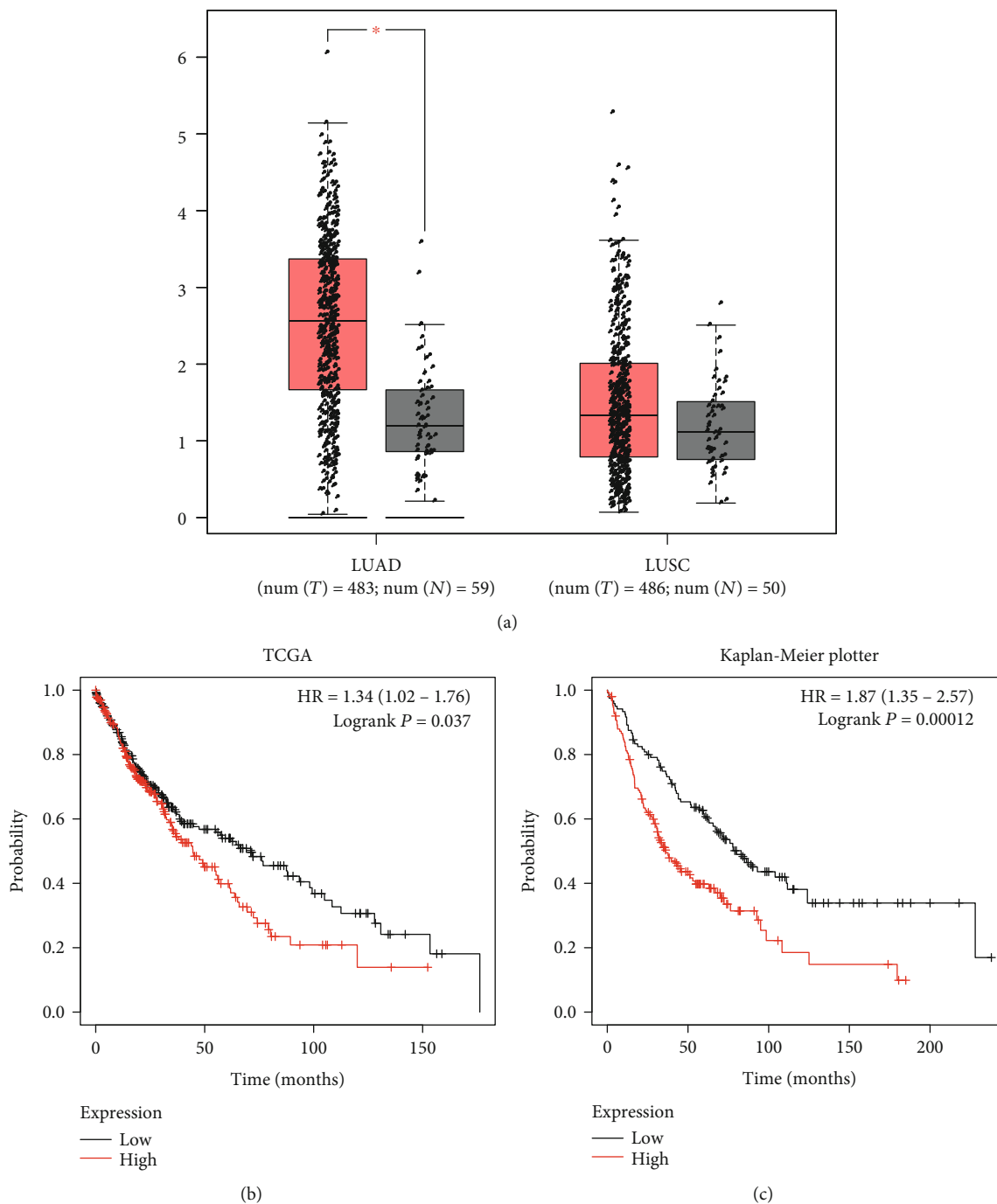


FIGURE 1: Continued.

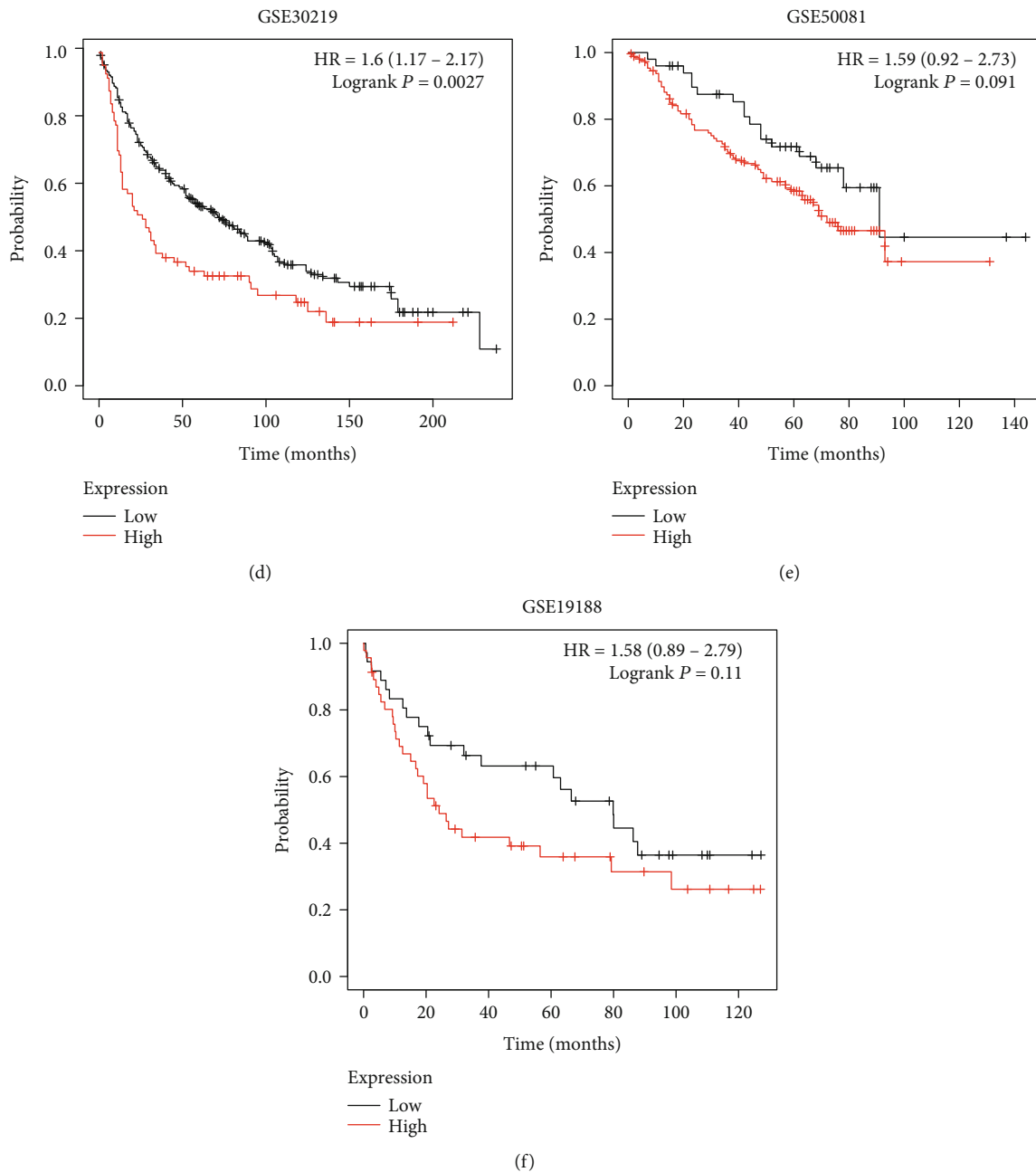


FIGURE 1: SRCIN1 was upregulated and correlated to shorter survival time in NSCLC. (a) SRCIN1 was upregulated in both lung adenocarcinoma and lung squamous cell carcinoma samples. (b) NSCLC patients with higher SRCIN1 expression had shorter overall survival (OS) time. (c-f) Overexpression of SRCIN1 was related to shorter OS time in NSCLC by analyzing the Kaplan-Meier Plotter (c), GSE30219 (d), GSE50081 (e), and GSE19188 databases (f). * $P < 0.05$.

4. Discussion

In our study, we firstly determined SRCIN1 expression and molecular functions. The data revealed that SRCIN1 increased in NSCLC tissues compared to that in normal ones. Higher expression levels of SRCIN1 correlated to shorter OS time in NSCLC patients. Secondly, we aimed to explore the association between SRCIN1 and circCCDC66 and their mechanism of involvement in tumor development. For that purpose, we performed cell viability detection. Our data

revealed that circCCDC66 knockdown and SRCIN1 silencing resulted in reduced abilities of NSCLC cell proliferation. The mechanism studies demonstrated that circCCDC66 sponged miR-211 to modulate SRCIN1 expression.

The molecular functions of SRCIN1 in cancer cells are controversial. According to previous reports, SRCIN1 could act as either an oncogene or a tumor suppressor. For example, Xu et al. revealed that SRCIN1 significantly inhibited gastric cancer cell viability, migration, and invasion [8]. Wang et al. demonstrated enhanced SRCIN1-repressed cell

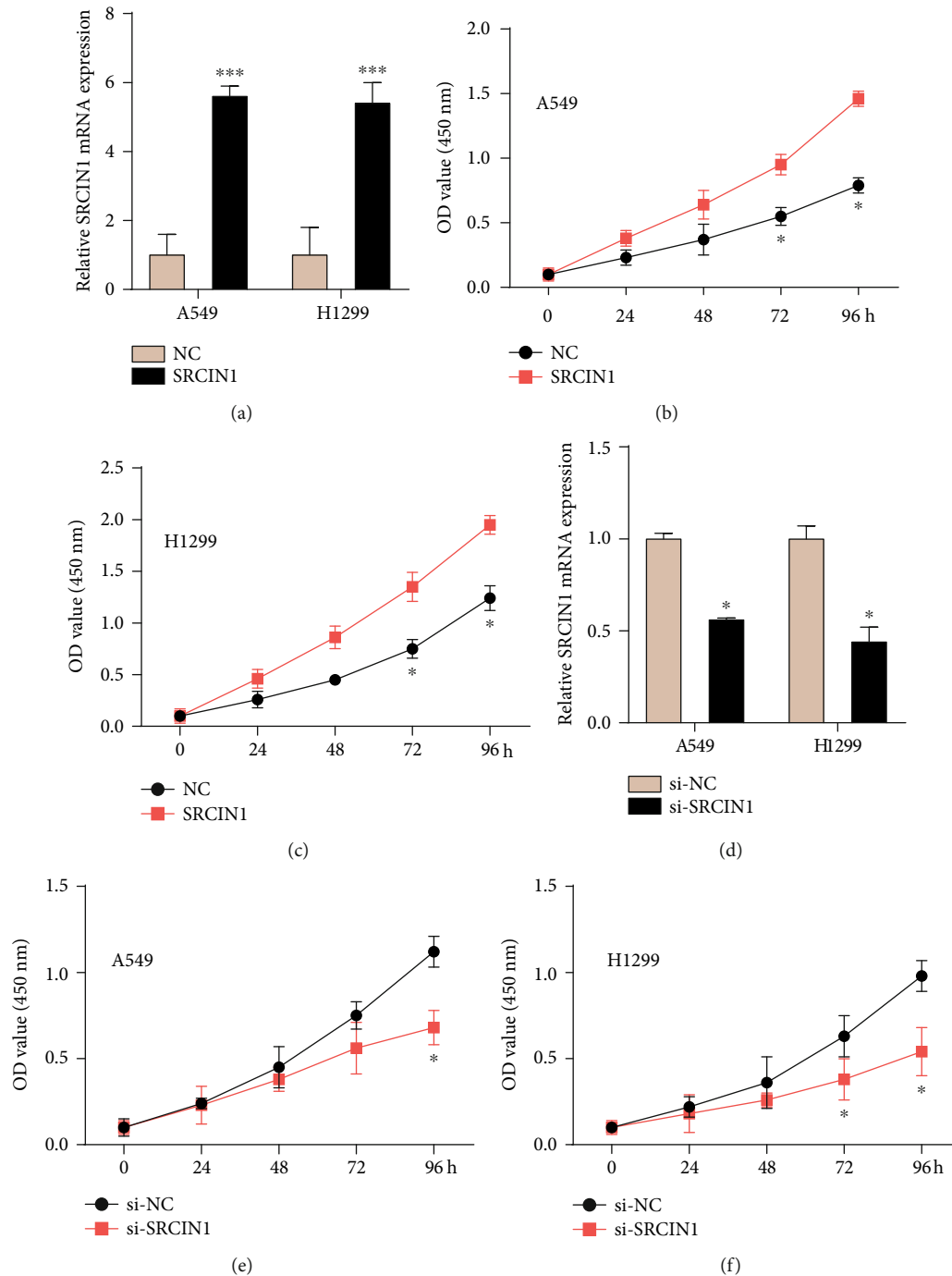


FIGURE 2: SRCIN1 acted as an oncogene in NSCLC cells. (a) Upregulated SRCIN1 could result in a higher level of SRCIN1. (b, c) Overexpressed SRCIN1 could induce A549 and H1299 cell growth. (e, f) Knockdown of SRCIN1 suppressed cell proliferation in A549 (e) and H1299 (f). * $P < 0.05$, *** $P < 0.001$.

viability and colony formation with the invasion of osteosarcoma [9]. Conversely, other studies confirmed that SRCIN1 could serve an oncogenic role. For instance, Zhang et al. revealed that SRCIN1 increased the capacity of colorectal cancer cell viability, migration, and invasion by activating Wnt/ β -catenin signaling routes [25]. The present study showed that SRCIN1 played as an oncogene in NSCLC. SRCIN1 was upregulated in NSCLC samples. Overexpres-

sion or knockdown of SRCIN1 significantly induced or reduced A549 and H1299 cell proliferation.

Emerging researches have discovered that miRNAs had an elementary function in carcinogenesis and NSCLC development [26–28]. For example, Gu et al. exhibited miR-940-mediated inhibition of the NSCLC cell proliferation by regulating its downstream target gene FAM83F [29]. Zhang et al. demonstrated that miRNA miR-21 promoted growth and

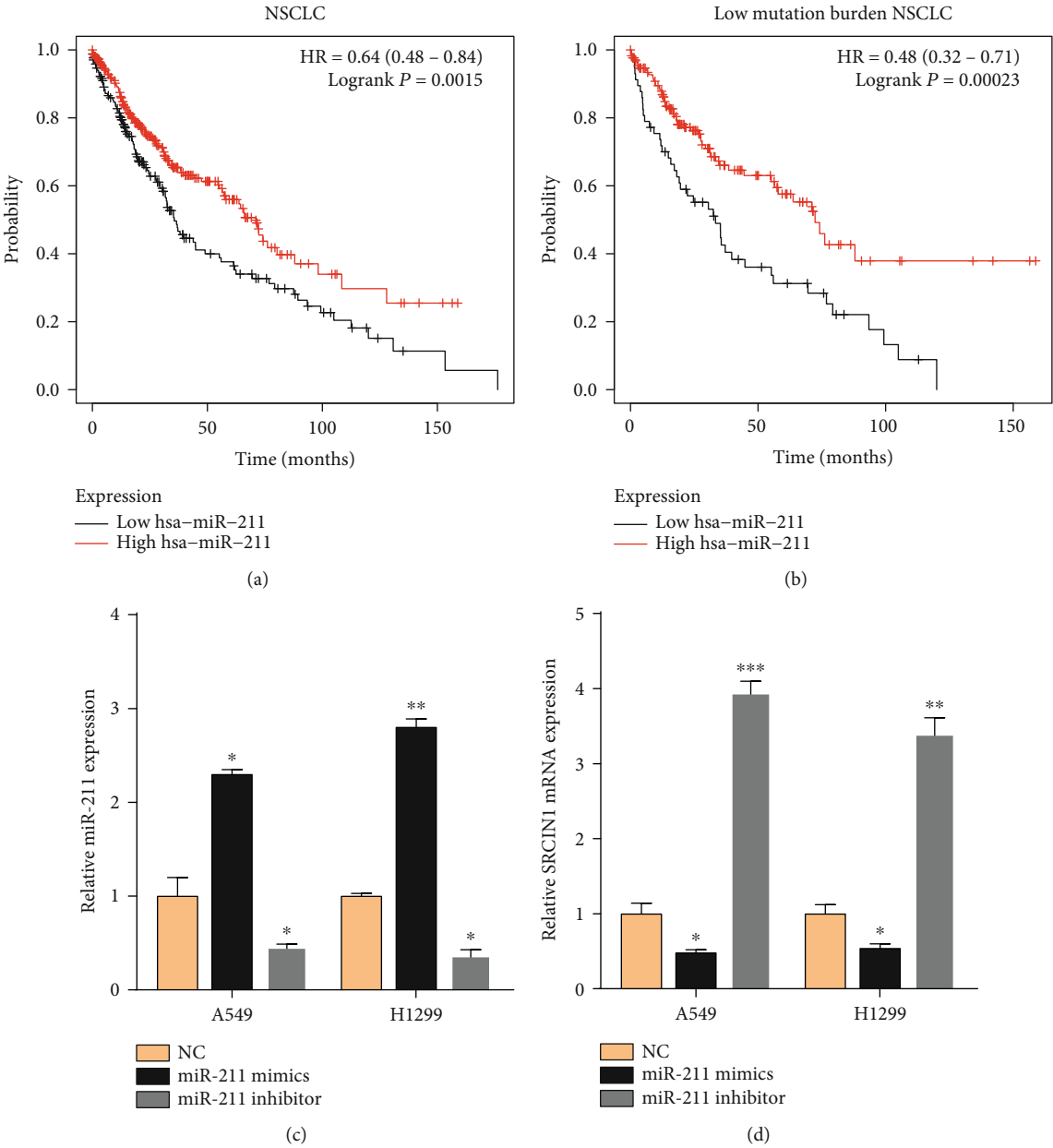


FIGURE 3: Continued.

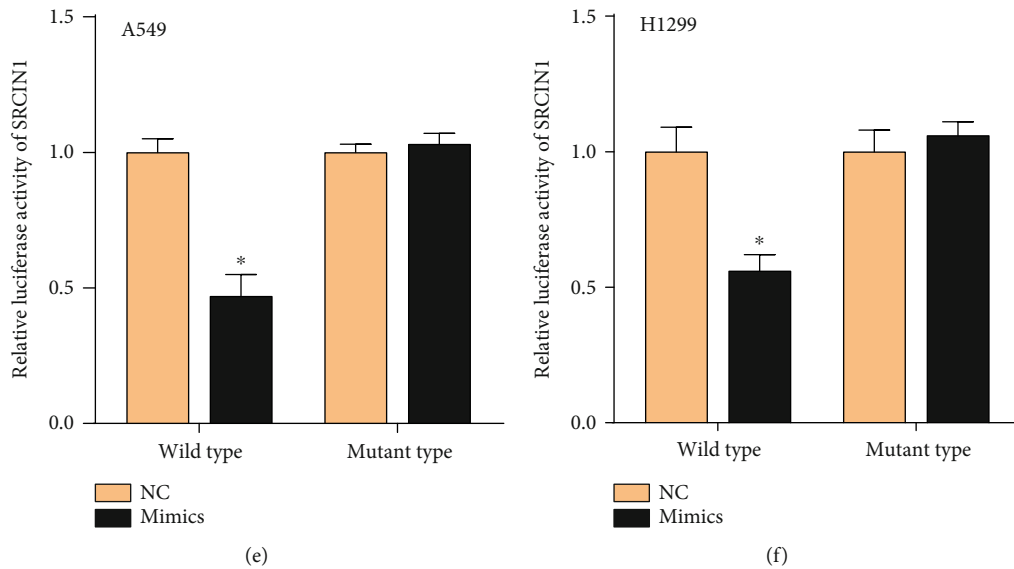


FIGURE 3: SRCIN1 was a target of miR-211 in NSCLC cells. (a, b) Higher expression of miR-211 was related to longer OS time in NSCLC, especially in low mutation-burden NSCLC. (c) Effects of transfected miRNA mimics or inhibitors on miR-211 levels in cells. (d) SRCIN1 was suppressed by the miR-211 mimics and induced by the inhibitors. (e, f) The cells cotransfected with miR-211 mimics and wild-type SRCIN1 showed lower luciferase activity. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

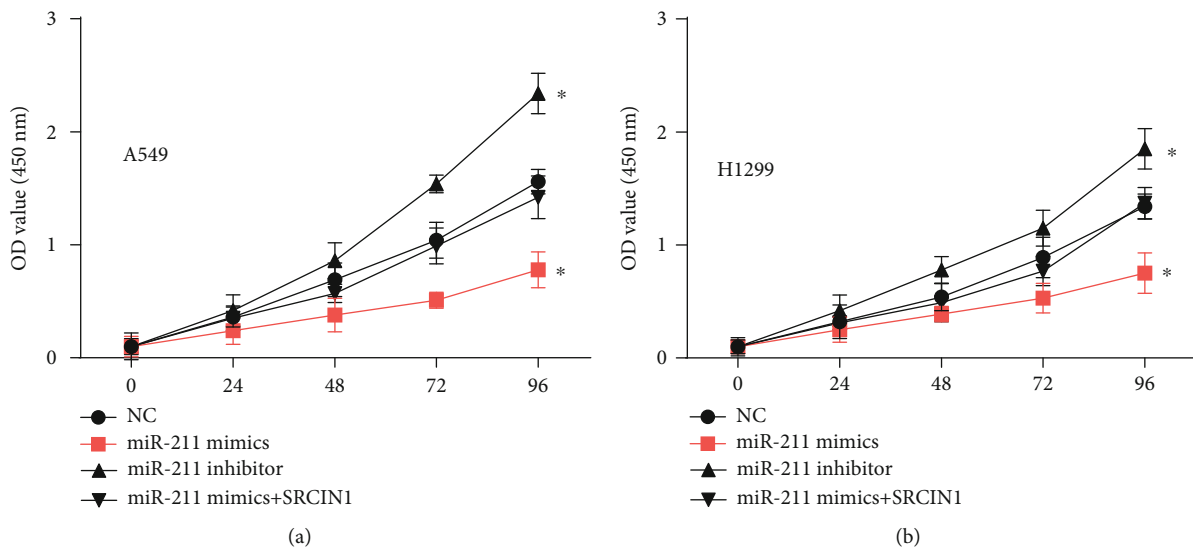


FIGURE 4: The effect of miR-211/SRCIN1 on cell growth in NSCLC cells. (a, b) In A549 and H1299, the upregulation of miR-211 inhibited the proliferation rate, while the downregulation of miR-211 increased the proliferation rate, and the overexpression of SRCIN1 weakened the inhibitory effect of miR-211 overexpression on cell proliferation. * $P < 0.05$.

invasion in NSCLC by controlling the expression of tumor suppressor PTEN [30]. Interestingly, Jiang et al. revealed that hsa-miR-125a-3p and hsa-miR-125a-5p had the same geographical origin but with the opposite function in NSCLC [31]. The former suppressed cell migration and invasion, while the latter enhanced tumor development. Our data suggested that the overexpression of miR-211 correlated to longer OS in NSCLC. Overexpression of miR-211 suppressed NSCLC cell viability. These functions indicate that miR-211 has potential as a new tumor marker for NSCLC.

The circRNA, a type of noncoding RNA, could extremely enrich microRNAs (miRNA) like a sponge [32]. They also could work on other kinds of RNAs by being base-paired [33]. In addition, circRNA could affect protein activity by binding with them [34]. Some researches show that circRNA can regulate the expression of host proteins, interact with RNA-binding proteins controlling transcription, play a role in transcriptional regulation in cis, and even regulate and control alternative splicing [35].

Our results showed that circCCDC66 promoted NSCLC cell proliferation in vitro. Besides, the concomitant

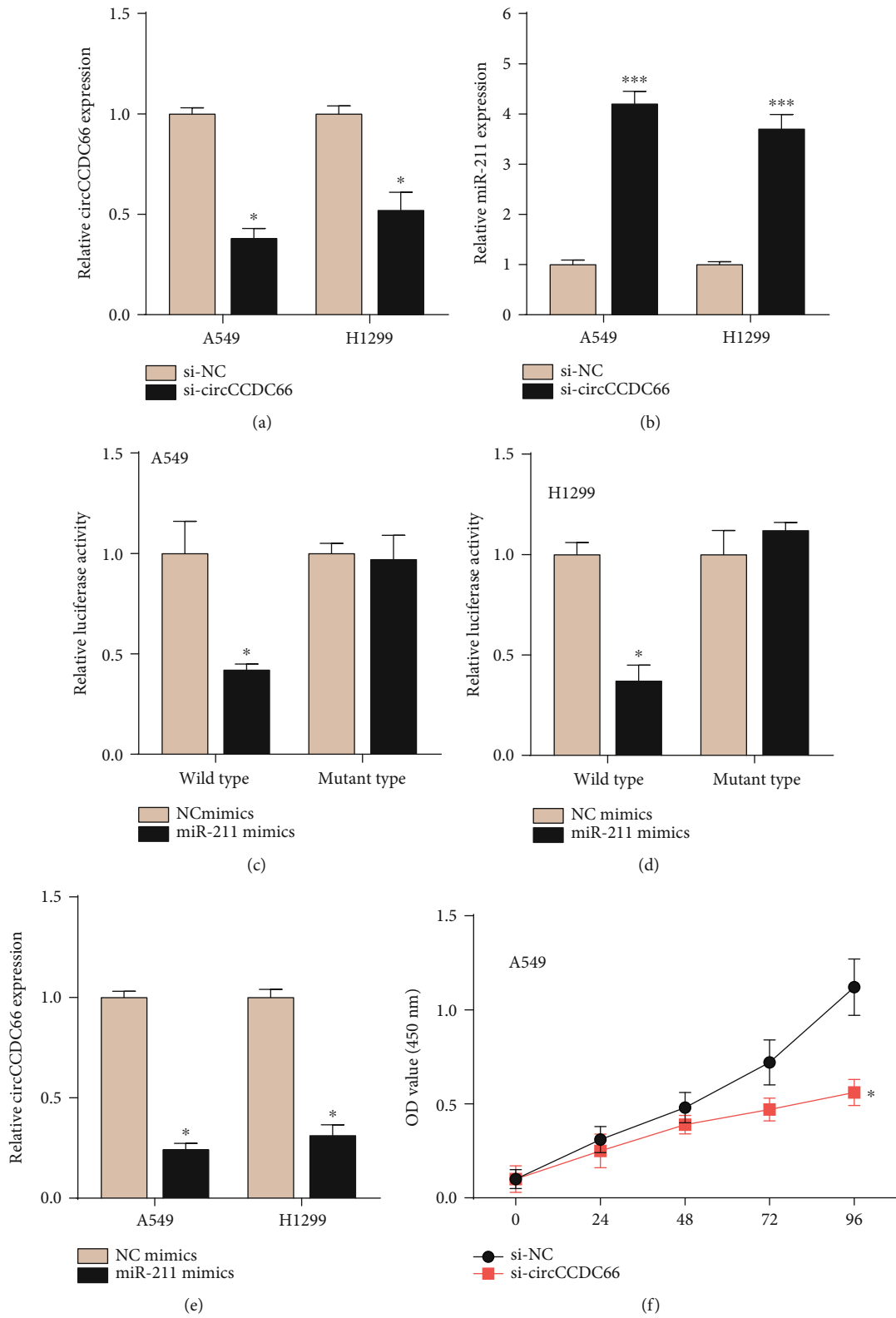


FIGURE 5: Continued.

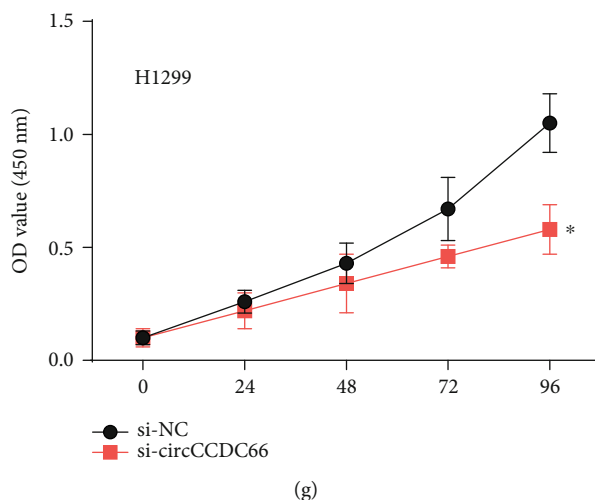


FIGURE 5: circCCDC66 targeted miR-211 and knockdown circCCDC66 suppresses cell proliferation in NSCLC. (a, b) Knockdown of circCCDC66 suppressed circCCDC66 levels and induced expression of miR-211. (c, d) miR-211 suppressed the luciferase activity of cells containing circCCDC66 sequences. (e) Overexpression of miR-211 in cells reduced circCCDC66 expression. (f, g) Inhibited circCCDC66 reduced cell proliferation. * $P < 0.05$, *** $P < 0.001$.

knockdown of circCCDC66 led to a significant increase in miR-211 expression and a decrease in SRCIN1 expression. Many studies implied that circRNAs were involved in many biological processes competing for endogenous RNA (ceRNA) of miRNA. For example, circ_104075 which was identified as ceRNA increased the YAP level through absorbing miR-582-3p which in turn affected cell growth in hepatocellular carcinoma [36]. Cheng et al. reported that circTP63 functioning as the ceRNA of miR-873-3p promoted lung cell proliferation by decreasing the level of FOXM1 [37]. In this study, circCCDC66 regulated the progression of NSCLC by functioning as a ceRNA-like competitive adsorbent to absorb miR-211 and thus control SRCIN1 expression.

More experimental verification of our findings in NSCLC is needed. Our work revealed that SRCIN1 and circCCDC66 are oncogenic growth factors for NSCLC. Meanwhile, miR-211-caused tumor suppression has been found in NSCLC. Mechanically, we showed that circCCDC66 acted as a miRNA sponge to affect the miR-211/SRCIN1 axis. Our findings provided a new viewpoint into the posttranscriptional regulation mechanism of SRCIN1 and suggested that circCCDC66, SRCIN1, and miR-211 might act as a potential diagnostic biomarker for NSCLC. This regulatory mechanism also helps us to explore the occurrence and development of NSCLC from the perspective of transcriptional regulatory pairs and clearly understand the important role of circRNA (circCCDC66) in this process. We will do our best to explore whether the regulatory axis plays a role in other cancers.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Weijun Hong and Suyun Yu contributed equally to this work.

Acknowledgments

This work is supported by the Shanghai Municipal Health Commission's Special Clinical Research Projects in the Health Industry (Grant Number 201940102) (applicant: Xiwen Gao); Industry, Education, Research and Medicine of Shanghai Science and Technology Commission (Grant Number 19DZ1920105) (applicant: Xiwen Gao); and Youth Fund of Minhang Hospital, Fudan University (Grant Number 2019MHJC08) (applicant: Weijun Hong).

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] B. P. Coe, W. W. Lockwood, L. Girard et al., "Differential disruption of cell cycle pathways in small cell and non-small cell lung cancer," *British Journal of Cancer*, vol. 94, no. 12, pp. 1927–1935, 2006.
- [3] G. Wolf, R. Elez, A. Doermer et al., "Prognostic significance of polo-like kinase (PLK) expression in non-small cell lung cancer," *Oncogene*, vol. 14, no. 5, pp. 543–549, 1997.
- [4] S. J. Swensen, J. R. Jett, T. E. Hartman et al., "Lung cancer screening with CT: Mayo Clinic experience," *Radiology*, vol. 226, no. 3, pp. 756–761, 2003.
- [5] J. She, P. Yang, Q. Hong, and C. Bai, "Lung cancer in China: challenges and interventions," *Chest*, vol. 143, no. 4, pp. 1117–1126, 2013.
- [6] R. Leidl, M. Wacker, and L. Schwarzkopf, "Better understanding of the health care costs of lung cancer and the

- implications," *Expert Review of Respiratory Medicine*, vol. 10, no. 4, pp. 373–375, 2016.
- [7] S. S. Ramalingam, T. K. Owonikoko, and F. R. Khuri, "Lung cancer: new biological insights and recent therapeutic advances," *CA: A Cancer Journal for Clinicians*, vol. 61, no. 2, pp. 91–112, 2011.
 - [8] X. Xu, W. Wang, N. Su et al., "miR-374a promotes cell proliferation, migration and invasion by targeting SRCIN1 in gastric cancer," *FEBS Letters*, vol. 589, no. 3, pp. 407–413, 2015.
 - [9] P. Wang, H. Wang, X. Li, Y. Liu, C. Zhao, and D. Zhu, "SRCIN1 suppressed osteosarcoma cell proliferation and invasion," *PLoS One*, vol. 11, no. 8, p. e0155518, 2016.
 - [10] R. Chen, J.-Y. Liao, J. Huang, W.-L. Chen, X.-J. Ma, and X.-D. Luo, "Downregulation of SRC kinase signaling inhibitor 1 (SRCIN1) expression by microRNA-32 promotes proliferation and epithelial-mesenchymal transition in human liver cancer cells," *Oncology Research*, vol. 26, no. 4, pp. 573–579, 2018.
 - [11] S. Gu, L. Jin, F. Zhang, P. Sarnow, and M. A. Kay, "Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs," *Nature Structural & Molecular Biology*, vol. 16, no. 2, pp. 144–150, 2009.
 - [12] R. Visone and C. M. Croce, "miRNAs and cancer," *The American Journal of Pathology*, vol. 174, no. 4, pp. 1131–1138, 2009.
 - [13] T. Yu, L. Liu, J. Li et al., "miRNA-10a is upregulated in NSCLC and may promote cancer by targeting PTEN," *Oncotarget*, vol. 6, no. 30, pp. 30239–30250, 2015.
 - [14] Z. Jiang, J. Yin, W. Fu et al., "miRNA 17 family regulates cisplatin-resistant and metastasis by targeting TGFbeta2 in NSCLC," *PLoS One*, vol. 9, no. 4, p. e94639, 2014.
 - [15] E. Cretney, A. Shanker, H. Yagita, M. J. Smyth, and T. J. Sayers, "TNF-related apoptosis-inducing ligand as a therapeutic agent in autoimmunity and cancer," *Immunology & Cell Biology*, vol. 84, no. 1, pp. 87–98, 2006.
 - [16] F. Yang, P. Shi, X. Xi et al., "Recombinant adenoviruses expressing TRAIL demonstrate antitumor effects on non-small cell lung cancer (NSCLC)," *Medical Oncology*, vol. 23, no. 2, pp. 191–204, 2006.
 - [17] M. Iaboni, V. Russo, R. Fontanella et al., "Aptamer-miRNA-212 conjugate sensitizes NSCLC cells to TRAIL," *Molecular Therapy - Nucleic Acids*, vol. 5, article e289, 2016.
 - [18] H. L. Sanger, G. Klotz, D. Riesner, H. J. Gross, and A. K. Kleinschmidt, "Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 73, no. 11, pp. 3852–3856, 1976.
 - [19] S. Meng, H. Zhou, Z. Feng et al., "CircRNA: functions and properties of a novel potential biomarker for cancer," *Molecular Cancer*, vol. 16, no. 1, p. 94, 2017.
 - [20] P. Li, S. Chen, H. Chen et al., "Using circular RNA as a novel type of biomarker in the screening of gastric cancer," *Clinica Chimica Acta*, vol. 444, pp. 132–136, 2015.
 - [21] J.-T. Yao, S.-H. Zhao, Q.-P. Liu et al., "Over-expression of circRNA_100876 in non-small cell lung cancer and its prognostic value," *Pathology - Research and Practice*, vol. 213, no. 5, pp. 453–456, 2017.
 - [22] Z. Zhong, M. Lv, and J. Chen, "Screening differential circular RNA expression profiles reveals the regulatory role of circTCF25-miR-103a-3p/miR-107-CDK6 pathway in bladder carcinoma," *Scientific Reports*, vol. 6, no. 1, p. 30919, 2016.
 - [23] J. Liu, H. Li, B. Zheng, L. Sun, Y. Yuan, and C. Xing, "Competitive endogenous RNA (ceRNA) regulation network of lncRNA-miRNA-mRNA in colorectal carcinogenesis," *Digestive Diseases and Sciences*, vol. 64, no. 7, pp. 1868–1877, 2019.
 - [24] J. Xu, Y. Li, J. Lu et al., "The mRNA related ceRNA-ceRNA landscape and significance across 20 major cancer types," *Nucleic Acids Research*, vol. 43, no. 17, pp. 8169–8182, 2015.
 - [25] M. Zhang, F. Ma, R. Xie et al., "Overexpression of Srcin1 contributes to the growth and metastasis of colorectal cancer," *International Journal of Oncology*, vol. 50, no. 5, pp. 1555–1566, 2017.
 - [26] Q. Wang, S. Jiang, A. Song et al., "HOXD-AS1 functions as an oncogenic ceRNA to promote NSCLC cell progression by sequestering miR-147a," *Oncotargets and Therapy*, vol. 10, pp. 4753–4763, 2017.
 - [27] G. Liu, H. Shi, L. Deng et al., "Circular RNA circ-FOXM1 facilitates cell progression as ceRNA to target PDPDF and MACC1 by sponging miR-1304-5p in non-small cell lung cancer," *Biochemical & Biophysical Research Communications*, vol. 513, no. 1, pp. 207–212, 2019.
 - [28] X. Chen, Z. Wang, F. Tong, X. Dong, G. Wu, and R. Zhang, "lncRNA UCA1 promotes gefitinib resistance as a ceRNA to target FOSL2 by sponging miR-143 in non-small cell lung cancer," *Molecular Therapy-Nucleic Acids*, vol. 19, pp. 643–653, 2020.
 - [29] G.-M. Gu, Y.-Y. Zhan, K. Abuduwaili et al., "miR-940 inhibits the progression of NSCLC by targeting FAM83F," *European Review for Medical and Pharmacological Sciences*, vol. 22, no. 18, pp. 5964–5971, 2018.
 - [30] J.-g. Zhang, J.-j. Wang, F. Zhao, Q. Liu, K. Jiang, and G.-h. Yang, "MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC)," *Clinica Chimica Acta*, vol. 411, no. 11–12, pp. 846–852, 2010.
 - [31] L. Jiang, Q. Huang, S. Zhang et al., "hsa-miR-125a-3p and hsa-miR-125a-5p are downregulated in non-small cell lung cancer and have inverse effects on invasion and migration of lung cancer cells," *BMC Cancer*, vol. 10, no. 1, p. 318, 2010.
 - [32] H. H. Kim, Y. Kuwano, S. Srikantan, E. K. Lee, J. L. Martindale, and M. Gorospe, "HuR recruits let-7/RISC to repress c-Myc expression," *Genes & Development*, vol. 23, no. 15, pp. 1743–1748, 2009.
 - [33] R. J. Webster, K. M. Giles, K. J. Price, P. M. Zhang, J. S. Mattick, and P. J. Leedman, "Regulation of epidermal growth factor receptor signaling in human cancer cells by microRNA-7," *The Journal of Biological Chemistry*, vol. 284, no. 9, pp. 5731–5741, 2009.
 - [34] W. W. Du, W. Yang, E. Liu, Z. Yang, P. Dhaliwal, and B. B. Yang, "Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2," *Nucl Acids Research*, vol. 44, no. 6, pp. 2846–2858, 2016.
 - [35] Y. Zhong, Y. Du, X. Yang et al., "Circular RNAs function as ceRNAs to regulate and control human cancer progression," *Molecular Cancer*, vol. 17, no. 1, p. 79, 2018.
 - [36] X. Zhang, Y. Xu, Z. Qian et al., "circRNA_104075 stimulates YAP-dependent tumorigenesis through the regulation of HNF4a and may serve as a diagnostic marker in hepatocellular carcinoma," *Cell Death & Disease*, vol. 9, no. 11, p. 1091, 2018.
 - [37] Z. Cheng, C. Yu, S. Cui et al., "circTP63 functions as a ceRNA to promote lung squamous cell carcinoma progression by upregulating FOXM1," *Nature Communications*, vol. 10, no. 1, p. 3200, 2019.

Research Article

A Novel RNA-Seq-Based Model for Preoperative Prediction of Lymph Node Metastasis in Oral Squamous Cell Carcinoma

Bo Qiao ^{1,2}, Min Zhao,³ Jing Wu,³ Huan Wu,³ Yiming Zhao,⁴ Fanhao Meng,² Yu Tian,² Situo Wang,² Jinlong Shi ³ and Haizhong Zhang ²

¹Medical School of Chinese PLA, Beijing 100853, China

²Department of Stomatology, The First Medical Centre, Chinese PLA General Hospital, Beijing 100853, China

³Medical Big Data Centre, Chinese PLA General Hospital, Beijing 100853, China

⁴Department of Gastroenterology and Hepatology, Hainan Hospital of Chinese PLA General Hospital, Sanya, 572013 Hainan, China

Correspondence should be addressed to Jinlong Shi; jinlong_301@163.com and Haizhong Zhang; zhanghaizhong301@aliyun.com

Received 3 July 2020; Revised 3 August 2020; Accepted 5 August 2020; Published 31 August 2020

Guest Editor: Tao Huang

Copyright © 2020 Bo Qiao 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.

Objective. To develop and validate a novel RNA-seq-based nomogram for preoperative prediction of lymph node metastasis (LNM) for patients with oral squamous cell carcinoma (OSCC). **Methods.** RNA-seq data for 276 OSCC patients (including 157 samples with LNM and 119 without LNM) were downloaded from TCGA database. Differential expression analysis was performed between LNM and non-LNM of OSCC. These samples were divided into a training set and a test set by a ratio of 9:1 while the relative proportion of the non-LNM and LNM groups was kept balanced within each dataset. Based on clinical features and seven candidate RNAs, we established a prediction model of LNM for OSCC using logistic regression analysis. Tenfold crossvalidation was utilized to examine the accuracy of the nomogram. Decision curve analysis was performed to evaluate the clinical utility of the nomogram. **Results.** A total of 139 differentially expressed RNAs were identified between LNM and non-LNM of OSCC. Seven candidate RNAs were screened based on FPKM values, including NEURL1, AL162581.1 (miscRNA), AP002336.2 (lncRNA), CCBE1, CORO6, RDH12, and AC129492.6 (pseudogene). Logistic regression analysis revealed that the clinical N stage ($p < 0.001$) was an important factor to predict LNM. Moreover, three RNAs including RDH12 (p value < 0.05), CCBE1 (p value < 0.01), and AL162581.1 (p value < 0.05) could be predictive biomarkers for LNM in OSCC patients. The average accuracy rate of the model was 0.7661, indicating a good performance of the model. **Conclusion.** Our findings constructed an RNA-seq-based nomogram combined with clinicopathology, which could potentially provide clinicians with a useful tool for preoperative prediction of LNM and be tailored for individualized therapy in patients with OSCC.

1. Introduction

Oral squamous cell carcinoma (OSCC) accounts for 95% of all oral malignancies, and its five-year survival rate is up to 50%-60% [1]. Lymph node metastasis (LNM) is considered to be an independent prognostic factor of OSCC, which is associated with tumor recurrence and prognosis [2–4]. Only 25-40% of OSCC patients with LNM at diagnosis will survive 5 years, compared to approximately 90% of those without LNM (non-LNM) [5]. Therefore, accurate assessment of the nodal status and decision about concurrent cervical lymph node dissection is of utmost importance for prognosis and therapy of OSCC.

Unfortunately, there is still no widely accepted method for noninvasive detection for preoperative prediction of LNM in OSCC currently. For OSCC patients with clinically negative neck (cN0), whether to perform cervical lymph node dissection remains a hot topic. Using current methods to predict LNM, approximately 70% of patients with cN0 OSCC who undergo elective neck dissection (END) are found to be pathologically node negative [6]. It is imminent to best select patients with LNM who will benefit from END and to decrease the cost and morbidity of neck dissection in those without LNM [7].

Emerging sequencing technologies in genomics and transcriptomics have characterized many types of human cancers

in specific molecules, which provide a critical relationship between cell phenotypes and their molecular characteristics, and new biomarkers or therapeutic strategies for patients [8, 9]. To our knowledge, no nomograms based on RNA-seq have been used to predict LNM in OSCC. We hypothesized that RNA-seq-based nomograms could improve the prediction of LNM in OSCC, so that patients who will benefit from END can be more accurately identified, while retaining END is unlikely to be beneficial. In this study, logistic regression analysis was used to screen for high risk factors for OSCC patients with LNM. We aimed to establish and verify a novel RNA-seq-based nomogram combined with clinicopathological factors to predict LNM in OSCC patients, which may provide an auxiliary tool for personalized precise treatment and assist the clinical therapy decision for OSCC patients.

2. Materials and Methods

2.1. Data Acquisition. OSCC RNA-seq were obtained from The Cancer Genome Atlas database (TCGA) (<https://portal.gdc.cancer.gov/>). Screening conditions were as follows: (a) primary sites: hard palate, lip, oral cavity, or oral tongue; (b) disease type: squamous cell neoplasms; (c) experimental strategy: RNA-seq; (d) sample type: primary tumor; and (e) clinical information was composed of AJCC pathological N status. Finally, a total of 276 OSCC samples with gene expression data and corresponding clinical information were utilized for this study, including 157 samples with lymph node metastasis (LNM) and 119 samples without LNM (non-LNM).

The GSE9844 microarray dataset was downloaded from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/gds/>), including 26 oral tongue squamous cell carcinoma (OTSCC) samples [10]. Among them, 11 samples had lymph node metastasis.

2.2. Differential Expression Analyses. Differential expression analyses were performed between LNM and non-LNM OSCC samples using the DESeq2 package in R (version 1.18.1) [11]. Differentially expressed RNAs (DERNAs) were identified in line with adjusted p value < 0.05 and $|\log_2$ fold change (FC) > 1 . The overall distribution of DERNAs was visualized into the volcano plot. A functionally grouped network of pathways was constructed using the ClueGO (version 2.5.1) [12, 13] of Cytoscape (version 3.6.1) [14]. The “load marker list” was set to differential gene sets for enrichment analysis, the “show only pathways with p value” was set to 0.05, and other settings were set to default.

2.3. Variable Selection Method. The chi-square (χ^2) test was used to analyze the difference of patient demography, risk exposure, clinical characteristics, and histopathological features between LNM and non-LNM OSCC samples by SPSS software (version 24.0). $p < 0.001$ was considered statistically significant. Characteristic genes related to LNM were screened utilizing the Boruta package in R [15] (version 6.0.0) based on FPKM values from the expression profile of

the GSE9844 dataset, which were plotted into a box plot by the ggplot2 package (version 3.2.1).

2.4. Classification Model Fitting and Validation. After removing the samples with incomplete clinical N stage (cN) information, a total of 265 samples were retained and randomly divided into the training set and the test set by a ratio of 9:1 while the relative proportion of non-LNM and LNM groups was kept balanced within each dataset. A classification model of the 10-fold crossvalidation was constructed utilizing the R language (version 3.4.4). In a training set, a RF model was built using the “random forest” package in R (version 4.6-14) [16], followed by support vector machine (SVM) model by the “e1071” package in R (version 1.7-2) [17]. Basic function “glm” in R (version 3.4.4) was used to fit the generalized logistic regression model, and the “XGBoost” package in R (version 0.90.0.2) [18] was utilized to implement the XGBoost model. The accuracy, sensitivity, and specificity of the four models were evaluated on the test set according to the calibration curve and ROC curve by the “rms” package (version 5.1-3.1) [19] and “pROC” package (version 1.15.3) in R. Area under the curve (AUC) was used to compare the diagnostic performance of the models. Furthermore, the nomogram of the logistic regression model was performed by the “regplot” package in R (version 0.2) [20]. Then, decision curve analysis (DCA) was conducted to estimate the clinical value of our nomogram between two groups using the “rmda” package (version 1.6) [21], which could analyze the net benefit of the cN and RNA scores for prediction of LNM for OSCC patients.

3. Results

3.1. Clinical Characteristics. Our study developed a model for the preoperative prediction of LNM. Figure 1 illustrates the work flowchart of the study. A total of 276 patients with OSCC were included in our study. Table 1 shows the demographic data and pathological characteristics of these patients. Among them, 157 (56.88%) patients had LNM. 265 patients possessed complete clinical N status; among them, 44 had no lymph node metastasis in clinical examination (cN-) but metastasis occurred in the pathological diagnosis (pN+). 22 patients' clinical examination indicated lymph node metastasis (cN+), but pathological examination results showed no lymph node metastasis (pN-). Furthermore, our data showed that clinical N was significantly different between the two cohorts ($p < 0.001$). However, there were no significant differences between the two groups in terms of neoplasm histologic grade, tobacco smoking history, and anatomic neoplasm subdivision.

3.2. Identification of Differentially Expressed RNAs between Non-LNM and LNM of OSCC. Volcano plots visualized that there were 139 DERNAs between non-LNM and LNM OSCC (Figure 2), including 104 upregulated and 35 downregulated genes. The specific DERNAs are listed in Supplementary Table 1. To analyze the underlying biological function of DERNAs, functional enrichment analysis was carried out by ClueGO and the database

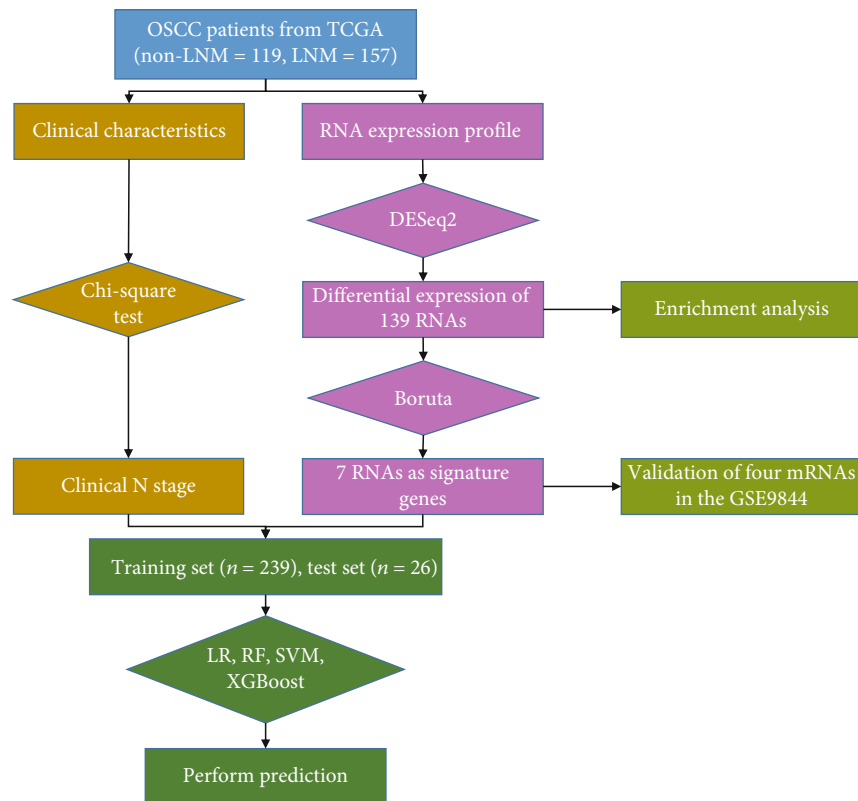


FIGURE 1: A work flowchart for this study. TCGA: The Cancer Genome Atlas; non-LNM: non lymph node metastasis; LNM: lymph node metastasis; LR: logistic regression; RF: random forest; SVM: support vector machine.

called by the GO biological process in ClueGO. Significant biological processes enriched by DERNA are shown in Figure 3. These DERNA were mainly involved in “regulation of striated muscle contraction,” “regulation of muscle system process,” and “muscle filament sliding.” These results indicated that a variety of biological processes of muscle could be involved in lymph node metastasis of OSCC.

3.3. Selection of Candidate DERNA to Predict LNM of OSCC.

The Boruta algorithm was used to screen out the signature genes to distinguish non-LNM and LNM of OSCC. Figure 4 shows the change in Z-score of 139 DERNA during Boruta’s operation. The blue boxplots indicated the minimum, average, and maximum Z-scores of the shadow gene. The red and green boxplots indicated the Z-score of the rejection attribute and confirmation genes, and yellow colors represented suggestive genes. These findings showed that the most important shadow attribute Z-score clearly separated important and nonimportant genes. Finally, a total of seven candidate DERNA were identified for distinguishing non-LNM and LNM of OSCC, including NEURL1, AL162581.1 (miscRNA), AP002336.2 (lncRNA), CCBE1, CORO6, RDH12, and AC129492.6 (pseudogene). Among them, expression profiles of four DERNA were obtained from the expression profile of the GSE9844 dataset. In Figure 5, we found that the expression levels of CCBE1, CORO6,

NEURL1, and RDH12 were significantly higher in N+ compared to N- OSCC patients.

3.4. Development and Validation of a Machine Learning Model to Predict LNM of OSCC.

Based on clinical N and seven candidate RNAs, we established a prediction model for LNM of OSCC. The machine learning was carried out, including LR, SVM, RF, and XGBoost. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy area under the curve (AUC) mean value, and 95% confidence interval (CI) are shown in Table 2. From the result, the average of the accuracy rate of the machine learning model was 0.79 and the AUC value was 0.84, indicating that the model had optimal performance. Based on the machine learning model, the receiver operating curves (ROCs) were depicted in the training set, test set, and entire data (Figure 6). The AUCs were 0.9773, 0.8441, and 0.8558, respectively, suggesting the good prediction efficiency of the model. As shown in Figure 7, a nomogram was established to predict the risk for LNM in OSCC. Logistic regression analysis revealed that the clinical N stage ($p < 0.001$) was an important factor to predict LNM of OSCC. Furthermore, three DERNA including RDH12 (p value < 0.05), CCBE1 (p value < 0.01), and AL162581.1 (p value < 0.05) possessed potential value as predictive biomarkers for LNM in OSCC patients.

The following two logistic regression models were established, with cN as the predictor and pN as the outcome; the

TABLE 1: The demographic characteristics of OSCC patients with or without lymph node metastasis.

| Characteristics | OSCC without LNM (<i>n</i> = 119) | OSCC with LNM (<i>n</i> = 157) | <i>p</i> value |
|-------------------------------|------------------------------------|---------------------------------|----------------|
| Gender | | | 0.017* |
| Male | 74 | 119 | |
| Female | 45 | 38 | |
| Age (years) | 62.92 ± 12.91 | 59.89 ± 12.31 | 0.049* |
| Clinical T | | | 0.017* |
| I-II | 53 | 44 | |
| III-IV | 63 | 108 | |
| Clinical N | | | <0.001*** |
| LN-negative | 94 | 43 | |
| LN-positive | 21 | 107 | |
| Neoplasm histologic grade | | | 0.195 |
| G1 | 23 | 17 | |
| G2 | 77 | 104 | |
| G3 | 18 | 31 | |
| G4 | 0 | 1 | |
| Tobacco smoking history | | | 0.267 |
| No | 36 | 34 | |
| Yes | 81 | 120 | |
| Margin status | | | 0.148 |
| Close | 15 | 19 | |
| Negative | 90 | 108 | |
| Positive | 9 | 26 | |
| Alcohol history | | | 0.024* |
| No | 43 | 46 | |
| Yes | 70 | 110 | |
| Anatomic neoplasm subdivision | | | 0.5 |
| Oral tongue | 49 | 63 | |
| Base of tongue | 4 | 9 | |
| Floor of mouth | 23 | 33 | |
| Buccal mucosa | 8 | 12 | |
| Alveolar ridge | 11 | 5 | |
| Oral cavity | 21 | 32 | |
| Hard palate | 1 | 2 | |
| Lip | 2 | 1 | |

Note: **p* value < 0.05, ****p* value < 0.001. OSCC: oral squamous cell carcinoma; LNM: lymph node metastasis.

other was a multiple logistic regression model (complex), with cN and characteristic RNA expression levels as predictors and pN as the outcome. In Figure 8(a), a decision curve showed that using the RNA nomogram in the current study to predict the LNM added was more beneficial than only using the clinical N stage. As shown in Figures 8(b) and 8(c), our clinical impact curves draw the clinical influence curves of the cN model and the complex model, respectively. The simple model to predict the risk stratification of 100 patients displayed the “cost:benefit ratio” *y*-axis, assigned 8 scales, and showed the confidence interval. The red curve (number of high risk) represented the number of people who were classified as positive (high risk) by the cN model

or the complex model at each threshold probability, and the blue curve (number of high risk with outcome) was the number of true positives at each threshold probability.

4. Discussion

LNM is the main reason for the failure of OSCC treatment, which will significantly reduce patients' survival rate [22, 23]. Once cervical LNM is detected during follow-up, cervical lymph node dissection does not always save the patient, and sometimes, the rescue rate is less than 40% [24]. At the same time, 60%-80% of END patients are evaluated as cN0 and no metastasis, but unnecessary

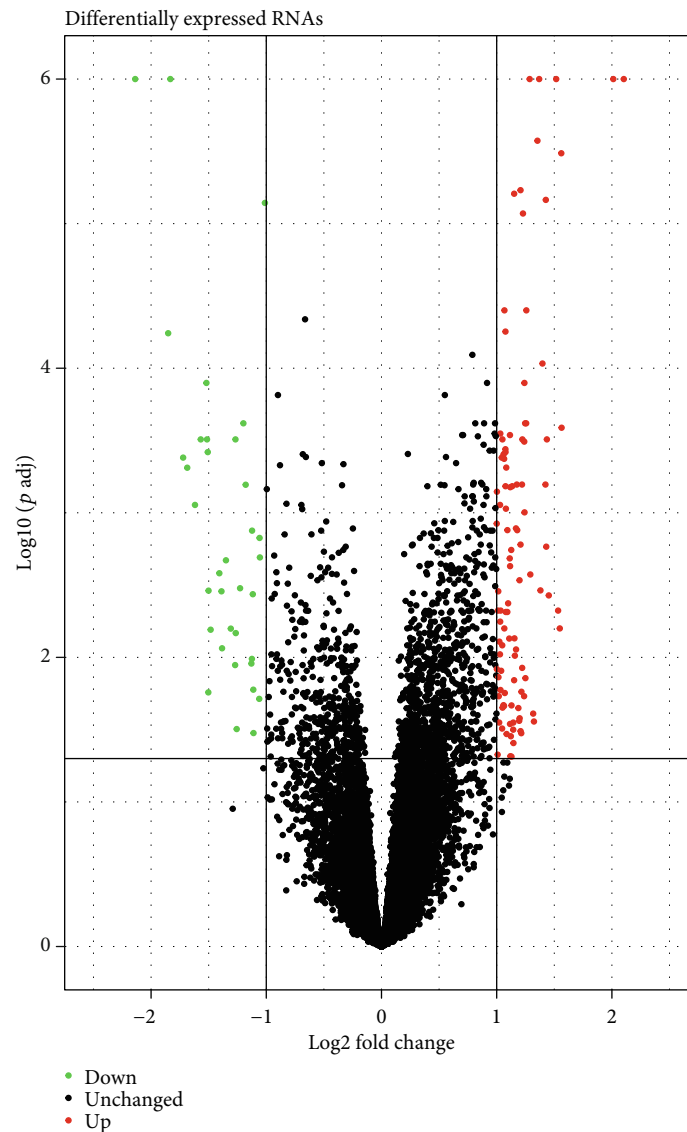


FIGURE 2: 139 differentially expressed RNAs between non-LNM and LNM of OSCC. Red dot represents up-regulation and green dot represents down-regulation.

cervical lymphadenectomy is required, which causes shoulder and neck pain and dysfunction, thereby affecting long-term quality of life (QOL) [25, 26]. Therefore, there is an urgent need to identify patients who can obtain greater benefits from END and to avoid unnecessary LN-related surgery in patients without LNM, especially in relatively early-stage T patients. The limitation of diagnostic imaging technology (including ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI)) is that LN status cannot be fully predicted. To date, many efforts have been made to develop diagnostic biomarkers for LNM in OCSS patients. However, most of these markers are limited by their detection potential.

Recent developments in technology of whole-transcriptome sequencing provide a possibility to develop new biomarkers and therapeutic strategies in most types of human cancers [8, 9]. In this study, we identified 139 DERNAs between non-LNM and LNM OSCC. These

DEmRNAs were distinctly involved in several key pathways such as “regulation of striated muscle contraction,” “regulation of muscle system process,” and “muscle filament sliding,” indicating that these genes could be involved in LNM of OSCC. Based on DEGs, Sonohara et al. proposed a novel gene-expression signature for prediction of lymph node metastasis in esophageal squamous cell carcinoma (ESCC) patients by analyzing RNA-sequencing profiles [27]. Daisuke et al. constructed a 15-gene signature for detection of lymph node metastasis in early-stage gastric cancer (GC) patients by using the genome-wide transcriptomic approach [28]. With regard to OSCC, Pasini et al. developed a four-gene expression model to predict LNM in OSCC, but their results showed a 22% false positive rate in pN0 cases, which may lead to over treatment [29]. In the present study, we constructed an RNA-seq-based nomogram combined with clinicopathological factors. This model was composed of NEURL1,

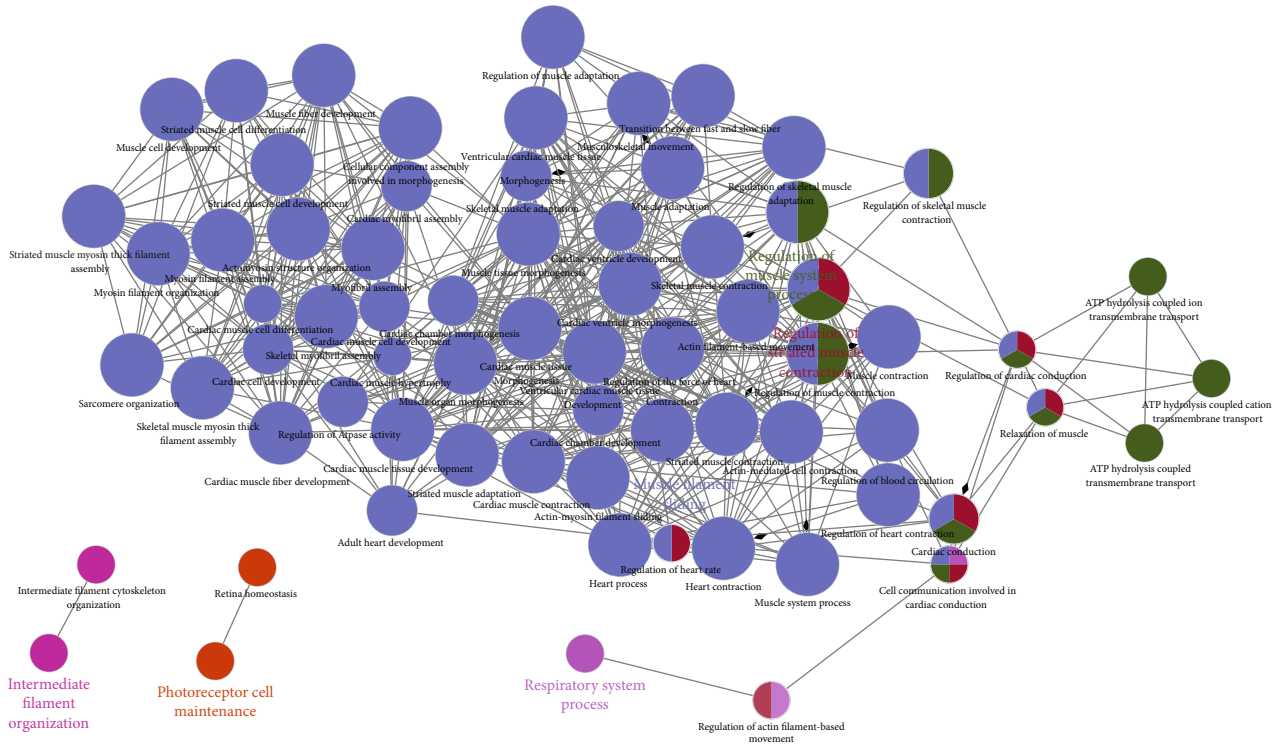


FIGURE 3: Network of functional enrichment results of differentially expressed RNAs between non-LNM and LNM of OSCC. Each node in the figure represents a term, the connection between the nodes reflects the correlation between the terms, and the color of the node reflects the enrichment classification of the node.

AL162581.1 (miscRNA), AP002336.2 (lncRNA), CCBE1, CORO6, RDH12, and AC129492.6 (pseudogene). These candidate DERNAs could significantly distinguish non-LNM and LNM of OSCC. Among them, CCBE1, CORO6, NEURL1, and RDH12 were significantly higher in N+ compared to N- OSCC patients in the GSE9844 dataset. More importantly, RDH12, CCBE1, and AL162581.1 were significantly associated with OSCC LNM. As per previous studies, CCBE1 is indispensable for the development of lymphatic vessels which have important roles in lymphangiogenesis and cancer metastasis [30]. Tumor lymphangiogenesis plays an important role in LNM of OSCC [31–33]. Hogan et al. suggested that CCBE1 may be an extracellular guidance molecule and an independent factor for normal lymphoblasts to germinate or even migrate [34]. Peng et al. concluded that CCBE1 had potential to be a biomarker for prediction of LNM in lung cancer patients because its expression was decreased in lung tumor tissue and further downregulated in patients with LNM [35]. Leong et al. found that CCBE1 as a direct target could promote invasion and metastasis of breast cancer [36]. Our research showed that analysis of CCBE1 expression in the OSCC tissues may help surgeons to evaluate the LNM risk, and CCBE1 might become a potentially therapeutic biomarker in OSCC. Further research is required to understand more about the function of CCBE1 in the LNM of OSCC.

Previous studies have shown that various clinicopathological predictors and genes are considered as risk factors for LNM in the patients with OSCC [37, 38]. However,

no study has combined a visual presentation nomogram with these risk factors. The nomogram is a visualization of a statistical model specifically developed to optimize the accuracy of individual prediction. The preoperative nomogram of estimated LNM can help surgeons identify patients who can obtain greater benefit from a more extensive operation [39–41]. Compared with traditional multiple regression models, the advantage of the nomogram is that all key predictors are displayed graphically. Therefore, we established and validated a novel clinically useful nomogram based on RNA-seq combined with clinicopathological factors to predict the LNM of patients with OSCC. The AUC of the model was up to 0.9773 in the training set, suggesting that the model exhibited a good performance to predict LNM of OSCC. After validation by the test and entire sets, the model still possessed high sensitivity and accuracy on prediction of LNM of OSCC. Thus, the model established by comprehensive use of clinical features had good performance, and candidate RNAs were superior to the use of certain indicators alone, indicating the clinical practicality and innovation of our research. To further validate the prediction efficiency of the model on OSCC LNM, we conducted logistic regression models. As expected, the RNA nomogram combining the clinical N stage showed higher accuracy to predict LNM compared to only the clinical N stage.

At present, some studies have identified some individual markers for LNM in OCSS patients based on microarrays, which require separate clinical tests and individual clinical tests, resulting in increased costs [39–41]. The

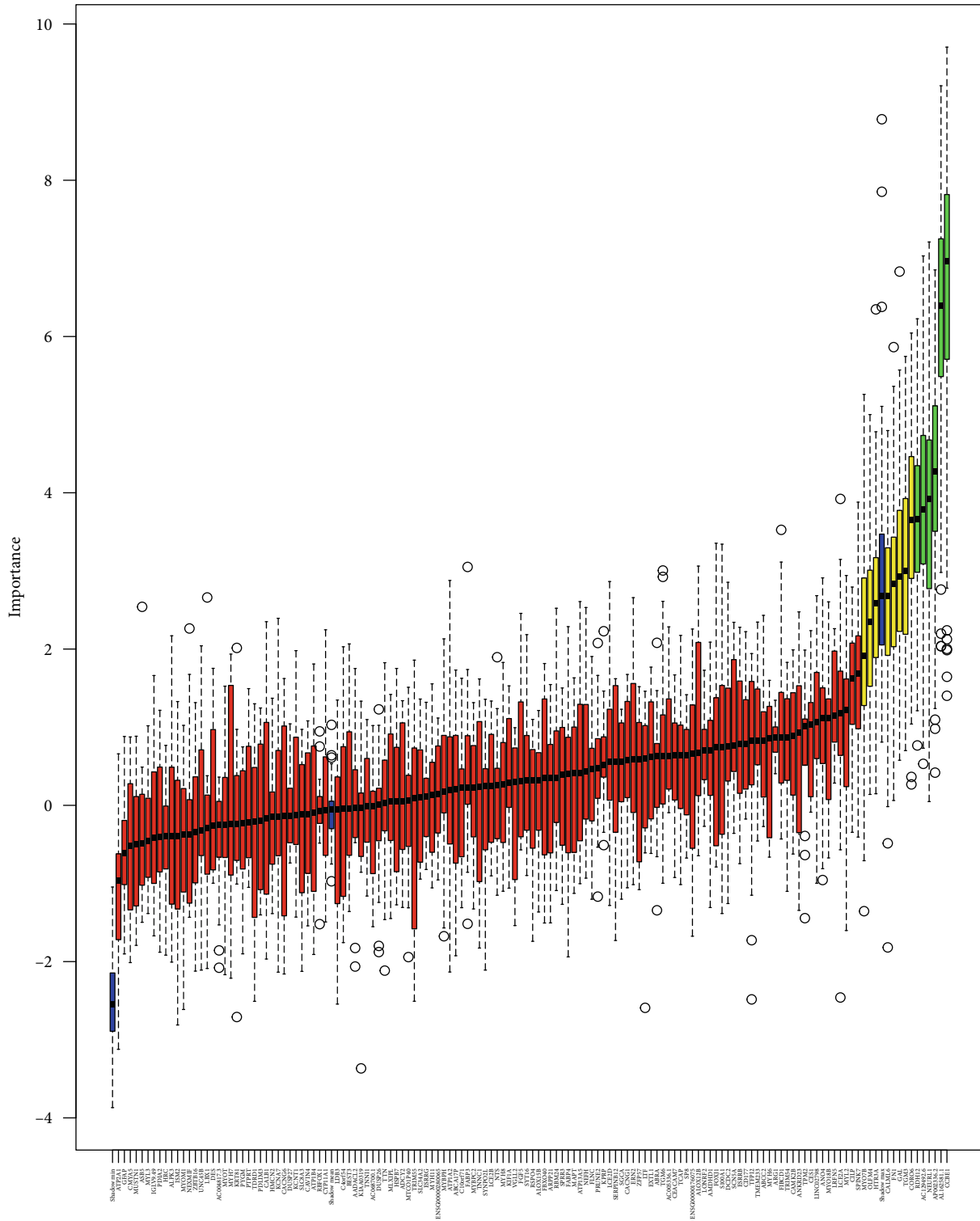


FIGURE 4: Predictive power of individual gene assessed by Boruta feature selection algorithm. Blue represents the minimum, average and maximum Z-score of the shadow gene. Red and green suggest the Z-Score of the rejection attribute and confirmation genes. Yellow expresses suggestive genes.

microarray technology used in those studies does not reflect current genomic views because it can only detect protein-coding genes. Combining protein-coding genes and noncoding genes may improve the robustness of molecular biomarkers. RNA-seq clinical tools have key

advantages over other platforms. The bias and limitations of microarray data are improved by RNA-seq, especially in the detection of low-abundance transcripts. This advantage of RNA-seq can be translated into a better correlation with qPCR data in laboratory and patient samples, which

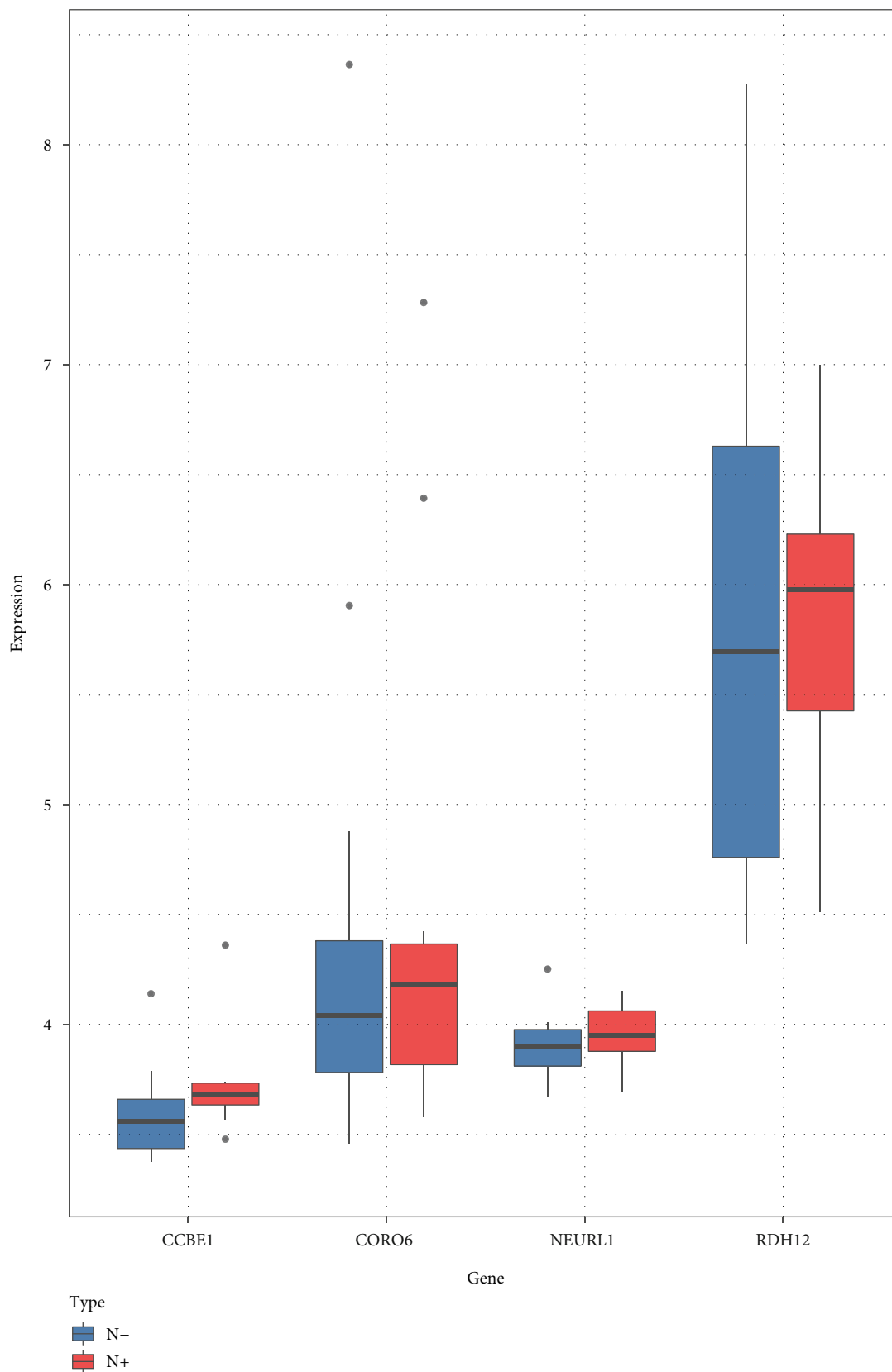


FIGURE 5: Box plots depicting the expression levels of CCBE1, RDH12, CORO6 and NEURL1 between non-LNM and LNM of OSCC using the GSE9844 dataset. N-: pathological examination without lymph node metastasis. N+: pathological examination with lymph node metastasis.

TABLE 2: Comparison of the predictive performances of the machine learning model in the test set.

| Model performances (HR, 95% CI) | RF | SVM | LR | XGBoost |
|---------------------------------|-------------------|-------------------|-------------------|-------------------|
| Sensitivity | 0.82 (0.75, 0.89) | 0.80 (0.75, 0.85) | 0.81 (0.74, 0.89) | 0.72 (0.63, 0.81) |
| Specificity | 0.67 (0.57, 0.78) | 0.76 (0.68, 0.84) | 0.76 (0.68, 0.84) | 0.67 (0.57, 0.77) |
| PPV | 0.77 (0.72, 0.83) | 0.82 (0.77, 0.87) | 0.82 (0.77, 0.88) | 0.74 (0.67, 0.81) |
| NPV | 0.76 (0.67, 0.84) | 0.75 (0.71, 0.79) | 0.78 (0.70, 0.83) | 0.65 (0.56, 0.75) |
| Accuracy | 0.75 (0.71, 0.80) | 0.78 (0.75, 0.81) | 0.79 (0.74, 0.83) | 0.70 (0.62, 0.77) |
| AUC | 0.82 (0.77, 0.88) | 0.84 (0.80, 0.87) | 0.84 (0.80, 0.89) | 0.77 (0.71, 0.83) |

RF: random forest; SVM: support vector machine; LR: logistic regression; HR: hazard ratio; CI: confidence interval; PPV: positive predictive value; NPV: negative predictive value; AUC: area under the curve.

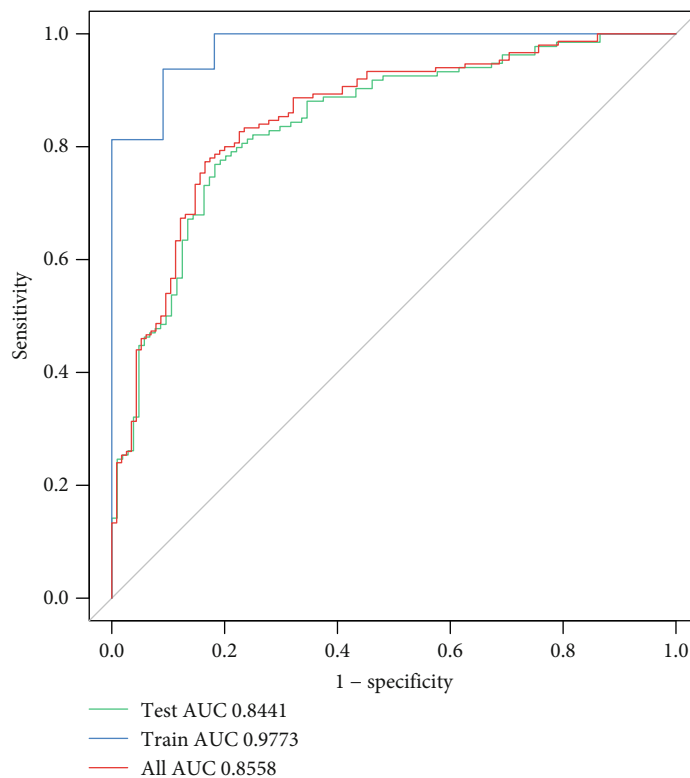


FIGURE 6: ROC analysis for comparison of the overall performance of the model using the different sets.

is especially important for genes that tend to be differentially expressed but have low absolute abundance. Moreover, RNA-seq provides comprehensive expression data, which will become increasingly important in understanding and predicting the therapeutic response of most tumors that lack classic targetable changes. Our prediction model is based on the RNA-seq dataset. This method is more economical and more clinically applicable than whole-genome sequencing. Our findings constructed an RNA-seq-based nomogram combined with clinical pathology, which may provide clinicians with useful tools for preoperative prediction of LNM and tailor-made personalized treatment for OSCC patients. It is easy to understand its graphical scoring system, which helps to customize treatment and medical decisions. To the best of our knowledge, the RNA-seq-based nomogram described in this article has not been reported previously, providing a powerful prognostic tool for precision oncology. Therefore,

this model may have important implications for clinical practice. Our RNA-seq data is based on the TCGA database and the study population is from different races, so the model can be extended to other races/ethnic populations. The RNA-seq-based nomogram combined with clinicopathology provides an opportunity for individualized adjuvant therapy based on biological factors and comprehensive change testing through the RNA-seq platform. Therefore, this model may be a clinically useful tool that can be easily incorporated into the RNA-seq clinical sequencing program to personalize OSCC treatment.

However, our study had certain limitations. We acknowledged that it was based on TCGA data and the sample size was relatively small. In addition, the nomogram had only been validated internally and still needed to be further validated by independent cohorts in a multi-center trial to investigate whether the results are applicable to other populations.

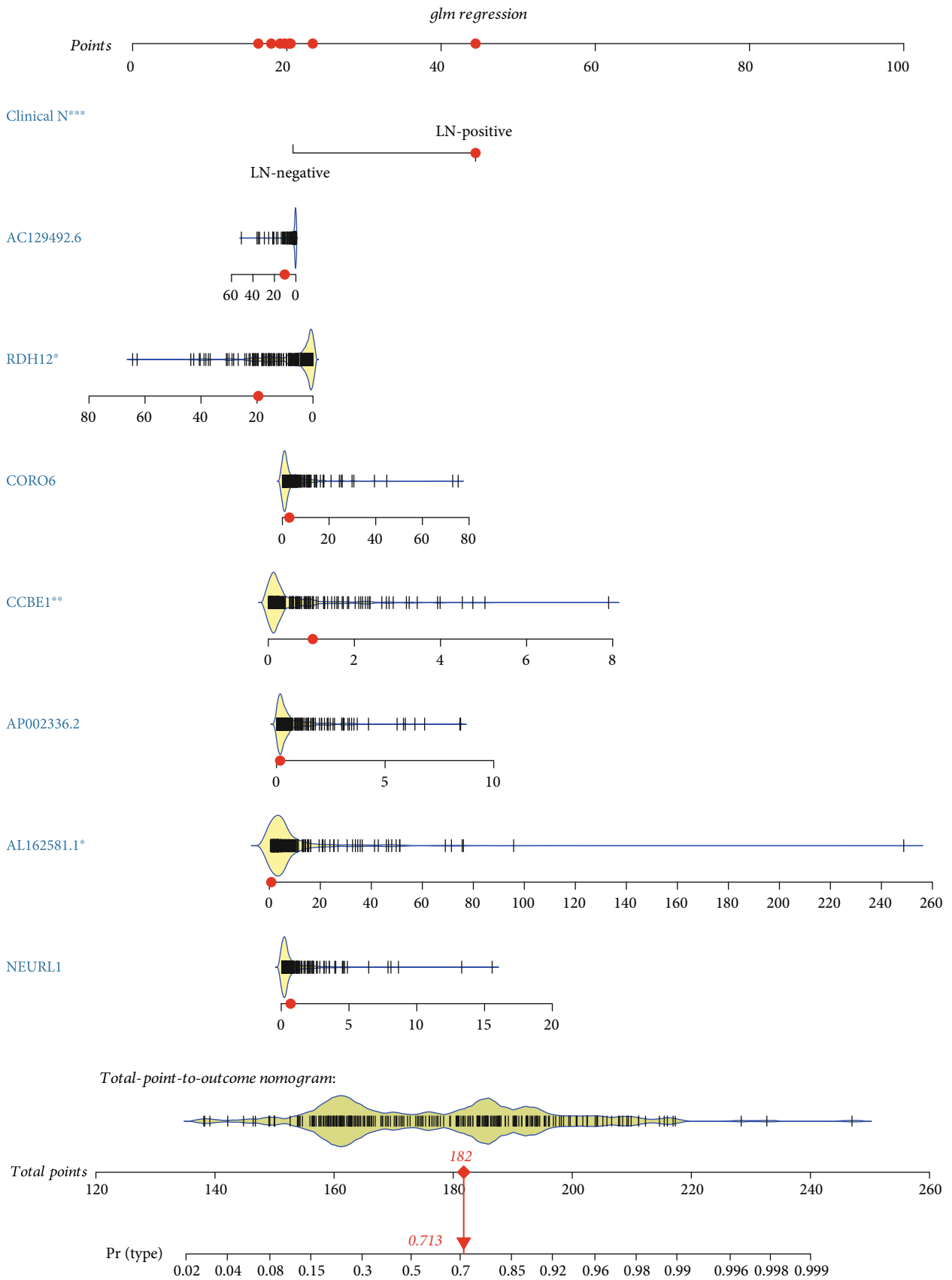


FIGURE 7: The nomogram for the prediction of lymph node metastasis.

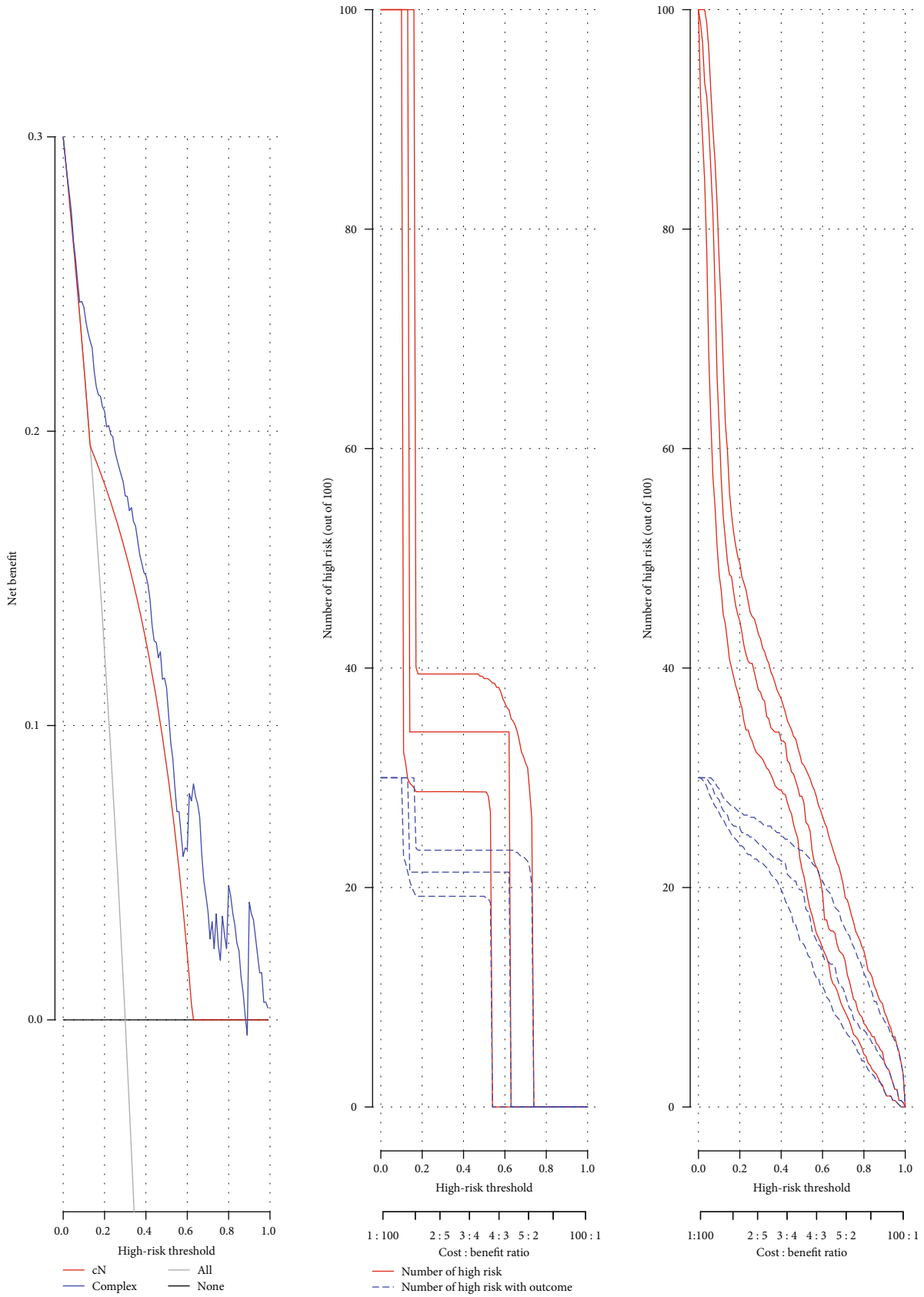


FIGURE 8: (a) The decision curve analyses (DCAs) for the clinical N stage and complex nomogram. (b) The clinical impact curve for the cN. (c) The clinical impact curve for the complex.

5. Conclusion

The RNA-seq-based nomogram combined with clinicopathology could potentially provide clinicians with a useful tool to best select patients with LNM who will benefit from neck dissection, while avoiding the cost and overtreatment of those without LNM. Ultimately, optimized individually tailored therapy will contribute to the management of OSCC patients based on the model.

Abbreviations

| | |
|----------|--|
| LNM: | Lymph node metastasis |
| OSCC: | Oral squamous cell carcinoma |
| non-LNM: | Without LNM |
| cN0: | Clinically negative neck |
| END: | Elective neck dissection |
| TCGA: | The Cancer Genome Atlas database |
| GEO: | Gene Expression Omnibus |
| OTSCC: | Oral tongue squamous cell carcinoma |
| DERNAs: | Differentially expressed RNAs |
| FC: | Fold change |
| cN: | Clinical N stage |
| SVM: | Support vector machine |
| AUC: | Area under the curve |
| LR: | Logistic regression |
| RF: | Random forest |
| N-: | Pathological examination without lymph node metastasis |
| N+: | Pathological examination with lymph node metastasis |
| HR: | Hazard ratio |
| CI: | Confidence interval |
| PPV: | Positive predictive value |
| NPV: | Negative predictive value. |

Data Availability

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

We have no conflicts of interest.

Authors' Contributions

Bo Qiao and Min Zhao contributed equally to this work.

Acknowledgments

The main results published here are based on data generated by TCGA (<http://cancergenome.nih.gov>). We are grateful to TCGA for the convenient source of data. This work was funded by the Chinese People's Liberation Army General Hospital Medical Big Data Project (No. 2019MBD-003).

Supplementary Materials

Supplementary Table 1: 139 differentially expressed mRNAs between non-LNM and LNM of OSCC. (*Supplementary materials*)

References

- [1] A. I. Lorenzo-Pouso, M. Gallas-Torreira, M. Pérez-Sayáns et al., "Prognostic value of CAIX expression in oral squamous cell carcinoma: a systematic review and meta-analysis," *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 35, no. 1, pp. 1258–1266, 2020.
- [2] Q. Fang, H. Gao, Q. Gao et al., "Elective neck dissection versus wait-and-see policy in cT1N0 buccal squamous cell carcinoma," *BMC Cancer*, vol. 20, no. 1, p. 537, 2020.
- [3] C. C. Stoop, R. Bree, A. J. W. P. Rosenberg, J. T. M. Gemert, T. Forouzanfar, and E. M. Van Cann, "Locoregional recurrence rate and disease-specific survival following marginal vs segmental resection for oral squamous cell carcinoma with mandibular bone invasion," *Journal of Surgical Oncology*, vol. 122, no. 4, pp. 646–652, 2020.
- [4] M. F. van der Kamp, F. O. W. Muntinghe, R. S. Iepsma et al., "Predictors for distant metastasis in head and neck cancer, with emphasis on age," *European Archives of Oto-Rhino-Laryngology*, 2020.
- [5] C.-T. Liao, J. T.-C. Chang, H.-M. Wang et al., "Survival in squamous cell carcinoma of the oral cavity," *Cancer*, vol. 110, no. 3, pp. 564–571, 2007.
- [6] A. K. D'Cruz, R. Vaish, N. Kapre et al., "Elective versus therapeutic neck dissection in node-negative oral cancer," *New England Journal of Medicine*, vol. 373, no. 6, pp. 521–529, 2015.
- [7] A. M. Bur, A. Holcomb, S. Goodwin et al., "Machine learning to predict occult nodal metastasis in early oral squamous cell carcinoma," *Oral Oncology*, vol. 92, pp. 20–25, 2019.
- [8] R. Akbani, P. K. S. Ng, H. M. J. Werner et al., "A pan-cancer proteomic perspective on The Cancer Genome Atlas," *Nature Communications*, vol. 5, no. 1, 2014.
- [9] M. Cieřlik and A. M. Chinnaiyan, "Cancer transcriptome profiling at the juncture of clinical translation," *Nature Reviews Genetics*, vol. 19, no. 2, pp. 93–109, 2018.
- [10] H. Ye, T. Yu, S. Temam et al., "Transcriptomic dissection of tongue squamous cell carcinoma," *BMC Genomics*, vol. 9, no. 1, p. 69, 2008.
- [11] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome Biology*, vol. 15, no. 12, p. 550, 2014.
- [12] G. Bindea, J. Galon, and B. Mlecnik, "CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data," *Bioinformatics*, vol. 29, no. 5, pp. 661–663, 2013.
- [13] G. Bindea, B. Mlecnik, H. Hackl et al., "ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks," *Bioinformatics*, vol. 25, no. 8, pp. 1091–1093, 2009.
- [14] P. Shannon, A. Markiel, O. Ozier et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.
- [15] M. B. Kursa and W. R. Rudnicki, "Feature selection with the BorutaPackage," *Journal of Statistical Software*, vol. 36, no. 11, 2010.

- [16] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [17] M. Awad and R. Khanna, "Support Vector Machines for Classification," in *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, pp. 39–66, Apress, Berkeley, CA, 2015.
- [18] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, NY, USA, 2016.
- [19] E. Nunez, E. W. Steyerberg, and J. Nunez, "Regression modeling strategies," *Revista espanola de cardiologia*, vol. 64, pp. 501–507, 2011.
- [20] D. J. Williams and J. D. Walker, "A nomogram for calculating the maximum dose of local anaesthetic," *Anaesthesia*, vol. 69, pp. 847–853, 2014.
- [21] A. J. Vickers and E. B. Elkin, "Decision curve analysis: a novel method for evaluating prediction models," *Medical Decision Making*, vol. 26, pp. 565–574, 2006.
- [22] Y. C. Lim, J. S. Lee, B. S. Koo, S.-H. Kim, Y.-H. Kim, and E. C. Choi, "Treatment of contralateral N0 neck in early squamous cell carcinoma of the oral tongue: elective neck dissection versus observation," *The Laryngoscope*, vol. 116, no. 3, pp. 461–465, 2006.
- [23] E. M. Iype, P. Sebastian, A. Mathew, P. G. Balagopal, B. T. Varghese, and S. Thomas, "The role of selective neck dissection (I–III) in the treatment of node negative (N0) neck in oral cancer," *Oral Oncology*, vol. 44, no. 12, pp. 1134–1138, 2008.
- [24] W. K. Wong and R. P. Morton, "Elective management of cervical and parotid lymph nodes in stage N0 cutaneous squamous cell carcinoma of the head and neck: a decision analysis," *European Archives of Oto-Rhino-Laryngology*, vol. 271, no. 11, pp. 3011–3019, 2014.
- [25] P. J. Bradley, A. Ferlito, C. E. Silver et al., "Neck treatment and shoulder morbidity: still a challenge," *Head & Neck*, vol. 33, no. 7, pp. 1060–1067, 2011.
- [26] S. N. Rogers, A. Ferlito, P. K. Pellitteri, A. R. Shaha, and A. Rinaldo, "Quality of life following neck dissections," *Acta Oto-Laryngologica*, vol. 124, no. 3, pp. 231–236, 2014.
- [27] F. Sonohara, F. Gao, N. Iwata et al., "Genome-wide discovery of a novel gene-expression signature for the identification of lymph node metastasis in esophageal squamous cell carcinoma," *Annals of Surgery*, vol. 269, no. 5, pp. 879–886, 2019.
- [28] D. Izumi, F. Gao, S. Toden et al., "A genomewide transcriptional approach identifies a novel gene expression signature for the detection of lymph node metastasis in patients with early stage gastric cancer," *EBioMedicine*, vol. 41, pp. 268–275, 2019.
- [29] F. S. Pasini, S. Maistro, I. Snitcovsky et al., "Four-gene expression model predictive of lymph node metastases in oral squamous cell carcinoma," *Acta Oncologica*, vol. 51, no. 1, pp. 77–85, 2011.
- [30] J. Song, W. Chen, X. Cui et al., "CCBE1 promotes tumor lymphangiogenesis and is negatively regulated by TGF β signaling in colorectal cancer," *Theranostics*, vol. 10, no. 5, pp. 2327–2341, 2020.
- [31] S. H. Wang, G. G. Liou, S. H. Liu et al., "Laminin γ 2-enriched extracellular vesicles of oral squamous cell carcinoma cells enhance in vitro lymphangiogenesis via integrin α 3-dependent uptake by lymphatic endothelial cells," *International Journal of Cancer*, vol. 144, no. 11, pp. 2795–2810, 2019.
- [32] Y. L. Chen, W. L. Wu, C. W. Jang et al., "Interferon-stimulated gene 15 modulates cell migration by interacting with Rac1 and contributes to lymph node metastasis of oral squamous cell carcinoma cells," *Oncogene*, vol. 38, no. 23, pp. 4480–4495, 2019.
- [33] A. C. Chang, M. Y. Lien, M. H. Tsai, C. H. Hua, and C. H. Tang, "WISP-1 promotes epithelial-mesenchymal transition in oral squamous cell carcinoma cells via the miR-153-3p/snail axis," *Cancers (Basel)*, vol. 11, no. 12, p. 1903, 2019.
- [34] B. M. Hogan, F. L. Bos, J. Bussmann et al., "ccbe1 is required for embryonic lymphangiogenesis and venous sprouting," *Nature Genetics*, vol. 41, no. 4, pp. 396–398, 2009.
- [35] P. Li, Z. Cong, Y. Qiang et al., "Clinical significance of CCBE1 expression in lung cancer," *Molecular Medicine Reports*, vol. 17, pp. 2107–2112, 2018.
- [36] A. Mesci, X. Huang, S. Taeb et al., "Targeting of CCBE1 by miR-330-3p in human breast cancer promotes metastasis," *British Journal of Cancer*, vol. 116, no. 10, pp. 1350–1357, 2017.
- [37] K. Miyaguchi, N. Uzawa, K. Mogushi et al., "Loss of NKX3-1 as a potential marker for an increased risk of occult lymph node metastasis and poor prognosis in oral squamous cell carcinoma," *International Journal of Oncology*, vol. 40, no. 6, pp. 1907–1914, 2012.
- [38] M. Seki, T. Sano, S. Yokoo, and T. Oyama, "Tumour budding evaluated in biopsy specimens is a useful predictor of prognosis in patients with cN0 early stage oral squamous cell carcinoma," *Histopathology*, vol. 70, no. 6, pp. 869–879, 2017.
- [39] Y. Hayashi, L. Xiao, A. Suzuki et al., "A nomogram associated with high probability of malignant nodes in the surgical specimen after trimodality therapy of patients with oesophageal cancer," *European Journal of Cancer*, vol. 48, no. 18, pp. 3396–3404, 2012.
- [40] A. M. Thompson, R. M. Turner, A. Hayen et al., "A preoperative nomogram for the prediction of ipsilateral central compartment lymph node metastases in papillary thyroid cancer," *Thyroid*, vol. 24, no. 4, pp. 675–682, 2014.
- [41] V. P. Balachandran, M. Gonen, J. J. Smith, and R. P. DeMatteo, "Nomograms in oncology: more than meets the eye," *The Lancet Oncology*, vol. 16, no. 4, pp. e173–e180, 2015.

Research Article

Development of an Immune Infiltration-Related Eight-Gene Prognostic Signature in Colorectal Cancer Microenvironment

Beilei Wu ¹, Lijun Tao ², Daqing Yang ³, Wei Li ¹, Hongbo Xu ² and Qianggui He ²

¹Medical Examination Center, Wenzhou Central Hospital, The Dingli Clinical College of Wenzhou Medical University, Wenzhou, 325000 Zhejiang, China

²Department of Traumatic Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, 325000 Zhejiang, China

³Department of Colorectal Surgery, Wenzhou Central Hospital, The Dingli Clinical College of Wenzhou Medical University, Wenzhou, 325000 Zhejiang, China

Correspondence should be addressed to Qianggui He; heqianggui163126@163.com

Received 22 February 2020; Revised 23 May 2020; Accepted 16 July 2020; Published 27 August 2020

Guest Editor: Shijia Zhu

Copyright © 2020 Beilei Wu 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.

Objective. Stromal cells and immune cells have important clinical significance in the microenvironment of colorectal cancer (CRC). This study is aimed at developing a CRC gene signature on the basis of stromal and immune scores. **Methods.** A cohort of CRC patients ($n = 433$) were adopted from The Cancer Genome Atlas (TCGA) database. Stromal/immune scores were calculated by the ESTIMATE algorithm. Correlation between prognosis/clinical characteristics and stromal/immune scores was assessed. Differentially expressed stromal and immune genes were identified. Their potential functions were annotated by functional enrichment analysis. Cox regression analysis was used to develop an eight-gene risk score model. Its predictive efficacies for 3 years, 5 years, overall survival (OS), and progression-free survival interval (PFI) were evaluated using time-dependent receiver operating characteristic (ROC) curves. The correlation between the risk score and the infiltrating levels of six immune cells was analyzed using TIMER. The risk score was validated using an independent dataset. **Results.** Immune score was in a significant association with prognosis and clinical characteristics of CRC. 736 upregulated and two downregulated stromal and immune genes were identified, which were mainly enriched into immune-related biological processes and pathways. An-eight gene prognostic risk score model was conducted, consisting of CCL22, CD36, CPA3, CPT1C, KCNE4, NFATC1, RASGRP2, and SLC2A3. High risk score indicated a poor prognosis of patients. The area under the ROC curves (AUC) s of the model for 3 years, 5 years, OS, and PFI were 0.71, 0.70, 0.73, and 0.66, respectively. Thus, the model possessed well performance for prediction of patients' prognosis, which was confirmed by an external dataset. Moreover, the risk score was significantly correlated with immune cell infiltration. **Conclusion.** Our study conducted an immune-related prognostic risk score model, which could provide novel targets for immunotherapy of CRC.

1. Introduction

CRC, as a heterogeneous disease, is a common cause of cancer-related deaths worldwide [1]. TNM staging is usually considered to be one of the main tools for CRC prognosis [2]. However, the prognosis varies greatly among CRC patients with the same TNM stage, suggesting that the current TNM stage does not well provide complete prognostic information

for CRC patients. Therefore, it is necessary to adopt a new strategy to increase the predictive efficiency of prognosis and survival outcomes of CRC patients.

Due to the considerable heterogeneity between CRCs, determination of the optimal treatment strategy at the individual level faces the large challenges. Thus, it is an urgent need to conduct robust models to identify high-risk CRC patients and to find novel molecular targets. In the tumor

microenvironment (TME), stromal and immune cells are involved in the development of CRC [3, 4]. Increasing evidence suggests that stromal and immune cells possess critical clinical significance for CRC. It has been reported that stromal cells can contribute to transcriptome and clinical features of CRC subtype [5]. Furthermore, stromal gene expression can more robustly predict the prognosis of CRC subtypes compared to epithelial tumor cells [6]. In a large cohort of CRC patients, infiltrating immune cell data could better predict patients' survival than histopathological methods [7]. Growing studies have found that infiltrating immune cells are involved in chemoresistance [8] and metastasis [9]. Thus, it is essential to further analyze the biological characteristics of stromal and immune genes and to determine their prognostic value for CRC patients. However, there is a lack of stromal and immune scores that can predict CRC patients' prognosis based on multiple clinical factors. Moreover, robust prognostic models on the basis of stromal and immune scores are also lacking.

In this study, we established a reliable prognostic immune-related risk score for CRC. Our results could offer novel insights for prediction of CRC patients' prognosis and development of individualized immunity therapy strategies.

2. Materials and Methods

2.1. CRC Datasets. TCGA RNA-seq data (including Counts and FPKM data) of GDC CRC (including 469 CRC tissue samples and 41 adjacent normal tissue samples) were downloaded from the xenabrowser website (<https://xenabrowser.net/>). Among all CRC samples, 433 samples contained complete clinical information, including gender, age, TNM stage, tumor grade, microsatellite instability (MSI), and mismatch repair (MMR). The clinical information of 433 CRC patients is listed in Table 1. Survival information including OS status, OS time, progression-free survival (PFS) status, and PFS time was derived from the pan-cancer on the GDC website (<https://gdc.cancer.gov/about-data/publications/PanCan-Clinical-2018>). Furthermore, mutation data (including BRAF, KRAS, and TP53) were from CRC MuTect. An overview of the workflow is shown in Figure 1.

2.2. Estimation of Stromal/Immune Scores. ESTIMATE algorithm was used to calculate the stromal/immune scores on the basis of unique expression profiles of stromal/immune cells by the ESTIMATE package in R (<https://bioinformatics.mdanderson.org/estimate/>) [10].

2.3. Kaplan-Meier Survival Analysis. According to the optimal cutoff of stromal/immune scores, CRC samples were classified into high and low stromal/immune score groups. Kaplan-Meier plot of overall survival between the two groups was analyzed, and the results were evaluated by log-rank test.

2.4. Correlation between Clinical Characteristics and Stromal/Immune Scores. To probe into the clinical significance of stromal/immune scores, we analyzed the correlation between clinical characteristics (including pathologic T,

TABLE 1: Clinical characteristics of CRC patients in TCGA datasets (overall = 433).

| Characteristics | Groups | N (%) |
|------------------|--------|------------|
| Age (%) | ≤60 | 136 (30.7) |
| | >60 | 297 (68.6) |
| Gender (%) | Female | 200 (46.2) |
| | Male | 233 (53.8) |
| Status (%) | Died | 338 (78.1) |
| | Alive | 95 (21.9) |
| Pathologic T (%) | T1 | 11 (2.5) |
| | T2 | 75 (17.3) |
| | T3 | 296 (68.4) |
| | T4 | 51 (11.8) |
| Pathologic N (%) | N0 | 254 (58.7) |
| | N1 | 102 (23.6) |
| | N2 | 77 (17.8) |
| Pathologic M (%) | M0 | 320 (75.1) |
| | M1 | 61 (14.3) |
| | Mx | 45 (10.6) |
| Tumor stage (%) | I | 73 (17.3) |
| | II | 165 (39.1) |
| | III | 123 (29.1) |
| | IV | 61 (14.5) |

pathologic N, pathologic M, and tumor stage) and stromal/immune scores.

2.5. Differential Expression Analysis. Differential expression analysis between high and low stromal/immune score groups was carried out using the edgeR package in R, following the screening criteria of $|\log_2 \text{fold change (FC)}| > 1$ and FDR (adjusted p value) < 0.05 . Then, up- or downregulated stromal/immune genes were intersected by the VennDiagram package in R, respectively.

2.6. Functional Enrichment Analysis. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of differentially expressed stromal and immune genes were carried out through the clusterProfiler package in R [11]. GO analysis contains three terms, cellular component (CC), molecular function (MF), and biological process (BP). p value after adjustment < 0.05 was significantly enriched.

2.7. Protein-Protein Interaction (PPI) Analysis. PPI analyses of differentially expressed stromal and immune genes were carried out via The Search Tool for the Retrieval of Interacting Genes (STRING, <https://string-db.org/>; version 11) [12]. Then, the PPI network was visualized through Cytoscape (version 3.7.2) [13].

2.8. Univariate and Multivariate Cox Regression Analyses. Univariate cox regression analysis of differentially expressed

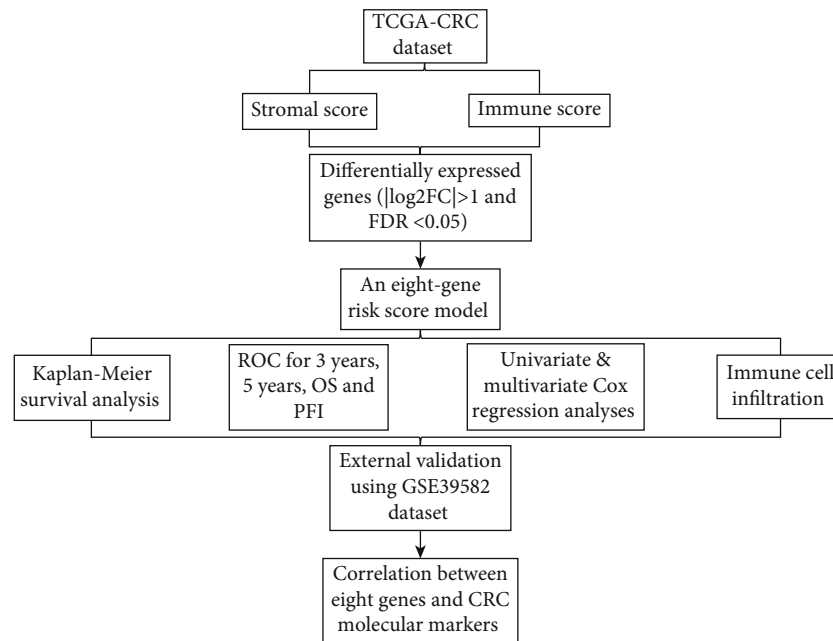


FIGURE 1: An overview of the workflow.

stromal and immune genes was carried out via the survival package in R. Genes with p value < 0.05 were screened for multivariate cox regression analysis. To validate the sensitivity and accuracy of the risk score for prediction of CRC, an ROC curve was drawn to evaluate the predictive performance of the risk score for 3 years, 5 years, OS, and PFI using the “tdROC” package in R. The results were visualized with the “ggplot2” package in R. The AUC was then calculated. The GSE39582 dataset from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov>) was used to validate the prognostic value of the risk score. The dataset was composed of 566 CRC samples.

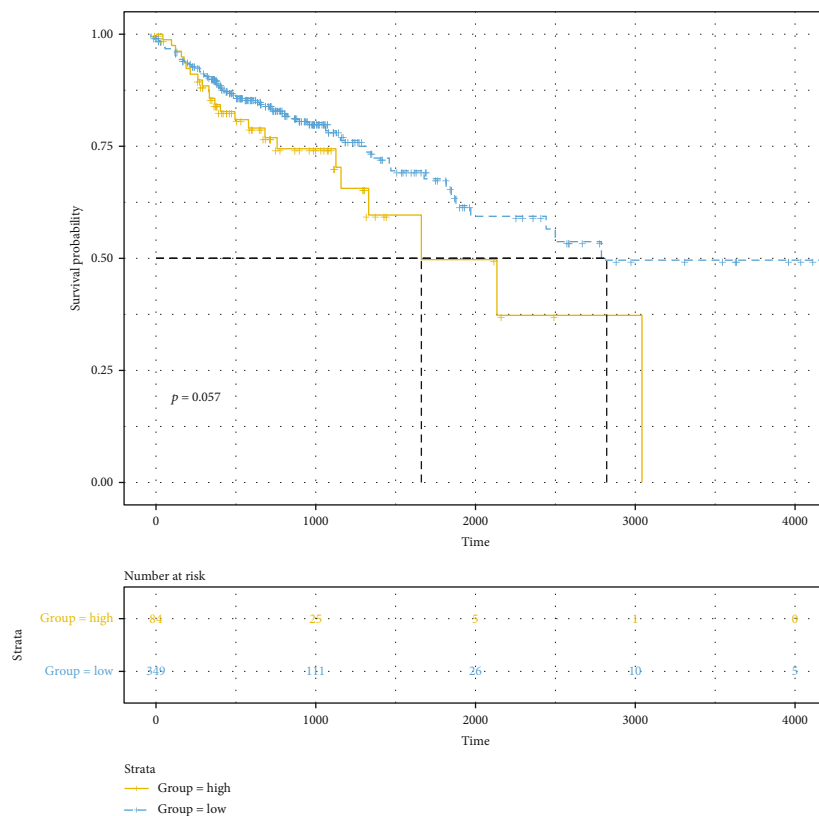
2.9. Immune Infiltration Analysis. The tumor-immune infiltration cells including B cells, CD4+T cells, CD8+T cells, macrophages, neutrophils, and dendritic cells were estimated via the TIMER (<https://cistrome.shinyapps.io/timer/>) [14]. Spearman’s correlation between the risk score and the infiltrating levels of immune cells was evaluated through the psych package in R. Furthermore, we also assessed the correlation between the genes in the risk score and marker genes of immune cells. The strength of correlation followed the criteria: $0.7 \leq |r| \leq 1$ suggested a high correlation, $0.3 \leq |r| < 0.7$ suggested a moderate correlation, and $0 < |r| < 0.3$ suggested a weak correlation [15].

3. Results

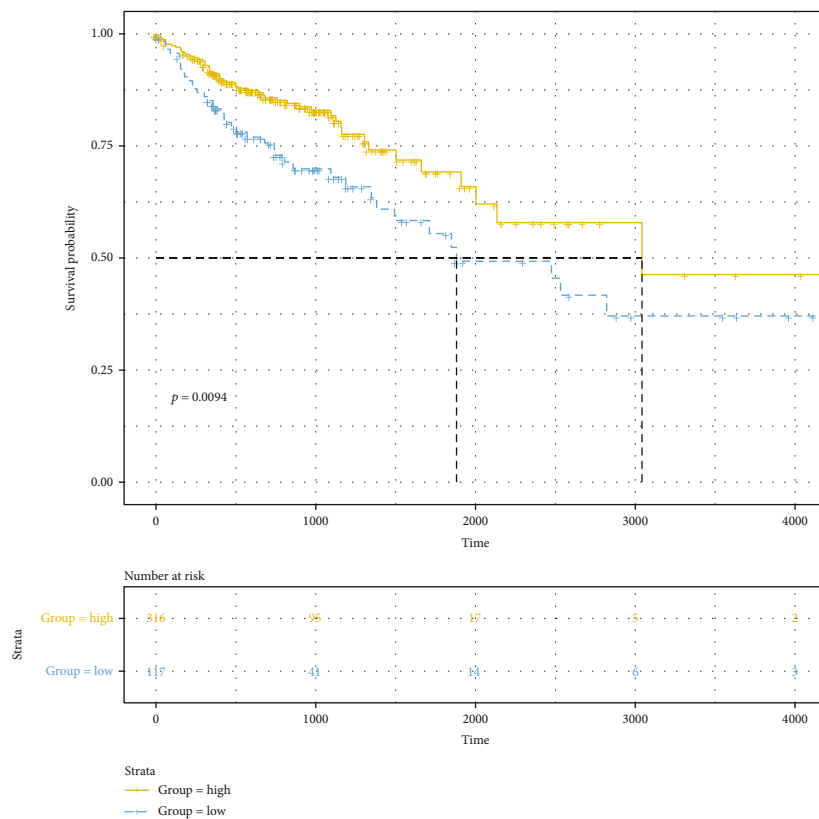
3.1. Immune Score Is in Significant Association with Prognosis and Clinical Features of CRC Patients. According to the optimal cutoff of stromal/immune scores, the CRC patients were divided into two groups. Kaplan-Meier OS analysis results showed that patients with high stromal score had shorter OS time than those with low stromal score; however, it was

not statistically significant (Figure 2(a); p value = 0.057). As depicted in Figure 2(b), we found that patients with low immune score implied a poor prognosis (p value = 0.0094). Furthermore, we analyzed the correlation between stromal/immune scores and clinical features. As depicted in the results, stromal score was not significantly associated with pathologic T (Figure 3(a); p value = 0.61), pathologic N (Figure 3(b); p value = 0.28), pathologic M (Figure 3(c); p value = 0.63), tumor stage (Figure 3(d); p value = 0.68), and age (Figure 3(e); p value = 0.76). Similarly, we also found that there was no statistical significance between immune score and pathologic T (Figure 3(f); p value = 0.88) and pathologic N (Figure 3(g); p value = 0.17). As expected, immune score was in significant association with pathologic M (Figure 3(h); p value = 0.0045) and tumor stage (Figure 3(i); p value = 0.0093). However, no significant correlation between immune score and age was found in Figure 3(j) (p value = 0.29). Furthermore, ESTIMATE scores were not correlated with pathologic T (Figure 3(k); p value = 0.98), pathologic N (Figure 3(l); p value = 0.73), pathologic M (Figure 3(m); p value = 0.095), tumor stage (Figure 3(n); p value = 0.28), and age (Figure 3(o); p value = 0.74). These findings indicated that immune score was in significant association with CRC patients’ prognosis and clinical features.

3.2. Identification of Differentially Expressed Stromal and Immune Genes for CRC. We analyzed differentially expressed genes (DEGs) with $|\log_2 \text{FC}| > 1$ and $\text{FDR} < 0.05$ between the high and low stromal/immune score groups. As volcano plots, there were 1197 up- and 28 downregulated stromal genes in the high stromal score group (Figure 4(a)). Furthermore, 899 immune genes were upregulated and eight immune genes were downregulated in the high immune score group (Figure 4(b)). Hierarchical clustering analysis

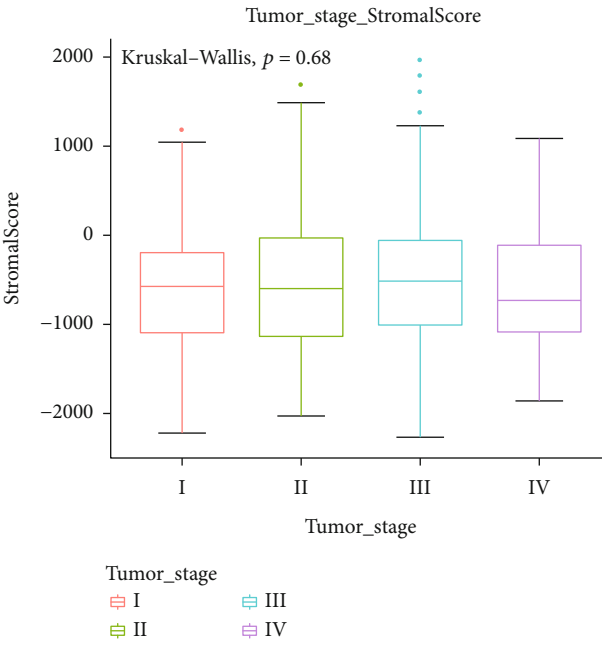
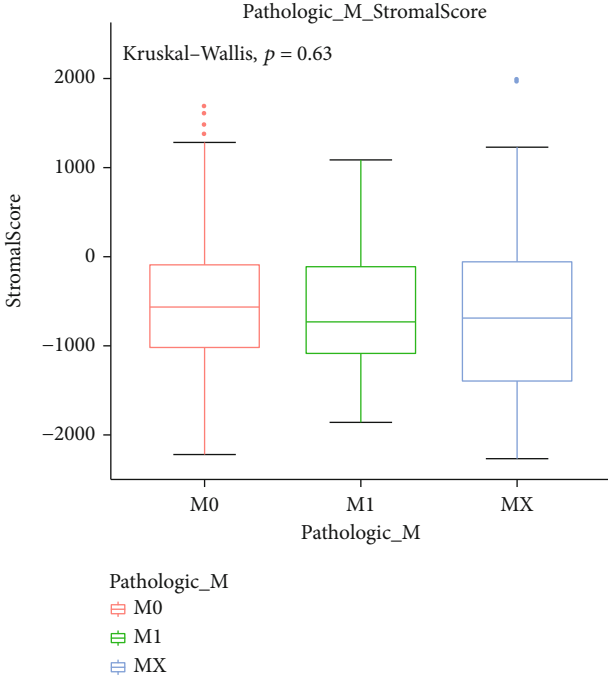
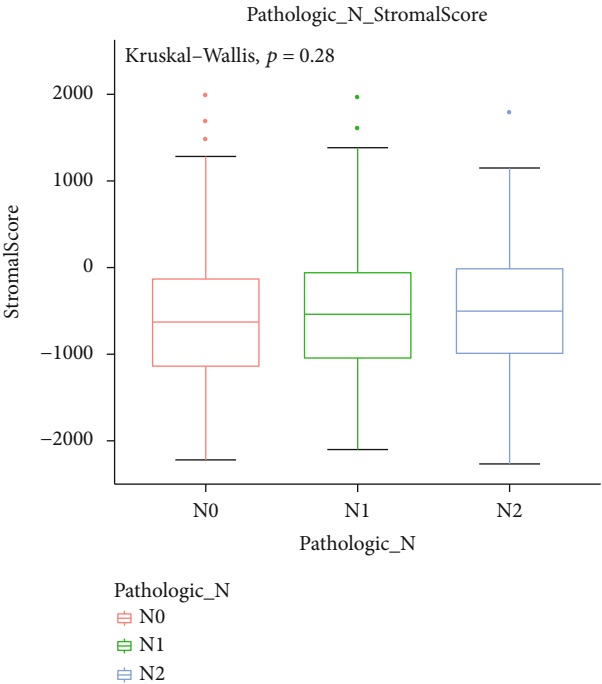
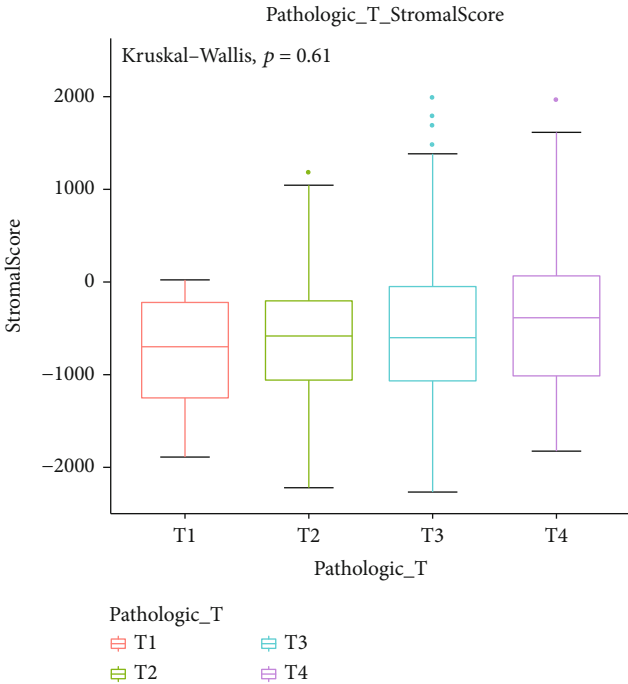


(a)



(b)

FIGURE 2: The correlation between stromal/immune scores and CRC patients' survival outcomes. (a) Stromal score. (b) Immune score. The x-axis suggests overall survival time and y-axis represents survival probability.



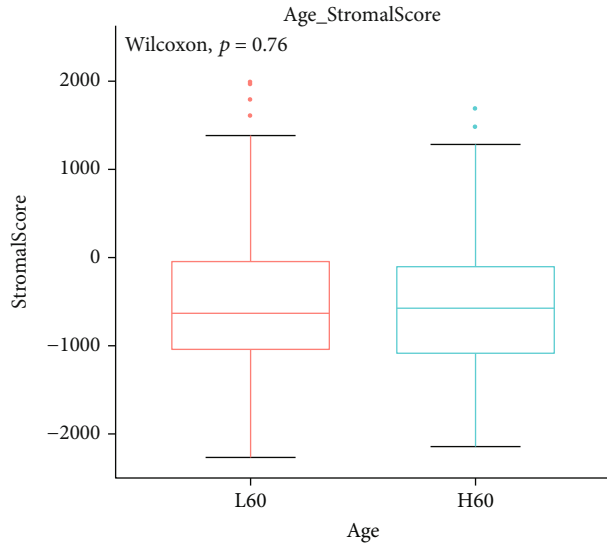
(a)

(b)

(c)

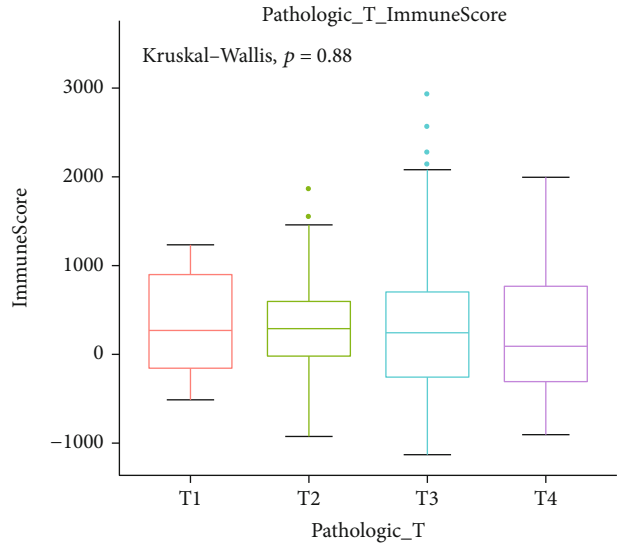
(d)

FIGURE 3: Continued.



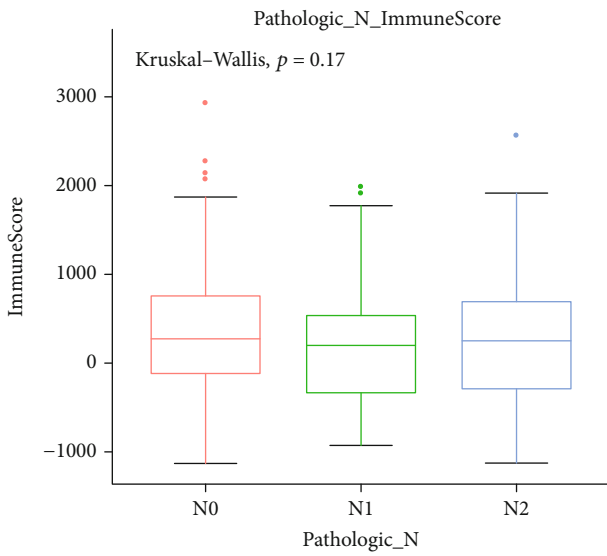
Age
L60
H60

(e)



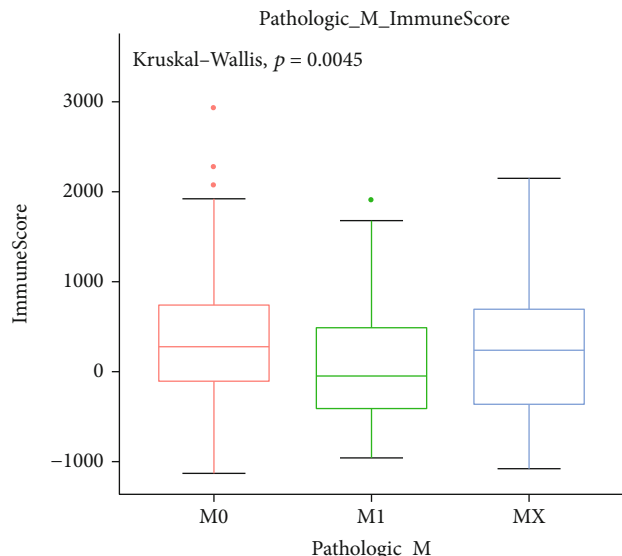
Pathologic_T
T1 T2 T3 T4

(f)



Pathologic_N
N0
N1
N2

(g)



Pathologic_M
M0
M1
MX

(h)

FIGURE 3: Continued.

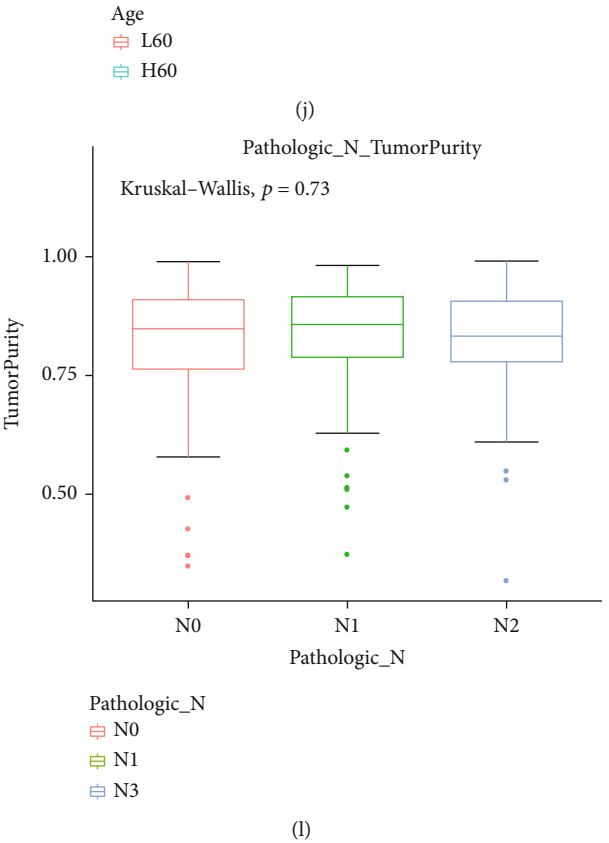
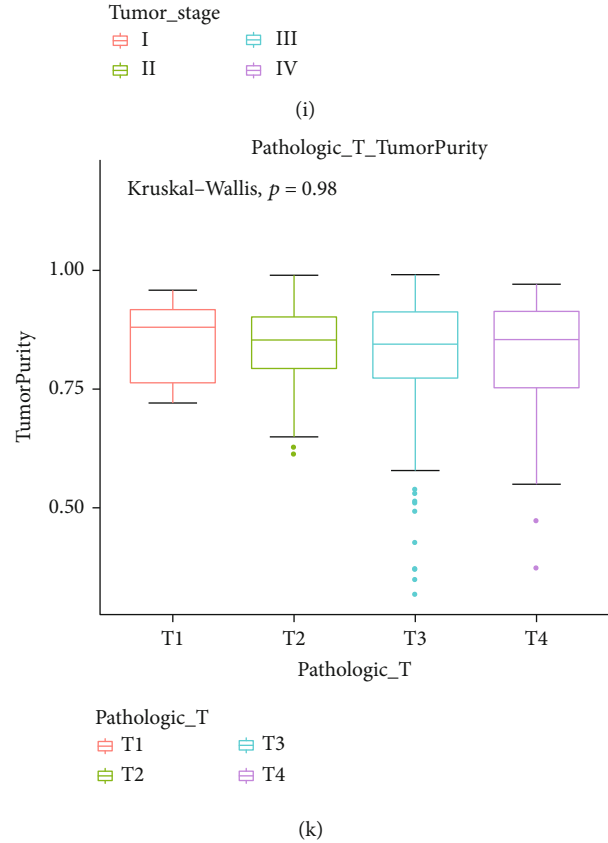
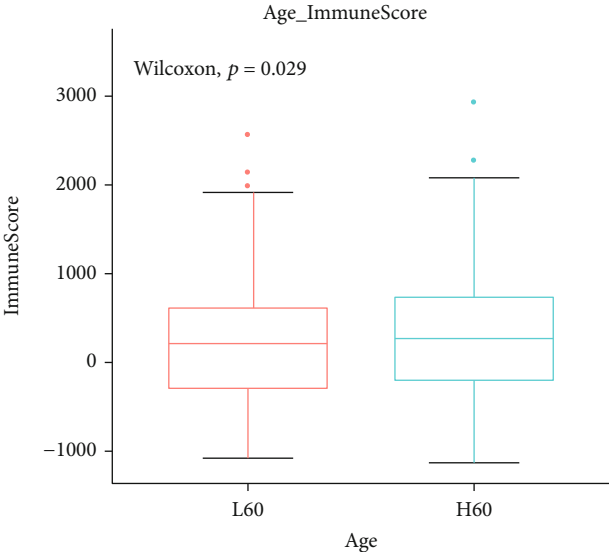
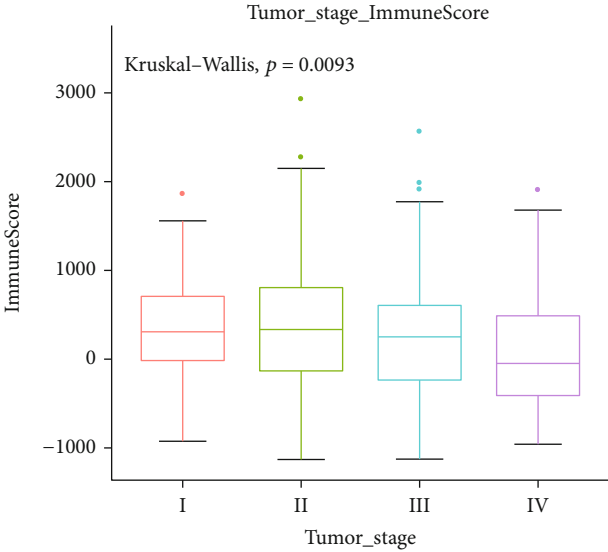


FIGURE 3: Continued.

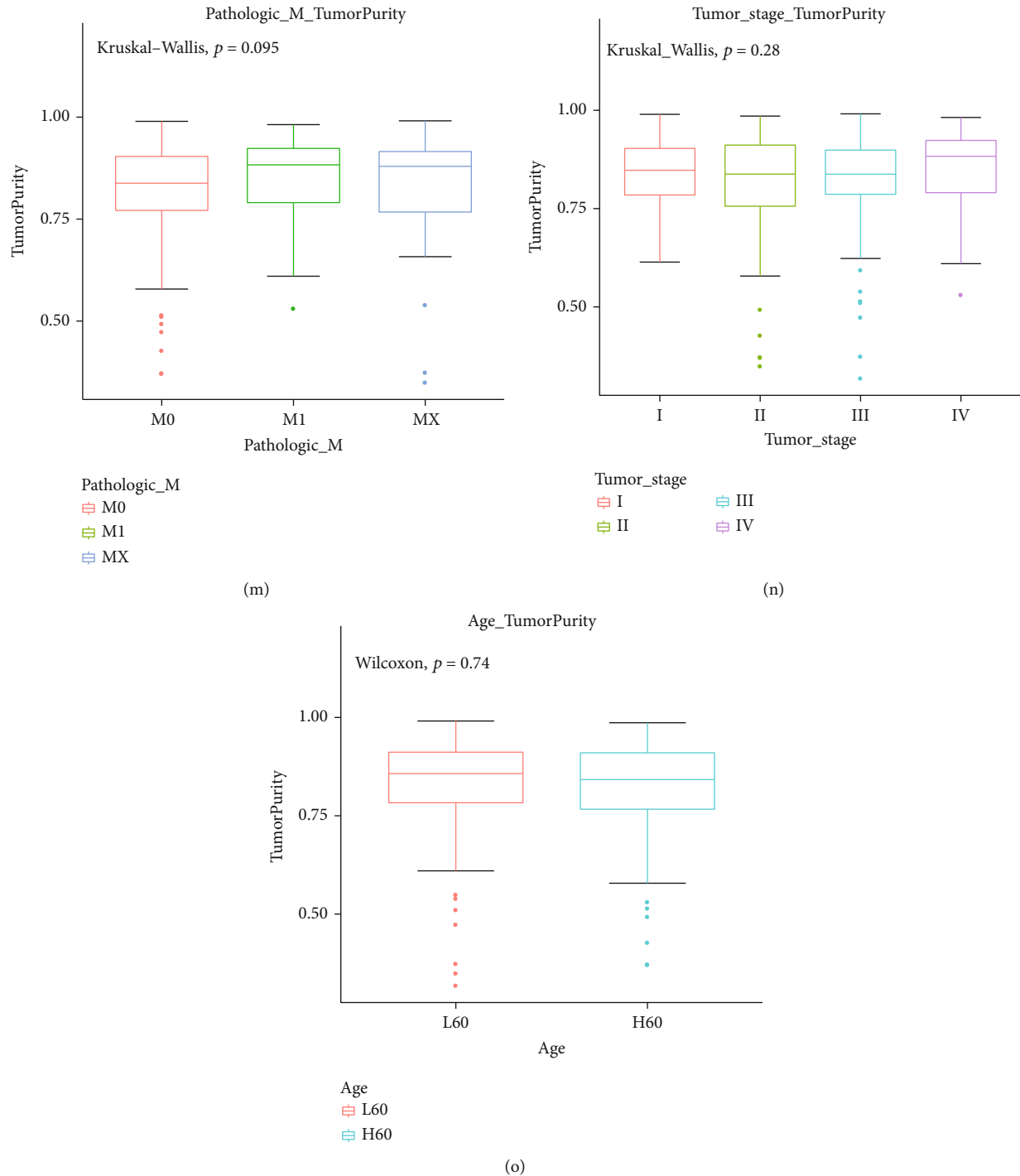


FIGURE 3: The correlation between stromal/immune/ESTIMATE scores and CRC patients' clinical features. Stromal score is not associated with (a) pathologic T (p value = 0.61), (b) pathologic N (p value = 0.28), (c) pathologic M (p value = 0.63), (d) tumor stage (p value = 0.68), and (e) age (p value = 0.76). Immune score is not correlated with (f) pathologic T (p value = 0.88) and (g) pathologic N (p value = 0.17). Immune score significantly associated with (h) pathologic M (p value = 0.0045) and (i) tumor stage (p value = 0.0093). Immune score is not correlated with (j) age (p value = 0.29). ESTIMATE scores are not correlated with (k) pathologic T (p value = 0.98), (l) pathologic N (p value = 0.73), (m) pathologic M (p value = 0.095), (n) tumor stage (p value = 0.28), and (o) age (p value = 0.74).

results showed that both stromal DEGs and immune DEGs could distinguish high stromal/immune score from low stromal/immune score (Figures 4(c) and 4(d)). 736 genes were

upregulated both in high stromal and immune scores (Figure 4(e)). Moreover, among eight downregulated immune genes, two genes were also downregulated in the

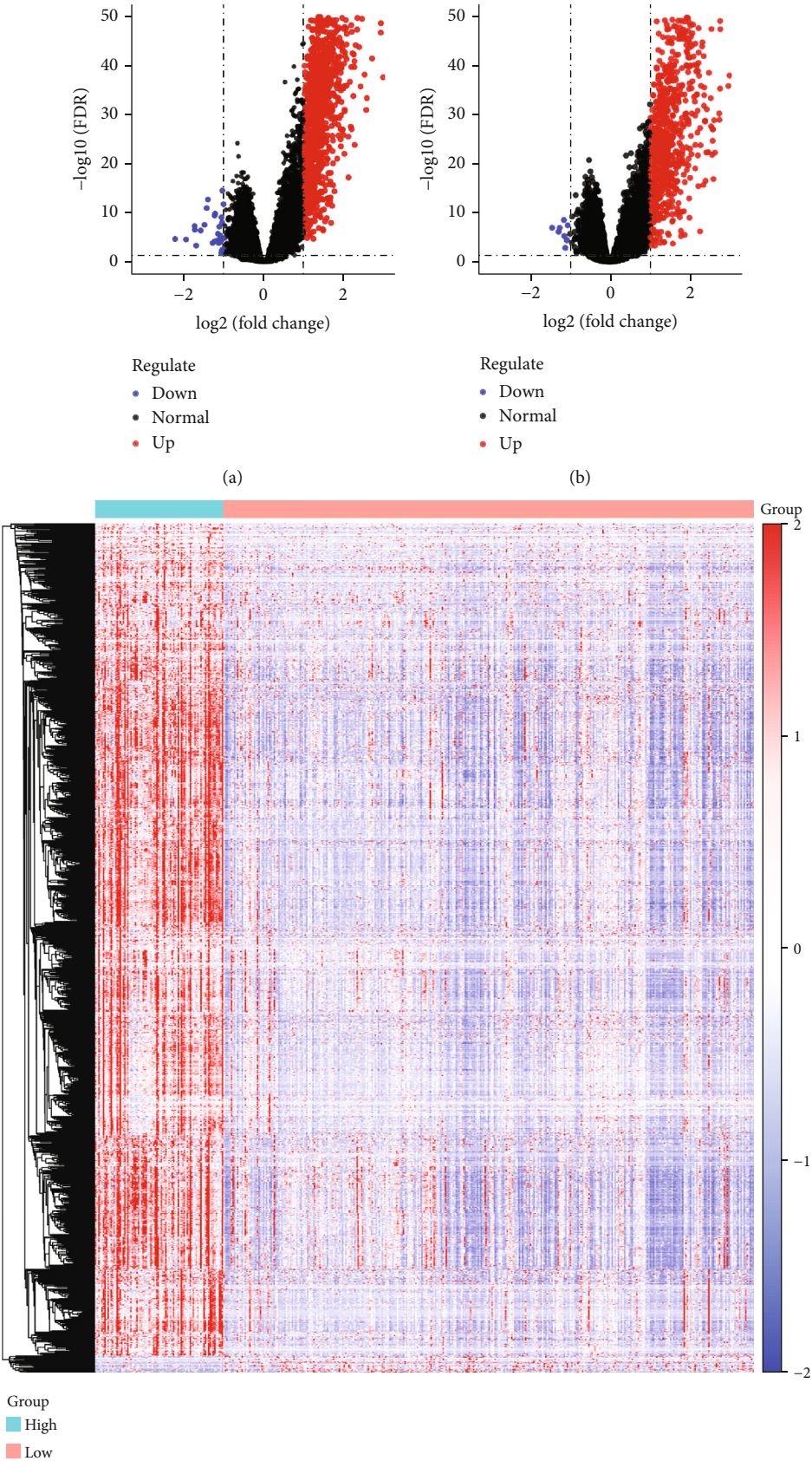


FIGURE 4: Continued.

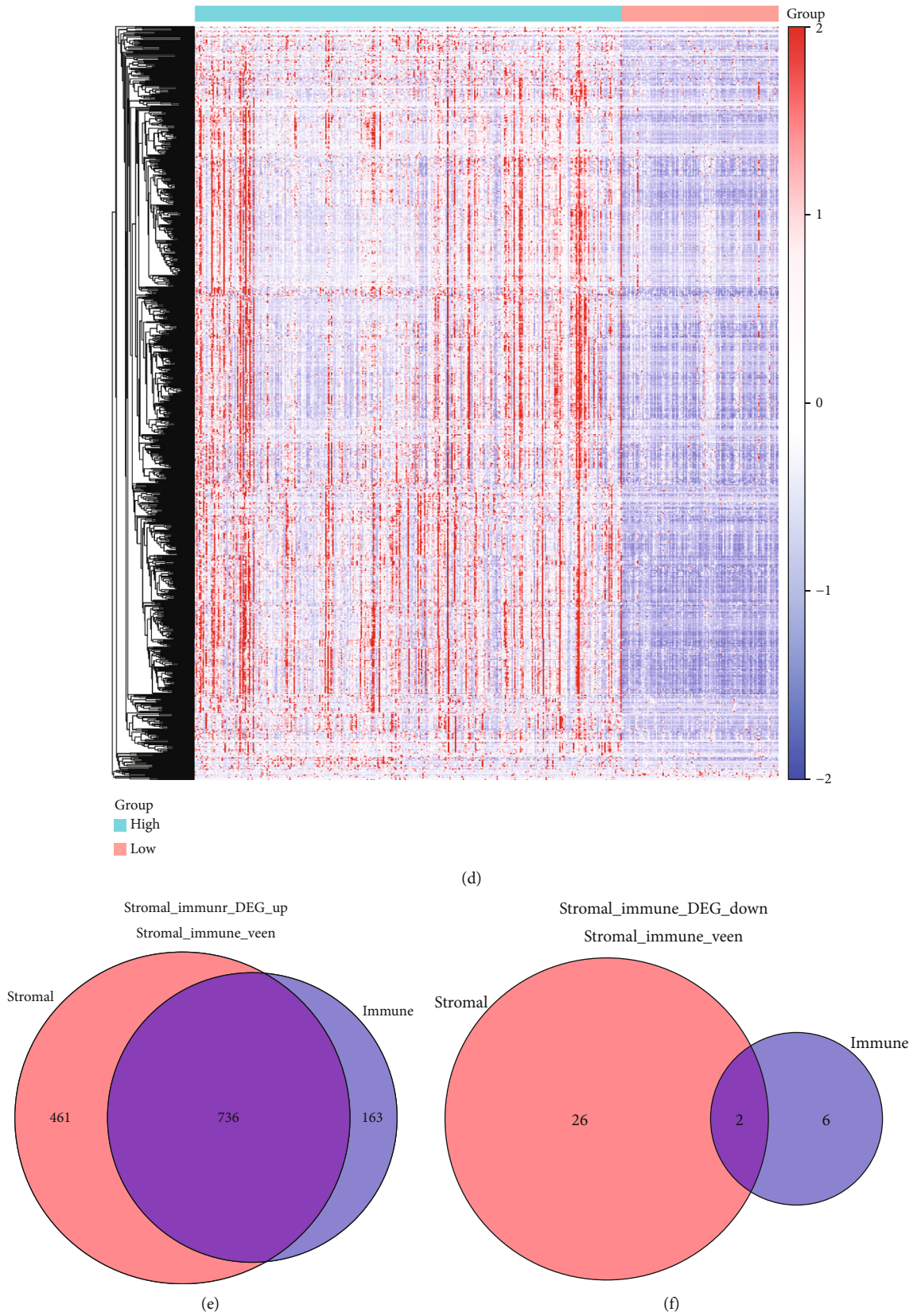
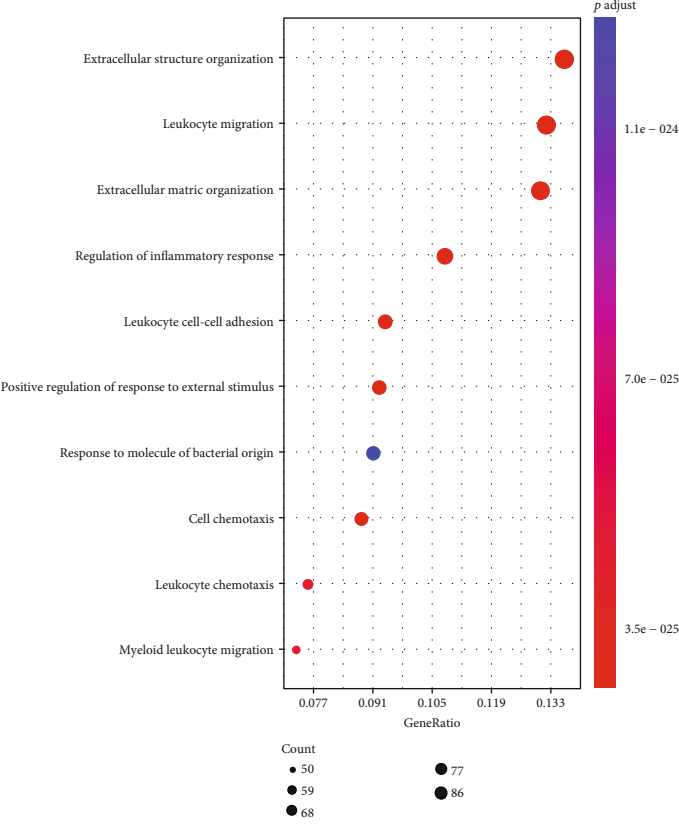
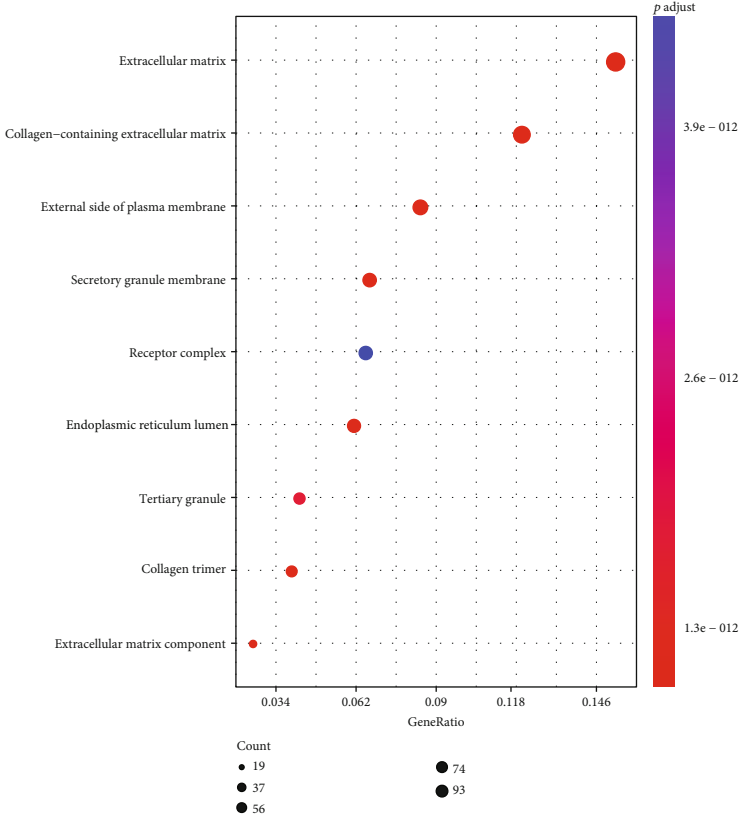


FIGURE 4: Identification of differentially expressed stromal and immune genes for CRC. Volcano plot showing up- (red) and downregulated (blue) stromal/immune genes in the high stromal score group (a) and immune score group (b). Heat maps depicting all DEGs in the high/low stromal score group (c) and high/low immune score group (d). Venn diagram showing common upregulated (e) and downregulated (f) stromal and immune genes.

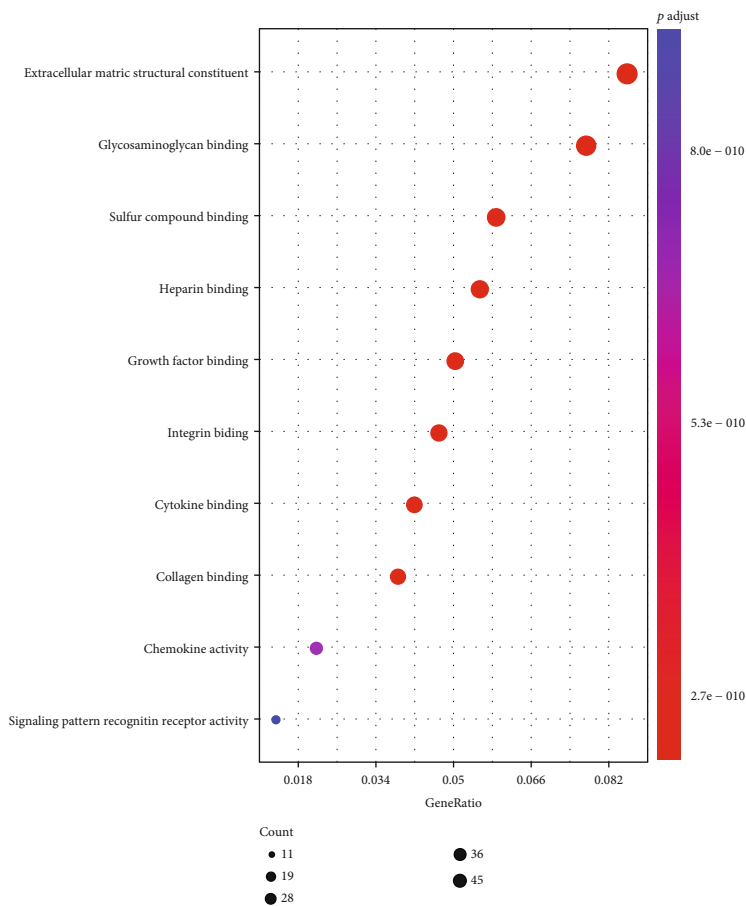


(a)

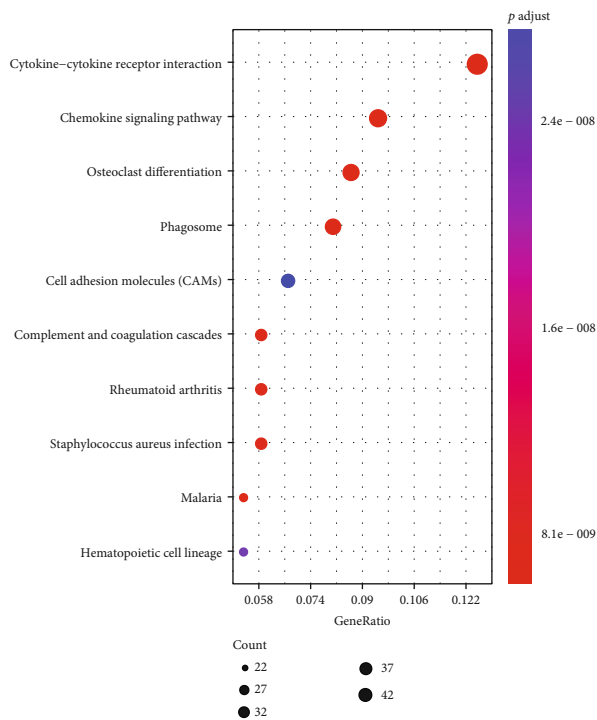


(b)

FIGURE 5: Continued.



(c)



(d)

FIGURE 5: Continued.

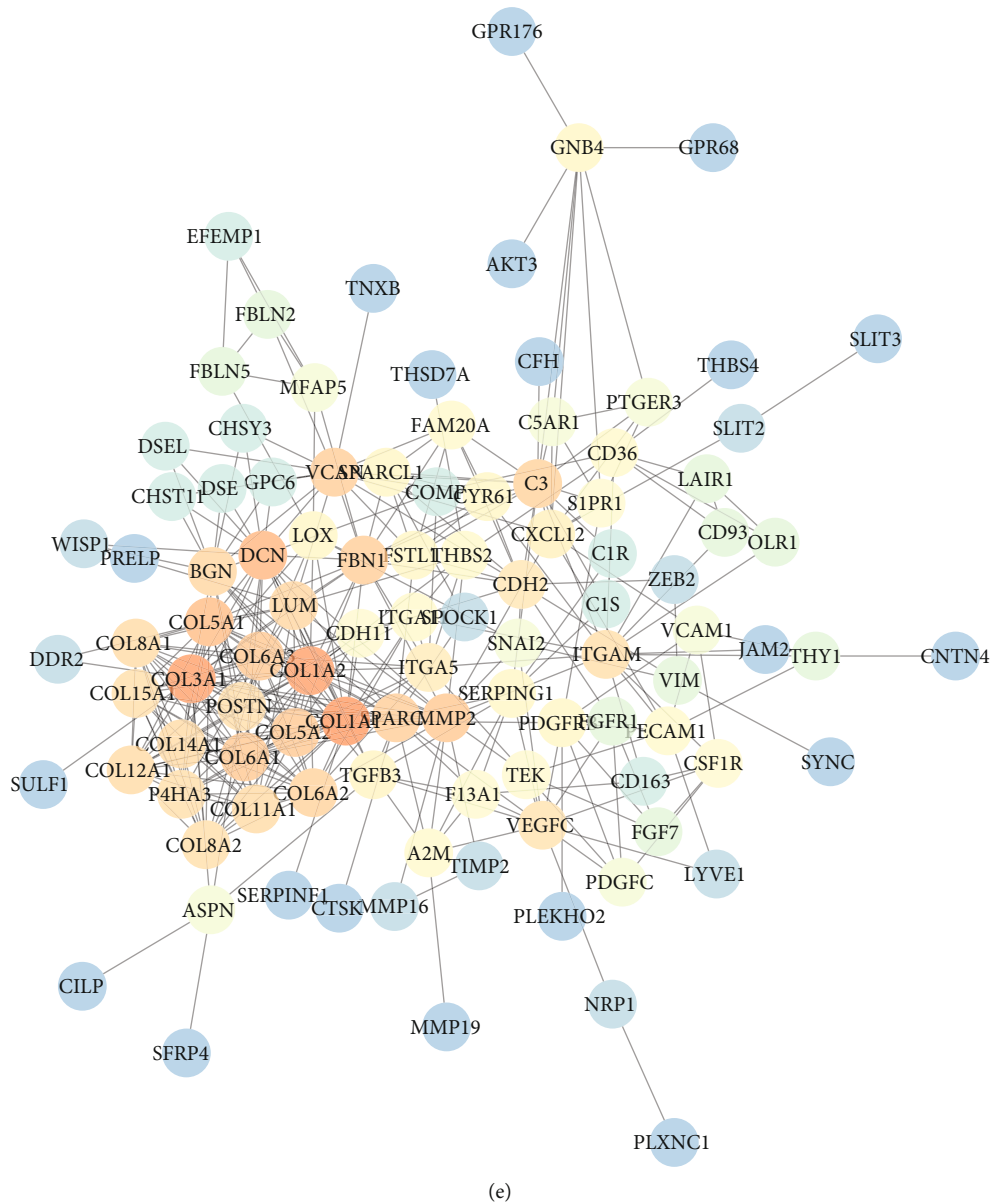


FIGURE 5: Functional enrichment and PPI analyses of common stromal and immune DEGs. (a) The top the BP terms of GO enrichment analysis. (b) The top ten CC terms of GO enrichment analysis. (c) The top ten MF terms of GO enrichment analysis. (d) The top ten KEGG pathway analysis. (e) PPI network construction.

high stromal score group (Figure 4(f)). We further performed functional enrichment analysis of these common stromal and immune genes. These genes were significantly correlated with inflammatory biological processes like regulation of inflammatory response and pathways such as cytokine-cytokine receptor interaction and chemokine signaling pathway (Figures 5(a)–5(d)). As shown in the PPI network, COL6A2, COL6A1, COL5A2, C1S, and C1R were the top five genes, which were considered hub genes (Figure 5(e)).

3.3. Construction of an Eight-Gene Prognostic Signature for CRC. Among 738 differentially expressed stromal and

immune genes, 23 genes were significantly associated with CRC patients’ prognosis according to univariate Cox regression analysis results. Of them, 20 genes were risk factors, and the remaining three (CCL22, CPA3, and MMP1) were protective factors (Table 2). These genes were used for multivariate Cox regression analysis. Finally, an eight-gene signature was constructed for CRC. The risk score was calculated on the basis of the coefficients and expression values of the eight genes. All CRC patients were divided into two groups in accordance with the median value of risk score (Figure 6(a)). Heat maps depicted the difference in expression patterns of the eight genes (including CD36, KCNE4, CPT1C, SLC2A3, RASGRP2, NFATC1, CCL22, and CPA3)

TABLE 2: Univariate Cox regression analysis results of 23 differentially expressed stromal and immune genes.

| Variables | HR | Lower 95% CI | Upper 95% CI | p value |
|-----------|-------|--------------|--------------|----------|
| CD36 | 1.381 | 1.046 | 1.823 | 0.02284 |
| KCNE4 | 1.346 | 1.004 | 1.805 | 0.047303 |
| VEGFC | 1.39 | 1.011 | 1.912 | 0.042907 |
| PDE1B | 2.242 | 1.361 | 3.693 | 0.001518 |
| BGN | 1.166 | 1.003 | 1.355 | 0.046135 |
| CPT1C | 2.561 | 1.476 | 4.446 | 8.26E-04 |
| GPX3 | 1.231 | 1.03 | 1.471 | 0.022285 |
| NGFR | 1.344 | 1.014 | 1.779 | 0.039388 |
| SERPINE1 | 1.173 | 1.01 | 1.362 | 0.037066 |
| CHST1 | 1.448 | 1.02 | 2.055 | 0.038579 |
| FBLN7 | 2.648 | 1.003 | 6.994 | 0.049317 |
| KCNJ8 | 1.372 | 1.003 | 1.876 | 0.047805 |
| SLC2A3 | 1.225 | 1.007 | 1.49 | 0.041957 |
| CD72 | 1.714 | 1.127 | 2.606 | 0.011724 |
| APLP1 | 1.631 | 1.031 | 2.58 | 0.036535 |
| SIGLEC1 | 1.348 | 1.019 | 1.783 | 0.036725 |
| RASGRP2 | 1.641 | 1.044 | 2.579 | 0.031925 |
| SPHK1 | 1.218 | 1.002 | 1.481 | 0.048069 |
| NFATC1 | 1.674 | 1.157 | 2.421 | 0.006228 |
| LRRN2 | 1.653 | 1.139 | 2.398 | 0.008126 |
| CCL22 | 0.686 | 0.51 | 0.923 | 0.012756 |
| CPA3 | 0.81 | 0.68 | 0.966 | 0.019315 |
| MMP1 | 0.902 | 0.816 | 0.996 | 0.042396 |

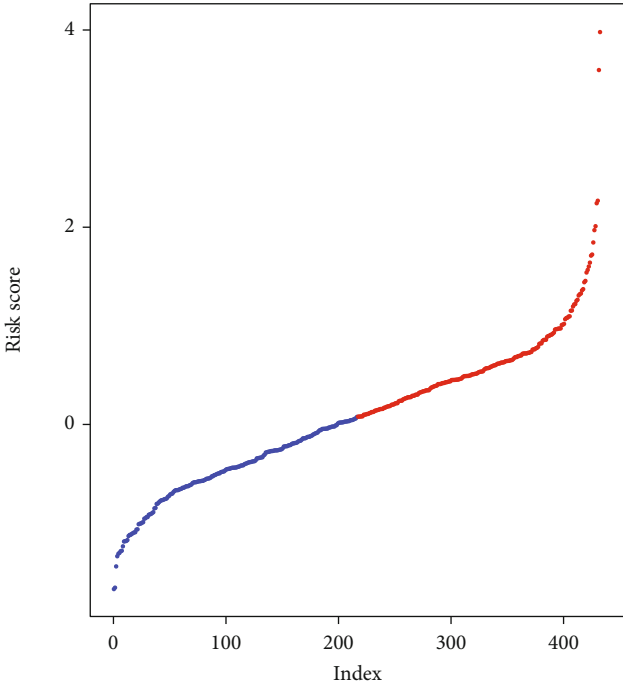
between the high and low risk scores (Figure 6(b)). As shown in Figure 6(c), the risk score was capable of predicting CRC patients' prognosis. High risk score implied a poor prognosis (p value < 0.0001). Among the eight genes, KCNE4 and CCL22 were protective factors of CRC, while CD36, CPT1C, SLC2A3, RASGRP2, NFATC1, and CPA3 were risk factors of CRC, as shown in the forest diagram (Figure 6(d)). We further validated the sensitivity and accuracy of the model. AUCs of the model for 3 years, 5 years, OS, and PFI were 0.71, 0.70, 0.73, and 0.66, respectively (Figures 6(e) and 6(f)). Thus, the risk score model could well predict CRC patients' prognosis, with high sensitivity and accuracy. As shown in the multivariate Cox regression analysis, the model and MMR could become independent factors for CRC prognosis after adjustment of other clinical characteristics (Table 3).

3.4. Eight Genes in the Risk Score Model Are Significantly Associated with CRC Patients' Prognosis. Box plot depicted the difference in expression patterns of CCL22 (Figure 7(a)), CD36 (Figure 7(b)), CPA3 (Figure 7(c)), CPT1C (Figure 7(d)), KCNE4 (Figure 7(e)), NFATC1 (Figure 7(f)), RASGRP2 (Figure 7(g)), and SLC2A3 (Figure 7(h)) between the high risk score and low risk score. Among them, CCL22 (p value $< 2.22e-16$), CPA3 (p value $< 2.22e-16$), CPT1C (p value = 0.00078), KCNE4 (p value =

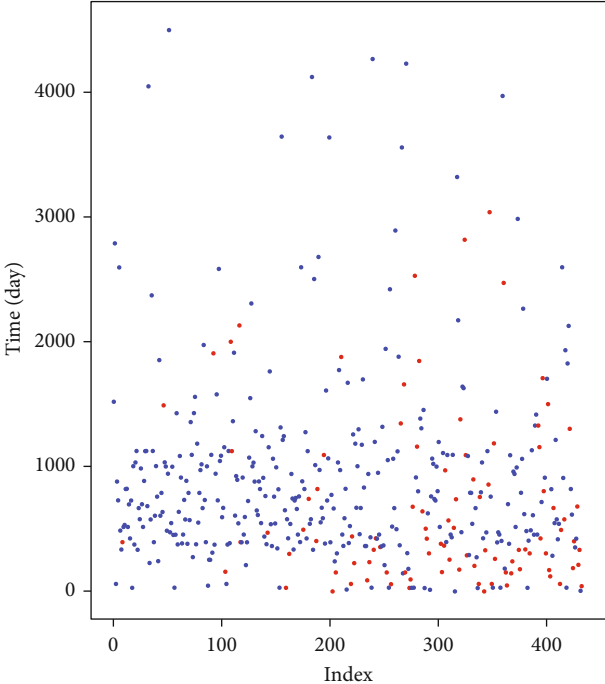
0.023), NFATC1 (p value = 0.00062), and SLC2A3 (p value = 0.00081) were differentially expressed between the high and low risk scores. Furthermore, the expression levels of these genes between CRC samples and normal samples were visualized (Figures 8(a)–8(g)). CD36 (p value $< 2.22e-16$), CPA3 (p value $< 2.22e-16$), NFATC1 (p value = $9.1e-08$), and RASGRP2 (p value $< 2.22e-16$) were highly expressed and SLC2A3 (p value = 0.0015) was lowly expressed in tumor samples. As shown in Figures 9(a)–9(h), low expression of CCL22 (p value = 0.0047) and CPA3 (p value = 0.035) indicated shorter OS time than high expression. Moreover, we found that highly expressed CPT1C (p value = 0.0017), KCNE4 (p value = 0.002), and SLC2A3 (p value = 0.0048) was significantly correlated with poor PFS (Figures 9(i)–9(p)).

3.5. The Eight-Gene Model Is in Significant Correlation with Immune Cell Infiltration. The correlation between the model and the infiltrating levels of six immune cells was analyzed. We found that the model was in significant association with the infiltrating levels of six immune cells, including B cell (Figure 10(a); $R = 0.13$, p value = 0.0064) and CD4+T cell (Figure 10(b); $R = 0.21$, p value = $4.3e-06$). However, no significant correlation between the model and CD8+T cell was found in Figure 10(c) ($R = 0.045$, p value = 0.34). Furthermore, there were distinct correlations between the model and dendritic cell (Figure 10(d); $R = 0.12$, p value = 0.0072), macrophage (Figure 10(e); $R = 0.19$, p value = $3.3e-05$), neutrophil (Figure 10(f); $R = 0.18$, p value = $9.4e-05$). We also found that the expression levels of the eight genes in the model were significantly correlated with the infiltrating levels of six immune cells, including CCL22 (Figure 11(a)), CD36 (Figure 11(b)), CPA3 (Figure 11(c)), CPT1C (Figure 11(d)), KCNE4 (Figure 11(e)), NFATC1 (Figure 11(f)), RASGRP2 (Figure 11(g)), and SLC2A3 (Figure 11(h)). Moreover, the eight genes were in significant association with markers of immune cells (Supplementary Table 1). These results suggested that the model could be in association with immune cell infiltration.

3.6. Validation of the Risk Score Using an Independent Dataset. Based on 566 CRC samples from the GSE39582 dataset, the prognostic value of the risk score was validated. The risk score distribution and survival status of CRC patients are shown in Figure 12(a). Heat maps showed the expression differences of CD36, KCNE4, CPT1C, SLC2A3, RASGRP2, NFATC1, CCL22, and CPA3 between the high and low risk scores (Figure 12(b)). As expected, CRC patients with high risk score had a poorer prognosis than those with low risk score (Figure 12(c)). Among the eight genes, CD36, NFATC1, and CCL22 were significantly associated with prognosis of CRC patients (Figure 12(d)). AUCs of the model for 3 years and 5 years were 0.65 and 0.66, respectively (Figure 12(e)), indicating that the risk score could well predict CRC patients' prognosis. The expression levels of CCL22 (Figure 13(a)), CD36 (Figure 13(b)), CPA3 (Figure 13(c)), CPT1C (Figure 13(d)), KCNE4 (Figure 13(e)), NFATC1 (Figure 13(f)), RASGRP2 (Figure 13(g)), and SLC2A3



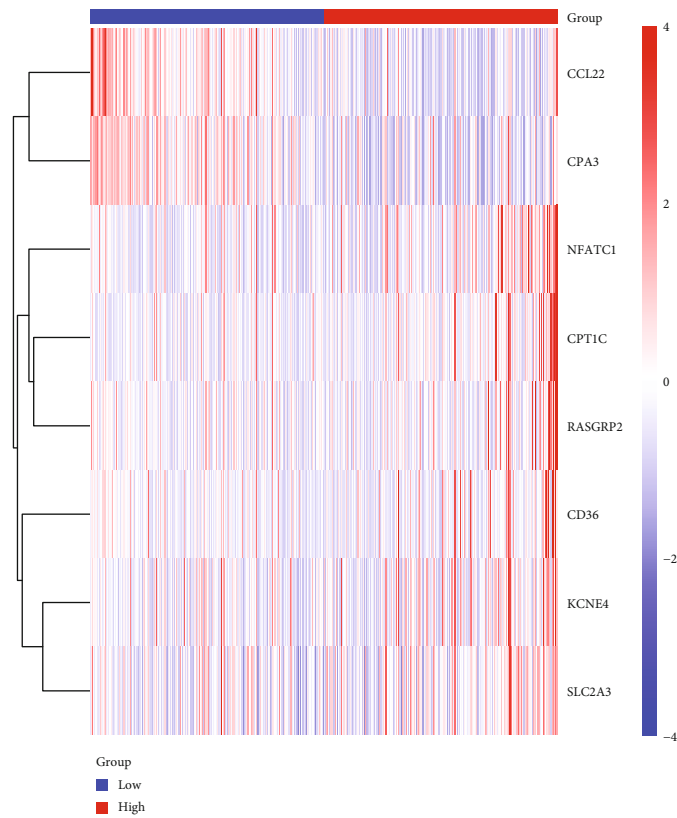
Group
● High
● Low



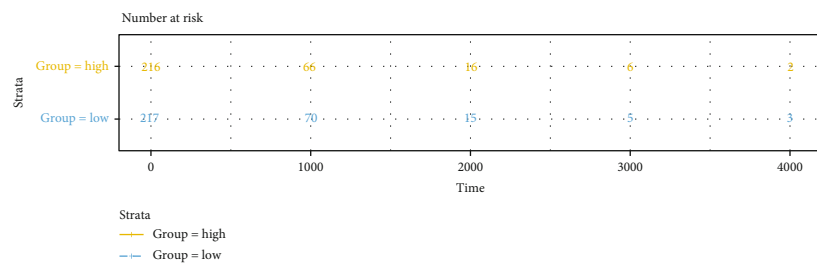
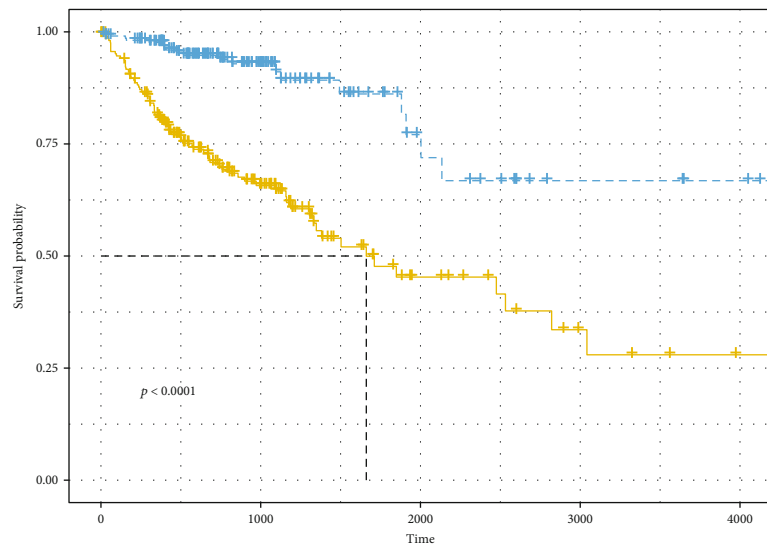
Status
● Alive
● Dead

(a)

FIGURE 6: Continued.



(b)



(c)

FIGURE 6: Continued.

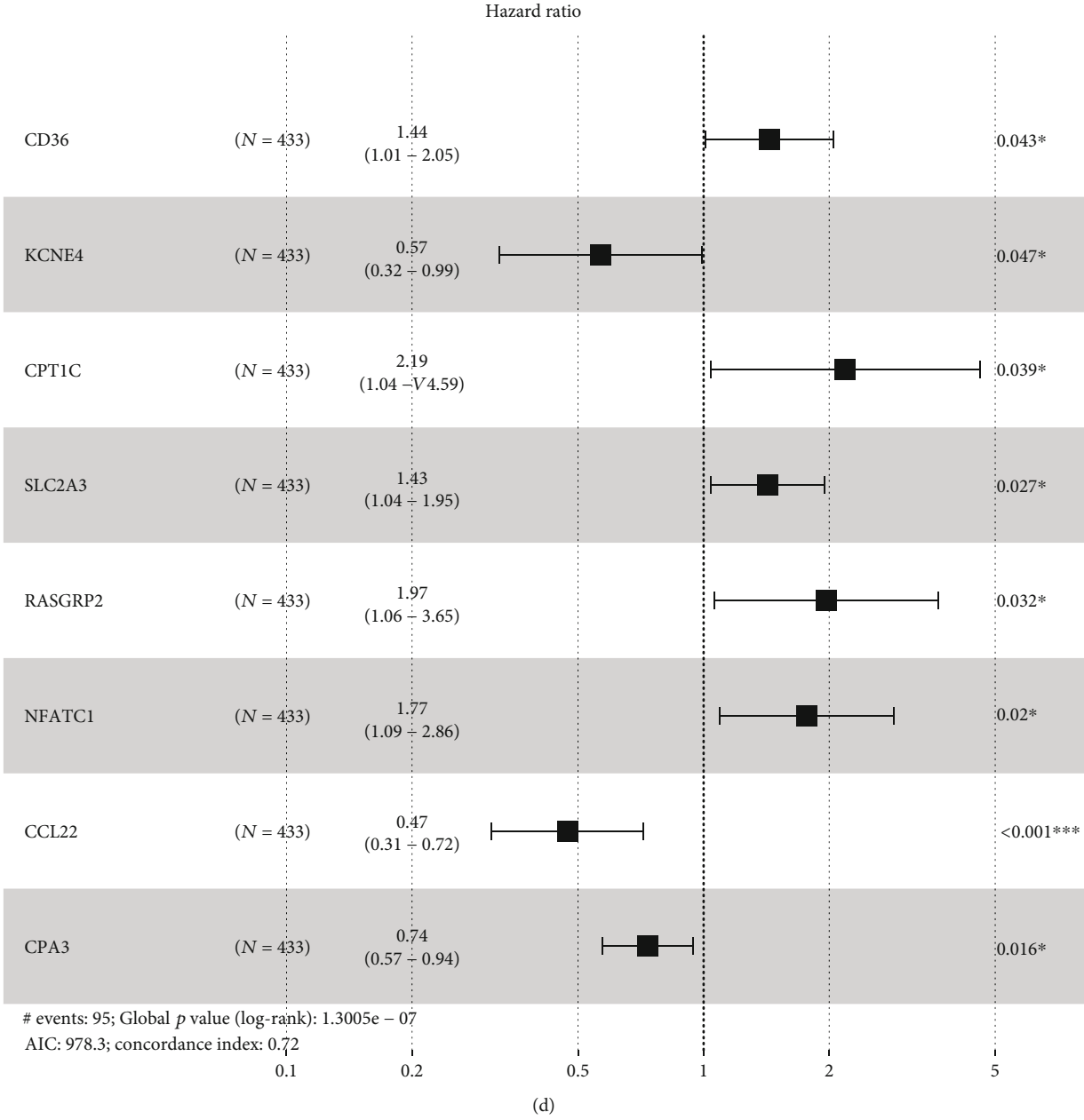


FIGURE 6: Continued.

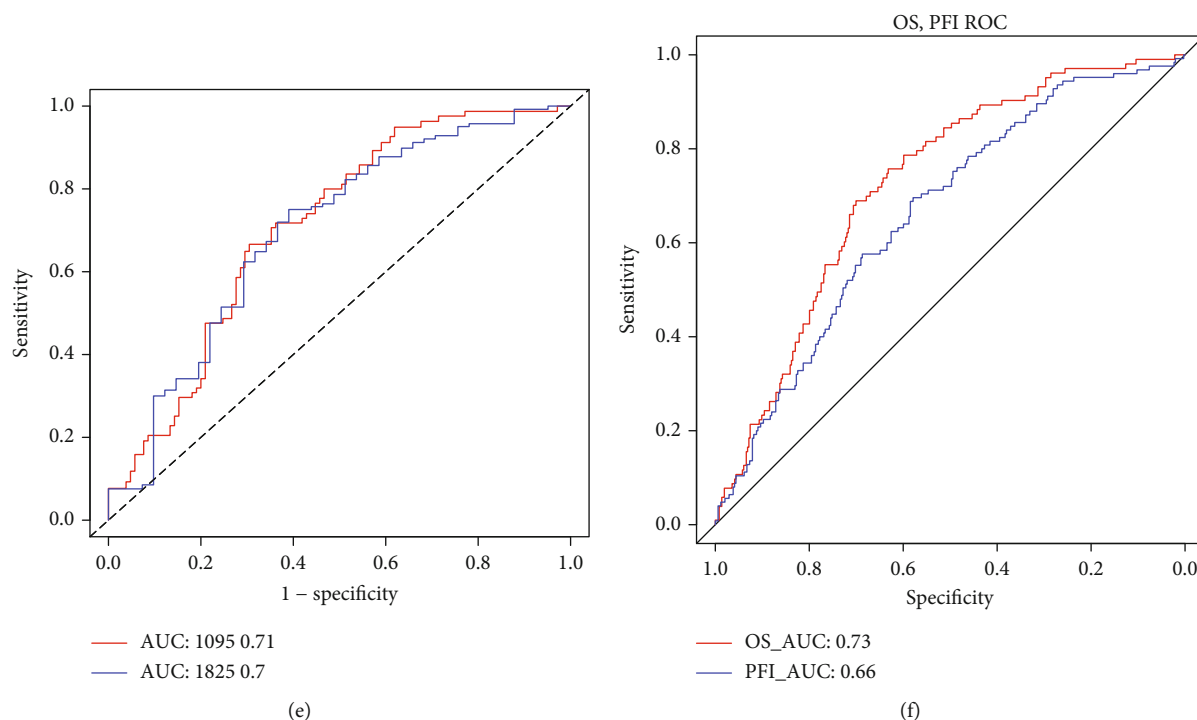


FIGURE 6: Construction of an eight-gene prognostic signature for CRC. (a) Risk score distribution and survival status. (b) Heat maps showing the expression patterns of the eight genes between high and low risk score. (c) Kaplan-Meier survival analysis of the model. (d) Forest plot of the eight genes for CRC. (e, f) ROC curve of the model for 3-year, 5-year, OS, and PFI.

TABLE 3: Univariate and multivariate Cox regression analyses in a TCGA-CRC cohort.

| Characteristics | Univariate analysis | | Multivariate analysis | |
|-----------------|---------------------|----------------|-----------------------|----------------|
| | HR (95% CI) | <i>p</i> value | HR (95% CI) | <i>p</i> value |
| Stromal score | 1 (1-1) | 0.653 | NA | NA |
| Immune score | 1 (1-1) | 0.941 | NA | NA |
| Age | 1.396 (0.878-2.22) | 0.158 | NA | NA |
| Gender | 1.127 (0.751-1.692) | 0.564 | NA | NA |
| Tumor stage | 3.064 (1.986-4.726) | <0.0001 | 3.320 (0.870-12.640) | 0.079 |
| Pathologic T | 3.204 (1.398-7.345) | 0.006 | 4.914 (0) | 0.996 |
| Pathologic N | 2.581 (1.705-3.909) | <0.0001 | 0.920 (0.290-2.890) | 0.880 |
| Pathologic M | 3.519 (2.312-5.356) | <0.0001 | 1.620 (0.84-3.13) | 0.151 |
| MMR | 0.181 (0.044-0.751) | 0.019 | 0.070 (0.010-0.500) | 0.009 |
| BRAF | 1.108 (0.620-1.980) | 0.729 | NA | NA |
| KRAS | 0.912 (0.580-1.434) | 0.691 | NA | NA |
| TP53 | 1.461 (0.884-2.417) | 0.139 | NA | NA |
| MSI | 0.907 (0.522-1.575) | 0.728 | NA | NA |
| Risk score | 2.718 (2.063-3.581) | <0.0001 | 2.420 (1.590-3.700) | <0.0001 |

NA: not available.

(Figure 13(h)) between the high risk score and low risk score were validated based on the 566 CRC samples. Univariate Cox regression analysis results showed that age, KRAS mutation, pathologic T, pathologic N, pathologic M, tumor stage, and risk score were notably associated with CRC patients' prognosis. After multivariate Cox

regression analysis, we found that age, KRAS mutation, pathologic M, and risk score could be independent prognostic factors for CRC (Table 4).

3.7. The Eight Genes in the Risk Score Are Distinctly Correlated with Molecular Markers of CRC Prognosis. In

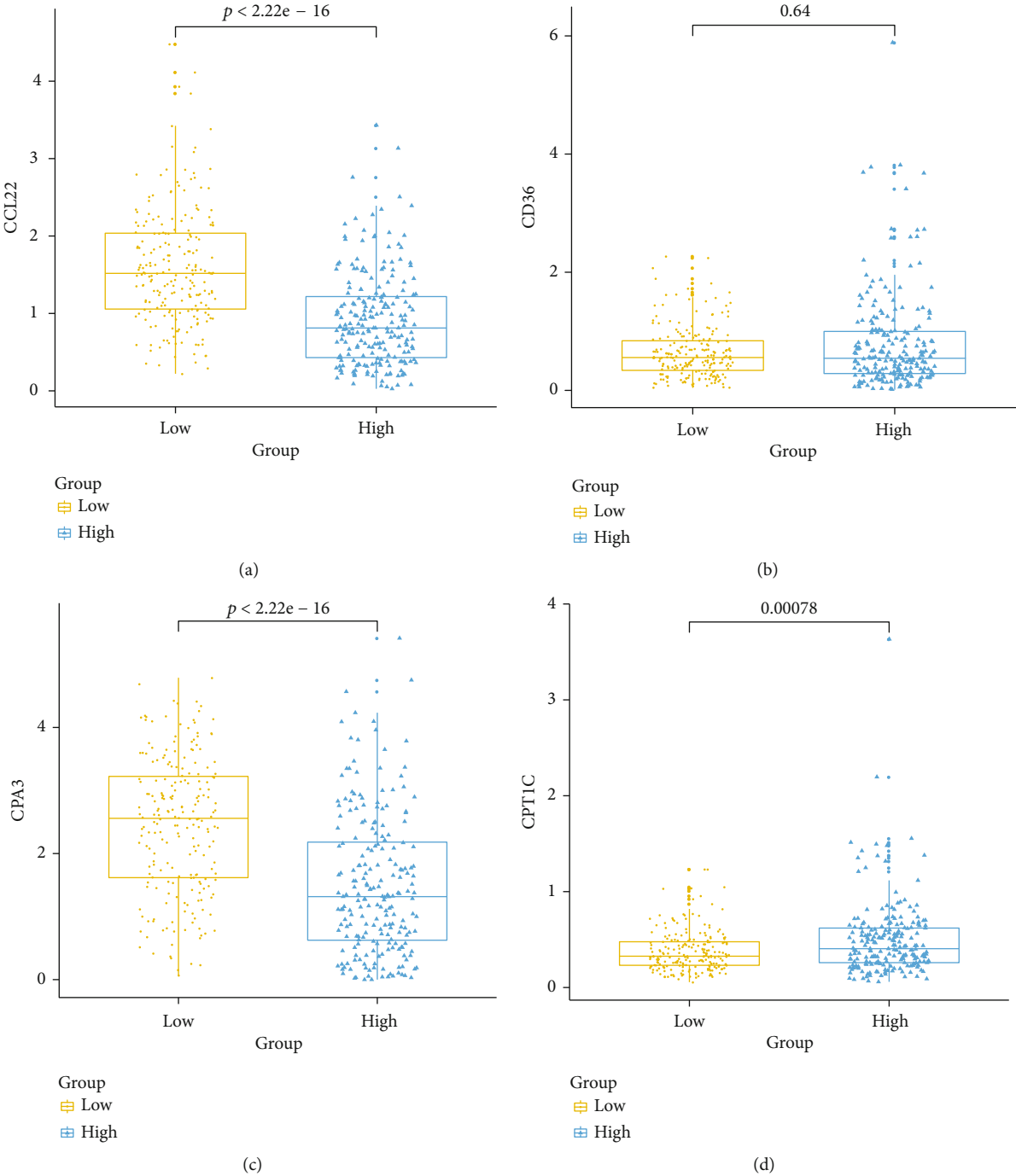


FIGURE 7: CONTINUED.

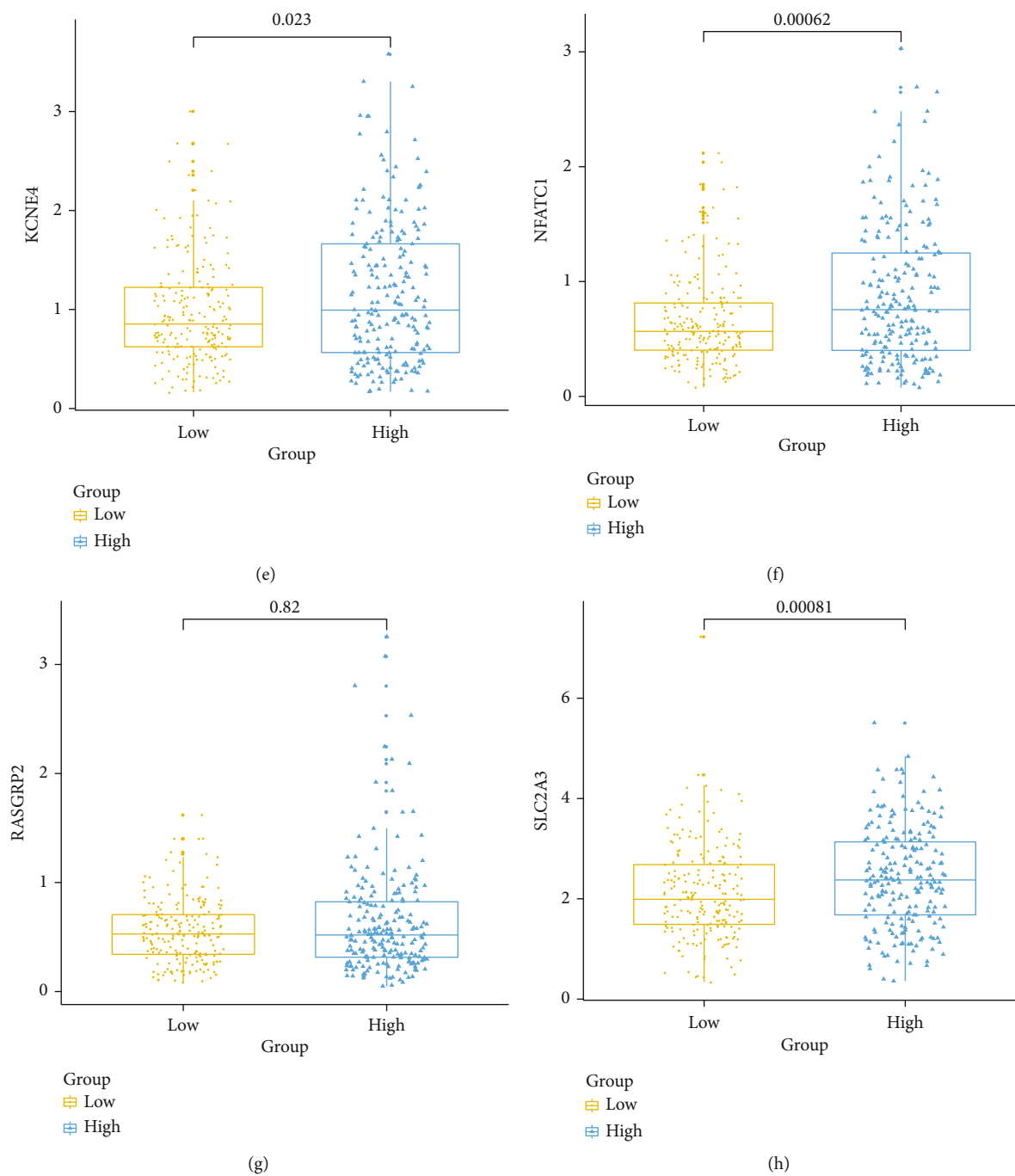
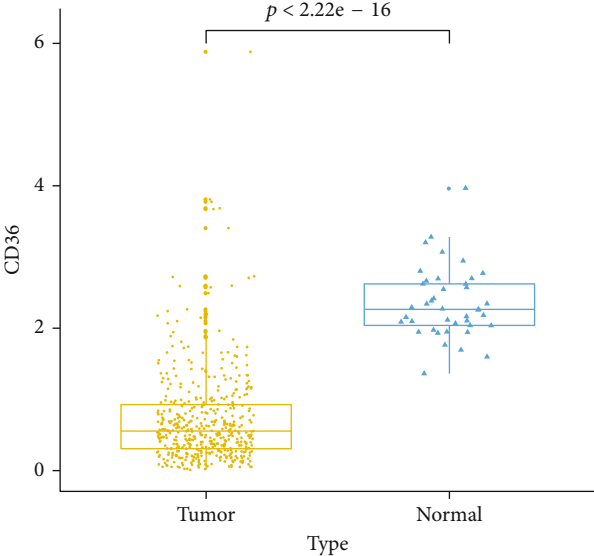
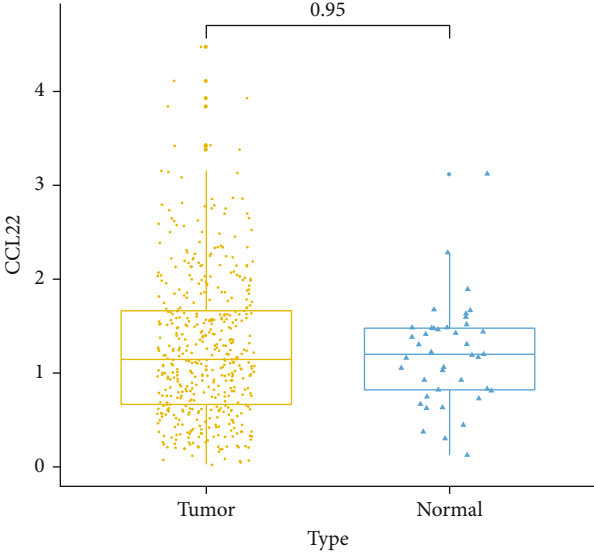


FIGURE 7: Box plots depicting the expression patterns of the eight genes in the risk score model between high and low risk score. (a) CCL22. (b) CD36. (c) CPA3. (d) CPT1C. (e) KCNE4. (f) NFATC1. (g) RASGRP2. (h) SLC2A3.

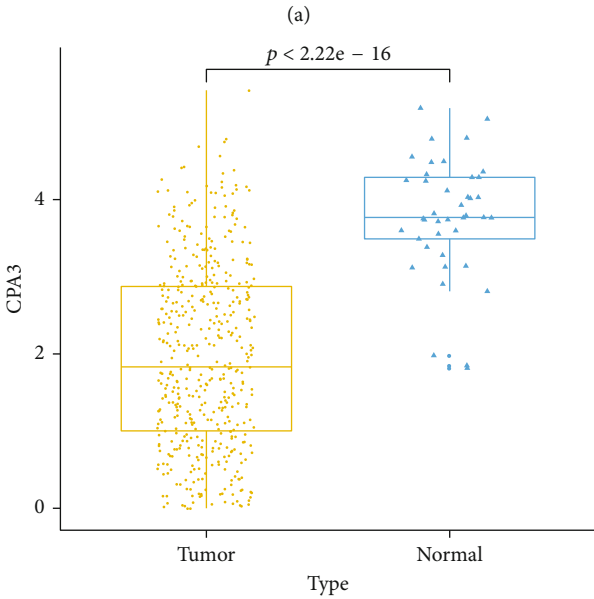
Figure 14(a), CCL22 was significantly correlated with BRAF mutation (p value = 0.014) and KRAS mutation (p value = 0.041). For CD36, there was a distinct correlation between its expression and KRAS mutation (p value = 0.00034) and MMR (p value = 0.025) in Figure 14(b). CPA3 (p value = 0.0066; Figure 14(c)) and CPT1C (p value = 0.005; Figure 14(d)) had higher expression levels in KRAS mutation. As shown in Figure 14(e), KCNE4 expression was in significant correlation with BRAF mutation (p value = 0.0014),

KRAS mutation (p value = 0.049), and MSI (p value = 0.05). NFATC1 expression was prominently correlated with BRAF mutation (p value = $2.2e-11$), KRAS mutation (p value = 0.00051) and MSI (p value = $1.1e-13$) in Figure 14(f). In Figure 14(g), RASGRP2 expression was significantly decreased in KRAS mutation. For SLC2A3, we found that there was a distinct correlation between its expression and BRAF mutation (p value = 0.0011) and MSI (p value = 0.00013) in Figure 14(h).

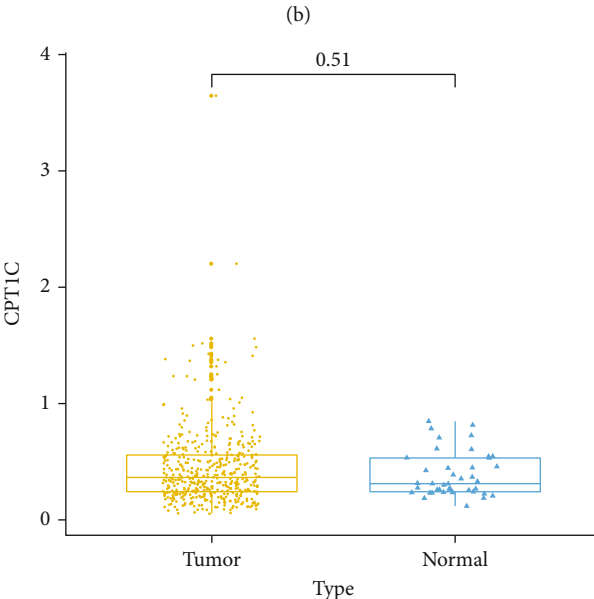


Type
Tumor
Normal

Type
Tumor
Normal



Type
Tumor
Normal



Type
Tumor
Normal

(c)

(d)

FIGURE 8: Continued.

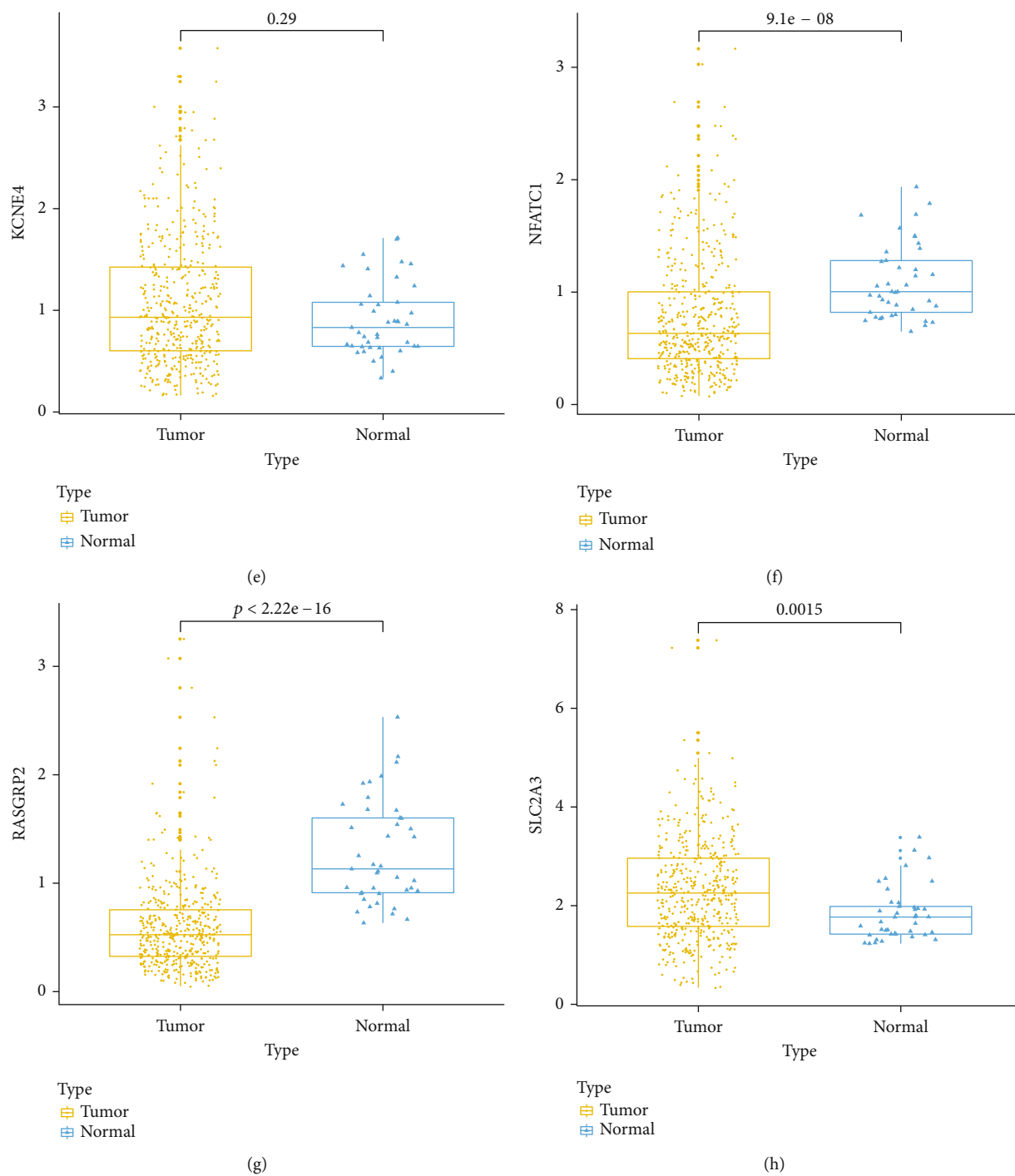
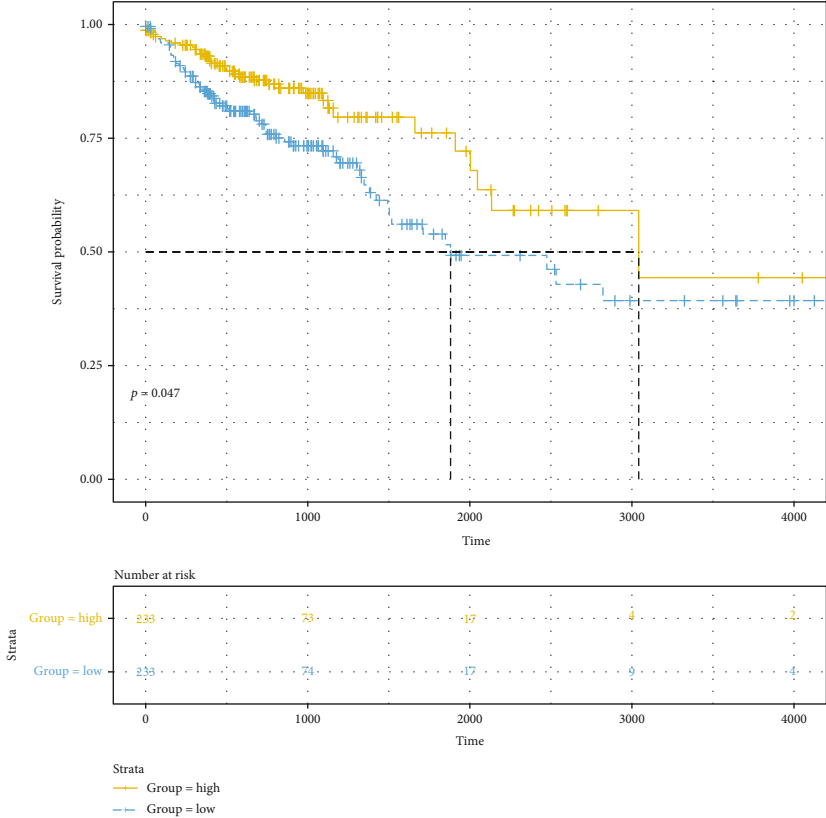


FIGURE 8: Box plots showing the expression patterns of the eight genes in the risk score model between CRC samples and normal samples. (a) CCL22. (b) CD36. (c) CPA3. (d) CPT1C. (e) KCNE4. (f) NFATC1. (g) RASGRP2. (h) SLC2A3.

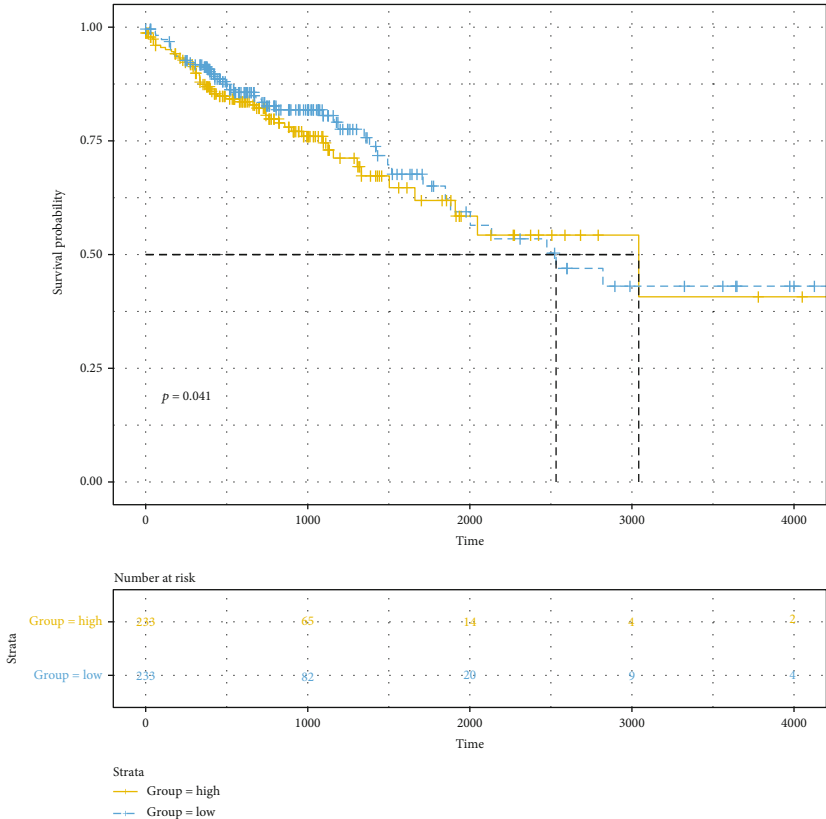
4. Discussion

In TME, stromal and immune cells are involved in the development of CRC. In this study; using the ESTIMATE algorithm, stromal and immune scores were calculated. A significant correlation between the immune score and CRC patients' prognosis was observed. Both the stromal score

and immune score were in significant correlation with clinical characteristics of CRC patients. Furthermore, we identified differentially expressed stromal and immune genes for CRC. Functional enrichment analysis results suggested that these genes were positively related with immune-related pathways like cytokine-cytokine receptor interaction [16, 17] and chemokine signaling pathway [18, 19].

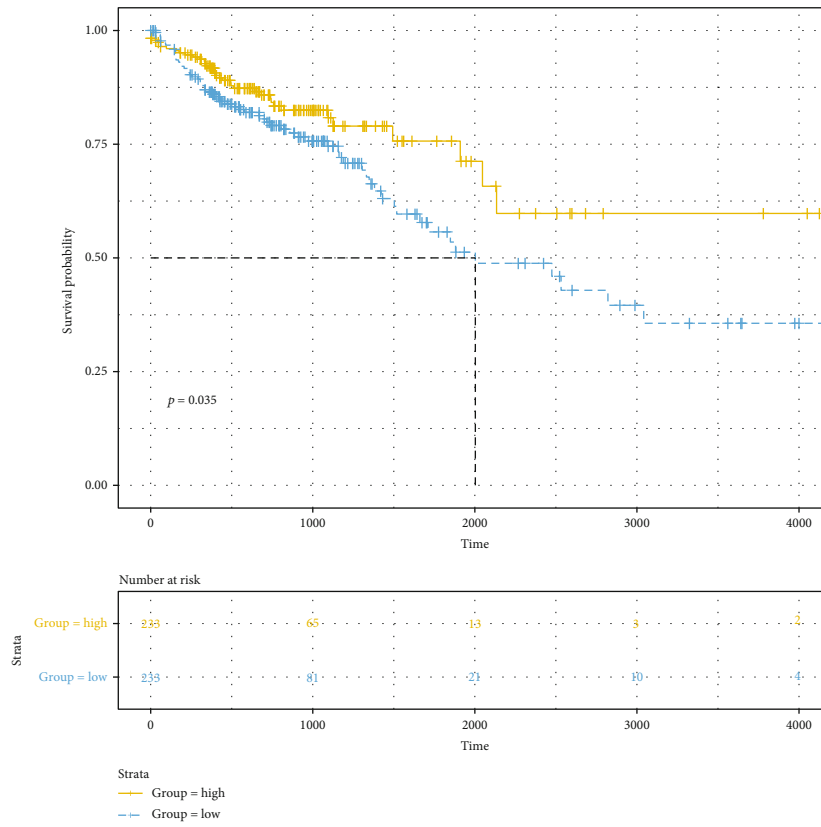


(a)

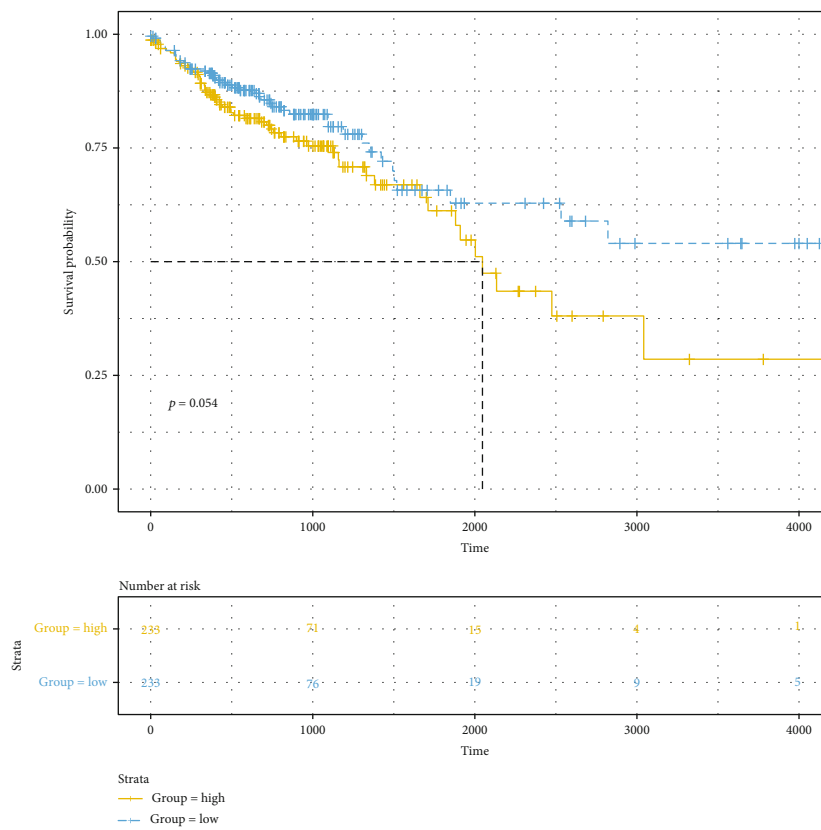


(b)

FIGURE 9: Continued.

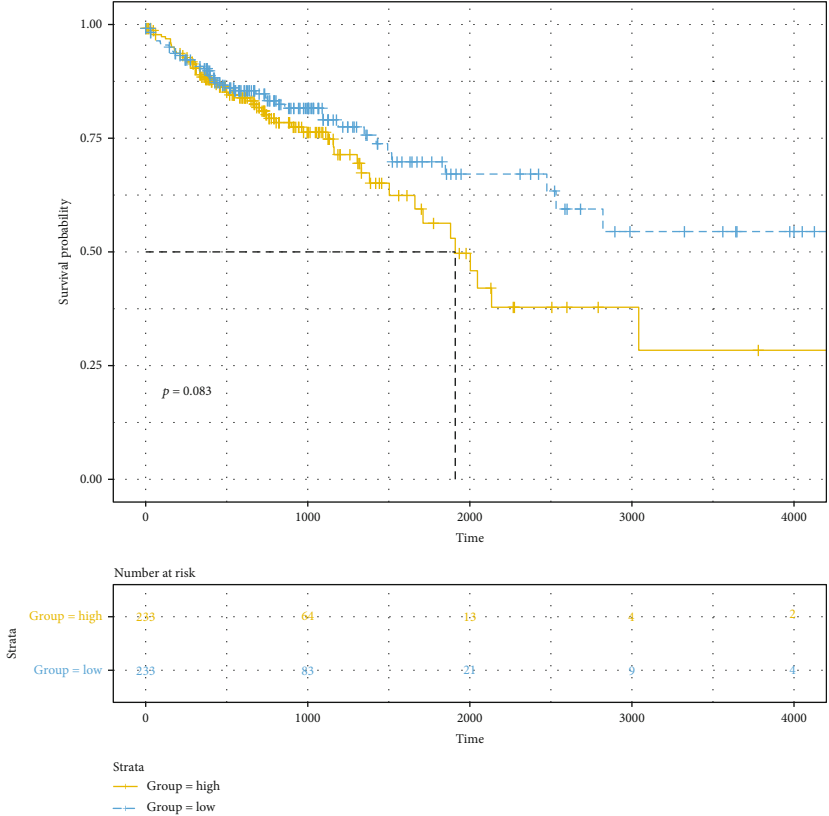


(c)

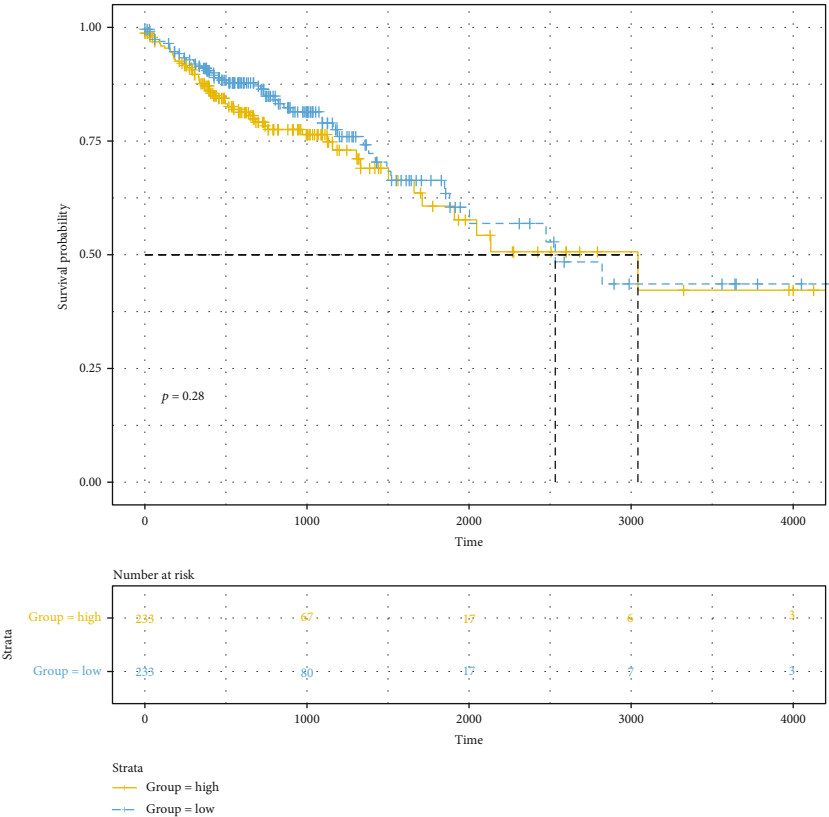


(d)

FIGURE 9: Continued.

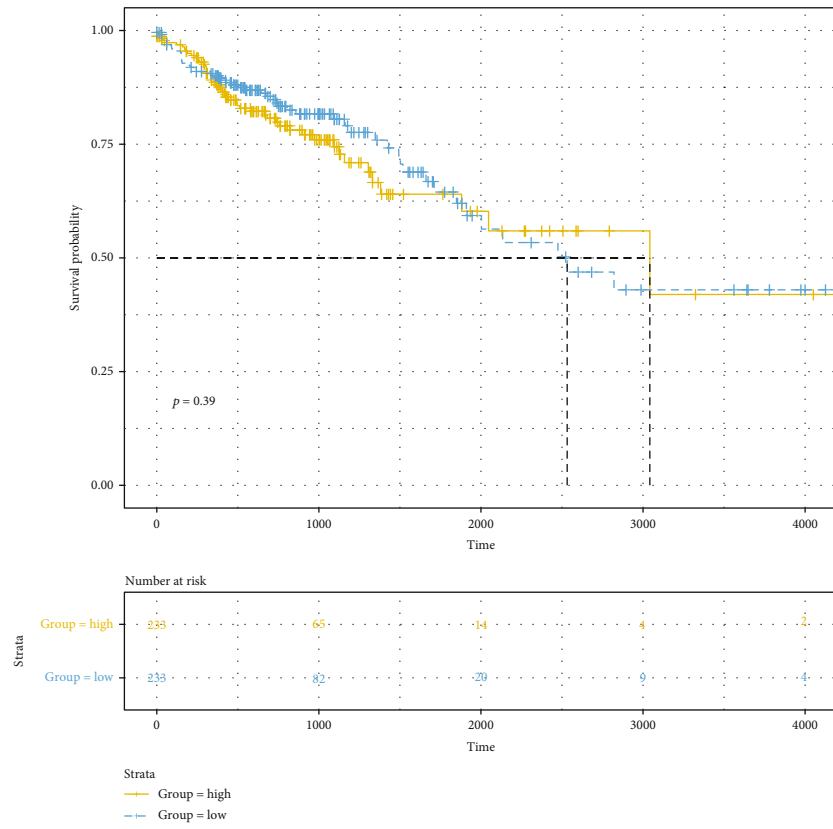


(e)

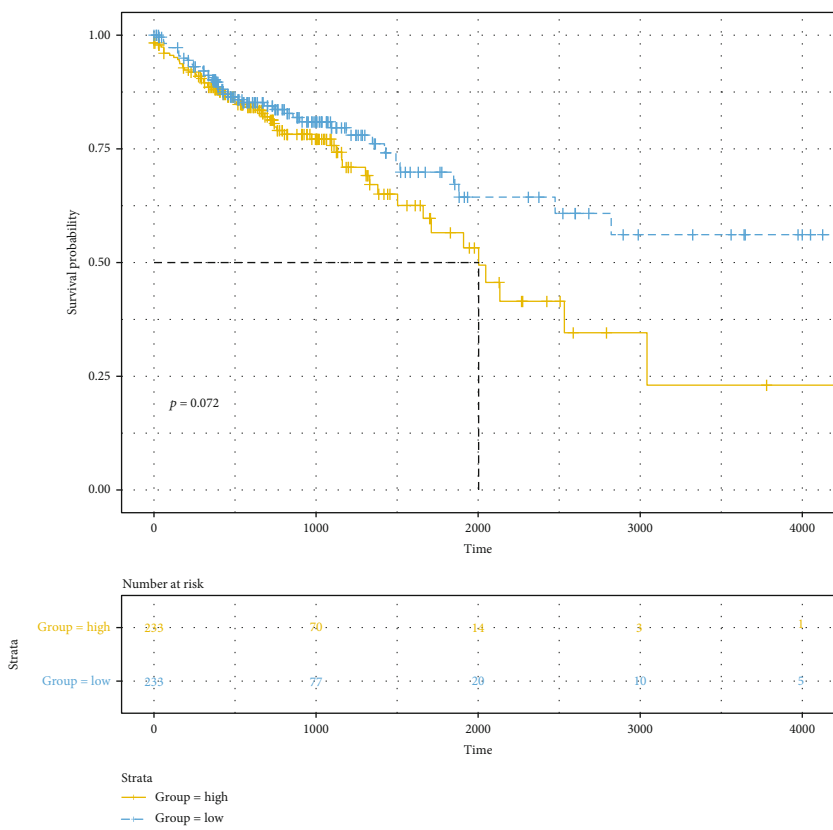


(f)

FIGURE 9: Continued.

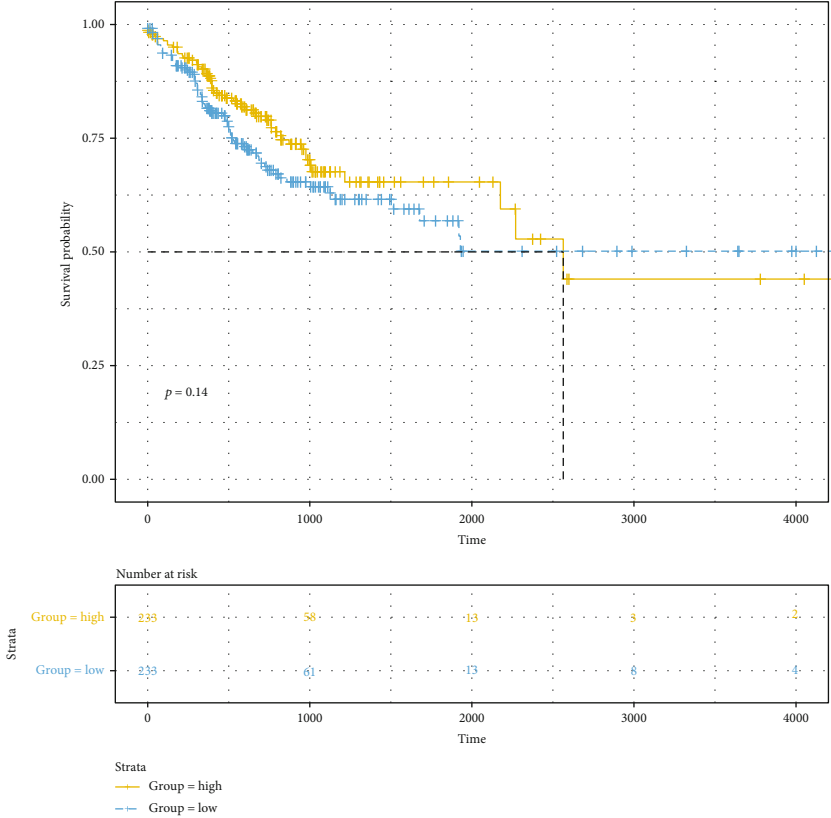


(g)

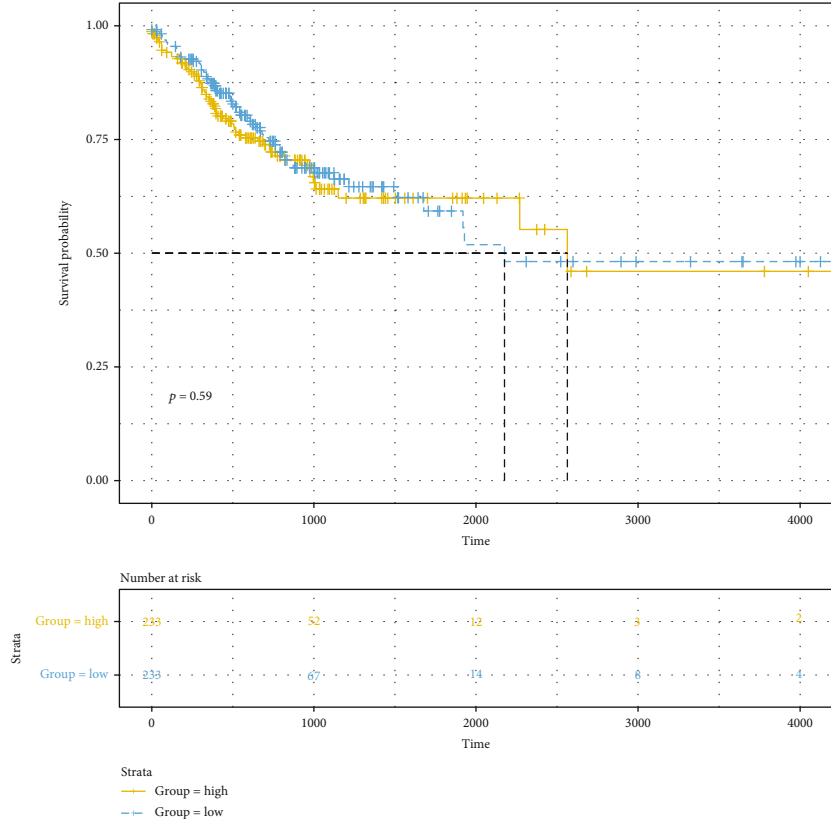


(h)

FIGURE 9: Continued.

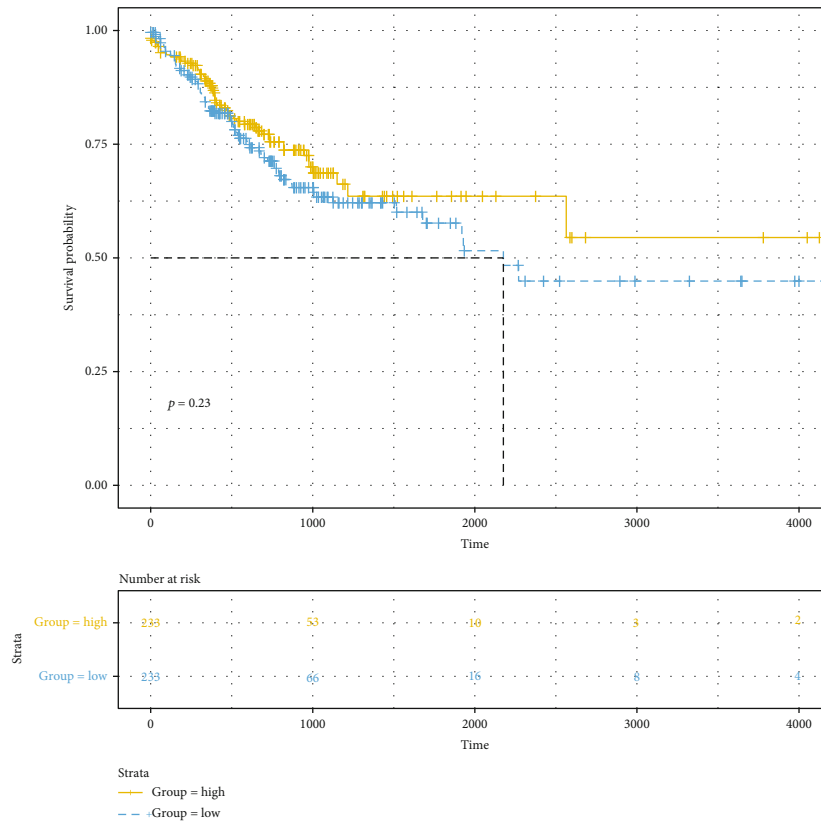


(i)

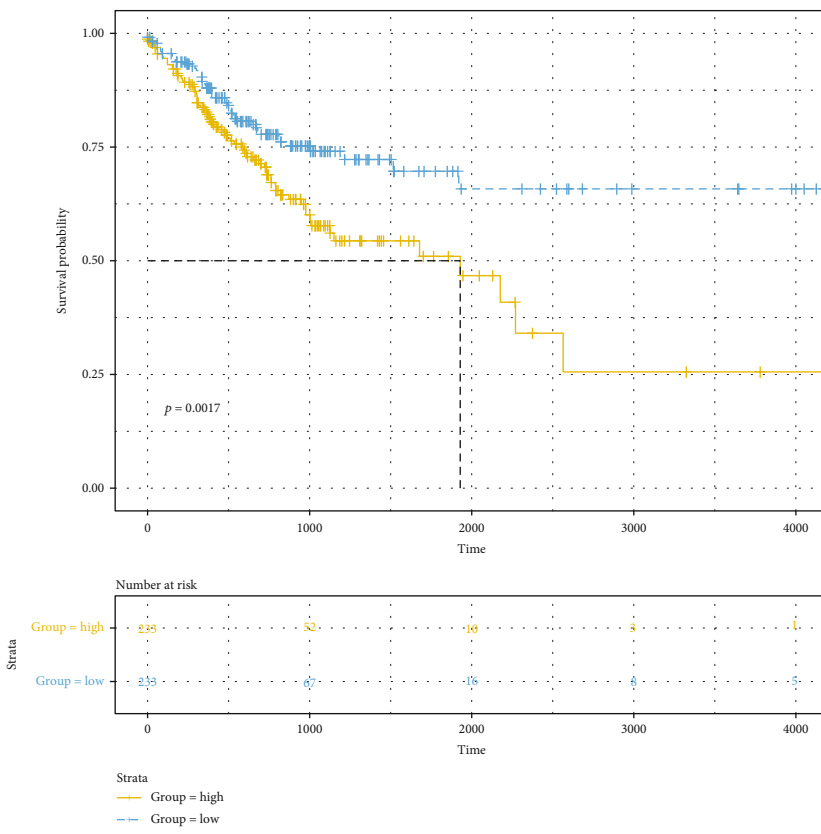


(j)

FIGURE 9: Continued.

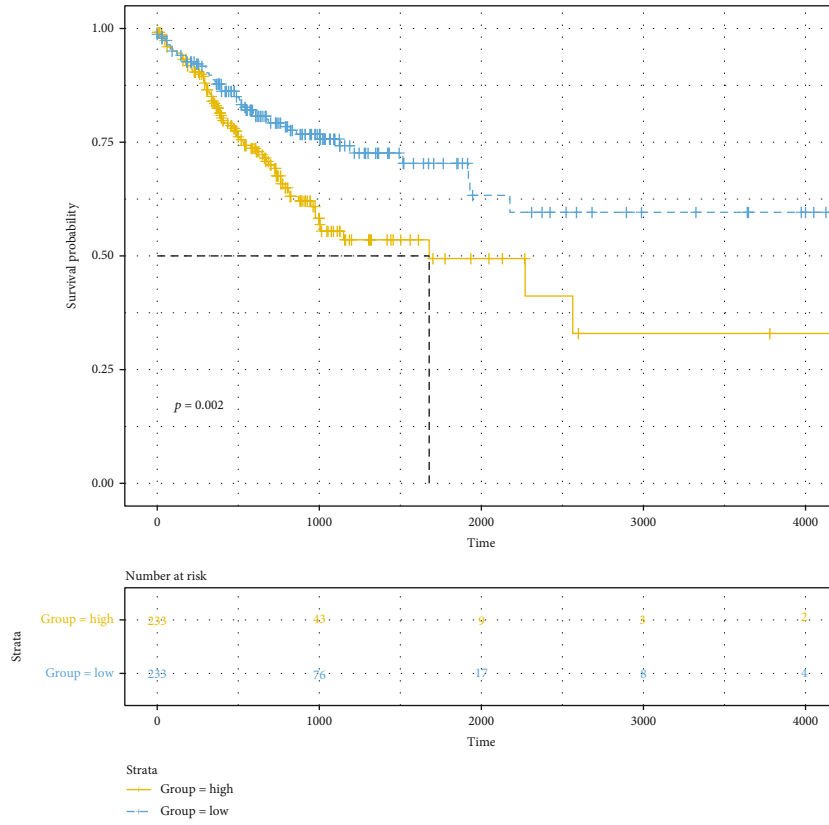


(k)

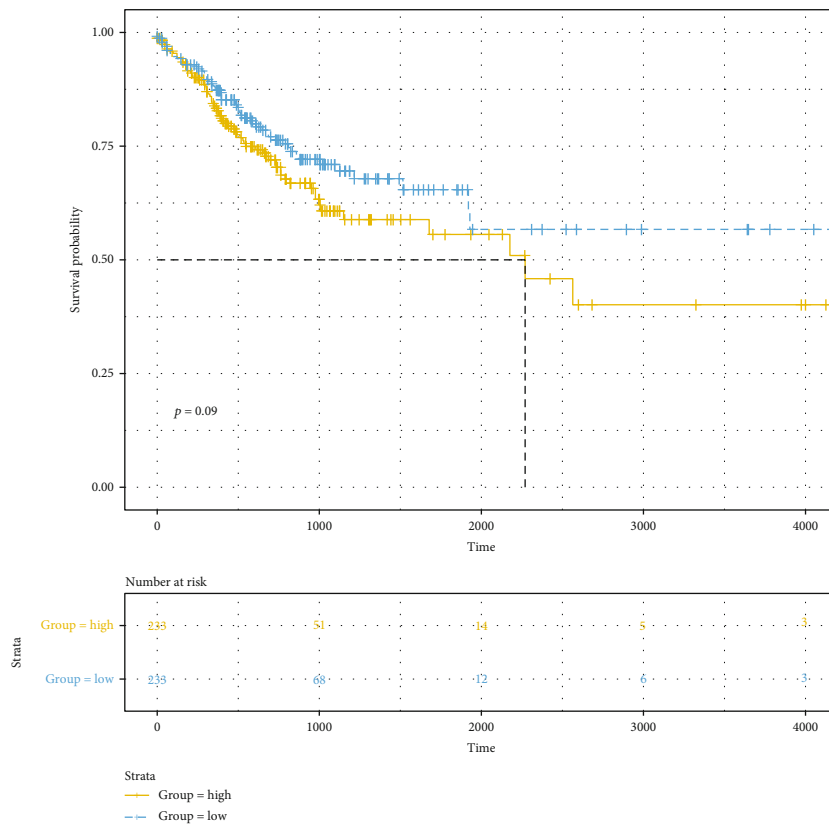


(l)

FIGURE 9: Continued.

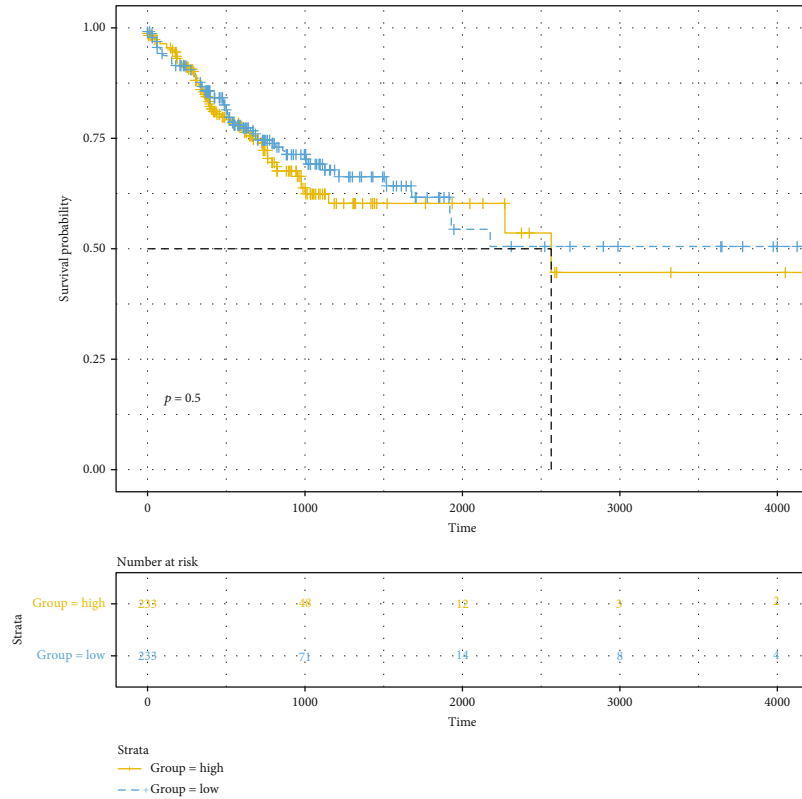


(m)

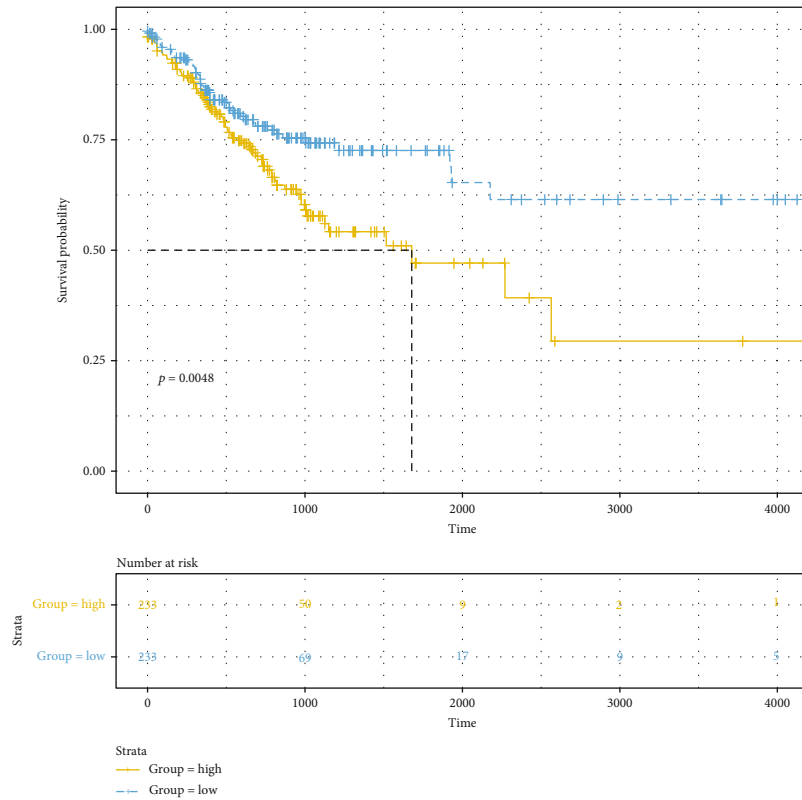


(n)

FIGURE 9: Continued.



(o)



(p)

FIGURE 9: Kaplan-Meier plot of OS and PFS for the eight genes in the risk score model between high and low risk score. Kaplan-Meier OS results for (a) CCL22, (b) CD36, (c) CPA3, (d) CPT1C, (e) KCNE4, (f) NFATC1, (g) RASGRP2, and (h) SLC2A3. Kaplan-Meier PFS results for (i) CCL22, (j) CD36, (k) CPA3, (l) CPT1C, (m) KCNE4, (n) NFATC1, (o) RASGRP2, and (p) SLC2A3.

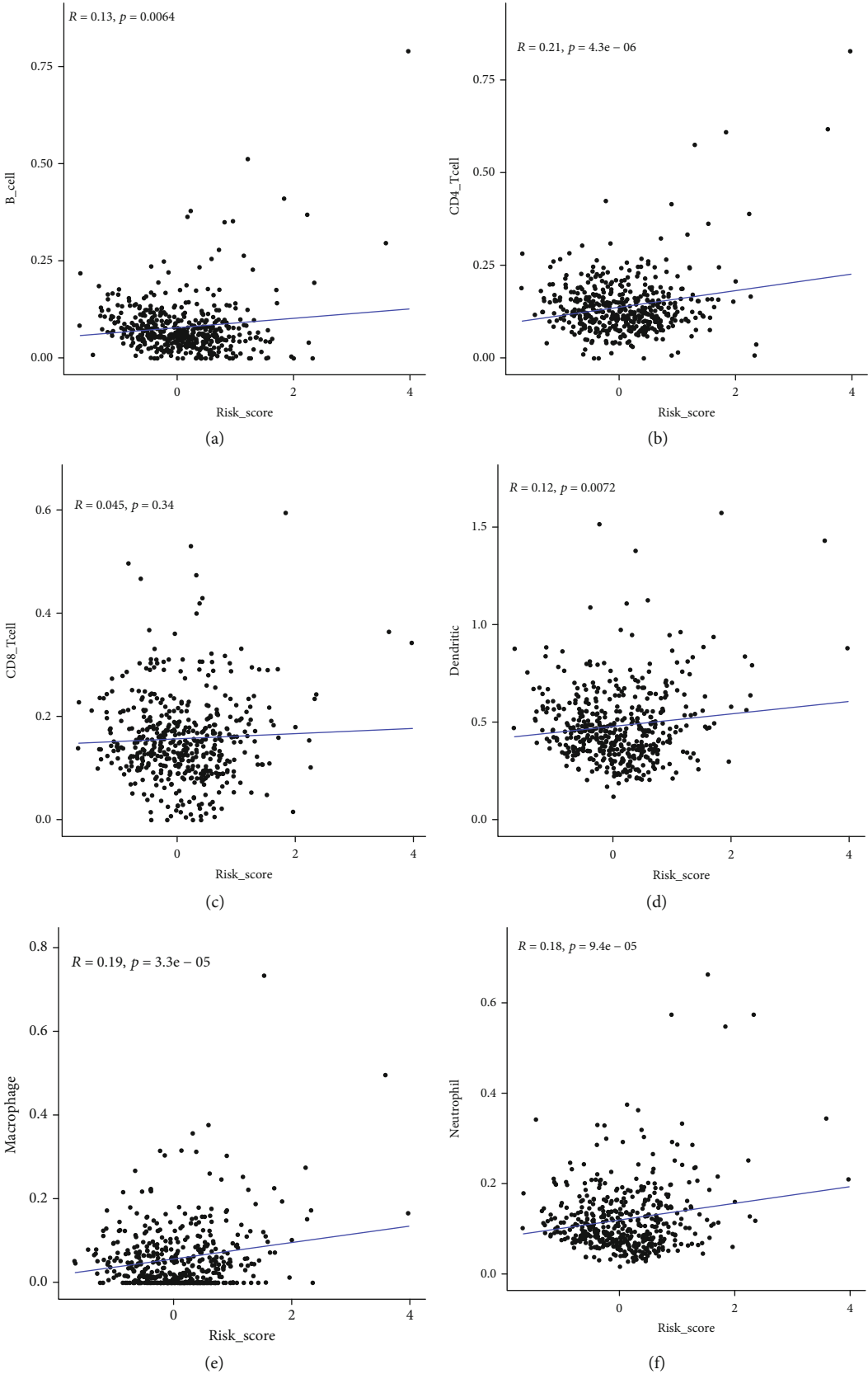
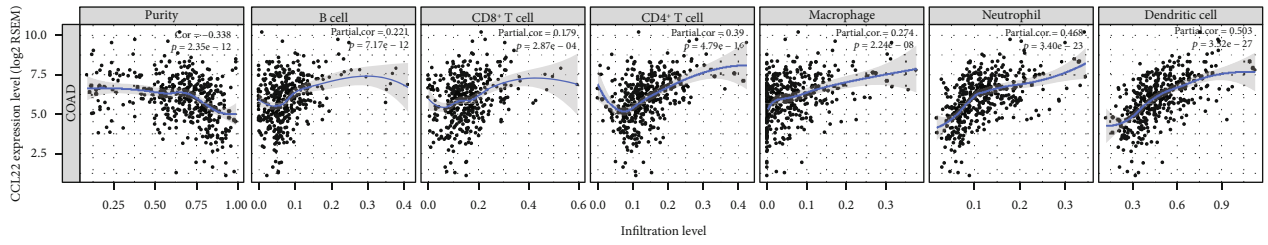
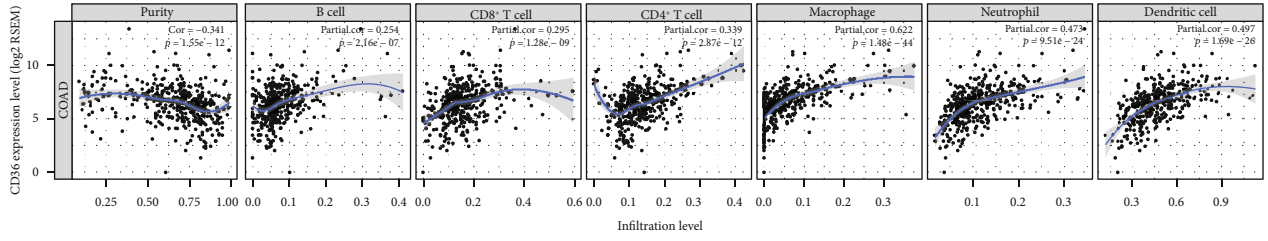


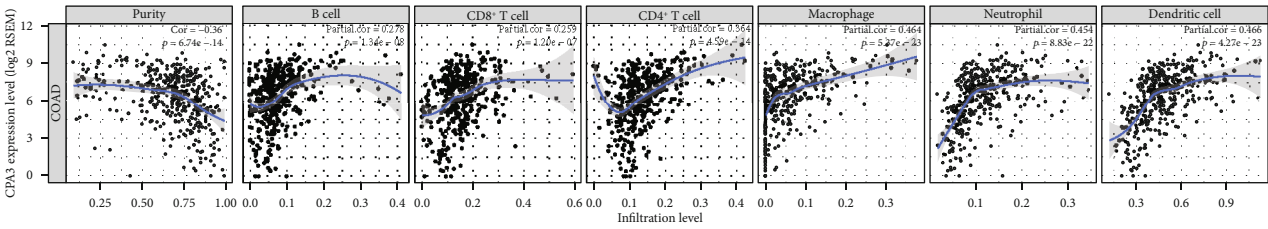
FIGURE 10: The correlation between the eight-gene risk score and immune cell levels. (a) B cell. (b) CD4+T cell. (c) CD8+T cell. (d) Dendritic cell. (e) Macrophage. (f) Neutrophil.



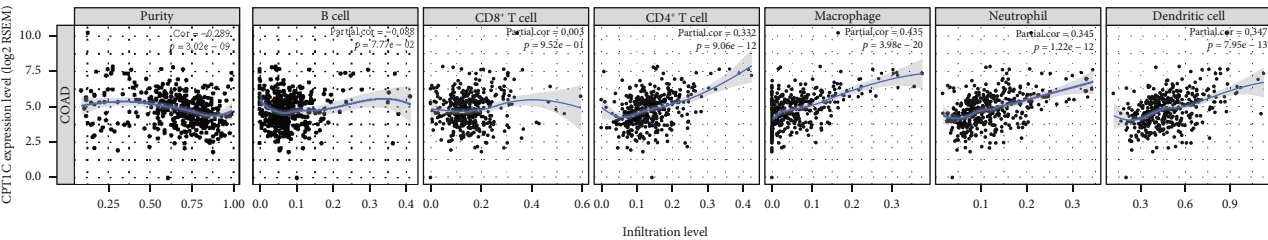
(a)



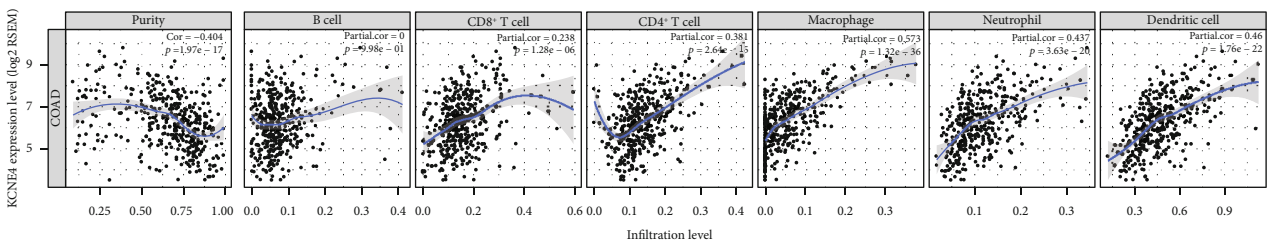
(b)



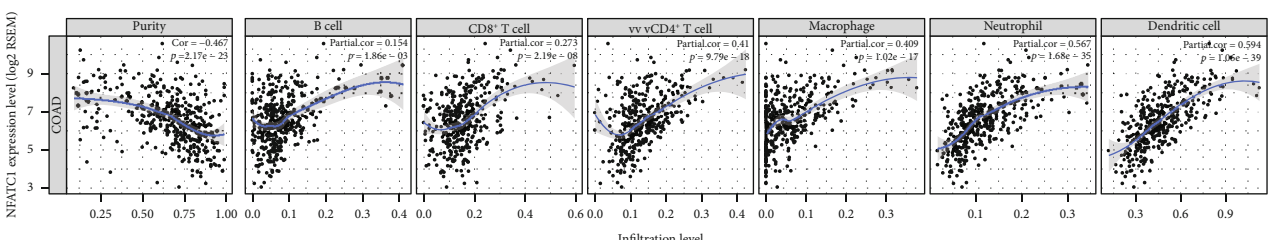
(c)



(d)



(e)



(f)

FIGURE 11: Continued.

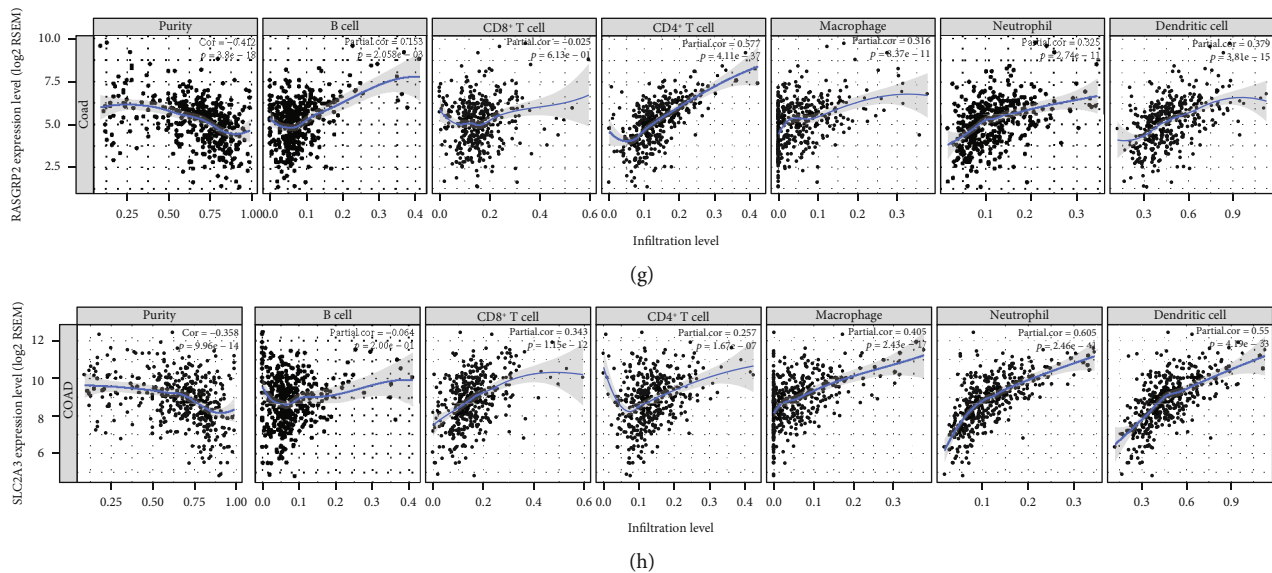


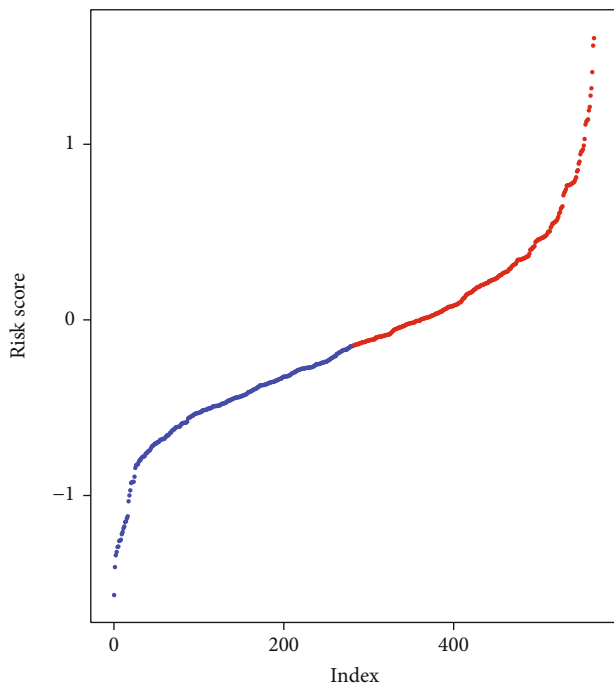
FIGURE 11: The correlation between the eight genes in the risk score model and immune cell levels. (a) CCL22. (b) CD36. (c) CPA3. (d) CPT1C. (e) KCNE4. (f) NFATC1. (g) RASGRP2. (h) SLC2A3.

Individual prognosis for CRC patients varies widely. Individual genes often cannot accurately predict the prognosis of patients with CRC. Genes in most prognostic risk scores are screened via differential expression analyses [20–22]. Yet, there are few prognostic models associated with CRC immune infiltration. Therefore, in this study, we selected eight differentially expressed stromal and immune genes related to prognosis for constructing a risk score model. However, focusing only on CRC-related immune-related genes may limit its clinical value. For this reason, through multivariate regression analysis, after adjustment of the clinical characteristics of CRC, we assessed the association between the risk score and CRC prognosis. The results showed that the model may become an independent prognostic factor for CRC. Our risk score exhibited well efficiency in predicting CRC patients' prognosis. Therefore, the risk score model possessed potential prognostic value, which was confirmed using an independent dataset. Among the eight genes in the model, both in the discovery and independent datasets, CCL22 was a protective factor of CRC, while CD36 and NFATC1 were two risk factors of CRC. However, other genes exhibited inconsistent results in the two datasets. This is partly due to the heterogeneity of the samples in the two datasets. Patients in the same pathological stage have different prognosis. Both in the discovery and independent datasets, CCL22 and CPA3 were lowly expressed and KCNE4, NFATC1, and SLC2A3 were highly expressed in the high-risk samples compared to the low-risk samples. However, there were inconsistent results about other genes between the high- and low-risk samples in the two datasets, partly due to the heterogeneity of the samples, different sequencing platforms, different background correction and normalization methods and so on. Thus, it is unreliable to predict CRC patients' prognosis by an individual

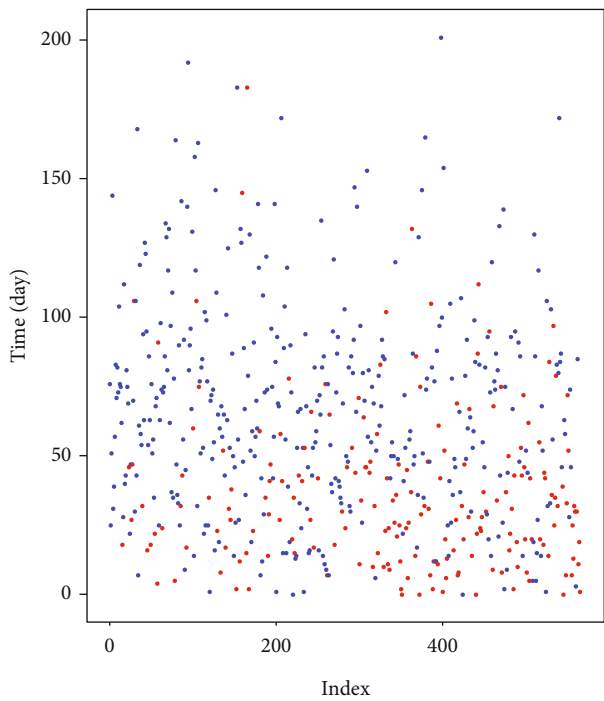
gene. However, our risk score composed of these genes may accurately suggest the patient's prognosis.

As described in a previous study, high CCL22 expression was found in CRC tissues [23]. Recent study has found that CCL22 secreted by M2 macrophages could mediate CRC 5-FU-mediated chemoresistance [24]. Furthermore, it has been reported that CCL22 was in significant correlation with the infiltrating levels of different T cell subsets for CRC [25]. Our results showed that CD36 was significantly downregulated in CRC tissues compared to normal tissues, which was validated in vitro and in vivo [26]. Genome-wide DNA methylation analysis revealed that hypermethylation of CD36 could contribute to its low expression [27]. Fang et al. found that CD36 expression gradually decreased from adenoma to cancer and CD36 loss implied a poor prognosis in patients with CRC [28]. NFATC1 was deregulated in CRC tissues, which was consistent with previous findings [29]. In vitro, its overexpression significantly promoted CRC cell invasion and metastasis [30]. Kumar et al. reported that NFATC1 indicated poor survival outcomes of CRC patients [31]. High SLC2A3 expression was observed in CRC tissues and its high expression indicated a poor prognosis, consistently with previous research [32, 33]. Furthermore, downregulated CPA3 and RASGRP2 and upregulated CPT1C and KCNE4 were found in CRC tissues, which implied poor prognosis.

As for immune cell infiltration, we found that the eight genes in the risk score model were moderately correlated with the infiltrating levels of CD4⁺T cell, dendritic cell, macrophage, and neutrophil. It has been confirmed that TME affects the efficacy of immunotherapy, and immune cells in TME possess predictive value for immunotherapy treatment [34–36]. Increasing genes have been shown to participate in the regulation of immune cells [37–40]. Therefore, our risk score model could possess potential value to predict CRC



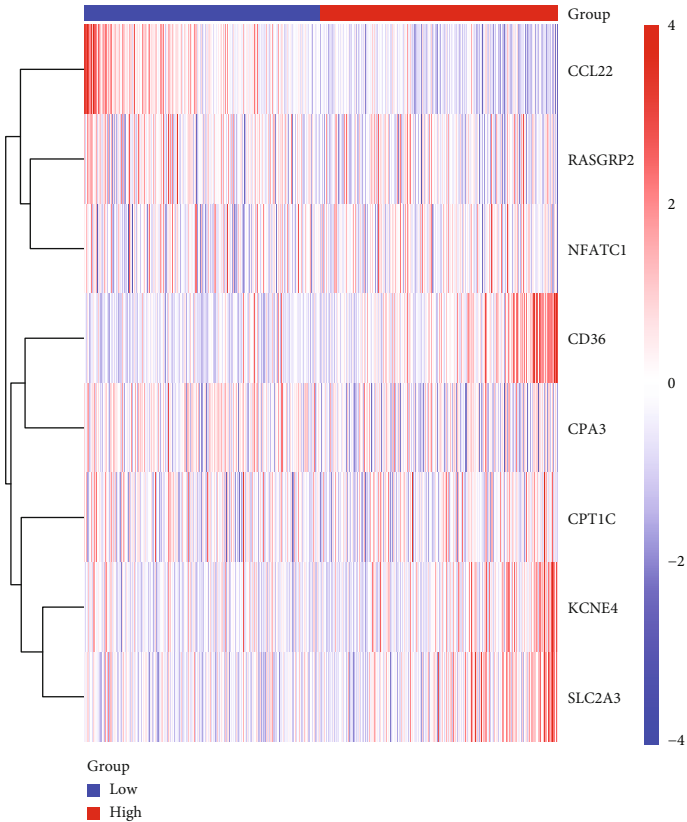
Group
● High
● Low



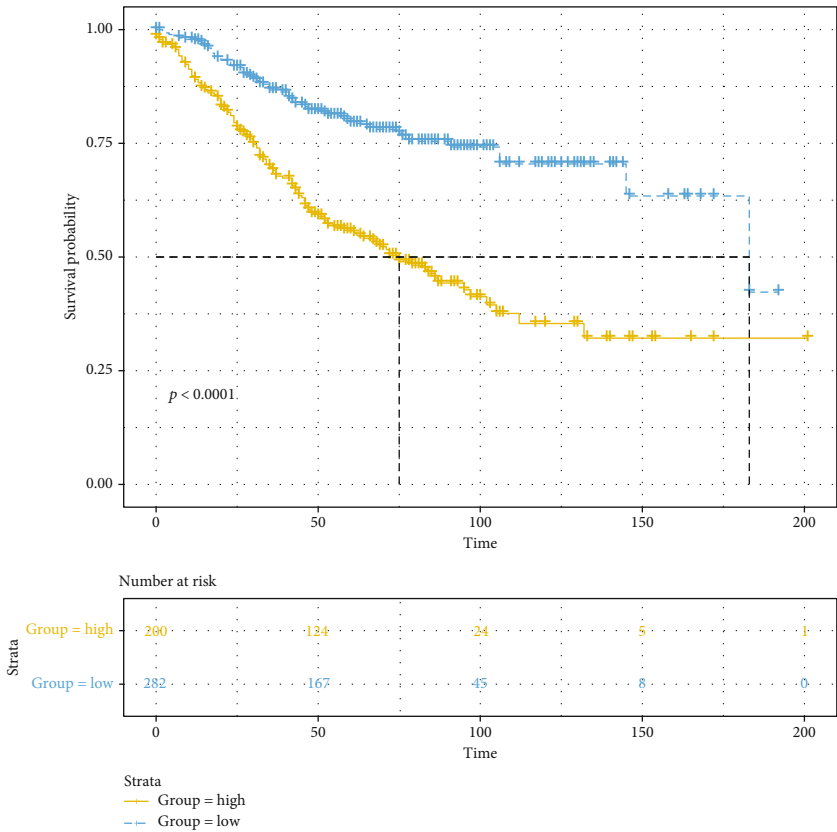
Status
● Alive
● Dead
■ Na

(a)

FIGURE 12: Continued.



(b)



(c)

FIGURE 12: Continued.

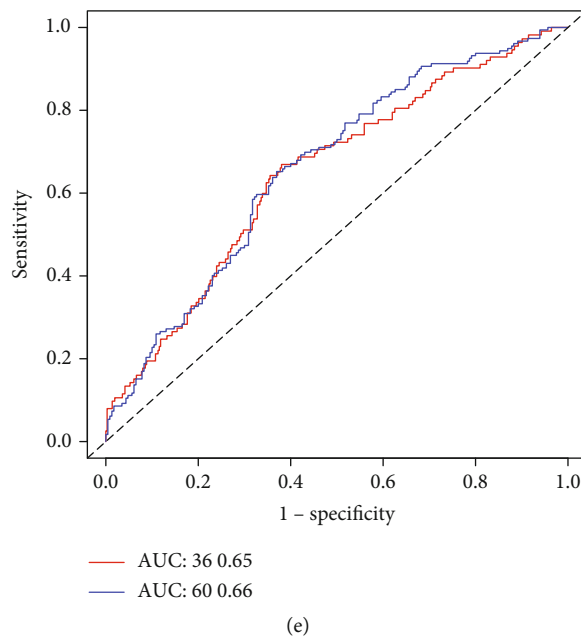
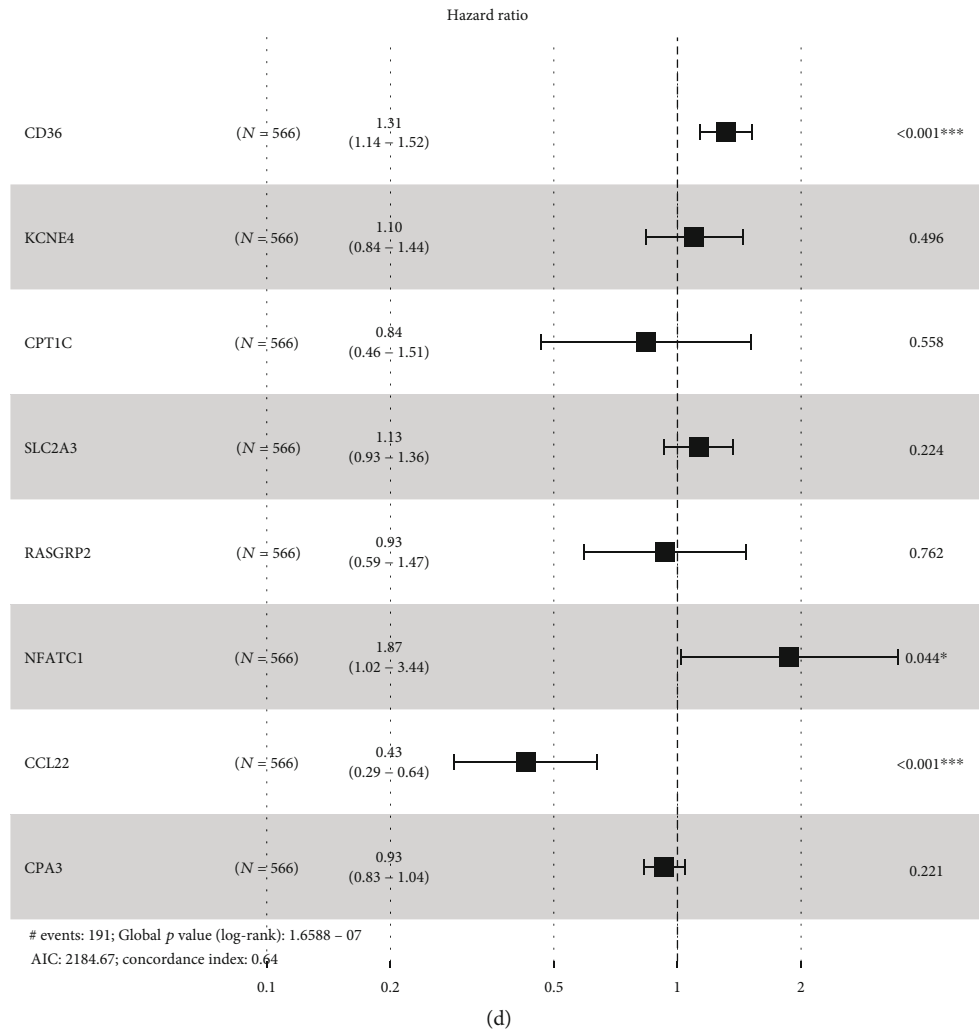


FIGURE 12: Validation of the eight-gene prognostic signature for CRC using an independent dataset. (a) Risk score distribution and survival status. (b) Heat maps showing the expression patterns of the eight genes between high- and low risk score. (c) Kaplan-Meier survival analysis of the model. (d) Forest plot of the eight genes for CRC. (e) ROC curve of the model for 3 years and 5 years.

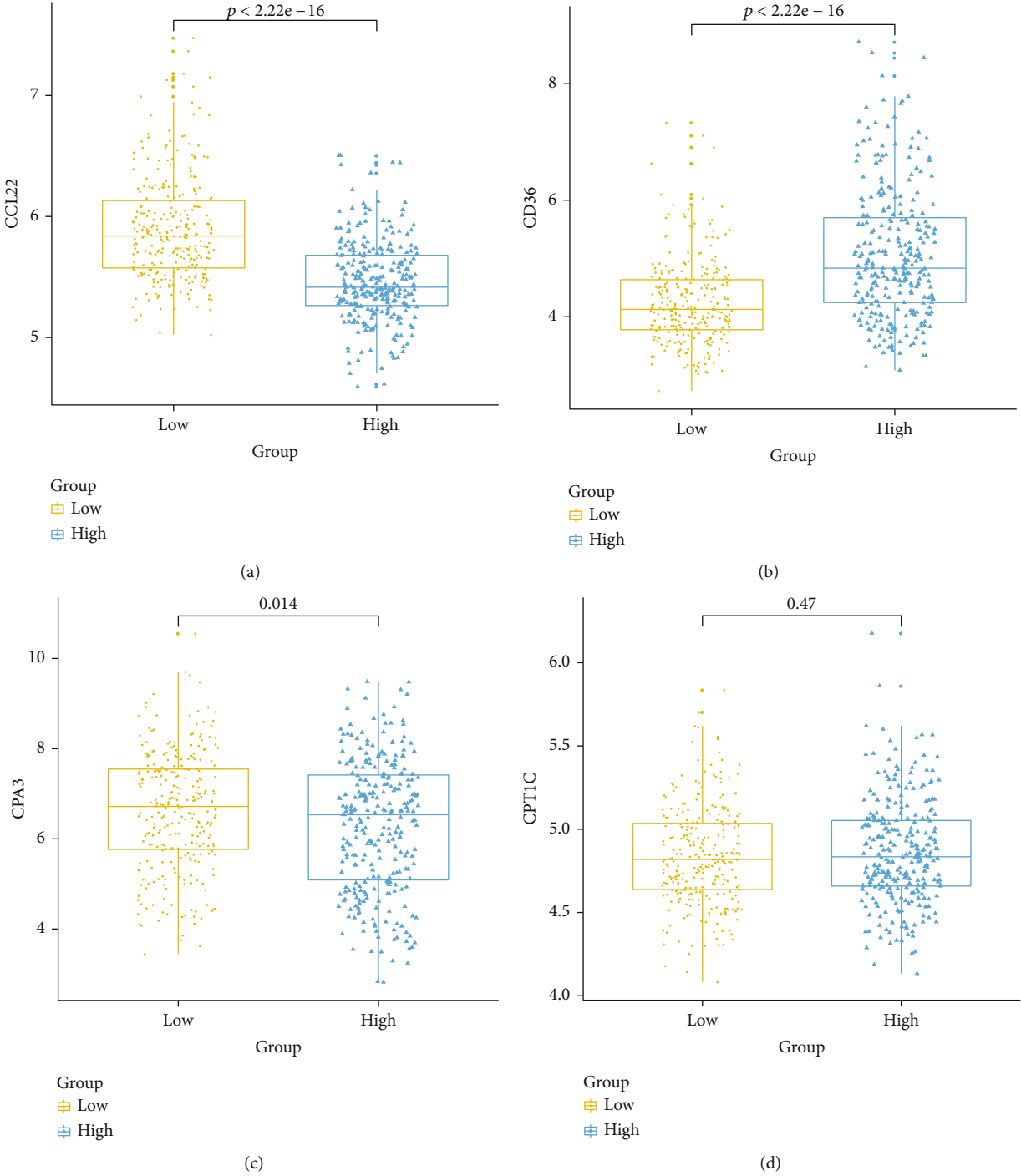


FIGURE 13: Continued.

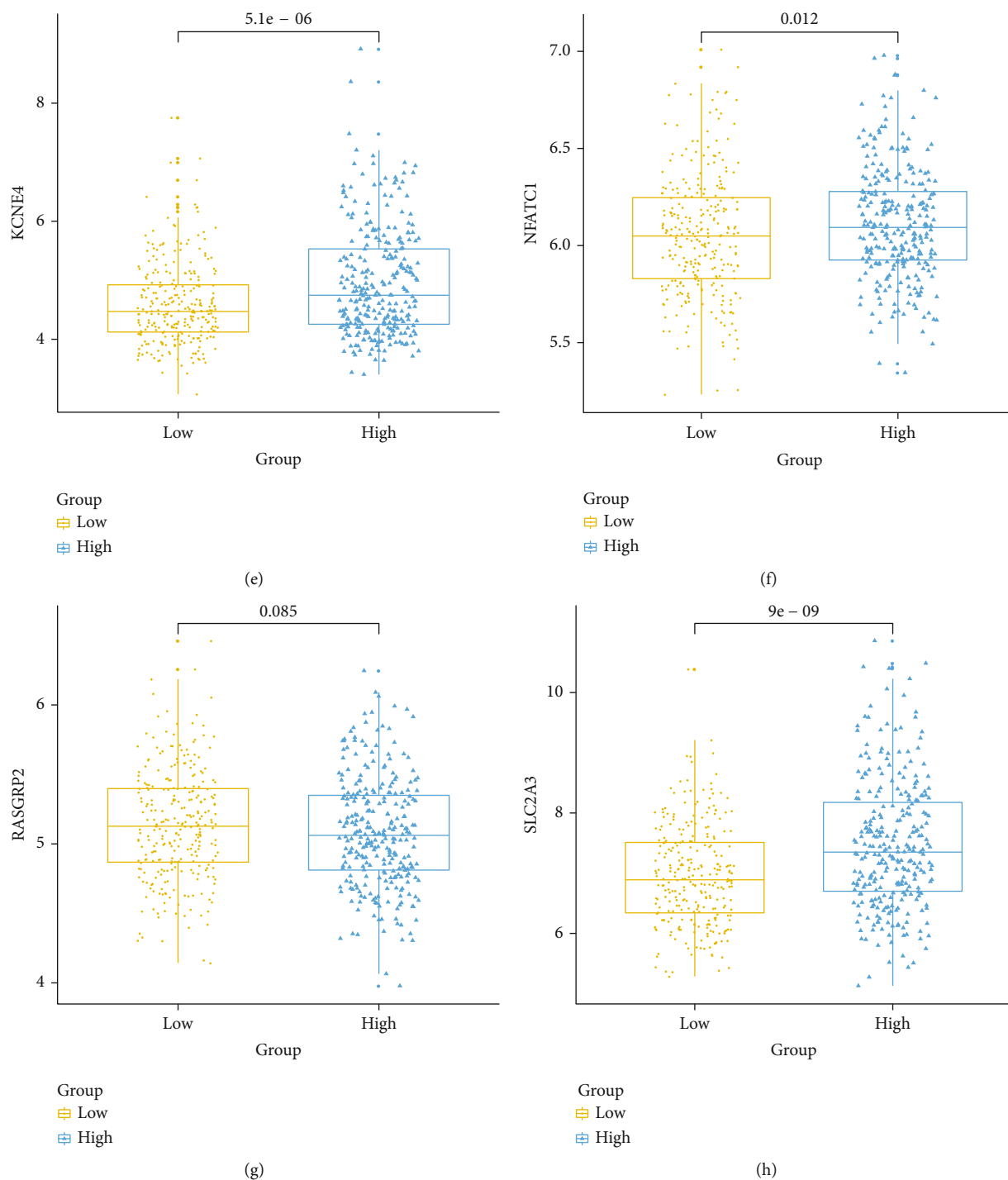


FIGURE 13: The expression patterns of the eight genes in the risk score model between high and low risk score were validated using an independent dataset. (a) CCL22. (b) CD36. (c) CPA3. (d) CPT1C. (e) KCNE4. (f) NFATC1. (g) RASGRP2. (h) SLC2A3.

patients' prognosis, and the eight genes could become promising immunotherapeutic targets, which deserve further study.

Our correlation analysis results confirmed that the eight genes in the risk score were distinctly correlated with molecular markers of CRC prognosis. However, our study has several limitations. First, our retrospective study limited the application of this risk score. Second, the heterogeneity of

the immune microenvironment would inevitably contribute to result bias. Therefore, it is necessary to validate our findings in a prospective clinical study.

5. Conclusion

In this study, we conducted an immune-related prognostic model for CRC on the basis of stromal and immune scores.

TABLE 4: Univariate and multivariate Cox regression analyses in the GSE39582 dataset.

| Characteristics | Univariate analysis | | Multivariate analysis | |
|-----------------|---------------------|---------|-----------------------|---------|
| | HR (95% CI) | p value | HR (95% CI) | p value |
| Age | 1.455 (1.033-2.051) | 0.031 | 1.590 (1.110-2.260) | 0.010 |
| Sex | 1.310 (0.980-1.750) | 0.068 | NA | NA |
| BRAF mutation | 1.111 (0.664-1.861) | 0.688 | NA | NA |
| KRAS mutation | 1.361 (1.018-1.818) | 0.037 | 1.420 (1.040-1.930) | 0.025 |
| TP53 mutation | 1.197 (0.844-1.697) | 0.312 | NA | NA |
| MMR status | 0.768 (0.471-1.252) | 0.290 | NA | NA |
| Pathologic T | 2.005 (1.025-3.921) | 0.042 | 1.510 (0.770-2.970) | 0.229 |
| Pathologic N | 1.648 (1.228-2.213) | 0.0008 | 1.090 (0.500-2.380) | 0.822 |
| Pathologic M | 5.175 (3.621-7.395) | <0.0001 | 3.880 (2.510-5.990) | <0.0001 |
| Tumor stage | 1.767 (1.326-2.354) | 0.0001 | 1.200 (0.520-2.800) | 0.670 |
| Risk score | 2.718 (2.056-3.595) | <0.0001 | 2.420 (1.800-3.260) | <0.0001 |

NA: not available.

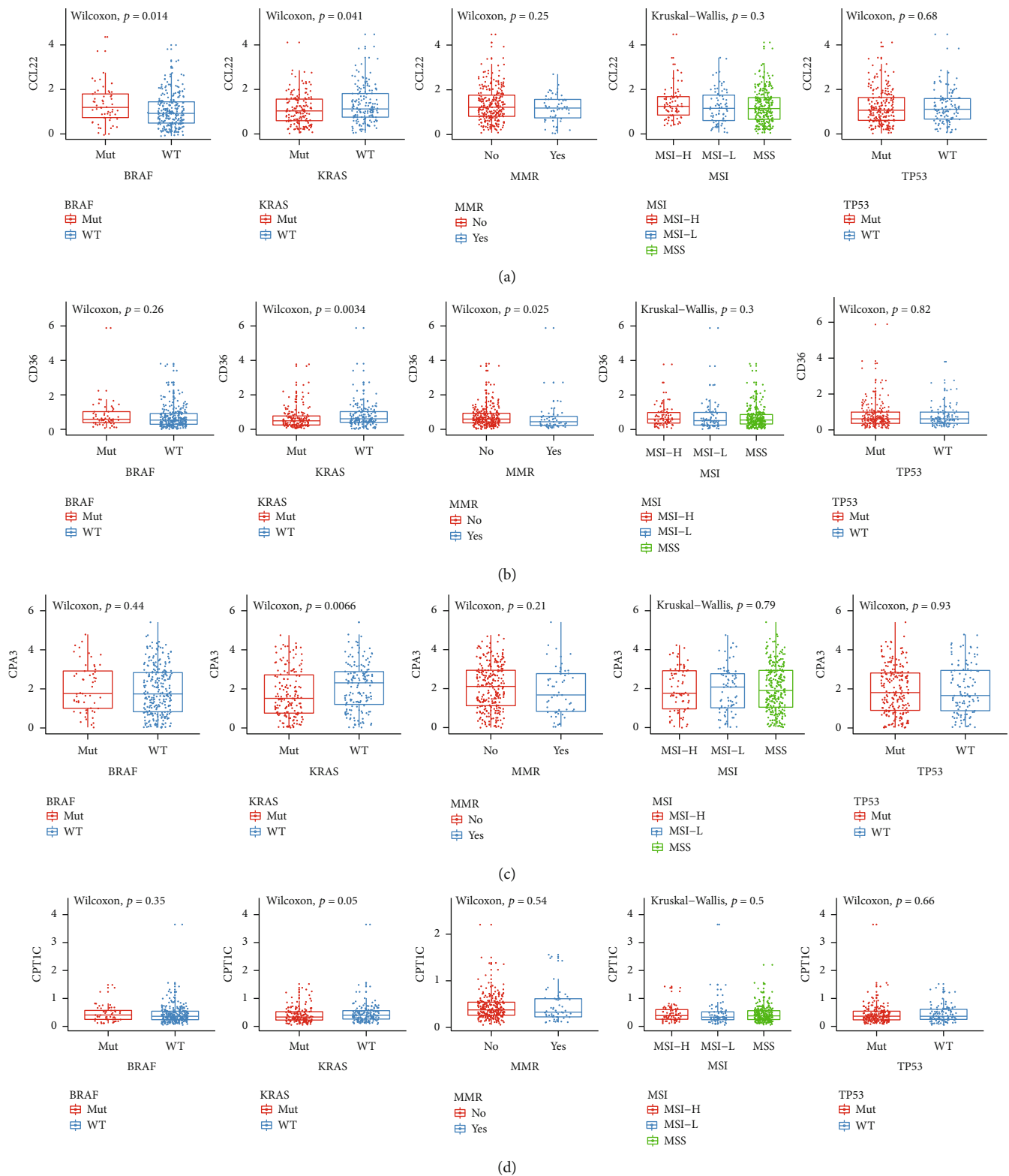


FIGURE 14: Continued.

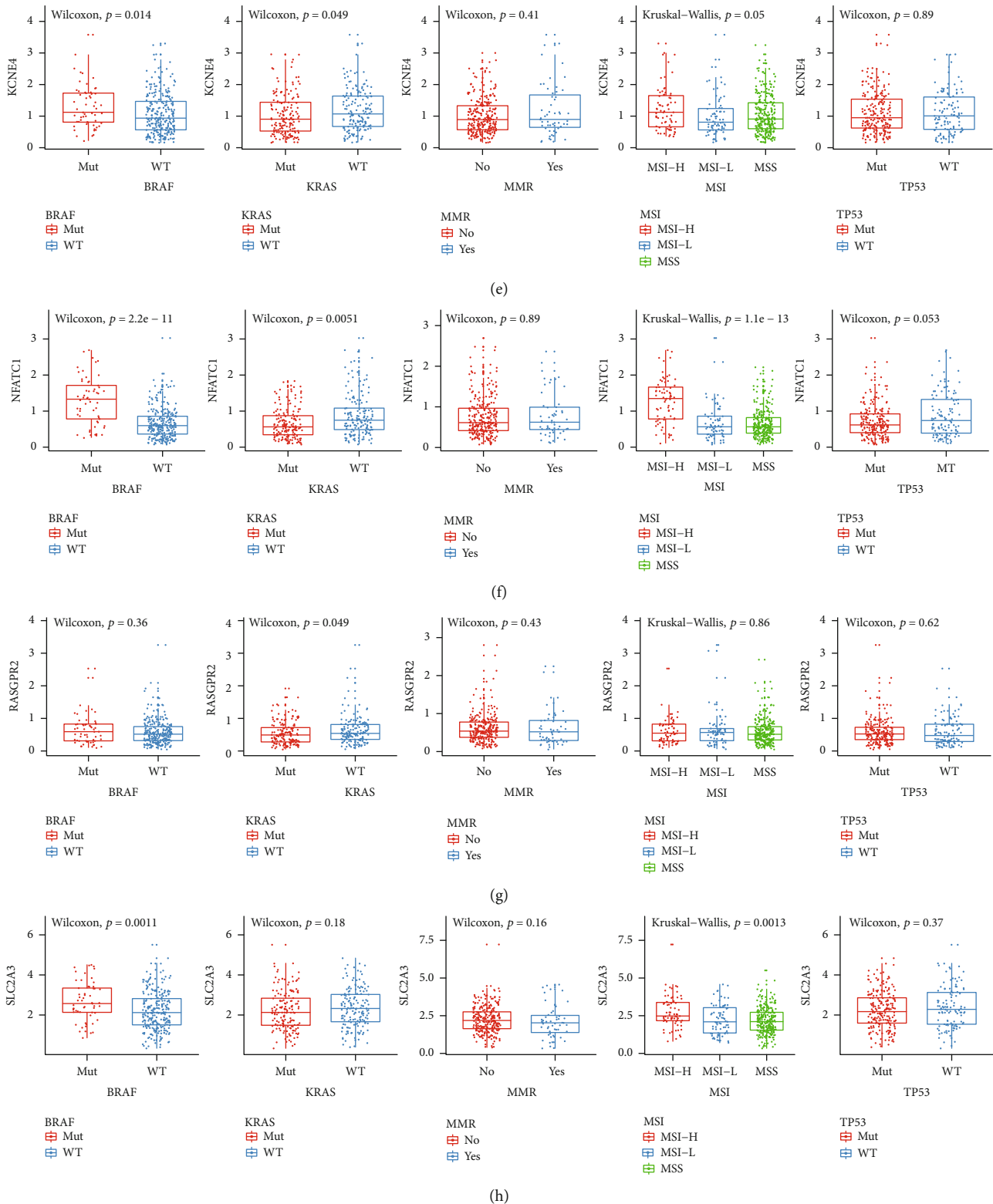


FIGURE 14: The eight genes in the risk score are distinctly correlated with molecular markers of CRC prognosis including BRAF mutation, KRAS mutation, MMR, MSI, and TP53. (a) CCL22. (b) CD36. (c) CPA3. (d) CPT1C. (e) KCNE4. (f) NFATC1. (g) RASGRP2. (h) SLC2A3.

The model had well predictive efficacy for CRC patients' prognosis. Our findings could provide novel biomarkers for predicting the prognosis of CRC patients and developing individualized immunity therapy strategies.

Abbreviations

CRC: Colorectal cancer
TCGA: The Cancer Genome Atlas

ROC: Receiver operating characteristic
 OS: Overall survival
 PFI: Progression-free survival interval
 AUC: Area under the ROC
 TME: Tumor microenvironment
 FC: Fold change
 GO: Gene Ontology
 KEGG: Kyoto Encyclopedia of Genes and Genomes
 CC: Cellular component
 MF: Molecular function
 BP: Biological process
 PPI: Protein-protein interaction
 STRING: The Search Tool for the Retrieval of Interacting Genes.

Data Availability

The (data type) data used to support the findings of this study are included within the supplementary information file(s).

Conflicts of Interest

The authors declare no conflicts of interest.

Supplementary Materials

Supplementary Table 1: the correlation between the eight genes in the risk score model and markers of immune cells. (*Supplementary Materials*)

References

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 7–30, 2018.
- [2] E. Becht, A. de Reyniès, N. A. Giraldo et al., "Immune and stromal classification of colorectal cancer is associated with molecular subtypes and relevant for precision immunotherapy," *Clinical Cancer Research*, vol. 22, no. 16, pp. 4057–4066, 2016.
- [3] A. Garcia-Gomez, J. Rodriguez-Ubreva, and E. Ballestar, "Epigenetic interplay between immune, stromal and cancer cells in the tumor microenvironment," *Clinical Immunology*, vol. 196, pp. 64–71, 2018.
- [4] D. Lambrechts, E. Wauters, B. Boeckx et al., "Phenotype molding of stromal cells in the lung tumor microenvironment," *Nature Medicine*, vol. 24, no. 8, pp. 1277–1289, 2018.
- [5] C. Isella, A. Terrasi, S. E. Bellomo et al., "Stromal contribution to the colorectal cancer transcriptome," *Nature Genetics*, vol. 47, no. 4, pp. 312–319, 2015.
- [6] A. Calon, E. Lonardo, A. Berenguer-Llargo et al., "Stromal gene expression defines poor-prognosis subtypes in colorectal cancer," *Nature Genetics*, vol. 47, no. 4, pp. 320–329, 2015.
- [7] J. Galon, A. Costes, F. Sanchez-Cabo et al., "Type, density, and location of immune cells within human colorectal tumors predict clinical outcome," *Science*, vol. 313, no. 5795, pp. 1960–1964, 2006.
- [8] X. Zhu, L. Chen, L. Liu, and X. Niu, "EMT-mediated acquired EGFR-TKI resistance in NSCLC: mechanisms and strategies," *Frontiers in Oncology*, vol. 9, p. 1044, 2019.
- [9] M. Van den Eynde, B. Mlecnik, G. Bindea et al., "The link between the multiverse of immune microenvironments in metastases and the survival of colorectal cancer patients," *Cancer Cell*, vol. 34, no. 6, pp. 1012–1026.e3, 2018.
- [10] K. Yoshihara, M. Shahmoradgoli, E. Martínez et al., "Inferring tumour purity and stromal and immune cell admixture from expression data," *Nature Communications*, vol. 4, no. 1, article 2612, 2013.
- [11] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS*, vol. 16, no. 5, pp. 284–287, 2012.
- [12] D. Szklarczyk, A. L. Gable, D. Lyon et al., "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Research*, vol. 47, no. D1, pp. D607–d613, 2019.
- [13] P. Shannon, A. Markiel, O. Ozier et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.
- [14] T. Li, J. Fan, B. Wang et al., "TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells," *Cancer Research*, vol. 77, no. 21, pp. e108–e110, 2017.
- [15] H. Akoglu, "User's guide to correlation coefficients," *Turkish Journal of Emergency Medicine*, vol. 18, no. 3, pp. 91–93, 2018.
- [16] G. Landskron, M. de la Fuente, P. Thuwajit, C. Thuwajit, and M. A. Hermoso, "Chronic inflammation and cytokines in the tumor microenvironment," *Journal of Immunology Research*, vol. 2014, Article ID 149185, 19 pages, 2014.
- [17] N. R. West, S. McCuaig, F. Franchini, and F. Powrie, "Emerging cytokine networks in colorectal cancer," *Nature Reviews Immunology*, vol. 15, no. 10, pp. 615–629, 2015.
- [18] V. O. Frick, C. Rubie, U. Keilholz, and P. Ghadjari, "Chemokine/chemokine receptor pair CCL20/CCR6 in human colorectal malignancy: an overview," *World Journal of Gastroenterology*, vol. 22, no. 2, pp. 833–841, 2016.
- [19] Y. Itatani, K. Kawada, S. Inamoto et al., "The role of chemokines in promoting colorectal cancer invasion/metastasis," *International Journal of Molecular Sciences*, vol. 17, no. 5, p. 643, 2016.
- [20] G. Sun, Y. Li, Y. Peng et al., "Identification of a five-gene signature with prognostic value in colorectal cancer," *Journal of Cellular Physiology*, vol. 234, no. 4, pp. 3829–3836, 2019.
- [21] X. Wang, J. Zhou, M. Xu et al., "A 15-lncRNA signature predicts survival and functions as a ceRNA in patients with colorectal cancer," *Cancer Management and Research*, vol. 10, pp. 5799–5806, 2018.
- [22] Z. Zhang, Q. Liu, P. Wang et al., "Development and internal validation of a nine-lncRNA prognostic signature for prediction of overall survival in colorectal cancer patients," *PeerJ*, vol. 6, article e6061, 2018.
- [23] Y. H. Huang, Y. F. Cao, Z. Y. Jiang, S. Zhang, and F. Gao, "Th22 cell accumulation is associated with colorectal cancer development," *World Journal of Gastroenterology*, vol. 21, no. 14, pp. 4216–4224, 2015.
- [24] C. Wei, C. Yang, S. Wang et al., "M2 macrophages confer resistance to 5-fluorouracil in colorectal cancer through the activation of CCL22/PI3K/AKT signaling," *Oncotargets and Therapy*, vol. 12, pp. 3051–3063, 2019.
- [25] E. Cremonesi, V. Governa, J. F. G. Garzon et al., "Gut microbiota modulate T cell trafficking into human colorectal cancer," *Gut*, vol. 67, no. 11, pp. 1984–1994, 2018.

- [26] X. Zhang, J. Yao, H. Shi, B. Gao, and L. Zhang, "LncRNA TINCR/microRNA-107/CD36 regulates cell proliferation and apoptosis in colorectal cancer via PPAR signaling pathway based on bioinformatics analysis," *Biological Chemistry*, vol. 400, no. 5, pp. 663–675, 2019.
- [27] M. Li, W. Chen, X. Sun et al., "Metastatic colorectal cancer and severe hypocalcemia following irinotecan administration in a patient with X-linked agammaglobulinemia: a case report," *BMC Medical Genetics*, vol. 20, no. 1, p. 157, 2019.
- [28] Y. Fang, Z. Y. Shen, Y. Z. Zhan et al., "CD36 inhibits β -catenin/c-myc-mediated glycolysis through ubiquitination of GPC4 to repress colorectal tumorigenesis," *Nature Communications*, vol. 10, no. 1, article 3981, 2019.
- [29] E. A. Pudova, A. V. Kudryavtseva, M. S. Fedorova et al., "HK3 overexpression associated with epithelial-mesenchymal transition in colorectal cancer," *BMC Genomics*, vol. 19, Supplement 3, p. 113, 2018.
- [30] M. K. Tripathi, N. G. Deane, J. Zhu et al., "Nuclear factor of activated T-cell activity is associated with metastatic capacity in colon cancer," *Cancer Research*, vol. 74, no. 23, pp. 6947–6957, 2014.
- [31] R. Kumar, R. Raman, V. Kotapalli et al., "Ca²⁺/nuclear factor of activated T cells signaling is enriched in early-onset rectal tumors devoid of canonical Wnt activation," *Journal of Molecular Medicine (Berlin, Germany)*, vol. 96, no. 2, pp. 135–146, 2018.
- [32] E. Kim, S. Jung, W. S. Park et al., "Upregulation of SLC2A3 gene and prognosis in colorectal carcinoma: analysis of TCGA data," *BMC Cancer*, vol. 19, no. 1, p. 302, 2019.
- [33] J. Martinez-Romero, S. Bueno-Fortes, M. Martín-Merino, A. Ramirez de Molina, and J. de Las Rivas, "Survival marker genes of colorectal cancer derived from consistent transcriptomic profiling," *BMC Genomics*, vol. 19, Supplement 8, p. 857, 2018.
- [34] B. Feng, F. Zhou, B. Hou et al., "Binary cooperative prodrug nanoparticles improve immunotherapy by synergistically modulating immune tumor microenvironment," *Advanced Materials*, vol. 30, no. 38, article 1803001, 2018.
- [35] W. Song, K. Tiruthani, Y. Wang et al., "Trapping of lipopolysaccharide to promote immunotherapy against colorectal cancer and attenuate liver metastasis," *Advanced Materials*, vol. 30, no. 52, article 1805007, 2018.
- [36] D. V. F. Tauriello, S. Palomo-Ponce, D. Stork et al., "TGF β drives immune evasion in genetically reconstituted colon cancer metastasis," *Nature*, vol. 554, no. 7693, pp. 538–543, 2018.
- [37] J. Zhao, B. Ou, D. Han et al., "Tumor-derived CXCL5 promotes human colorectal cancer metastasis through activation of the ERK/Elk-1/Snail and AKT/GSK3 β / β -catenin pathways," *Molecular Cancer*, vol. 16, no. 1, p. 70, 2017.
- [38] Y. Yin, S. Yao, Y. Hu et al., "The immune-microenvironment confers chemoresistance of colorectal cancer through macrophage-derived IL6," *Clinical Cancer Research*, vol. 23, no. 23, pp. 7375–7387, 2017.
- [39] J. Gil, D. Ramsey, P. Pawlowski et al., "The influence of tumor microenvironment on ATG4D gene expression in colorectal cancer patients," *Medical Oncology*, vol. 35, no. 12, p. 159, 2018.
- [40] L. Chen, G. Wang, X. Qiao et al., "Downregulated miR-524-5p participates in the tumor microenvironment of ameloblastoma by targeting the Interleukin-33 (IL-33)/suppression of tumorigenicity 2 (ST2) Axis," *Medical Science Monitor*, vol. 26, article e921863, 2020.

Research Article

Comprehensive Characterization of Prognostic Long Noncoding RNAs in Osteosarcoma

Hua Gao,¹ Yuanyuan Guo ,² Miaomiao Zhang,¹ and Zuqiang Yi ³

¹Department of Oncology, Henan Provincial Hospital, Zhengzhou Airport Economy Zone, Zhengzhou City, Henan Province, China 450000

²Department of Pharmacy, The First Affiliated Hospital of Zhengzhou University, Zhengzhou City, Henan Province, China 450000

³Department of Osteology, Henan Provincial Hospital, Zhengzhou City, Henan Province, China 450000

Correspondence should be addressed to Yuanyuan Guo; guoyuanyuan@gs.zzu.edu.cn and Zuqiang Yi; 15936275945@163.com

Received 27 May 2020; Revised 17 July 2020; Accepted 22 July 2020; Published 25 August 2020

Guest Editor: Tao Huang

Copyright © 2020 Hua 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.

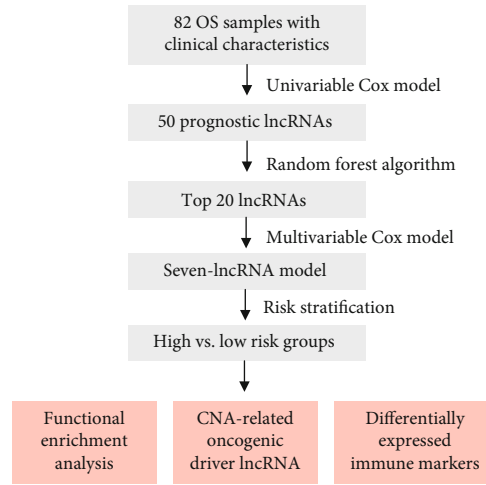
The molecular mechanism of osteosarcoma (OS) based on protein-coding genes has largely been studied in the past decades. However, much remains to be explored when it comes to the role that long noncoding RNAs (lncRNAs) play in the pathogenesis and progression of OS and how they are associated with OS metastasis. In the present study, we collected RNA-seq-based gene expression data of 82 OS samples from the Therapeutically Applicable Research To Generate Effective Treatments (TARGET) database, along with their clinical information. We found that 50 lncRNAs were significantly associated with patients' survival by univariable Cox regression model. Moreover, we built multivariable Cox regression model based on 7 lncRNAs and successfully stratified patients into two risk groups, which exhibited significantly different prognostic outcomes. Significantly enriched Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways detected by differential expression analysis on DEGs between the two groups with different prognostic outcomes were both immune-related, indicating that such GO terms and pathways are critical for OS survival. Among the seven lncRNA signatures, *AC011442.1* was predicted to act as an oncogenic driver in OS by correlation analysis of copy number alteration (CNA) and lncRNA expression, and it was predicted to regulate AMPK and hedgehog signaling pathways. In summary, the identification of novel prognostic lncRNAs in OS could not only improve our understanding of the lncRNAs involved in OS tumorigenesis or progression but also assist the diagnosis and development of molecularly targeted therapies for OS, which in turn benefit patients' survival.

1. Introduction

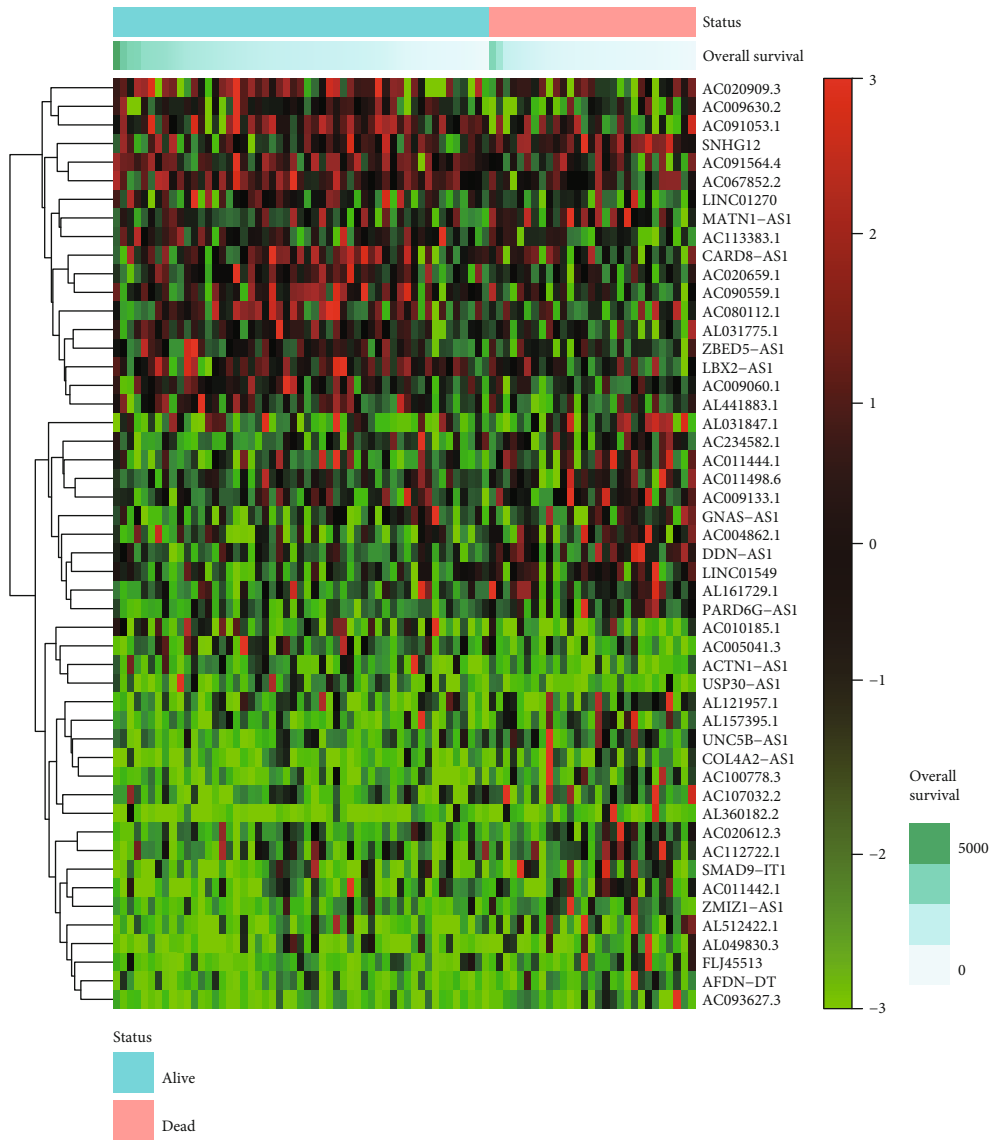
Osteosarcoma (OS) is among the most prevalent malignancies in children and adolescents [1]. According to previous study, it takes up approximately 20% of all bone cancers, which also makes it one of the most common primary skeletal tumors [2, 3]. For example, according to the American Cancer Society, the estimated number of newly diagnosed cases of skeletal malignancies in 2017 would reach 3,260 in the United States, among them there would be roughly over 600 OS patients [4]. Unfortunately, over one-fifth of osteosarcoma patients exhibit lung metastasis at the time of diagnosis, which often results in unsatisfactory prognosis [5]. No significant improvement in 10-year overall survival of OS

patients has been observed since the 1990s [6]. The magical effect of traditional tumor resection surgery and chemotherapy seems to encounter a bottleneck as they had once improved overall 10-year survival of OS from 30% to about 50% in the 1970s, and with the advances in molecular biology and related techniques, molecularly targeted therapies have since emerged as a new option in the management strategy of various cancers, including OS.

It is crucial for the development of molecularly targeted therapies to identify metastatic-related biomarkers and underlying mechanism in OS, in order to deliver a more accurate prognosis prediction and therapeutic decisions [7]. Long noncoding RNAs (lncRNAs) attract researchers' keen attention worldwide as they play a critical role through



(a)



(b)

FIGURE 1: The study design and the expression profiles of the 50 prognostic lncRNAs in OS. (a) The workflow of the present study. (b) The lncRNAs were clustered by hierarchical clustering algorithm, and the samples were ordered by survival status and survival time.

TABLE 1: The summary for seven prognostic lncRNAs in univariable and multivariable Cox regression model.

| Features | Univariable Cox regression | | | Multivariable Cox regression | | |
|------------|----------------------------|--------------|----------|------------------------------|--------------|----------|
| | Coefficient | Hazard ratio | P value | Coefficient | Hazard ratio | P value |
| USP30-AS1 | -1.25 | 0.29 | 1.95E-06 | -2.10 | 0.122 | 4.22E-03 |
| AC113383.1 | -0.09 | 0.91 | 8.34E-03 | -0.11 | 0.89 | 4.47E-03 |
| LINC01549 | 0.02 | 1.02 | 1.38E-04 | 0.02 | 1.02 | 5.08E-03 |
| AC093627.3 | 0.12 | 1.13 | 2.66E-04 | 0.16 | 1.18 | 1.35E-05 |
| DDN-AS1 | 0.35 | 1.42 | 1.95E-06 | 0.22 | 1.25 | 4.22E-03 |
| GNAS-AS1 | 0.46 | 1.58 | 7.49E-03 | 0.68 | 1.98 | 7.02E-04 |
| AC011442.1 | 0.39 | 1.48 | 1.54E-02 | 0.72 | 2.06 | 3.22E-03 |

epigenetic, transcriptional, and posttranscriptional mechanisms in diverse biological processes, such as tumor initiation, growth, and metastasis [8]. Though lncRNAs are not to be translated into proteins, they can function as key regulators through interacting with miRNAs, mRNAs, and proteins [8]. Many lncRNAs are identified to exert oncogenic or tumor suppressor functions in OS, such as *ZEB1-AS1* [9], *SPRY4-IT1* [10], *BCAR4* [11], and *MFI2* [12]. For example, previous studies have reported that lncRNA DANCR could function as a competitive endogenous RNA in OS, thereby promoting ROCK1-mediated proliferation and metastasis [13]. CEBPA-AS1, an antisense RNA of CEBPA, has the capability of inhibiting proliferation and migration and promoting apoptosis in OS via Notch signaling [14]. These studies demonstrate that lncRNAs can regulate the progression, metastasis, and prognosis of OS [15].

In this study, RNA-seq data and clinical information of patients with osteosarcoma from the TARGET database were processed with univariable Cox regression and random forest algorithm, and we selected seven long noncoding RNAs (lncRNAs); all of them have the potential to affect the survival of osteosarcoma patients and to construct a prognosis risk model. Based on the stratification offered by our model, the corresponding biological differences among osteosarcoma patients and how these characteristics would result in varied prognostic outcomes were further explored and explained.

2. Materials and Methods

2.1. Data Resources. We downloaded RNA-seq-based gene expression data (TPM, transcript per million), somatic copy number alteration (SCNA) data, and clinical data of 82 corresponding osteosarcoma patients from the TARGET (Therapeutically Applicable Research to Generate Effective Treatments) database [16]. The segmented SCNA was annotated by Ensembl gene annotation v37.75 [17]. The SCNA status for each was called as gain or loss only if the log₂ ratio (tumor/normal copy numbers) was more than 0.6 or less than -0.6. To meet the requirement for data analysis, we only collected 82 osteosarcoma samples with matched SCNA, gene expression, and clinical data.

2.2. Selection of lncRNAs in OS for Prognostic Risk Model Construction. First, based on 9 biotypes for lncRNAs (which were 3prime_overlapping_ncRNA, antisense, lincRNA, macro_lincRNA, non_coding, sense_intronic, sense_overlapping, bidirectional_promoter_lincRNA, and retained intron) in Ensembl, we obtained a total of 3,159 lncRNAs that exhibited TPM (transcript per million) > 0.1 in more than half of the samples. The expression status of lncRNAs were firstly classified into high and low expression, respectively, based on the median of the expression levels. Combined with the clinical information, univariable Cox regression analysis was then performed with package Survival v3.1-11 in R v3.6.3 to pick up lncRNAs significantly related to the survival of the patients (log-rank test, $P < 0.05$). Utilizing the random forest algorithm in R package randomForestSRC with default options, we evaluated ranked those lncRNAs and built a multivariable Cox model based on the top 20 prognostic lncRNAs. Subsequently, we only retained the prognostically insignificant lncRNAs in the initial multivariable Cox model ($P > 0.05$) and built the optimal multivariable Cox model based on these prognostic lncRNAs.

2.3. Model Construction for Evaluating Osteosarcoma Prognostic Risk. Taking into consideration the expression of qualified lncRNAs in each patient and the patient's survival status, we applied multivariable Cox regression with survival package in R v3.6.3 to build our osteosarcoma prognosis risk model, and lncRNAs with significant contribution to the model were selected. These lncRNAs were used to construct a risk-scoring method, which assigned a score that reflected the risk of death to each osteosarcoma patient. The patients were then divided by the median score into the high-risk and low-risk groups, accordingly. We visualized the survival curves of the two groups of patients by the Kaplan-Meier method and assessed the differences between the two groups by log-rank test.

2.4. Functional Enrichment Analysis of the Dysregulated Genes in the Two Risk Groups. As osteosarcoma patients were categorized, their gene expression profiles fell into two groups, accordingly. Utilizing the screening criteria of |log₂ (fold change)| > 1 and P value < 0.05, genes with significant differential expression between the two groups were selected. Subsequently, Gene Ontology (GO) [18] and Kyoto

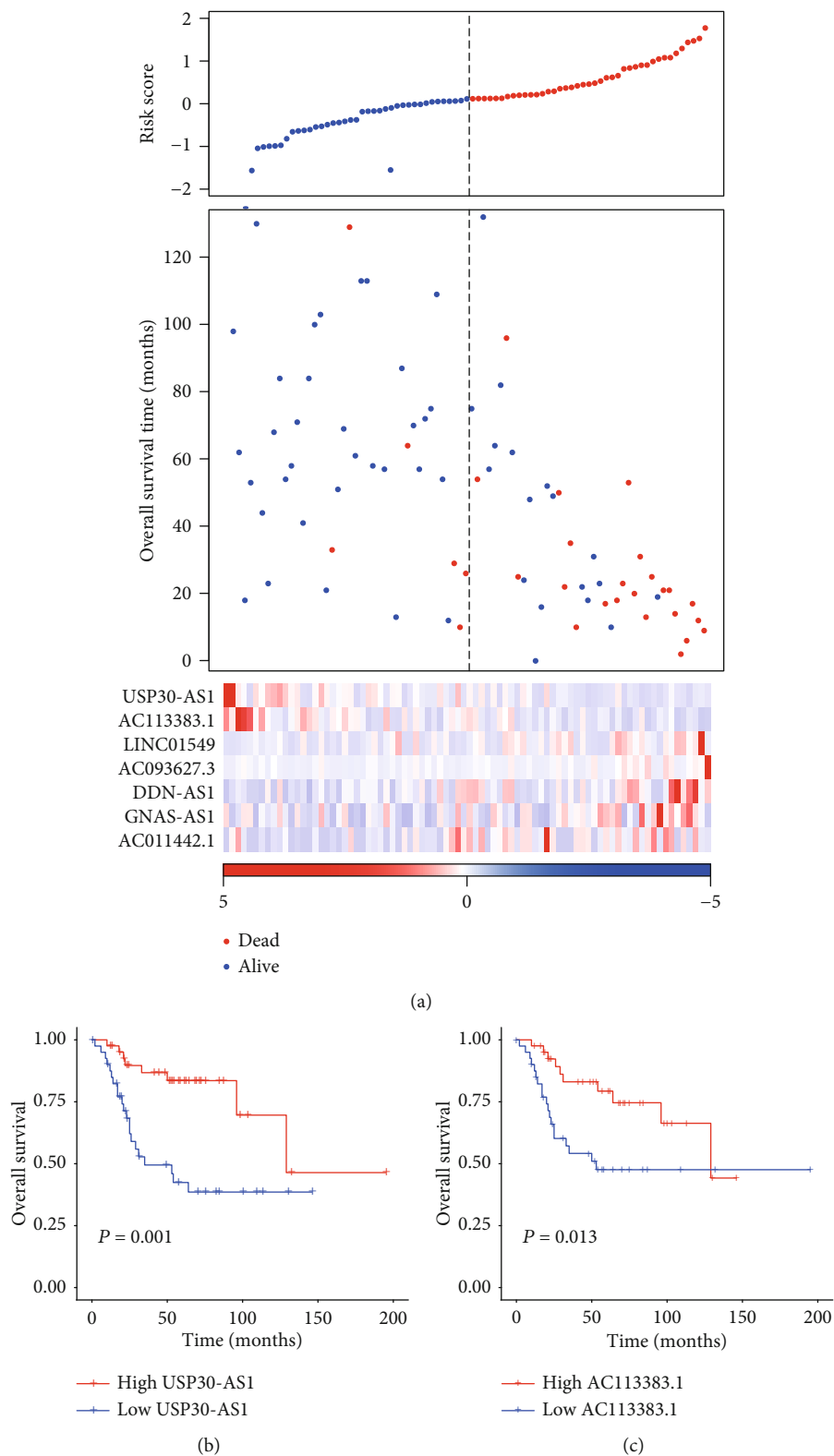


FIGURE 2: Continued.

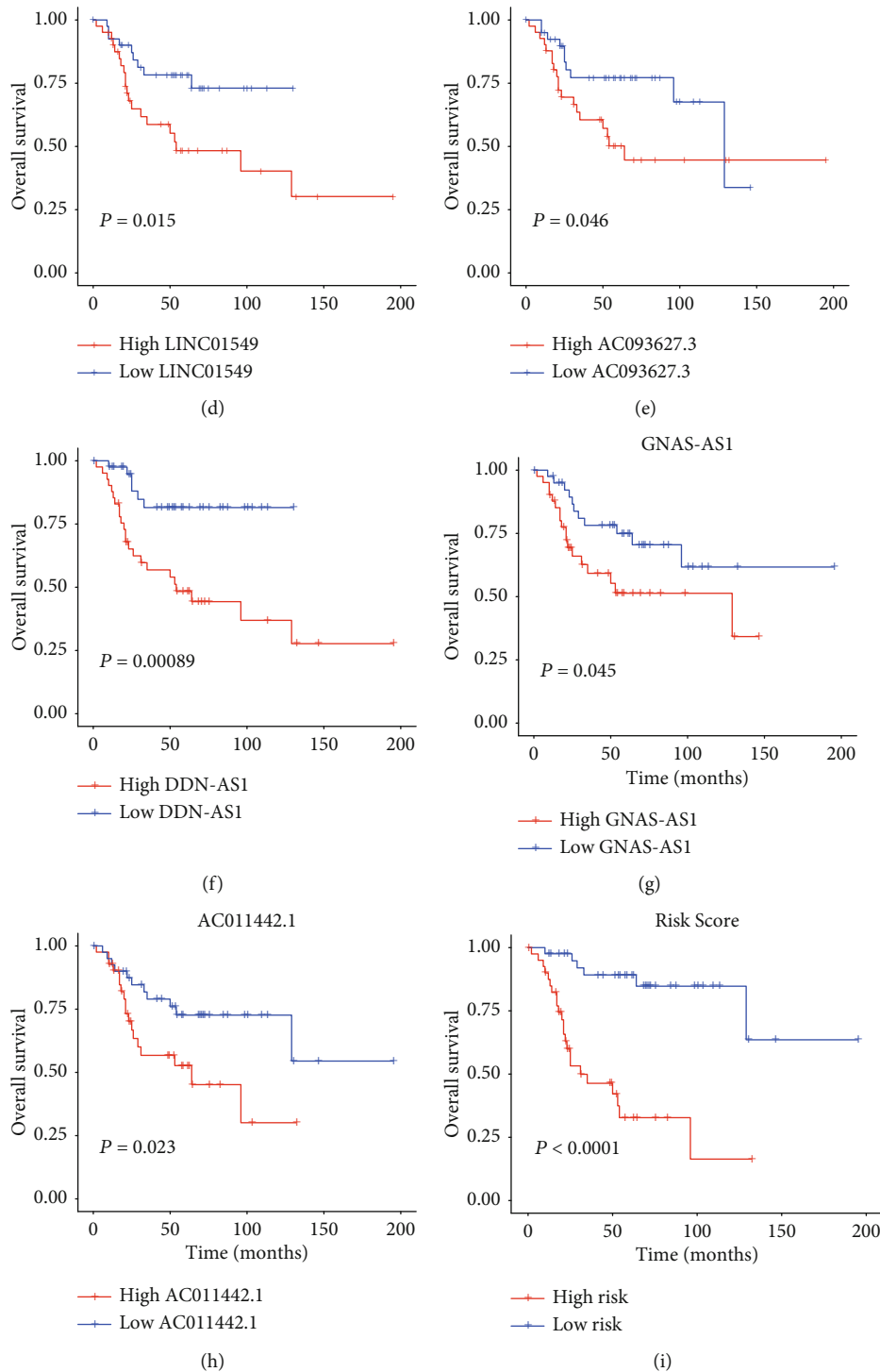


FIGURE 2: The performance of the seven lncRNAs in OS survival prediction. (a) Risk scores for each patient in different groups, where the blue points represent low-risk patients, and red points represent high-risk patients at the top panel. In the middle panel, the distribution of survival time and survival status of two groups of patients, of which the y-axis stands for survival time, blue points represent living patients, and red dots represent the dead patients. The expression patterns of the selected lncRNAs in each OS patient were displayed at the bottom. The Kaplan-Meier curves for survival of patients stratified by the seven lncRNAs were displayed in (b)–(h), respectively. (i) The Kaplan-Meier curve for the samples stratified by the risk score of multivariable Cox regression model.

Encyclopedia of Genes and Genomes (KEGG) pathway [19] enrichment analysis were performed on identified differentially expressed genes with the package clusterProfiler v3.12.0 in R v3.6.3.

2.5. Estimation of Immune Cell Infiltrating Levels. The infiltrating levels of immune cells were estimated based on the gene expression profiles and marker genes of immune cells. Single-sample gene set enrichment analysis (ssGSEA) was

TABLE 2: The comparative analysis of the risk score with other clinical factors in univariable and multivariable Cox regression models.

| Features | P value | Univariable Cox regression | | | P value | Multivariable Cox regression | | |
|----------------------|--------------|----------------------------|--------------|--------------|--------------|------------------------------|--------------|--------------|
| | | HR | Lower 95% CI | Upper 95% CI | | HR | Lower 95% CI | Upper 95% CI |
| Risk score | $6.82E - 12$ | 19.7 | 8.41 | 46.2 | $7.33E - 12$ | 19.7 | 8.41 | 46.4 |
| Gender (female/male) | 0.30 | 0.68 | 0.33 | 1.41 | 0.22 | 0.60 | 0.27 | 1.35 |
| Race (white/other) | 0.23 | 0.64 | 0.30 | 1.34 | 0.47 | 0.75 | 0.35 | 1.64 |
| Age | 0.82 | 1 | 1 | 1 | 0.72 | 1 | 1 | 1 |

employed in this study. This analysis was implemented in R GSVA v1.32.0 package [20].

3. Results

3.1. Identification of Prognostic lncRNAs in OS. As shown in Figure 1(a), the present study conducted a series of data analysis to build a predictive model for OS risk. The gene expression and clinical information of 84 osteosarcoma patients were obtained using the TARGET database, among which, two samples were excluded due to a lack of overall survival time. Based on the gene annotation from the Ensembl database and criteria regarding TPM, we selected 3,159 long noncoding RNA for later establishment of the prognostic risk model (see Materials and Methods). Among these 3,159 lncRNAs, we identified 50 lncRNAs significantly associated with patients' overall survival by univariable Cox regression model (log-rank test, P values < 0.05 , Supplementary Table S1). As illustrated in Figure 1(b), the expression of the prognostic lncRNAs were significantly differentially expressed between the alive and deceased OS patients. These results indicated that the prognostic lncRNAs identified by the univariable Cox regression model may be essential for OS tumorigenesis and/or progression.

3.2. Construction of lncRNA-Based Multivariable Cox Model for Risk Prediction in OS Patients. To build a lncRNA-based Cox regression model for OS risk prediction, we first ranked the prognostic lncRNAs by random forest algorithm, and the top 20 lncRNAs were considered candidates for the construction of an OS prognostic risk model. We then built our model with multivariable Cox regression on the samples with clinic information and expression data of these lncRNAs and obtained seven lncRNAs that significantly contributed to the model (Table 1). Based on the multivariable Cox model, the OS patients were divided into two risk groups using the median risk score. As shown in Figure 2(a), the proportion of deceased samples in the high-risk group (high-risk) was much greater than that in the low-risk group (low-risk) (25/41 vs. 4/41, test of proportion, $P < 0.05$). Moreover, compared with the low-risk group, patients in the high-risk group exhibited significantly lower overall survival time (33.6 vs. 68.7, log-rank test, $P = 1.54E - 5$). Furthermore, patients were then divided into the high-/low-expression groups based on the expression profiles of these seven lncRNAs, respectively. The Kaplan-Meier curves showed a significant association of the seven lncRNAs with overall survival of patients with OS (Figures 2(b)–2(h)). Consistently, the risk

score was observed to have a higher statistical significance than any of the seven prognostic lncRNAs (Figure 2(i)).

In addition, to assess the independence of this scoring in predicting patients' prognosis, we performed both univariable and multivariable Cox regression for samples using the precalculated risk scores and their clinical information such as gender, race, and age. We found that this risk score was an independent indicator for OS patients' survival (Table 2), further suggesting that the risk score by the seven-lncRNA-based Cox model had the potential to predict the risk of OS patients.

3.3. Functional Characterization of Dysregulated Genes in High-Risk and Low-Risk Groups. To investigate dysregulated genes in the two risk groups, we compared the gene expressions of these two risk groups. With thresholds at $|\log_2(\text{fold change})| > 1$ and P value < 0.05 , we identified 864 significant differentially expressed genes (DEGs), and when compared with the low-risk group, the expression of 728 gene was significantly upregulated in the high-risk group, and the expression of another 136 genes was downregulated (Figure 3(a)).

The GO and KEGG pathway enrichment analyses proved that the immune microenvironment of osteosarcoma patients played a crucial role in OS progression. It can be learned that the top 10 GO terms exhibited close association with immunity, including inflammatory responsive response T cell activation, humoral immune response, lymphocyte-mediated immunity, axonemal dynein complex assembly, positive regulation of T cell activation, and regulation of leukocyte cell-cell adhesion (Figure 3(b)), suggesting that the varied immune environment between the high- and low-risk groups may result in their prognostic differences. What is more, from the KEGG pathway enrichment analysis, we observed that a majority of the pathways, where these differentially expressed genes were significantly enriched, consisted of immune-related ones, such as NK cell-mediated cytotoxicity, staphylococcus aureus infection, Th1 and Th2 cell differentiation, antigen processing, and presentation (Figure 3(c)). The consistence between the GO and KEGG enrichment analyses further demonstrated the immune-related biological process may play a key role in OS progression.

3.4. AC011442.1 May Act as an Oncogenic Driver lncRNA in OS. As lncRNAs upregulated or downregulated by copy number alterations (CNA) probably acted as driver lncRNAs in cancer, we performed correlation analysis of the expression level and the corresponding copy number status for the seven prognostic lncRNAs in the multivariable Cox

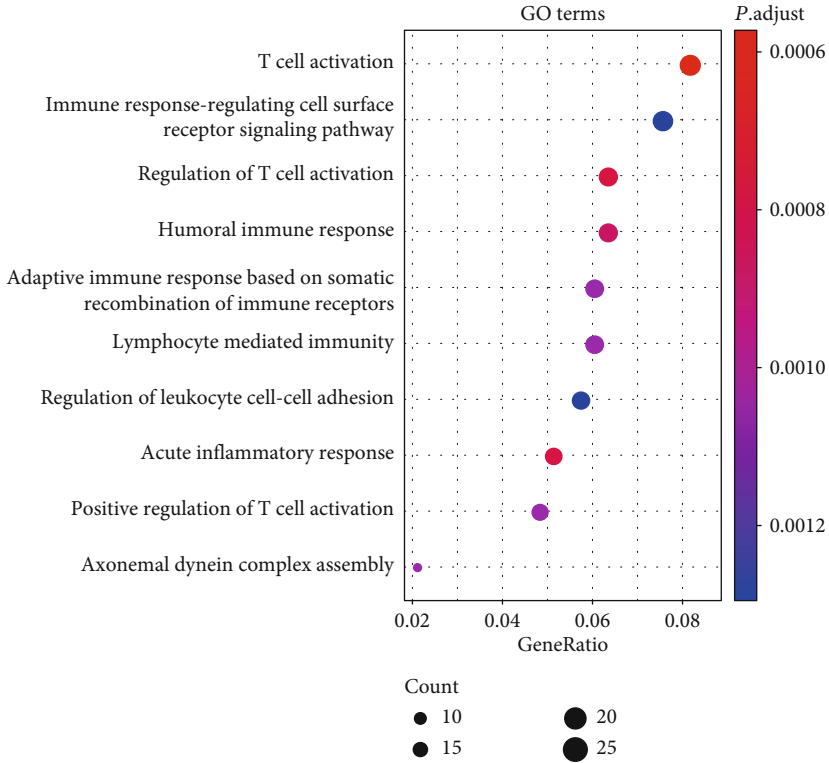
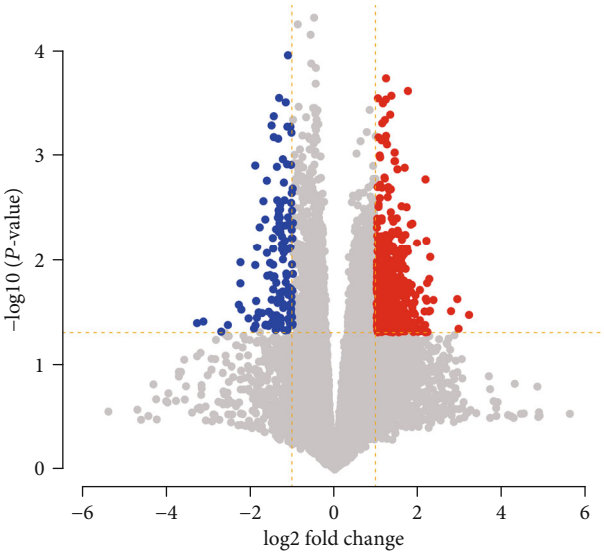


FIGURE 3: Continued.

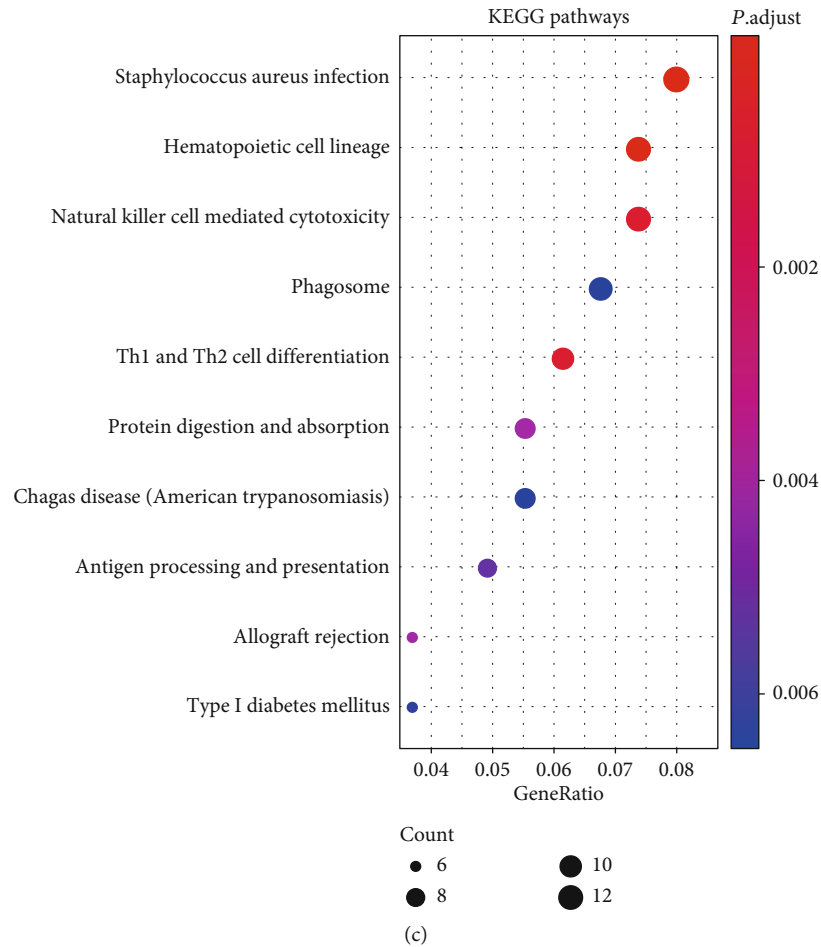


FIGURE 3: The biological differences between the high-risk and low-risk groups stratified by the multivariable Cox regression model. (a) The overview of the differentially expressed genes between the two risk groups. The red and blue points represent the upregulated and downregulated genes in the high-risk group compared with low-risk group. The differentially expressed genes were significantly enriched in GO terms (b) and KEGG pathways (c).

model. We observed that *AC011442.1* was highly upregulated in samples with CNA as compared with wild-type samples ($P < 0.001$, Figure 4(a)). Notably, the copy numbers of the four genes were frequently gained in OS samples (frequency > 10%).

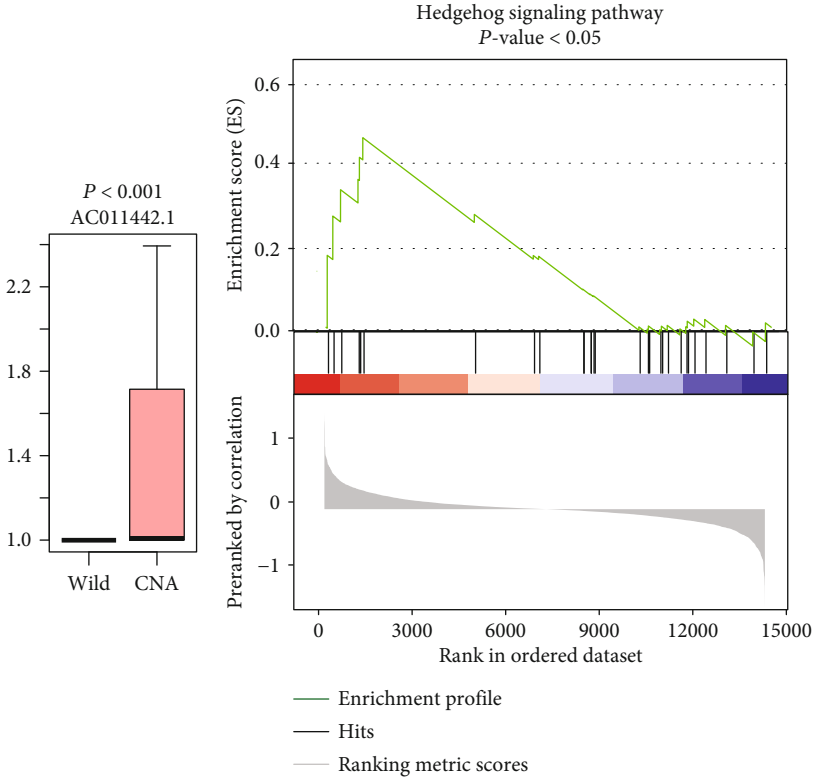
To further investigate the biological function of the four lncRNAs, we conducted gene set enrichment analysis on protein-coding genes that highly correlated with identified lncRNAs. We found that *AC011442.1* was significantly and positively correlated with genes involved in the AMPK signaling pathway and hedgehog signaling pathway, respectively (Figures 4(b) and 4(c), P value < 0.05). These results indicated that *AC011442.1* may enhance the activities of the AMPK signaling pathway and hedgehog signaling pathway.

3.5. The Immune Markers Associated with OS Prognosis. To further explore the immune cells and related markers associated with OS prognosis, we first examined the expression patterns of the immune markers. Specifically, the immune inhibitory genes such as *BTN3A1*, *CD48*, *HAVCR2*, *LAG3*, and *TIGIT* were significantly upregulated in the low-risk group (Figure 5(a), $P < 0.01$), suggesting that the anticancer

activity of the immune cells might be suppressed by these inhibitory genes. Furthermore, we also observed that the relative infiltrating levels of CD8 T cells and activated natural killer cells were attenuated in the high-risk group (Figure 5(b), $P < 0.01$), suggesting that the worse survival in the high-risk group of OS may be caused by the lack of CD8 and NK cells. Consistently, the marker genes of CD8 and NK cells, *CD8A*, *CD8B*, *GZMA*, and *NCR3* were also downregulated in the high-risk group. These findings indicated that the immune cells and related markers were highly associated with OS prognosis.

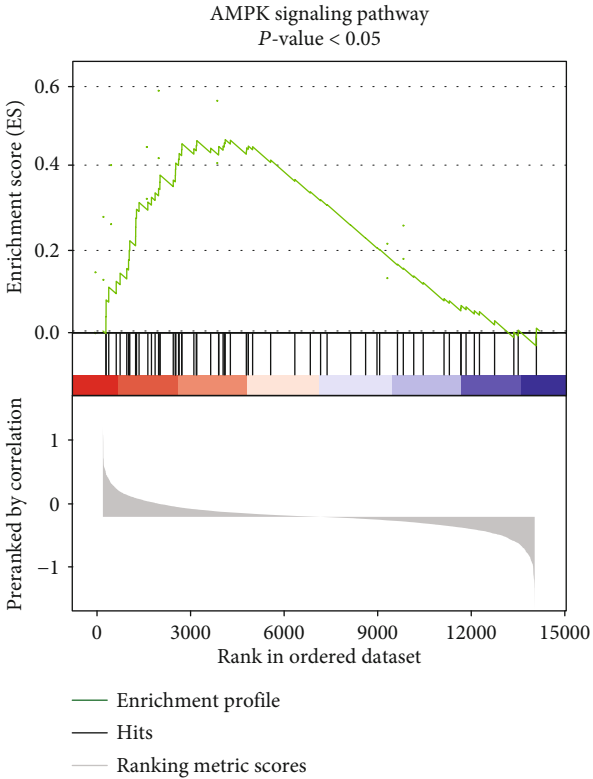
4. Discussion

The molecular mechanism of OS based on protein-coding genes has largely been studied in the past decades. Despite extensive researches about the molecular mechanism of OS, there is still a lack of understanding of the lncRNAs' role in OS tumorigenesis, progression, and metastasis. Meanwhile, the identification of the prognostic lncRNAs involved in OS can facilitate the development of new diagnostic or therapeutic biomarkers.



(a)

(b)



(c)

FIGURE 4: The oncogenic driver lncRNA AC011442.1 and its functionality. (a) The expression patterns of AC011442.1 in OS patients with CNAs and wild type. (b) The predicted pathways that the AC011442.1 might participate in.

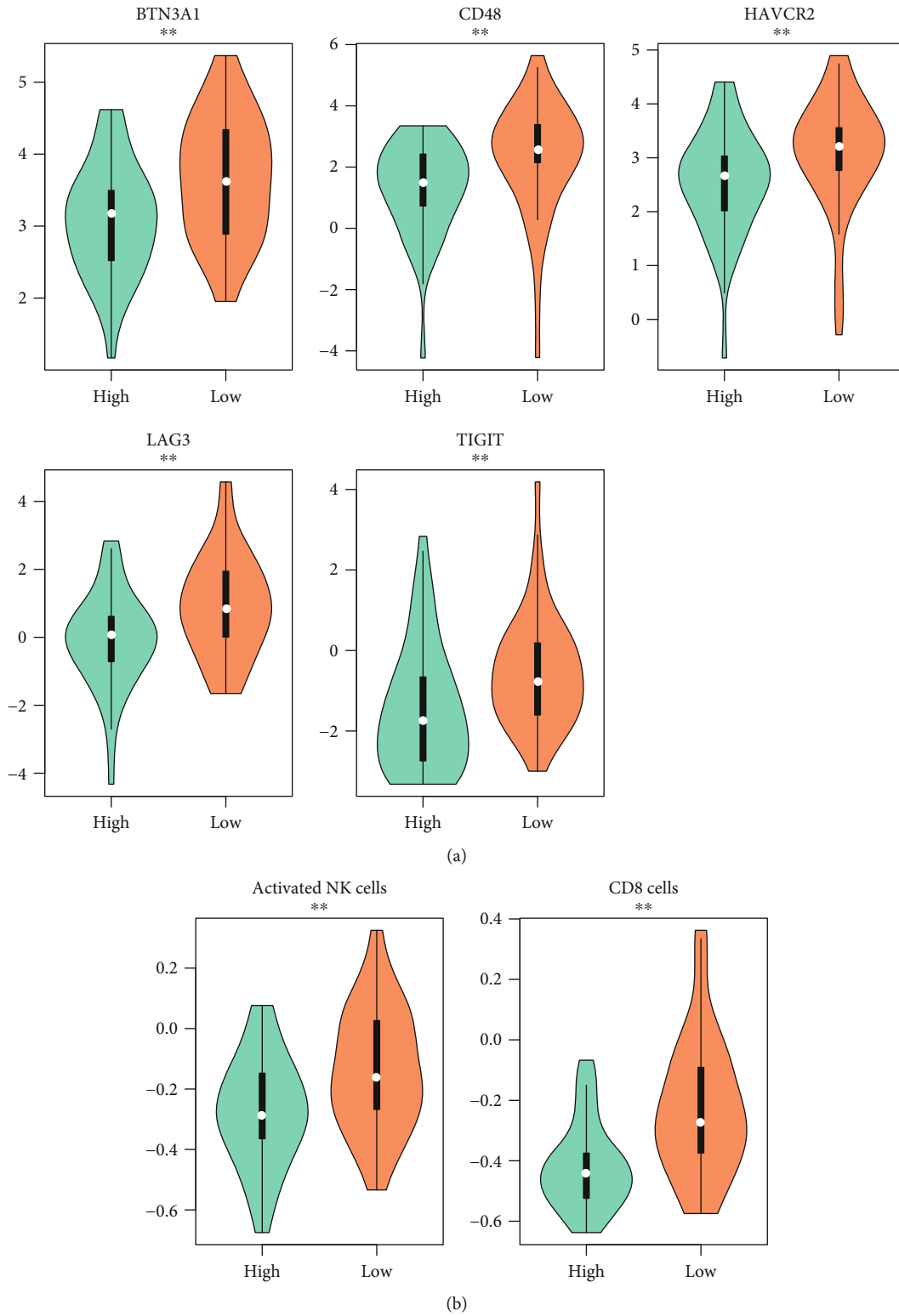


FIGURE 5: Continued.

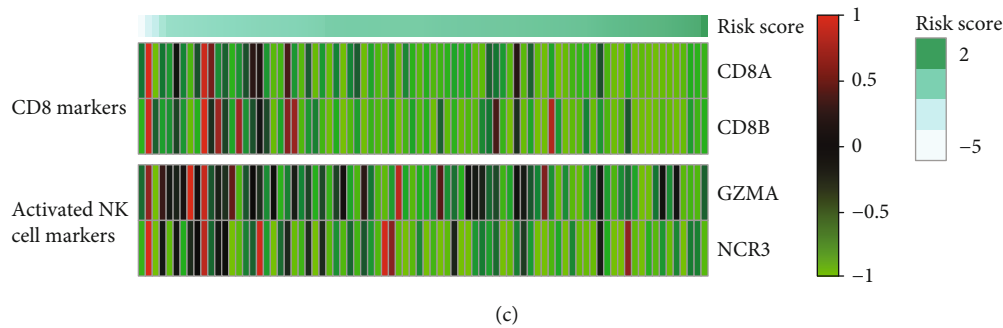


FIGURE 5: OS prognosis-related immune signatures. (a) The expression patterns of immune inhibitory receptors in the two risk groups. (b) The differential infiltrating levels of CD8 T and activated NK cells. The high-risk and low-risk groups are represented by the labels of “high” and “low” and colored by green and orange, respectively. (c) The expression profiles of the marker genes of the CD8 and NK cells. The samples are ordered by the risk scores.

In the present study, we collected 82 OS samples with RNA-seq-based gene expression data and their clinical information from the TARGET database. We found that 50 lncRNAs were significantly associated with patients’ survival by a univariable Cox regression model (P values < 0.05). Through the selection of prognostic lncRNAs, we identified 7 lncRNAs with significant performance in OS survival prediction, built multivariable Cox regression model under the 7 lncRNAs, and successfully stratified patients into different risk groups with distinctive survival outcomes. Notably, *DDN-AS1*, one of the seven lncRNAs used by the multivariable Cox regression model, has been reported to act as competing endogenous RNA (ceRNA) that promoted the expression of TCF3 through competitively binding miR-15a and miR-16 [21], suggesting that *DDN-AS1* may promote OS progression in a similar manner. We further analyzed gene expression profiles of patients in different risk groups and obtained a list of DEGs. Functional enrichment analysis revealed that significantly enriched GO terms and pathways were associated with many aspects of immunity, indicating that immune-related functions are critical for OS survival, which is consistent with previous studies [22].

As dysregulated lncRNAs caused by copy number alterations (CNA) may act as driver lncRNAs in cancer, correlation analysis of the expression level and the corresponding copy number status for the seven prognostic lncRNAs was performed to identify the driver lncRNAs. Notably, *AC011442.1* was also predicted as one of the four-driver lncRNA ($P < 0.001$). Interestingly, *AC011442.1* was predicted to participate in cancer-related pathways [23–25], including the AMPK signaling pathway and hedgehog signaling pathway. Hedgehog signaling pathways have been frequently observed to drive tumorigenesis and metastasis of OS [26]. These results further demonstrated that the driver lncRNAs played a key role in OS, which could be used for further research of molecular mechanism.

As the exploration into varied molecular patterns between the two risk groups revealed that the immune-related pathways were enriched by DEGs in OS, we then examined whether the abundance of immune cells and markers were associated with OS prognosis. Specifically, the immune inhibitors such as *BTN3A1*, *CD48*, *HAVCR2*,

LAG3, and *TIGIT*; CD8 T and activated NK cells; and related markers were significantly downregulated in the high-risk group. Particularly, *CD48*, *HAVCR2*, *LAG3*, and *TIGIT* were identified as novel immunotherapeutic targets of several cancers [27–30], suggesting that the low-risk OS patients might benefit from their candidate inhibitors.

In addition, the limitations of this study should be pointed out. Firstly, the multivariable Cox regression model needs an independent gene expression data for the validation of its robustness. Secondly, though a list of dysregulated lncRNAs associated with OS survival was identified, but future experimental verification is still needed. Moreover, detailed molecular functions of identified dysregulated lncRNAs had not been thoroughly discussed in this study. We hope that, when validation datasets become available in the near future, we can further confirm our findings and perform experimental validation. In summary, the identification of novel prognostic lncRNAs in OS would not only improve our understanding of the lncRNAs involved in OS tumorigenesis or progression but also assist the prediction of OS survival and development of molecularly targeted therapies to some extent, which in turn benefit patients’ survival.

Data Availability

Previously reported gene expression and clinical data were used to support this study and are available at TARGET (Therapeutically Applicable Research To Generate Effective Treatments) database (<https://ocg.cancer.gov/programs/target/data-matrix>). These prior studies (and datasets) are cited at relevant places within the text as references [14].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

Y.G. and Z.Y. designed this study. H.G., Y. G., and M. Z. conducted the experiments. H.G., Y. G., and Y. Z. contributed to the writing of the paper and setting of figures. Hua Gao and Yuanyuan Guo contributed equally to this work.

Supplementary Materials

Supplementary Table S1: the 50 prognostic long noncoding RNAs and statistical significance (*P* value) in osteosarcoma. (*Supplementary Materials*)

References

- [1] A. Longhi, C. Errani, M. De Paolis, M. Mercuri, and G. Bacci, "Primary bone osteosarcoma in the pediatric age: state of the art," *Cancer Treatment Reviews*, vol. 32, no. 6, pp. 423–436, 2006.
- [2] C. Munoz-Bendix, P. J. Sloty, H. J. Steiger, S. A. Ahmadi, R. Bostelmann, and J. F. Cornelius, "Primary bone tumors of the spine revisited: a 10-year single-center experience of the management and outcome in a neurosurgical department," *Journal of Craniovertebral Junction and Spine*, vol. 6, no. 1, pp. 21–29, 2015.
- [3] L. H. Chung, P. K. Wu, C. F. Chen, H. K. Weng, T. H. Chen, and W. M. Chen, "Pathological fractures in predicting clinical outcomes for patients with osteosarcoma," *BMC Musculoskeletal Disorders*, vol. 17, no. 1, p. 503, 2016.
- [4] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2017," *CA: A Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7–30, 2017.
- [5] M. E. Anderson, "Update on survival in osteosarcoma," *The Orthopedic Clinics of North America*, vol. 47, no. 1, pp. 283–292, 2016.
- [6] E. Simpson and H. L. Brown, "Understanding osteosarcomas," *Journal of the American Academy of Physician Assistants*, vol. 31, no. 8, pp. 15–19, 2018.
- [7] Z. Wang, M. Tan, G. Chen, Z. Li, and X. Lu, "LncRNA SOX2-OT is a novel prognostic biomarker for osteosarcoma patients and regulates osteosarcoma cells proliferation and motility through modulating SOX2," *IUBMB Life*, vol. 69, no. 11, pp. 867–876, 2017.
- [8] G. Yang, X. Lu, and L. Yuan, "LncRNA: a link between RNA and cancer," *Biochimica et Biophysica Acta*, vol. 1839, no. 11, pp. 1097–1109, 2014.
- [9] C. Liu and J. Lin, "Long noncoding RNA ZEB1-AS1 acts as an oncogene in osteosarcoma by epigenetically activating ZEB1," *American Journal of Translational Research*, vol. 8, no. 10, pp. 4095–4105, 2016.
- [10] J. Xu, R. Ding, and Y. Xu, "Effects of long non-coding RNA SPRY4-IT1 on osteosarcoma cell biological behavior," *American Journal of Translational Research*, vol. 8, no. 12, pp. 5330–5337, 2016.
- [11] F. Chen, J. Mo, and L. Zhang, "Long noncoding RNA BCAR4 promotes osteosarcoma progression through activating GLI2-dependent gene transcription," *Tumor Biology*, vol. 37, no. 10, pp. 13403–13412, 2016.
- [12] Z. Yin, H. Ding, E. He, J. Chen, and M. Li, "Overexpression of long non-coding RNA MFI2 promotes cell proliferation and suppresses apoptosis in human osteosarcoma," *Oncology Reports*, vol. 36, no. 4, pp. 2033–2040, 2016.
- [13] Y. Wang, X. Zeng, N. Wang et al., "Long noncoding RNA DANCR, working as a competitive endogenous RNA, promotes ROCK1-mediated proliferation and metastasis via decoying of miR-335-5p and miR-1972 in osteosarcoma," *Molecular Cancer*, vol. 17, no. 1, p. 89, 2018.
- [14] P. Xia, R. Gu, W. Zhang, and Y. F. Sun, "lncRNA CEBPA-AS1 overexpression inhibits proliferation and migration and stimulates apoptosis of OS cells via Notch signaling," *Molecular Therapy - Nucleic Acids*, vol. 19, pp. 1470–1481, 2020.
- [15] Z. Yang, X. Li, Y. Yang, Z. He, X. Qu, and Y. Zhang, "Long noncoding RNAs in the progression, metastasis, and prognosis of osteosarcoma," *Cell Death & Disease*, vol. 7, no. 9, p. e2389, 2016.
- [16] X. Ma, Y. Liu, Y. Liu et al., "Pan-cancer genome and transcriptome analyses of 1, 699 paediatric leukaemias and solid tumours," *Nature*, vol. 555, no. 7696, pp. 371–376, 2018.
- [17] P. Flicek, M. R. Amode, D. Barrell et al., "Ensembl 2011," *Nucleic Acids Research*, vol. 39, pp. D800–D806, 2010.
- [18] The Gene Ontology Consortium, "The Gene Ontology resource: 20 years and still GOing strong," *Nucleic Acids Research*, vol. 47, no. D1, pp. D330–D338, 2019.
- [19] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.
- [20] S. Hanzelmann, R. Castelo, and J. Guinney, "GSVA: gene set variation analysis for microarray and RNA-seq data," *BMC Bioinformatics*, vol. 14, no. 1, p. 7, 2013.
- [21] Z. Liu, M. Wu, H. Shi, C. Huang, S. Luo, and X. Song, "DDN-AS1-miR-15a/16-TCF3 feedback loop regulates tumor progression in cervical cancer," *Journal of Cellular Biochemistry*, vol. 120, no. 6, pp. 10228–10238, 2019.
- [22] 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 103711, 2019.
- [23] S. M. Jeon and N. Hay, "The double-edged sword of AMPK signaling in cancer and its therapeutic implications," *Archives of Pharmacal Research*, vol. 38, no. 3, pp. 346–357, 2015.
- [24] M. Evangelista, H. Tian, and F. J. de Sauvage, "The hedgehog signaling pathway in cancer," *Clinical Cancer Research*, vol. 12, no. 20, pp. 5924–5928, 2006.
- [25] Y. Xia, S. Shen, and I. M. Verma, "NF- κ B, an active player in human cancers," *Cancer Immunology Research*, vol. 2, no. 9, pp. 823–830, 2014.
- [26] Z. Yao, L. Han, Y. Chen et al., "Hedgehog signalling in the tumourigenesis and metastasis of osteosarcoma, and its potential value in the clinical therapy of osteosarcoma," *Cell Death & Disease*, vol. 9, no. 6, p. 701, 2018.
- [27] Y. He, J. Cao, C. Zhao, X. Li, C. Zhou, and F. R. Hirsch, "TIM-3, a promising target for cancer immunotherapy," *Onco Targets and Therapy*, vol. 11, pp. 7005–7009, 2018.
- [28] L. P. Andrews, A. E. Marciscano, C. G. Drake, and D. A. A. Vignali, "LAG3 (CD223) as a cancer immunotherapy target," *Immunological Reviews*, vol. 276, no. 1, pp. 80–96, 2017.
- [29] B. L. Solomon and I. Garrido-Laguna, "TIGIT: a novel immunotherapy target moving from bench to bedside," *Cancer Immunology, Immunotherapy*, vol. 67, no. 11, pp. 1659–1667, 2018.
- [30] N. Hosen, H. Ichihara, A. Mugitani et al., "CD48 as a novel molecular target for antibody therapy in multiple myeloma," *British Journal of Haematology*, vol. 156, no. 2, pp. 213–224, 2012.

Research Article

Construction of Protein-related Risk Score Model in Bladder Urothelial Carcinoma

Qizhan Luo^{1,2} and Xiaobo Zhang¹ 

¹Xiangya International Medical Center, Department of Geriatrics, Xiangya Hospital, Central South University, Changsha, China

²Department of Urology, RWTH Aachen University, Pauwelsstrasse 30, 52072 Aachen, Germany

Correspondence should be addressed to Xiaobo Zhang; 549651566@qq.com

Received 24 May 2020; Accepted 6 July 2020; Published 3 August 2020

Guest Editor: Tao Huang

Copyright © 2020 Qizhan Luo and Xiaobo Zhang. 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. Though there are several prognostic models, there is no protein-related prognostic model. The aim of this study is to identify possible prognostic-related proteins in bladder urothelial carcinoma and to try to predict the prognosis of bladder urothelial carcinoma based on these proteins. **Methods.** Profile data and corresponding clinical traits were obtained from The Cancer Proteome Atlas (TCPA) and The Cancer Genome Atlas (TCGA) expression. Survival-associated protein in bladder urothelial carcinoma patients were estimated with Kaplan-Meier (KM) test and COX regression analysis. The potential molecular mechanisms and properties of these bladder urothelial carcinoma-specific proteins were also explored with the help of computational skills. The risk score model was validated in different clinical traits. Sankey diagram representation is for protein correlation. A new prognostic-related risk model based on proteins was developed by using multivariable COX analysis. Next, the alteration of the corresponding genes to the 6 prognostic-related proteins was analyzed. Finally, the relation between the corresponding genes and the immune infiltration was analyzed using the TIMER. **Results.** Six proteins were identified to be associated with the prognosis of bladder urothelial carcinoma. A prognostic signature based on proteins (BECLIN, EGFR, PKCALPHA, SRC, ANNEXIN1, and AXL) performed moderately in prognostic predictions. The alteration of corresponding genes was in 31(24%) sequenced cases. ANXA1, AXL, and EGFR were positively related to CD8+ T cell. **Conclusion.** Our results screened six proteins of clinical significance. The importance of a personalized protein signature model in the recognition, surveillance. The abnormal expression of six prognostic-related proteins may be caused by corresponding gene alteration. Furthermore, these proteins may affect survival via the immune infiltration.

1. Introduction

One of the most common urological carcinomas is bladder urothelial carcinoma. It is a complex biological mechanism. It is well known for its rapid metastasis to another part of the body also it has high recurrence rates [1]. The genesis of bladder urothelial carcinoma is highly associated with smoking, sex, age, schistosomiasis infection, and chemical contact [2, 3]. Current treatment and prognosis still heavily

rely on clinical and pathologic staging that does not always reflect the individual condition of the patient.

The mechanism of bladder cancer is found as being a result of alterations in different molecular and pathways. Based on alterations in these molecular or pathways, different assessment of biomarkers that give novel insights into bladder urothelial carcinoma mechanism [4]. Therefore, future treatment and prognosis will add prognostic molecular models for risk stratification and management personalization.

TABLE 1: 17 proteins were significantly correlated to prognosis in BC.

| Gene | KM | HR | HR.95 L | HR.95H | P value |
|----------|-------------|-------------|-------------|-------------|-------------|
| ANNEXIN1 | 1.15E-05 | 1.362261548 | 1.166818065 | 1.590442059 | 9.13E-05 |
| TAZ | 0.00013013 | 1.886722131 | 1.12970026 | 3.151030876 | 0.015266466 |
| SF2 | 0.000454652 | 0.510368312 | 0.2776481 | 0.938150894 | 0.030348532 |
| BAK | 0.000708446 | 0.588931975 | 0.353990253 | 0.979803449 | 0.041497671 |
| SRC | 0.000839191 | 0.638626026 | 0.484164999 | 0.842364074 | 0.001502422 |
| GATA3 | 0.000860212 | 0.781102448 | 0.66153008 | 0.922287668 | 0.003565311 |
| EGFR | 0.001732026 | 1.369395712 | 1.119307856 | 1.6753609 | 0.002247458 |
| ARID1A | 0.001860228 | 0.545250791 | 0.297422122 | 0.999584103 | 0.049842967 |
| AXL | 0.002309345 | 1.744438807 | 1.024848166 | 2.969285454 | 0.040324605 |
| GATA6 | 0.003401439 | 2.020152915 | 1.085751303 | 3.75870403 | 0.026441307 |
| CABL | 0.00431513 | 1.522389409 | 1.073197575 | 2.159592573 | 0.018474699 |
| SMAC | 0.014092504 | 0.603575018 | 0.425783275 | 0.855606182 | 0.004569768 |
| RICTOR | 0.01732931 | 1.157220752 | 1.040793571 | 1.286671927 | 0.00695473 |
| BECLIN | 0.027199215 | 2.272028814 | 1.160499759 | 4.448182682 | 0.016655813 |
| ADAR1 | 0.03454304 | 0.731078894 | 0.543672655 | 0.983084848 | 0.038185369 |
| SMAD3 | 0.036176387 | 0.464509845 | 0.254776118 | 0.846898045 | 0.0123406 |
| PKCALPHA | 0.044446818 | 1.362756846 | 1.001230588 | 1.854823697 | 0.049093018 |

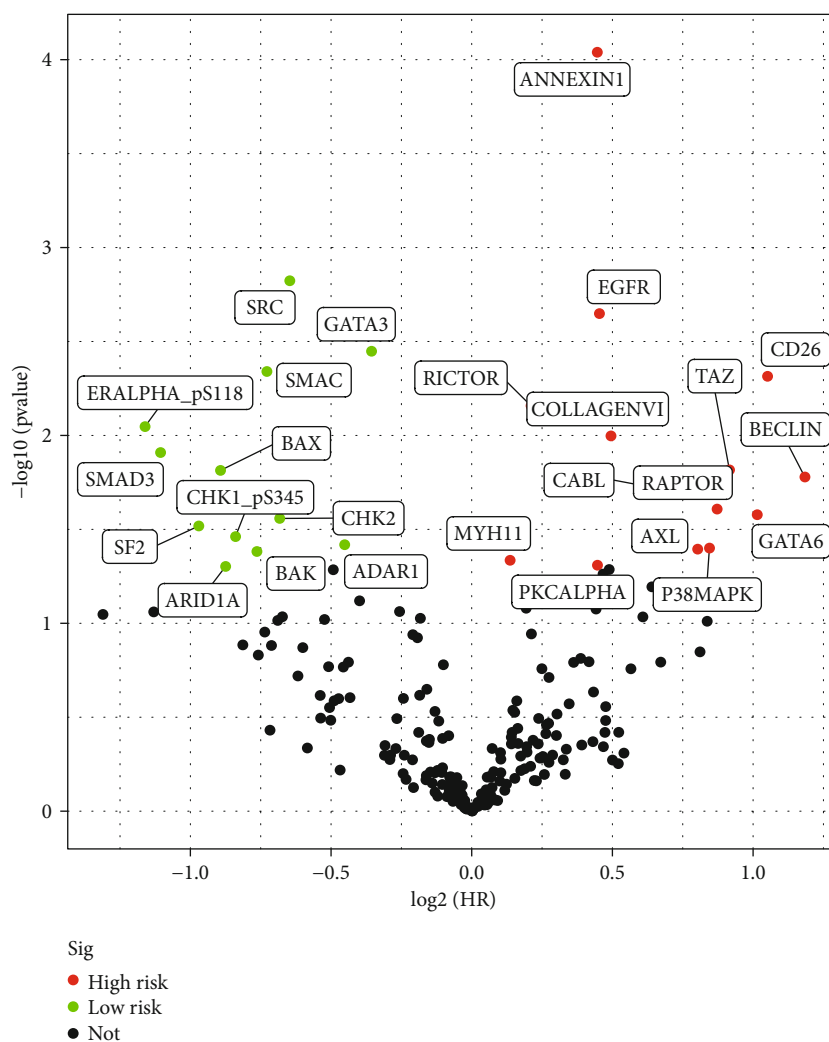


FIGURE 1: Volcano plots show significant prognostic proteins in the high-risk and low-risk groups.

TABLE 2: 6 highly optimal prognostic-related proteins and their coefficient.

| Id | Coef | HR | HR.95L | HR.95H | P value |
|----------|--------------|-------------|-------------|-------------|-------------|
| BECLIN | 1.284587895 | 3.613178642 | 1.827081627 | 7.145307415 | 0.000222126 |
| EGFR | 0.286373326 | 1.331589479 | 1.055793909 | 1.679428652 | 0.015585541 |
| PKCALPHA | 0.242223245 | 1.274078593 | 0.949494745 | 1.70962111 | 0.106413498 |
| SRC | -0.240328123 | 0.786369793 | 0.579236289 | 1.067573739 | 0.123376435 |
| ANNEXIN1 | 0.293178172 | 1.340681641 | 1.122426065 | 1.60137698 | 0.001221101 |
| AXL | 0.555691215 | 1.743145457 | 1.01996076 | 2.979091162 | 0.042128919 |

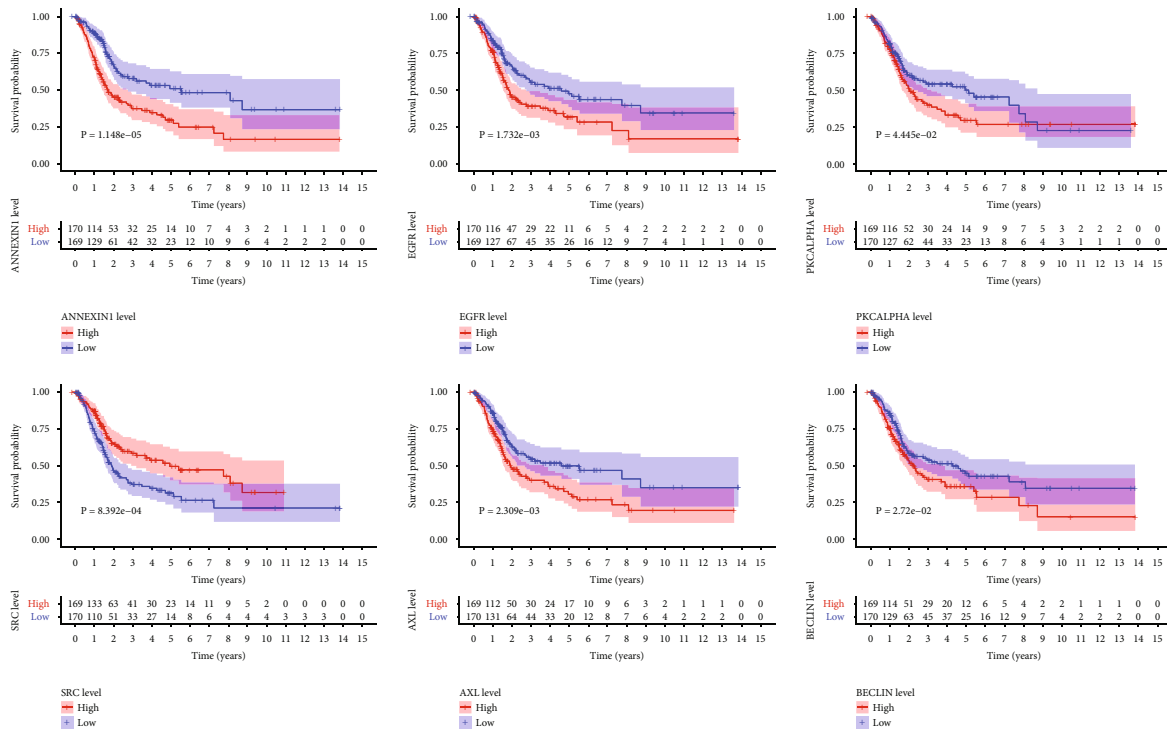


FIGURE 2: Kaplan-Meier survival curve among high-risk and low-risk groups based on the 6 highly optimal prognostic-related protein. Patients were assigned to high-risk and low-risk groups according to the median level of each highly optimal prognostic-related protein expression.

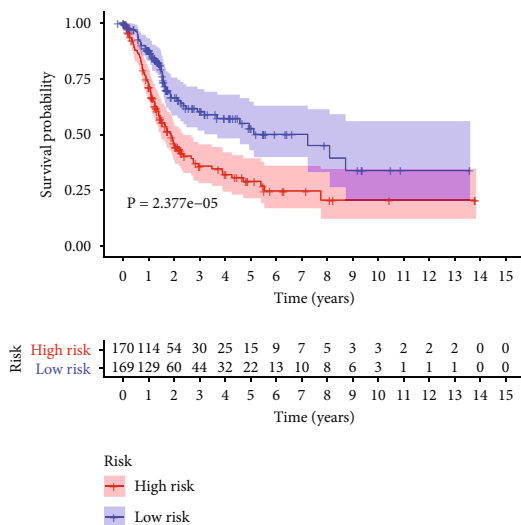


FIGURE 3: Kaplan-Meier survival curve among two groups based on risk score.

TCGA and TCPA have a great advantage to help our understanding of the complex mechanism. There is no study that is a protein-related risk score model. We downloaded the data from both databases and developed a prognostic signature model based on bladder urothelial carcinoma proteins.

2. Method

2.1. *The Protein Data BC from TCPA and the Clinical Data from TCGA Were Analysed.* The clinical data of 409 bladder cases were obtained from TCGA dataset on the 9th of Feb 2020. The protein profiles of 344 samples were obtained from TCPA dataset on the 9th of Feb 2020. The clinical data was level 4. Because our data was downloaded directly from TCGA and TCPA databases, and we strictly observe the publishing guidelines provided by TCGA and TCPA; there was no requirement for ethical approvals.

2.2. *Identification of Survival-Associated Proteins.* Missing protein data was supplied and was performed with R software

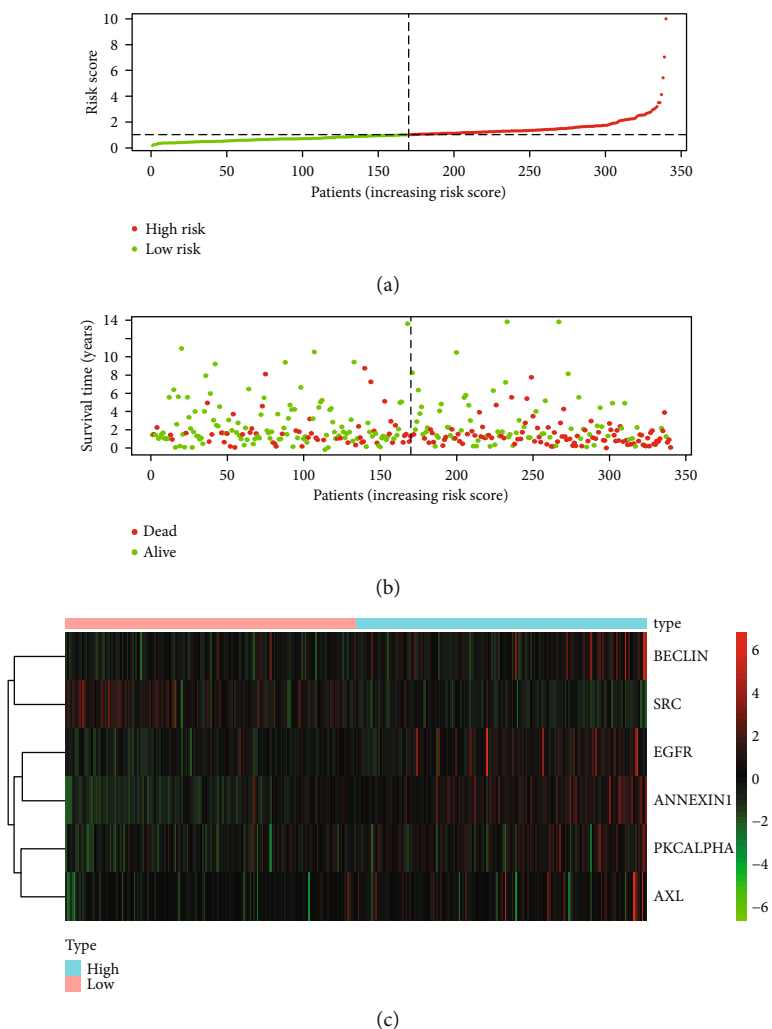


FIGURE 4: Risk curve shows the relationship between risk scores and survival time and status. (a). Rank of risk score and distribution of groups. (b). Survival status of patients in different groups. (c). Heatmap of expression profiles of included proteins.

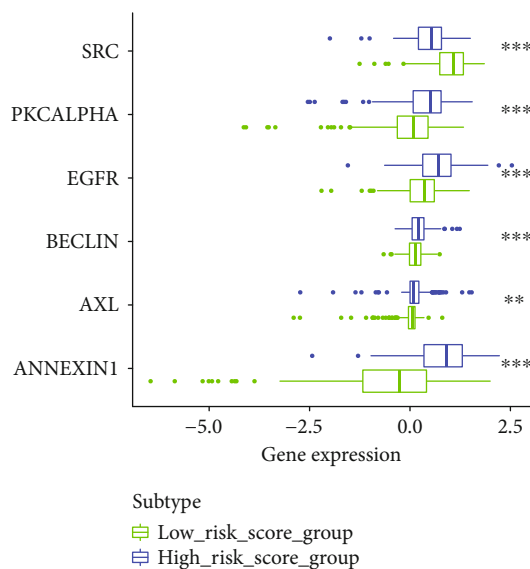


FIGURE 5: Highly prognostic-related proteins in low-risk score group and high-risk score group.

package “impute.” Survival time and survival status were extracted from the clinical characteristic which was corresponding to the expression data of proteins. To investigate the prognostic value of proteins in bladder urothelial carcinoma patients, single-factor Cox regression analysis and Kaplan-Meier analysis with the “survival” package and the P value <0.05 . Kaplan-Meier analysis was performed based on the mediate of each protein expression. Patients have been divided into high-risk and low-risk groups. When the P value of single factor Cox regression analysis and Kaplan-Meier analysis both were <0.05 , the significant prognostic-related proteins were obtained to build the protein-related prognosis model. Volcano plots were conducted with “dplyr,” “ggplot2,” and “ggrepel” R software to show significant proteins in the high risk and low-risk groups, two groups based on mediate of each protein expression.

2.3. Construction of a Prognostic Signature. The prognosis-related proteins were analyzed via the multifactor Cox regression model with R software “survival” package. Both direction selections were carried out based on the significant prognostic-related proteins. Kaplan-Meier survival curve was

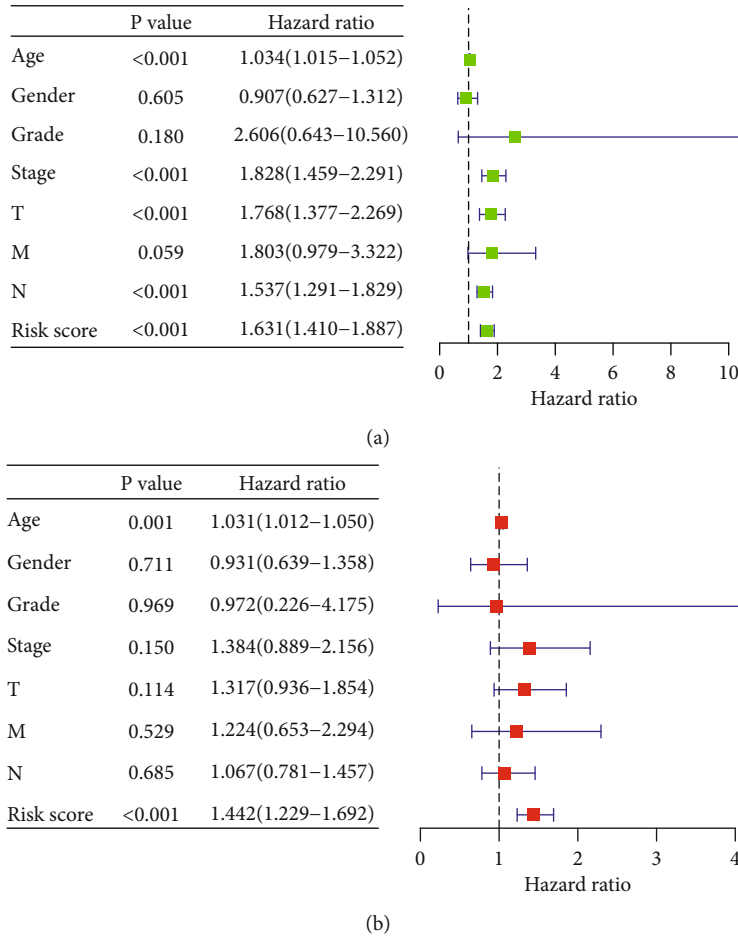


FIGURE 6: (a) Univariate analyses. (b) Multivariate analyses to verify risk scores were an independent prognostic factor.

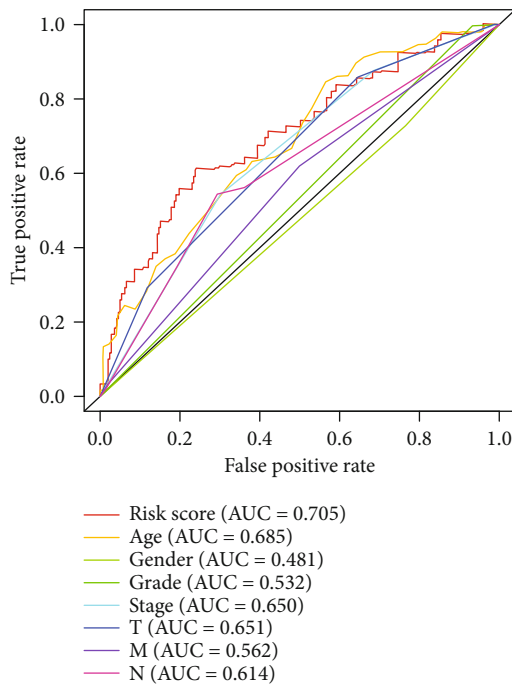


FIGURE 7: The prognostic accuracy of risk score.

performed by “survival” and “survminer” package based on the level of highly optimal prognostic-related protein expression. Patients were assigned to high-risk and low-risk groups according to the median value of highly optimal prognostic-related protein expression. Risk scores were obtained based on the expression level of each highly optimal prognostic-related protein multiplied by its corresponding regression coefficient. Using the above methods, it has been calculated the risk scores of each case. Patients were assigned to high-risk and low-risk groups according to the median risk scores.

2.4. Validation of the Protein Signature. Kaplan-Meier survival curve was performed by “survival” and “survminer” package to compare the survival time of two groups. The risk curve was performed with the R software “heatmap” package to show the relationship between risk scores and survival time. Highly prognostic-related proteins were compared between two groups with Wilcox. test. It has been combined risk score with the available pathologic and the clinical traits in univariate and multivariate analysis to verify whether the risk score was an independent prognostic factor. The prognostic accuracy of risk score was measured using the AUC of the ROC curve with 0.5 indicating random chance and 1.0 indicating perfect classification. AUC of ROC curve was carried out by the R software package survival ROC. For further

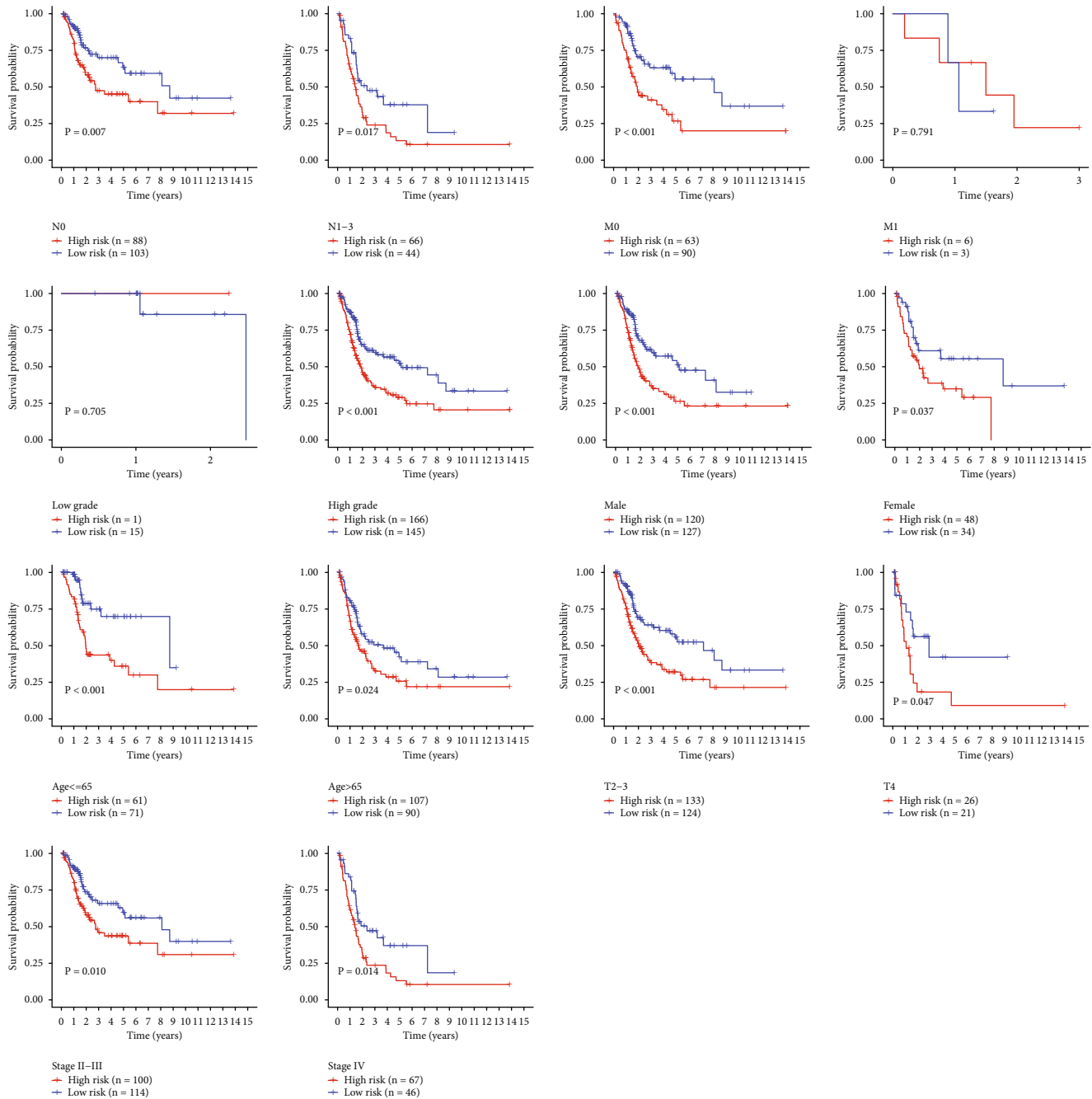


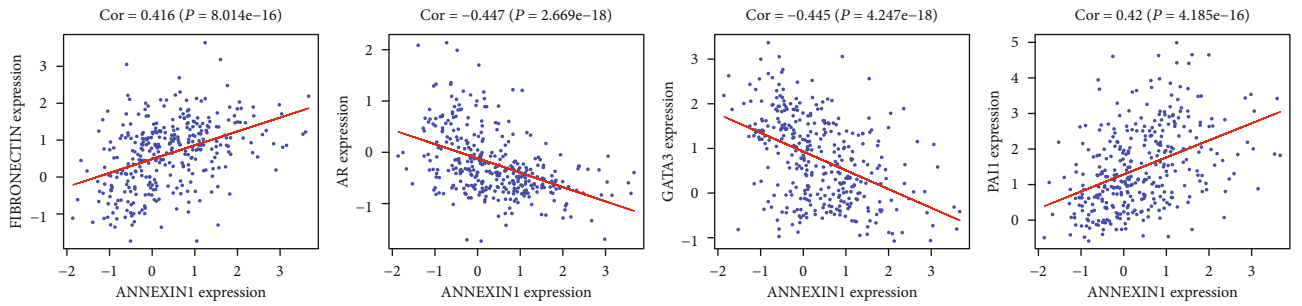
FIGURE 8: Survival time in high-risk score group and low-risk score group in different clinical traits.

validation of the risk score model, survival time was compared between the high-risk group and low-risk group in different clinical traits.

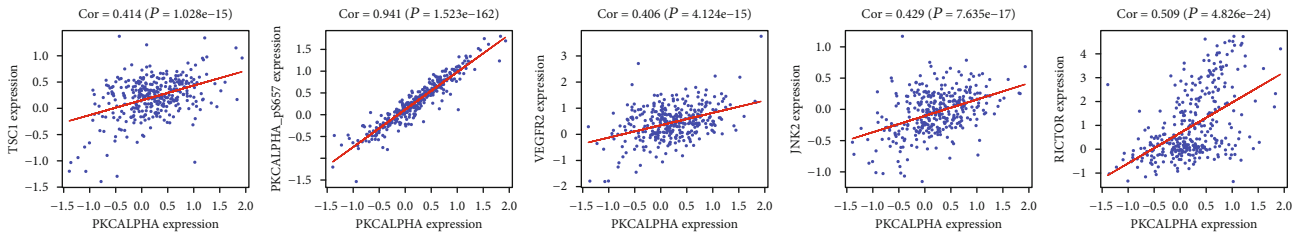
2.5. Correlation between the Highly Prognostic-Related Proteins and Expression Protein. To explore the regulatory mechanisms of prognostic proteins, it has been chosen the highly prognostic-related proteins among the expression protein of BC patients. The correlation was analyzed using Pearson correlation analysis. CorFilter equals 0.4 and P value is < 0.001 . Sankeycharts were performed with R software ggplot2, ggalluvial, and dplyr package.

2.6. The Alteration of the Corresponding Genes to the 6 Prognostic-Related Proteins. We transferred the protein ID to gene ID. The corresponding genes to the 6 prognostic-related proteins were analyzed in the cBioPortal database. The alteration in selected genes was shown website (<http://www.cbioportal.org/>). The styles of alteration of genes also were analyzed in the cBioPortal database.

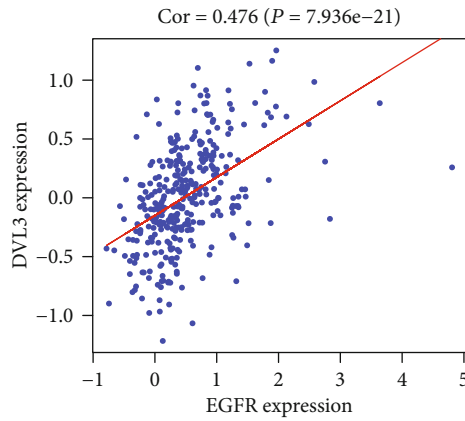
2.7. Correlations between the Corresponding Genes Expression and Immune Cells in TIMER. The corresponding genes to the 6 prognostic-related proteins were analyzed in TIMER. The relationship between the corresponding genes and the



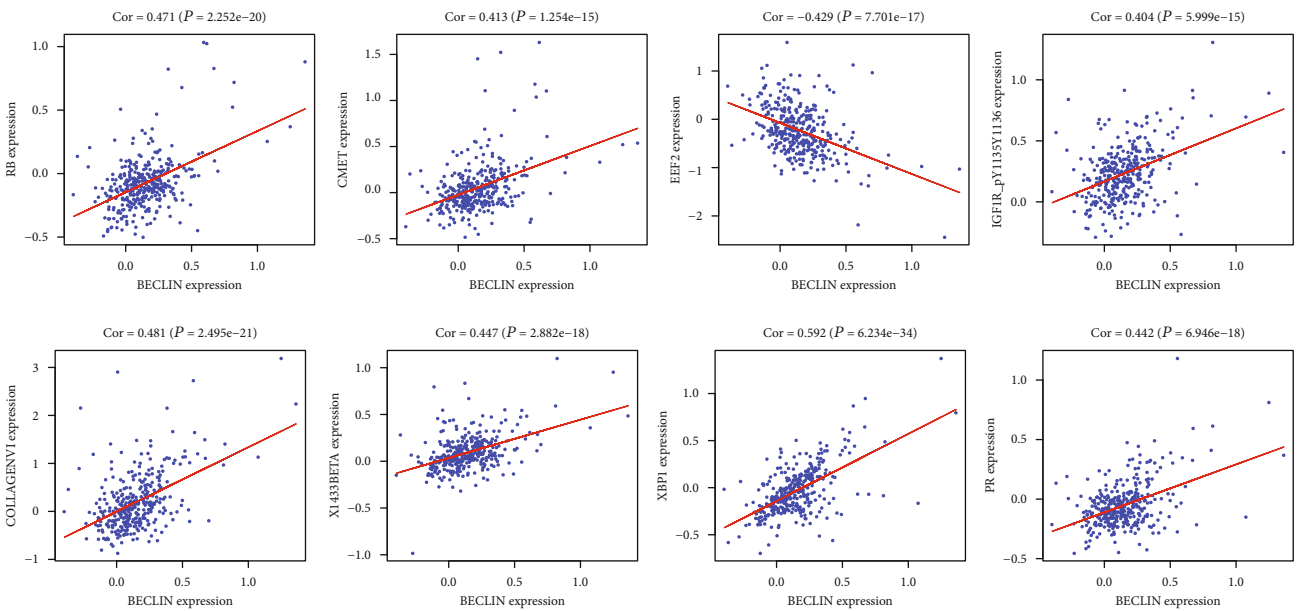
(a)



(b)

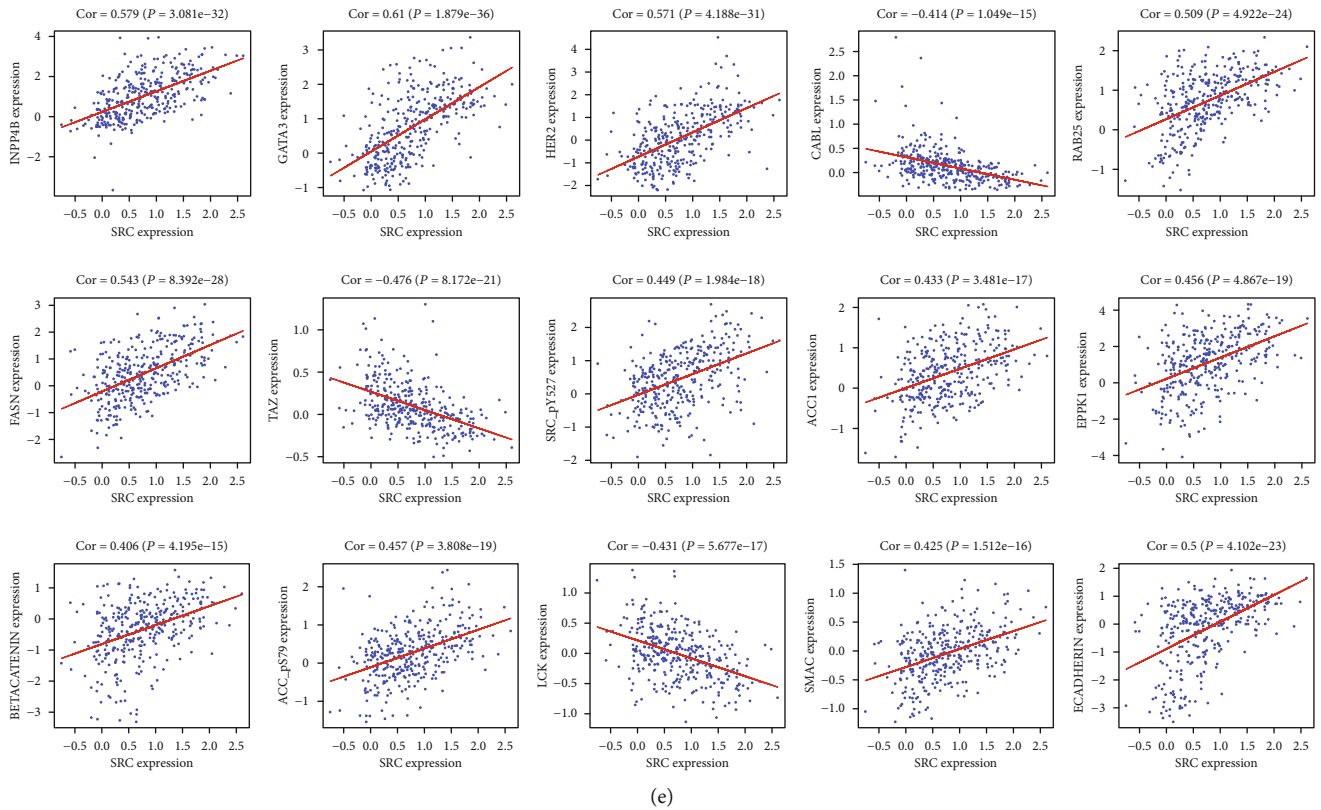


(c)



(d)

FIGURE 9: Continued.



(e)

FIGURE 9: The correlation of protein. (a). Correlation between ANNEXIN1 and (FIBRONECTIN, PAI1, AR, GATA3). (b). Correlation between PKCALPHA and (PKCALPHA_pS657, VEGFR2, RICTOR, TSC1). (c). Correlation between EGFR and DVL3. (d). Correlation between BECLIN and (COLLAGENVI, PR, XBPI, X1433BETA, CMET, RB, IGF1R_pY1135Y1136, EEF2). (e). Correlation between SRC and (ACC_pS79, ACC1, BETACATENIN, ECADHERIN, GATA3, HER2, INPP4B, SRC_pY527, FASN, RAB25, EPPK1, SMAC, LCK, SRC, TAZ).

immune infiltration was analyzed using the TIMER (<http://cistrome.org/TIMER/>). CorFilter equals 0.3 and P value is <0.001 . TIMER also draws Kaplan-Meier plots for immune infiltrates to visualize the survival differences.

3. Result

3.1. After Screening. 17 proteins were significantly correlated to prognosis in BC. The names of proteins are BAK, BECLIN, EGFR, GATA3, PKCALPHA, SMAD3, SRC, ARID1A, RICTOR, SF2, TAZ, ANNEXIN1, ADAR1, SMAC, AXL, GATA6, and CABL (Table 1). Volcano plots show significant proteins in the high-risk and low-risk groups (Figure 1).

3.2. Construction of a Prognostic Signature. Both direction selections were carried out based on the 17 significant prognostic-related proteins, and 6 highly optimal prognostic-related proteins were found to be the final prognostic-related proteins (Table 2). Kaplan-Meier survival curve showed based on the 6 highly optimal prognostic-related protein expression (Figure 2). The risk score was calculated as the following formula: Risk scores = $1.285 \times \text{BECLIN} + 0.286 \times \text{BGFR} + 0.242 \times \text{PKCALPHA} - 0.240 \times \text{SRC} + 0.293 \times \text{ANNEXIN1} + 0.556 \times \text{AXL}$. Kaplan-Meier survival curve was performed by “survival” and “survminer” package with the highly optimal prognostic-related protein expression.

3.3. Validation of the Protein Signature. Kaplan-Meier survival curve was performed to compare the survival time of the two groups based on the risk score (Figure 3). The risk curve was performed to show the relation between risk scores and survival rates (Figure 4). Highly prognostic-related proteins were compared between two groups (Figure 5). Three asterisks mean that P value is less than 0.001. Two asterisks mean that P value is less than 0.01. One asterisk means that P value is less than 0.05. Ns means that there is no significance difference between the two groups. It has been found the risk score was an independent prognostic factor (Figure 6). The prognostic accuracy of risk score was measured using the AUC of the ROC curve, which is 0.705 (Figure 7). Survival time was compared between high-risk score group and low-risk score group in different clinical traits (Figure 8).

3.4. Correlation between the Highly Prognostic-Related Proteins and Expression Protein. The correlation was analyzed using Pearson’s correlation analysis (Figure 9). Sankey-charts were performed (Figure 10). This regulatory network revealed the regulatory relationships among these proteins.

3.5. The Corresponding Genes to the 6 Prognostic-Related Proteins in the cBioPortal Database. The corresponding genes to the 6 prognostic-related proteins were analyzed in

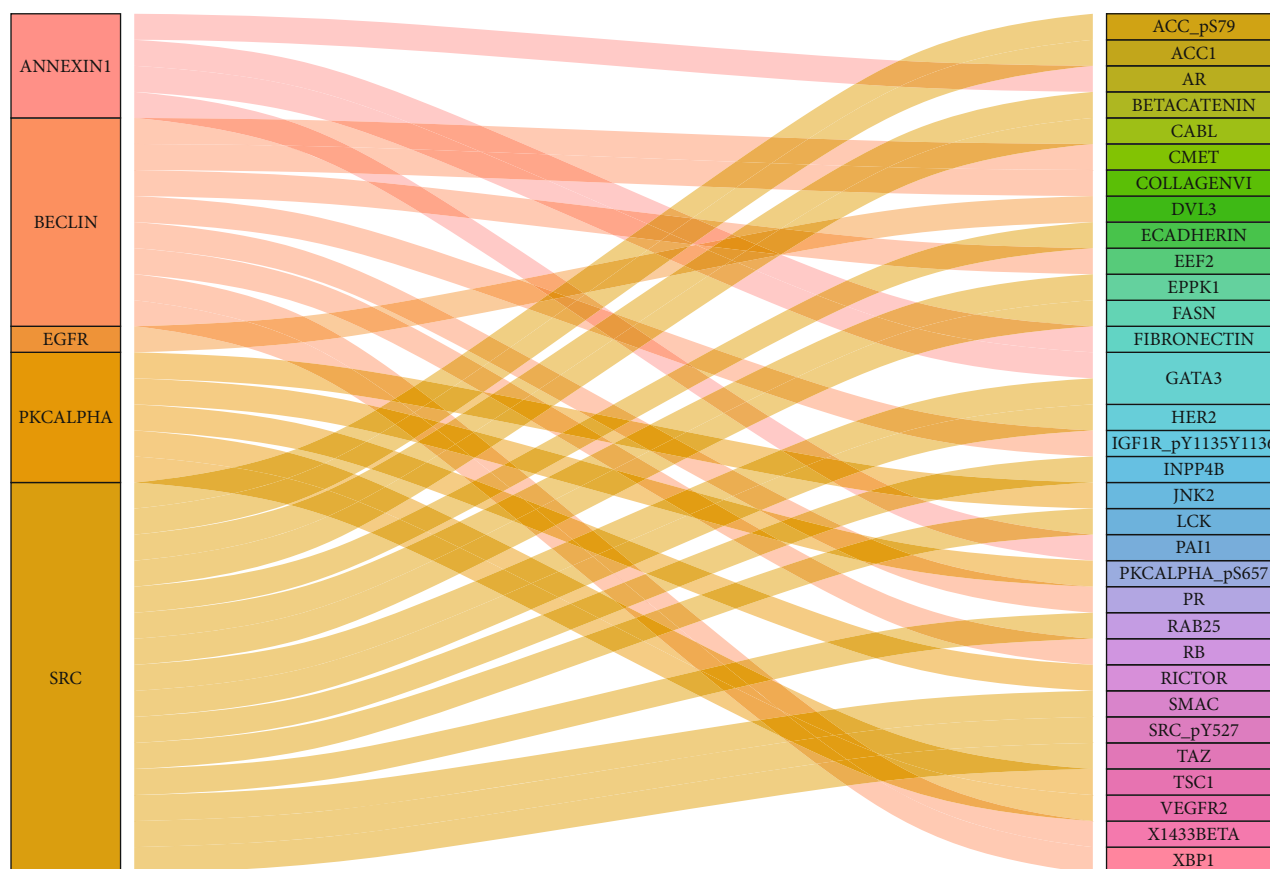


FIGURE 10: Sankeycharts were performed among highly optimal prognostic-related proteins and expression proteins.

the cBioPortal database. The rate of 6 genes altered in all cases as shown in Figure 11(a). The types of alteration of genes were shown in Figure 11(b).

3.6. Correlations between the Corresponding Genes Expression and Immune Cells in TIMER. The corresponding genes to the 6 prognostic-related proteins were analyzed in TIMER. The relationship between genes and immune infiltration was shown in Figure 12(a). TIMER drew Kaplan-Meier plots for CD8+ T cell to visualize the survival significantly differences in Figure 12(b).

4. Discussion

Careful assessing sensitive and novel biomarkers could monitor the progress and prognosis of carcinoma outcomes [5–11]. Based on alterations in these molecular or pathways, different assessment of biomarkers that give novel insights into bladder urothelial carcinoma mechanism. In an effort to bolster the clinical tool in bladder urothelial carcinoma, it has been developed a prognostic model to predict the prognosis in order to bolster the clinical tool in bladder carcinoma.

In the current study, we obtained seventeen significant prognostic candidate proteins and established a novel six-protein prognostic model. According to the further analysis, the gender, stage, T status, and M status had statistical significance in univariate analyses. These factors did not have statistical significance in multivariate analyses. However, the

six-protein prognostic model had a statistical significance, which was proved to be an independent prognostic indicator (Figure 6). The six-protein expressions were significantly different in the two groups (Figure 5). Furthermore, the prognostic significance of the six genes was performed by multivariate analysis. EGFR, SRC, ANNEXIN1, and AXL were independent prognostic proteins ($P = 0.012, 0.003, 0.004, \text{ and } 0.041$, respectively). The AUC of ROC curve was a moderate classification (Figure 7). In order to validate the accuracy of the risk score, the survival time was compared between high-risk score group and low-risk score group in different clinical traits (Figure 8). The risk score model was accurate except the risk model in M0 status and low grade. Because in M0 status and low grade, the cases were too little. The six-protein signature significantly stratified patients into high-risk and low-risk groups based on the intermediate value of risk score independent of clinical and pathologic factors. High-risk group had significantly poor prognosis (Figures 3 and 4). All the above results showed that the six-protein prognostic could play as an effective marker for bladder urothelial carcinoma prognosis prediction. This model could help clinicians make an accurate decision and avoid describing unnecessary medications and adverse drug effects, while it is strongly suggested other patients were high-risk to make treatment.

Kaplan-Meier survival curve among high-risk and low-risk groups based on the 6 highly optimal prognostic-related protein. Patients were divided into high risk and

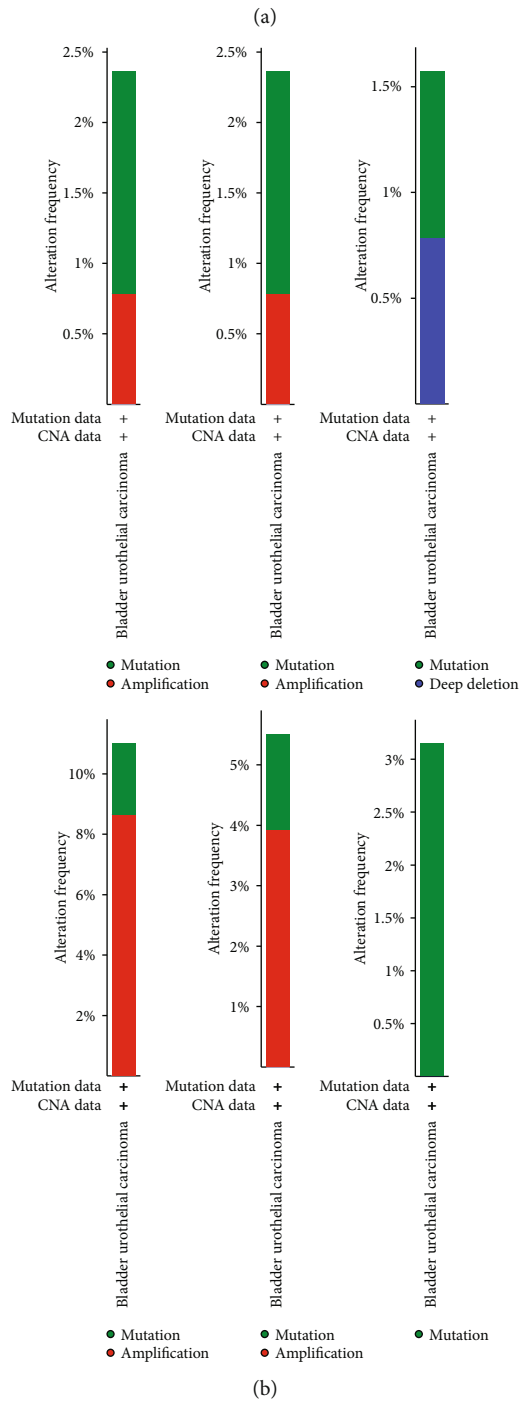
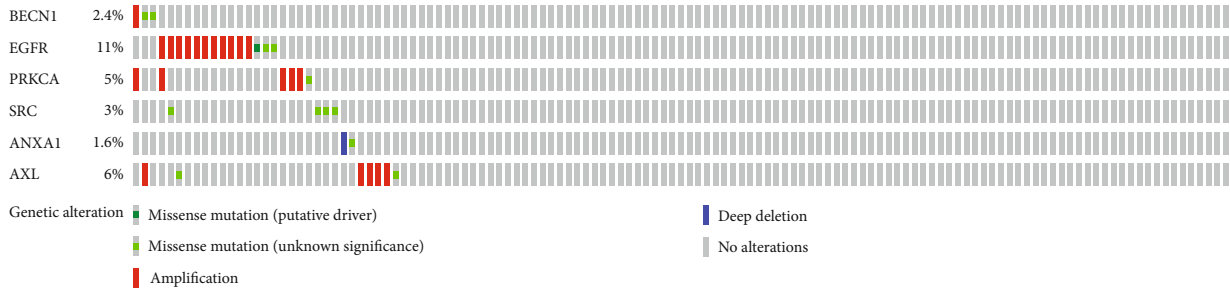


FIGURE 11: (a) The rate of 6 genes altered. B1.The types of alteration of BECN1 B2. The types of alteration of PRKCA. B3.The types of alteration of ANXA1. B4.The types of alteration of EGFR. B5.The types of alteration of AXL. B6.The types of alteration of SRC.

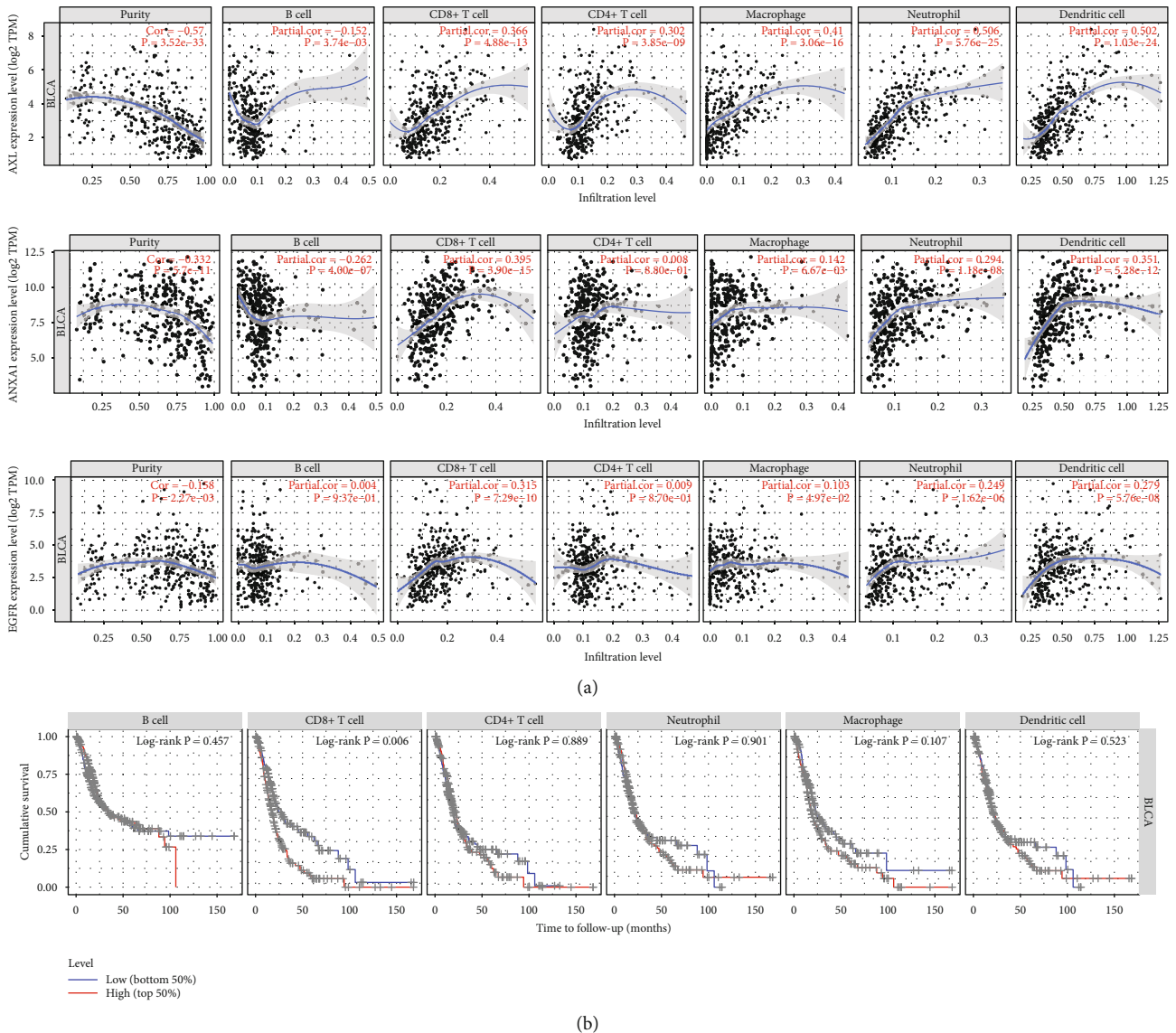


FIGURE 12: (a) The relationship between genes and immune infiltration. (b) TIMER drew Kaplan-Meier plots for CD8+ T cell.

low-risk groups based on the median level of each high optimal prognostic-related protein expression. The survival time of high-risk and low-risk groups of each protein had statistical significance (Figure 2).

Since the correlation network among the protein is one of the most common and useful statistics, it has been elucidated the correlation and the mechanism of protein. To date, several studies have been published about PKCALPHA in bladder disease. Apart from the fact that the function of PKCALPHA is not clear. PRKCA is the gene, and the gene-synonym is PKC-alpha. PKC-alpha is one subtype of classical protein kinase C which is a candidate for PDK-2 in T cells on TCR stimulation [12]. In addition, it is a proinflammatory [13]. In the bladder, PKC-alpha is closely associated with the recurrence of bladder cancer [14]. PKC-alpha can promote proliferation, migration, and the survival rate of carcinoma cells via the downstream signal transduction pathways ERK1/2 and NF-κB [15]. PKC alpha regulated the trin-1/UNC5B-mediated survival pathway in bladder cancer [16]. In the present study,

the overexpression PKC-alpha was a high-risk factor in bladder urothelial carcinoma (Figures 1 and 2). PKC-alpha between PKCALPHA and PKCALPHA_pS657, VEGFR2, RICTOR, and TSC1 were positive correlations. However, there has been no publication about the correlation between the PKC-alpha, JNK2, RICTOR, and TSC-1 in bladder tumor (Figures 11 and 12).

Though PKC-alpha and RICTOR are in the mTOR signaling pathway [17–19] in the bladder tumor, it is uncertain that the PKC-alpha and RICTOR are the mTOR signaling pathway. Therefore, we suggest a lab experiment to verify it. The miR-200a overexpression causes low expression, which results in the upregulation of JNK2 expression and promotes the bladder cancer invasion [20]. Nevertheless, predicting the positive correlation between PKC-alpha and JNK2, still, there is no experimental study to prove this hypothesis. Furthermore, the function of PKC-alpha and JNK2 in the bladder tumor, we need further research to explore them.

Then, the alteration of the corresponding genes to the 6 prognostic-related proteins was analyzed in 127 Firehose legacy sample in the cBioPortal database. The alteration of genes in 31 (24%) sequenced cases was shown in Figure 11(a). The style of alteration of genes was different (Figure 11(b)). The alteration styles of BECN1, EGFR, PRKCA, and ALX were mutation and amplification. The alteration style of SRC was only the mutation. Moreover, the alteration styles of ANXA1 were mutation and deep deletion.

Finally, the corresponding genes to the 6 prognostic-related proteins were analyzed to reveal the correlation between genes and immune infiltration. ANXA1 is crucial to provide immunity as it helps CD8+ T cell stimulation dendritic cell present antigen [21]. Figure 12(a) shows the expression of ANXA1 is strongly and positively related to CD8+ T cell. Maybe ANXA1 had a similar function in bladder. A study reported that Axl knockout tumors had more infiltration of CD8+ T cells after radiation [22]. However, in the present study, AXL was strong and positively correlated to CD8+ T cell from Figure 12(a). CD8+ T cells expressing EGFR could benefit from EGFR ligands produced by the tumor [23]. EGFR is beneficial to CD8+ T cells in bladder cancer. From Figure 12(a), the expression of EGFR was strongly and positively related to CD8+ T cell. Figure 12(a) revealed TIMER drew Kaplan-Meier plots for CD8+ T cell to visualize the survival significantly differences in Figure 12(b). All the figures above suggested ANXA1, AXL, and EGFR may affect the prognosis via transcribing the proteins and medicating CD8+ T cell.

There are still some limitations in our research, and our data are downloaded from the analysis of publicly expressed data. We lack the lab or clinical experiment to verify these results. So, further research on these proteins will be the focus of our next step.

In this study, our results screened six proteins of clinical significance. The importance of a personalized protein signature model in the recognition, surveillance. The abnormal expression of six prognostic-related proteins may be caused by corresponding gene alteration. Furthermore, these proteins may affect survival via the immune infiltration.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no competing interests.

Acknowledgments

We would like to acknowledge the supportive research environments at the Xiangya International Medical Center, Department of Geriatrics, Xiangya Hospital, Central South University, and RWTH university and Department of Urology, RWTH Aachen University, Pauwelsstrasse 30, 52072 Aachen, Germany. China Schorship Council and Xiaobo Zhang.

References

- [1] J. Alfred Witjes, T. Lebre, E. M. Compérat et al., "Updated 2016 EAU guidelines on muscle-invasive and metastatic bladder cancer," *European Urology*, vol. 71, no. 3, pp. 462–475, 2017.
- [2] S. Letašiova, A. Medvedová, A. Šovčíková et al., "Bladder cancer, a review of the environmental risk factors," *Environmental Health*, vol. 11, article S11, Supplement 1, 2012.
- [3] M. H. Mostafa, S. A. Sheweita, and P. J. O'Connor, "Relationship between schistosomiasis and bladder cancer," *Clinical Microbiology Reviews*, vol. 12, no. 1, pp. 97–111, 1999.
- [4] A. P. Mitra and S. Daneshmand, "Molecular Prognostication in Bladder Cancer," in *Genitourinary Cancers. Cancer Treatment and Research*, S. Daneshmand and K. Chan, Eds., vol. 175, pp. 165–191, Springer, Cham, 2018.
- [5] P. Liao, W. Li, R. Liu et al., "Genome-scale analysis identifies SERPINE1 and SPARC as diagnostic and prognostic biomarkers in gastric cancer," *Oncotargets and Therapy*, vol. - Volume 11, pp. 6969–6980, 2018.
- [6] J. L. Wang, M. Y. Yang, S. Xiao, B. Sun, Y. M. Li, and L. Y. Yang, "Downregulation of castor zinc finger 1 predicts poor prognosis and facilitates hepatocellular carcinoma progression via MAPK/ERK signaling," *Journal of Experimental & Clinical Cancer Research*, vol. 37, no. 1, p. 45, 2018.
- [7] J. Li, Z. Yang, Q. Zou et al., "PKM2 and ACVR1C are prognostic markers for poor prognosis of gallbladder cancer," *Clinical & Translational Oncology*, vol. 16, no. 2, pp. 200–207, 2014.
- [8] M. J. Ke, L. D. Ji, and Y. X. Li, "Explore prognostic marker of colorectal cancer based on ceRNA network," *Journal of Cellular Biochemistry*, vol. 120, no. 12, pp. 19358–19370, 2019.
- [9] Dhruva Biswas, TRACERx Consortium, N. J. Birkbak, R. Rosenthal et al., "A clonal expression biomarker associates with lung cancer mortality," *Nature Medicine*, vol. 25, no. 10, pp. 1540–1548, 2019.
- [10] F. Willenbrock, C. Cox, C. Wilhelm-Benartzi et al., "CCL5 is associated with poor prognosis in locally advanced pancreatic cancer (LAPC): biomarker analysis from the randomised phase II SCALOP trial," *Annals of Oncology*, vol. 30, Supplement 4, pp. iv110–iv121, 2019.
- [11] G. Li, M. Li, X. Liang et al., "Identifying DCN and HSPD1 as Potential Biomarkers in Colon Cancer Using 2D-LC-MS/MS Combined with iTRAQ Technology," *Journal of Cancer*, vol. 8, no. 3, pp. 479–489, 2017.
- [12] L. Yang, G. Qiao, H. Ying, J. Zhang, and F. Yin, "TCR-induced Akt serine 473 phosphorylation is regulated by protein kinase C- α ," *Biochemical and Biophysical Research Communications*, vol. 400, no. 1, pp. 16–20, 2010.
- [13] L. F. Yang, H. M. Kong, X. Q. Zhang, and F. Yin, "Roles of PKC α on the biological functions of T cells," *Zhongguo Dang Dai Er Ke Za Zhi*, vol. 17, no. 12, pp. 1354–1359, 2015.
- [14] C. Kong, Y. Zhu, D. Liu et al., "Role of protein kinase C- α in superficial bladder carcinoma recurrence," *Urology*, vol. 65, no. 6, pp. 1228–1232, 2005.
- [15] B. Wu, H. Zhou, L. Hu, Y. Mu, and Y. Wu, "Involvement of PKC α activation in TF/VIIa/PAR2-induced proliferation, migration, and survival of colon cancer cell SW620," *Tumour Biology*, vol. 34, no. 2, pp. 837–846, 2013.
- [16] J. Liu, C.-z. Kong, D.-x. Gong, Z. Zhang, and Y.-y. Zhu, "PKC α regulates netrin-1/UNC5B-mediated survival pathway in bladder cancer," *BMC Cancer*, vol. 14, no. 1, article 93, 2014.

- [17] M. Meghan, C. D. Young, S. Wang et al., “Abstract B49: mTORC2 directs breast morphogenesis through Rictor-dependent PKC α /Rac1 signaling independent of Akt,” *Molecular Cancer Therapeutics*, vol. 14, no. 7, pp. 165–191, 2015.
- [18] G. Chandrika, K. Natesh, D. Ranade, A. Chugh, and P. Shastry, “Suppression of the invasive potential of Glioblastoma cells by mTOR inhibitors involves modulation of NF κ B and PKC- α signaling,” *Scientific Reports*, vol. 6, no. 1, article 22455, 2016.
- [19] M. M. Morrison, C. D. Young, S. Wang et al., “mTOR directs breast morphogenesis through the PKC-alpha-Rac1 signaling axis,” *PLoS Genetics*, vol. 11, no. 7, p. e1005291, 2015.
- [20] R. Yang, J. Xu, X. Hua et al., “Overexpressed miR-200a promotes bladder cancer invasion through direct regulating Dicer/miR-16/JNK2/MMP-2 axis,” *Oncogene*, vol. 39, pp. 1983–1996, 2020.
- [21] M. H. Sheikh and E. Solito, “Annexin A1: Uncovering the Many Talents of an Old Protein,” *International Journal of Molecular Sciences*, vol. 19, no. 4, article 1045, 2018.
- [22] T. A. Aguilera, M. Rafat, L. Castellini et al., “Reprogramming the immunological microenvironment through radiation and targeting Axl,” *Nature Communications*, vol. 7, no. 1, article 13898, 2016.
- [23] T. Lozano, S. Chocarro, C. Martin et al., “Genetic modification of CD8⁺ T cells to express EGFR: potential application for adoptive T cell therapies,” *Frontiers in Immunology*, vol. 10, p. 2990, 2019.

Research Article

miR-1258 Regulates Cell Proliferation and Cell Cycle to Inhibit the Progression of Breast Cancer by Targeting E2F1

Xianbao Zhao 

Department of Oncology, Yiwu Central Hospital, No. 699 Nanmen Road, 322000 Yiwu, China

Correspondence should be addressed to Xianbao Zhao; zhaoxianbao_234@163.com

Received 1 April 2020; Revised 8 June 2020; Accepted 17 June 2020; Published 15 July 2020

Guest Editor: Shijia Zhu

Copyright © 2020 Xianbao Zhao. 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.

Objective. This study is designed to clarify that miR-1258 targets E2F1 to regulate the proliferation and cell cycle of breast cancer (BC) cells and consequently suppress the progression of BC. **Methods.** Bioinformatics analysis was used to analyze the differentially expressed genes in BC. The expression of miR-1258 and E2F1 mRNA in BC cell lines and immortalized breast epithelial cell lines were detected by qRT-PCR. The proliferation and growth activity of BC cells were detected by MTT and colony formation assays. The apoptosis and cell cycle of BC cells were detected by flow cytometry and the targeting relationship between miR-1258 and E2F1 was identified by dual-luciferase assay. **Results.** The expression of miR-1258 was decreased while that of E2F1 was increased in BC cells. Overexpression of miR-1258 and silencing E2F1 could inhibit the cell proliferation and growth, block cells in the G0/G1 phase, and promote cell apoptosis. Besides, miR-1258 inhibited cell proliferation and growth, block cells in the G0/G1 phase, and promote cell apoptosis by downregulating E2F1. **Conclusion.** miR-1258 regulates the proliferation and cell cycle to inhibit the progression of BC by targeting and downregulating E2F1.

1. Introduction

Breast cancer (BC) is a hormone-dependent cancer most frequently diagnosed in women, and it poses a serious threat to women's life and health [1, 2]. There are many pathogenic factors leading to BC, including age, overweight, alcohol abuse, and smoking. Intensive studies and improved treatments have diminished the mortality of BC in recent years, but the mortality still accounts for 9.6% of global cancer-related deaths [3, 4]. Therefore, in-depth discussion on the molecular mechanism underlying BC occurrence and progression and identification of potential molecular therapeutic targets for BC are of great significance for reducing BC mortality.

MicroRNAs (miRNAs), small non-coding RNA molecules expressed in different tissue and cell types, are key regulators inhibiting the expression of target genes, and the dysregulation of miRNAs tends to initiate various diseases [5]. miR-1258 regulates the occurrence and development of multiple cancers, such as oral squamous cell carcinoma, liver cancer, and gastric cancer [5–7], and it also shows a relationship with BC to some extent with its expression lowly expressed [8]. This study examined the effect of miR-1258

overexpression on BC cells, as well as predicted and validated the target gene of miR-1258 to state the mechanism of miR-1258 regulating the progression of BC.

As a member of the E2F family, E2F1 encodes the transcription factor E2F1 protein, which plays an important role in cell proliferation and apoptosis by regulating the expression of various genes [9, 10]. In this study, bioinformatics analysis was used to predict the downstream target gene of miR-1258, finding that there was a binding site of miR-1258 on E2F1 3'UTR. Meanwhile, published literature has indicated that E2F1 is related to the prognosis of BC. The targeting relationship between miR-1258 and E2F1 was verified, and the effects of miR-1258 and E2F1 on BC cells were observed.

This article is aimed at studying the role of miR-1258 in BC and predicting its target gene to provide a theoretical basis for the diagnostic and therapeutic values of miR-1258 in BC.

2. Methods

2.1. Bioinformatics Analysis. The miRNA and mRNA expression profiles of BC were downloaded from the TCGA-BRCA dataset (<https://portal.gdc.cancer.gov/>), and differential

analysis was conducted by edgeR package with $|\log_{2}FC| > 2$ and $\text{padj} < 0.05$ as threshold. Survival analysis of the differentially expressed miRNAs (DEmiRNAs) was conducted combined with the clinical information of the samples to determine the target miRNA. Thereafter, the target genes for the miRNA were predicted by TargetScan (http://www.targetscan.org/vert_71/), miRDB (<http://www.mirdb.org/miRDB/policy.html>), and miRDIP (<http://ophid.utoronto.ca/mirDIP/index.jsp>) databases, and then, the candidate differentially expressed mRNAs (DEmRNAs) with targeting binding sites of the target miRNA were obtained from the intersection of DEmRNAs and predicted target genes. GSEA software was used to perform pathway enrichment analysis to study the mechanism of the target miRNA and its target gene involved in BC.

2.2. Cell Culture. Human BC cell lines HBL100, 4T1, MDA-MB-231, MDA-MB-361, MDA-MB-435, MDA-MB-468, T47D, and immortalized mammary epithelial cell lines MCF-10A and 184A1 were all obtained from the American Type Culture Collection (ATCC; Manassas, VA, USA). Human BC cell lines were cultured in RPMI 1640 (Invitrogen, Carlsbad, CA, USA) medium containing 10% fetal bovine serum (FBS; Invitrogen, Carlsbad, CA, USA), 100 U/mL penicillin, and 100 $\mu\text{g}/\text{mL}$ streptomycin. The MCF-10A cell line was incubated in M-171 medium supplemented with breast epithelial growth factor (Invitrogen, China). The 184A1 cell line was grown in Mammary Epithelial Cell Basal Medium (MEBM; Invitrogen, Carlsbad, CA, USA). All these cell lines were placed in an incubator at 37°C with 5% CO_2 .

2.3. Cell Transfection. miR-1258 mimic, E2F1-shRNA, and corresponding negative control (NC) were purchased from GenePharma (Shanghai, China). E2F1-shRNA and E2F1 were subcloned into pcDNA3.1 (Sangon Biotech, Shanghai, China) to construct E2F1 silencing vector (sh-E2F1) and E2F1 overexpression vector (oe-E2F1), respectively. miR-1258 mimic was transfected into cells at a final concentration of 50 nM to overexpress miR-1258. For rescue experiments, miR-1258 mimic or NC mimic was cotransfected with oe-E2F1 or an empty vector (oe-NC) into BC cells. The Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) kit was used for all the above transfections. 48 h after transfection, the cells were collected for subsequent experiments.

2.4. qRT-PCR. Total RNA was extracted from cells by TRIzol reagent (Life Technologies, Grand Island, NY, USA), and complementary DNA (cDNA) was synthesized using the Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's method. The transcription levels of miR-1258 and E2F1 mRNA were determined on the 7500 real-time PCR system using the TaqMan gene expression assay (Applied Biosystems, Foster City, CA, USA). miR-1258 and E2F1 took U6 and GAPDH as internal regulators, respectively. The sequences for the forward and reverse primers synthesized by Sangon (Sangon Biotech, Shanghai, China) were shown in Table 1.

TABLE 1: Primer sequences.

| Gene | Primer | Sequences |
|----------|---------|------------------------------|
| miR-1258 | Forward | 5'-CTGCGAGTCCCTGGAGTTAG-3' |
| | Reverse | 5'-CGGTGCCCTAACTACCCATT-3' |
| U6 | Forward | 5'-CTCGCTTCGGCAGCACA-3' |
| | Reverse | 5'-AACGCTTACGAATTTGCGT-3' |
| E2F1 | Forward | 5'-CCGTGGACTCTTCGGAGAAC-3' |
| | Reverse | 5'-ATCCCACCTACGGTCTCCTC-3' |
| GAPDH | Forward | 5'-GACTCATGACCACAGTCCATGC-3' |
| | Reverse | 5'-AGAGGCAGGGATGATGTTCTG-3' |

2.5. Western Blot (WB). The transfected cells were seeded at 2×10^5 cells/well in a well plate (Corning, NY, USA). After culture for 72 h, the cells were lysed on ice using RIPA lysis buffer (Beyotime, Shanghai, China) to collect proteins. BCA protein assay kit (Beyotime, Shanghai, China) was used for assessment of the protein concentration according to the instructions. After denaturation at a high temperature, the proteins were separated from a sample loading buffer by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and sequentially transferred onto a polyvinylidene fluoride membrane (PVDF; Millipore, Billerica, MA, USA). Subsequently, the membrane was blocked with 5% skim milk powder for 2 h and then incubated with primary antibodies including rabbit anti-E2F1 (ab179445, 1:1000, Abcam, Cambridge, MA, USA) and rabbit anti-GAPDH (ab181602, 1:10,000, Abcam, Cambridge, MA, USA) at 4°C overnight, followed by the addition of peroxidase-conjugated secondary antibody (ab6721, 1:10000, Abcam, Cambridge, UK) at room temperature for 2 h. Protein signals were detected using an enhanced chemiluminescence kit (GE Healthcare, Chicago, IL, USA).

2.6. MTT Assay. Transfected cells were seeded into 96-well plates (Corning, NY, USA) at 5×10^3 cells/well, with a final volume of 200 μL . After 24, 48, 72, and 96 h of incubation, the original medium was replaced with 200 μL of fresh medium, and 25 μM MTT solution (5 g/L^{-1} in phosphate buffer saline) was added to each well to achieve a final concentration of 1 g/L^{-1} . The cells were incubated for another 4 h, and then, dimethyl sulfoxide (DMSO; Sigma, St. Louis, MO, USA) was added to dissolve the formed crystal. Detection of the absorbance at 490 nm in the wavelength was carried out with a microplate reader (SpectraMax M2, Molecular Devices, CA, USA).

2.7. Colony Formation Assay. Transfected cells were inoculated into 6-well plates (Corning, NY, USA) at 4×10^2 cells/well, with a final volume of 2 mL. The medium was replaced every 4 days. On day 8, cells were fixed with 4% paraformaldehyde (Invitrogen, Carlsbad, CA, USA) and then stained with crystal violet (Invitrogen, Carlsbad, CA, USA) to count the number of stained cells.

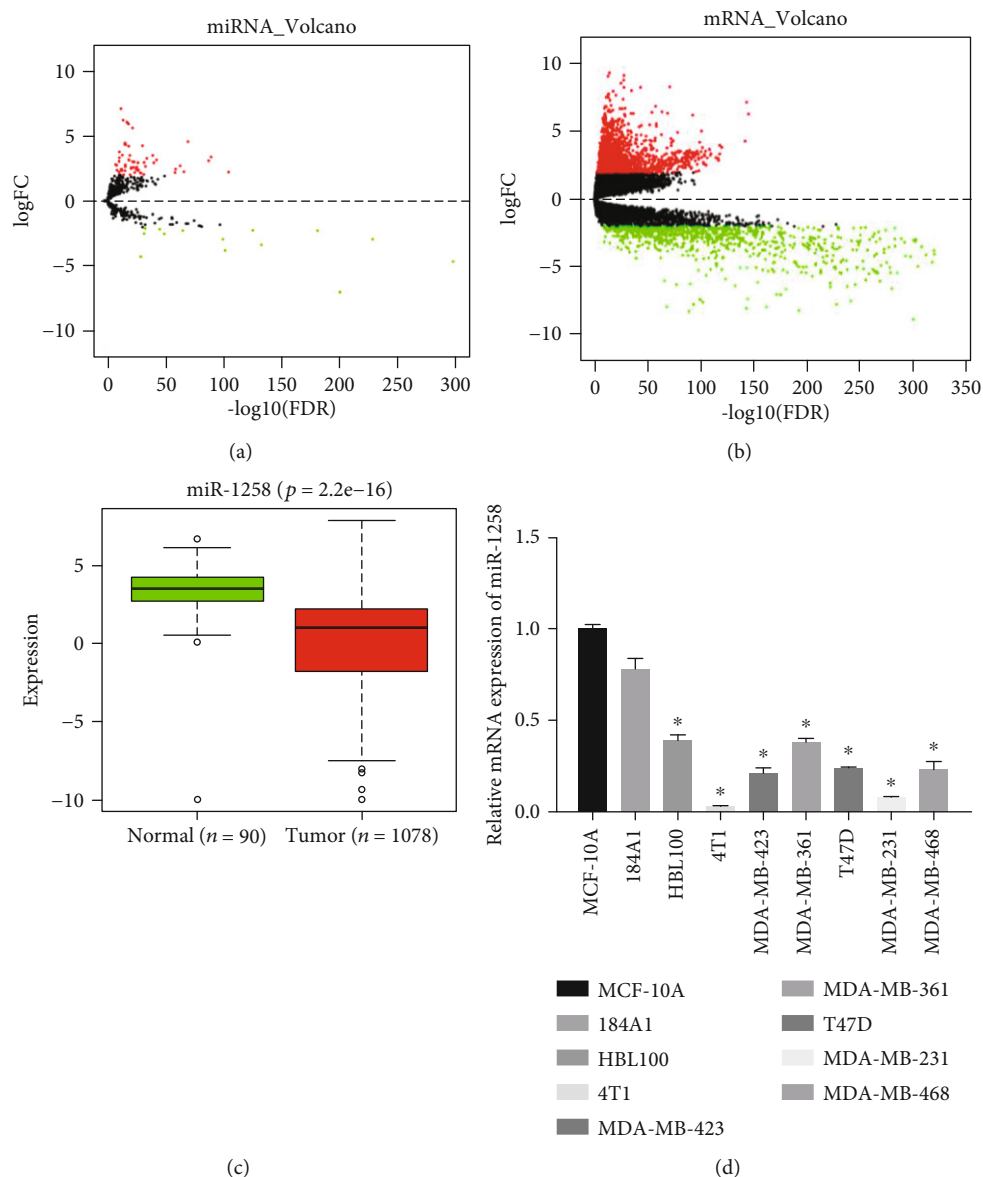


FIGURE 1: miR-1258 is poorly expressed in BC. (a, b) Volcano plots of DE miRNAs and DE mRNAs in normal and tumor groups in the TCGA-BRCA dataset. Red dots indicate the upregulated miRNAs and mRNAs in BC, while green dots indicate the downregulated miRNAs and mRNAs. (c) Expression of miR-1258 in the normal and tumor groups in the TCGA-BRCA dataset. (d) Expression of miR-1258 in BC cells and breast epithelial cells. * is compared to MCF-10A, $p < 0.05$.

2.8. Cell Apoptosis Assay. Transfected cells were seeded in a 6-well plate (Corning, NY, USA) with 1×10^5 cells/well for 72 h and then collected by centrifugation at $100 \times g$ for 3 min. After being washed with cold PBS, the cells were stained with FITC-AnnexinV and propidium iodide (PI) according to the instructions of the BD apoptosis assay kit (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) and then analyzed by flow cytometry (FCM, Becton, Dickinson and Company, Franklin Lakes, NJ, USA). Cell QuestPro software (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) was used to analyze FCM data, and the percentage of apoptotic cells was calculated.

2.9. Cell Cycle Assay. Transfected cells were seeded in 6-well plates (Corning, NY, USA) at 1×10^5 cells/well for 72 h. After

centrifugation at $100 \times g$ for 3 min, the cells were collected and fixed with 70% ethanol at $4^\circ C$ overnight. After being washed with PBS, the cells were stained with $20 \mu g/mL$ PI (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) and $200 \mu g/mL$ RNaseA (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) at $37^\circ C$ for 30 min in dark and immediately analyzed by FCM. ModFit Software (Verity Software House, Topsham, ME, USA) was used to analyze FCM data.

2.10. Dual-Luciferase Reporter Gene Assay. Vectors containing wild-type (WT) or mutant (MUT) E2F1 3'-UTR (E2F1-WT, E2F1-MUT) and the control vector pRL-TK (Promega, Madison, Wis. USA) encoding Renilla luciferase were cotransfected with miR-1258 mimic or NC mimic into

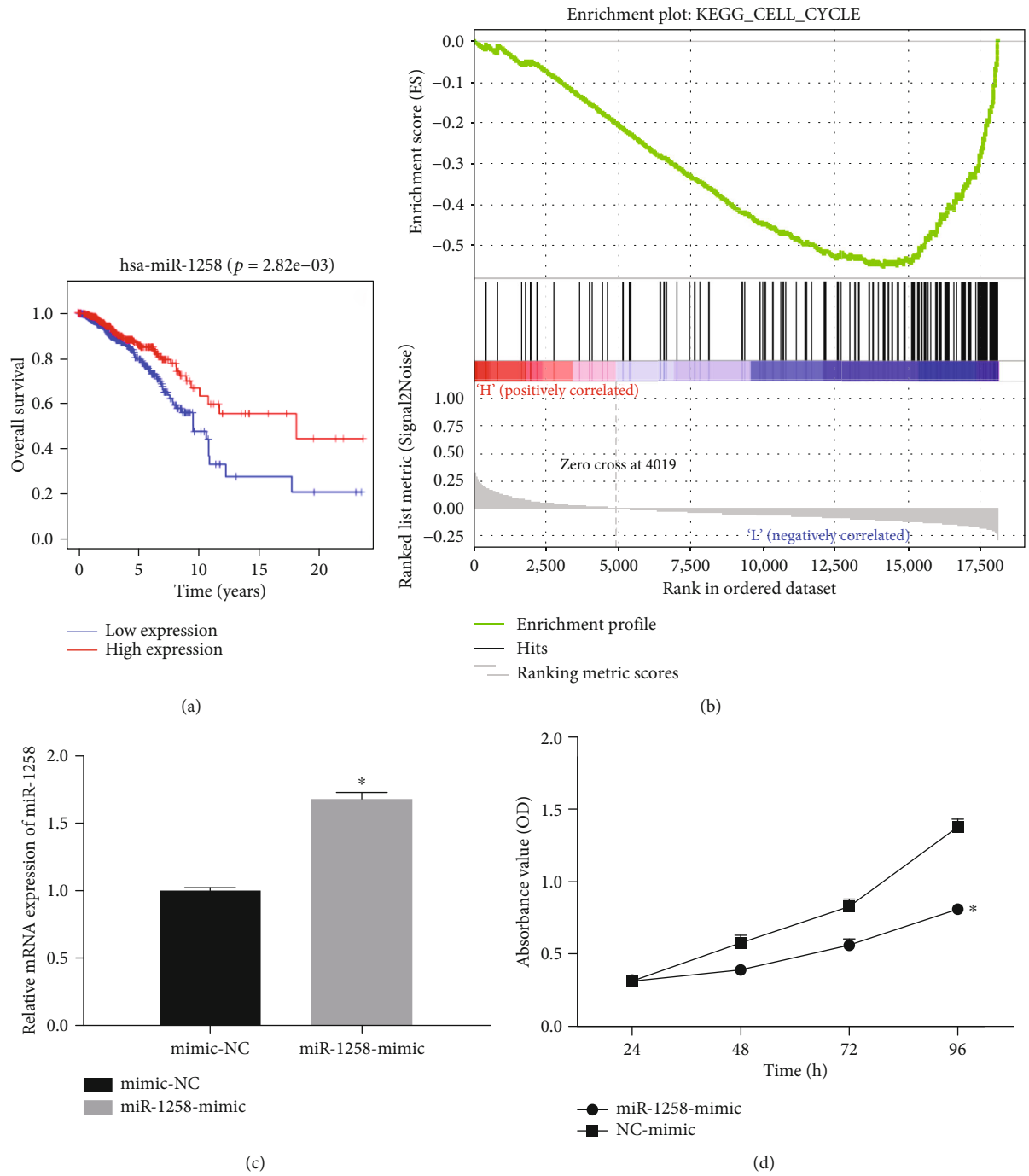


FIGURE 2: Continued.

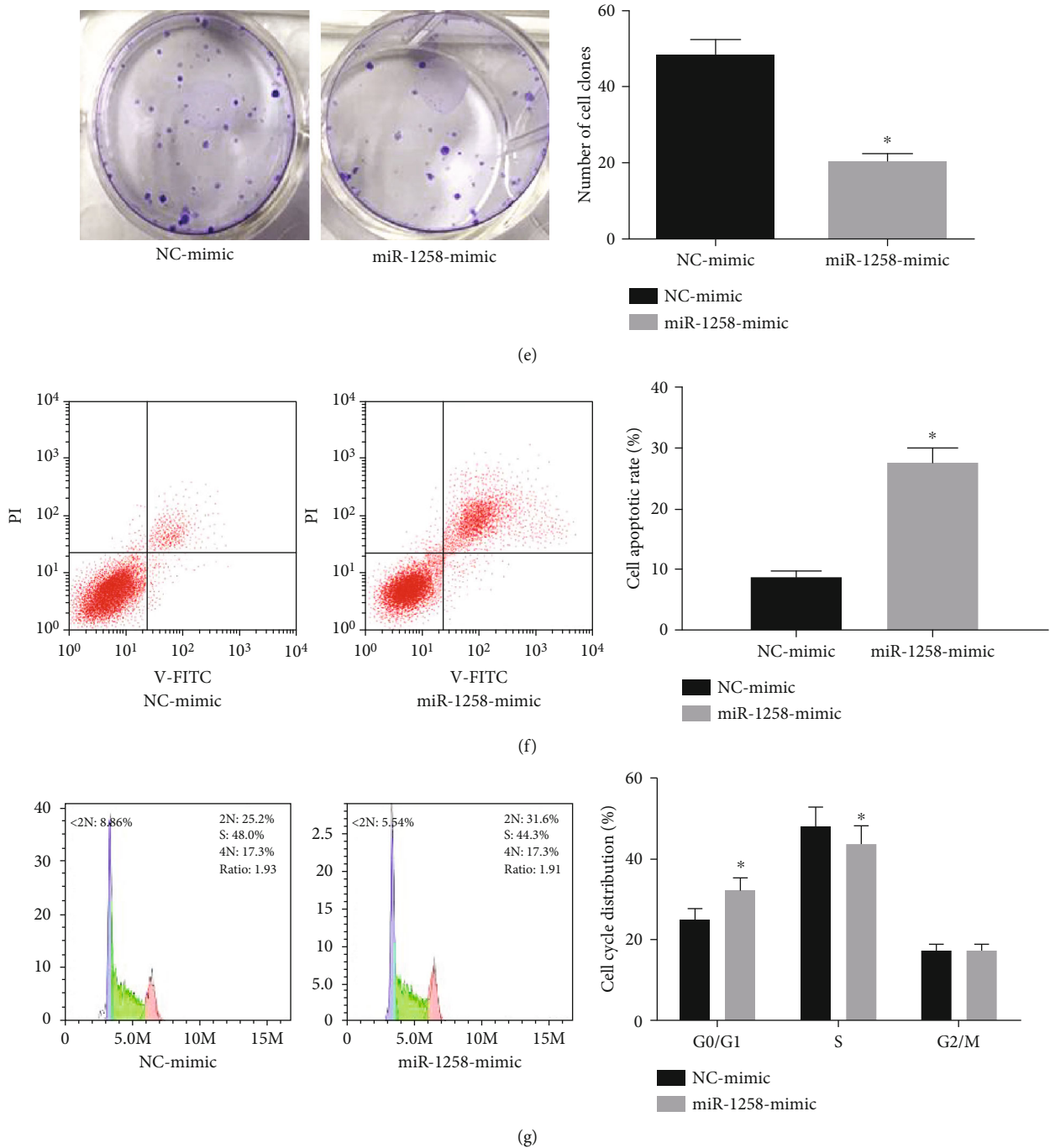


FIGURE 2: Overexpression of miR-1258 affects the cell cycle and proliferation of BC cells. (a) Survival curves show the prognosis of patients with high expression of miR-1258 (red) and low expression of miR-1258 (blue). (b) GSEA enrichment analysis of miR-1258. (c) Expression of miR-1258 in BC cells after transfection with miR-1258 mimic. (d–g) Effects of miR-1258 overexpression on BC (d) cell proliferation, (e) growth, (f) apoptosis, and (g) cell cycle were detected. The experiments were repeated 3 times, * $p < 0.05$.

cells using the Lipofectamine 2000 kit (Invitrogen, Carlsbad, CA, USA). Cells were harvested and lysed 48 h after transfection, and luciferase activity was determined using a Dual-Glo luciferase assay kit (Promega, Madison, WI, USA). The relative ratio of Firefly/Renilla activity was calculated.

2.11. Statistical Analysis. All data were processed by SPSS 22.0 statistical software (IBM, SPSS, Chicago, IL, USA), and

the measurement data were exhibited as mean \pm standard deviation. Differences between two groups were compared by t -test, and the differences of more than two groups were analyzed by one-way analysis of variance (ANOVA). The Spearman correlation analysis was used to describe the correlation between miR-1258 and E2F1 expression in BC cells. Each experiment was carried out in triplicate. * $p < 0.05$ indicated that the difference was significant between groups.

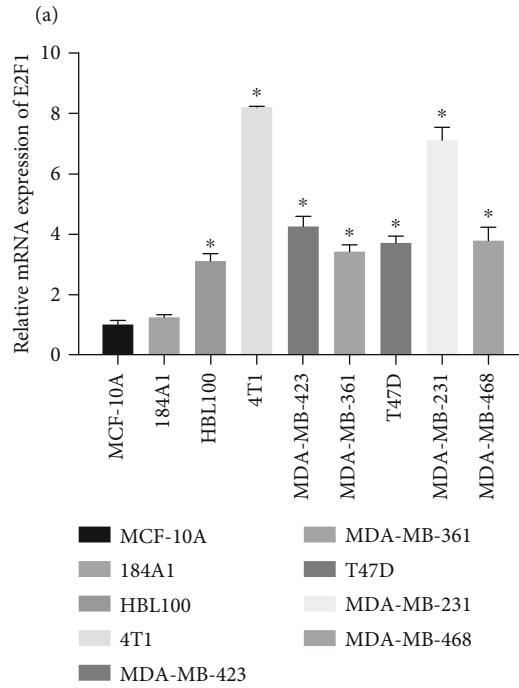
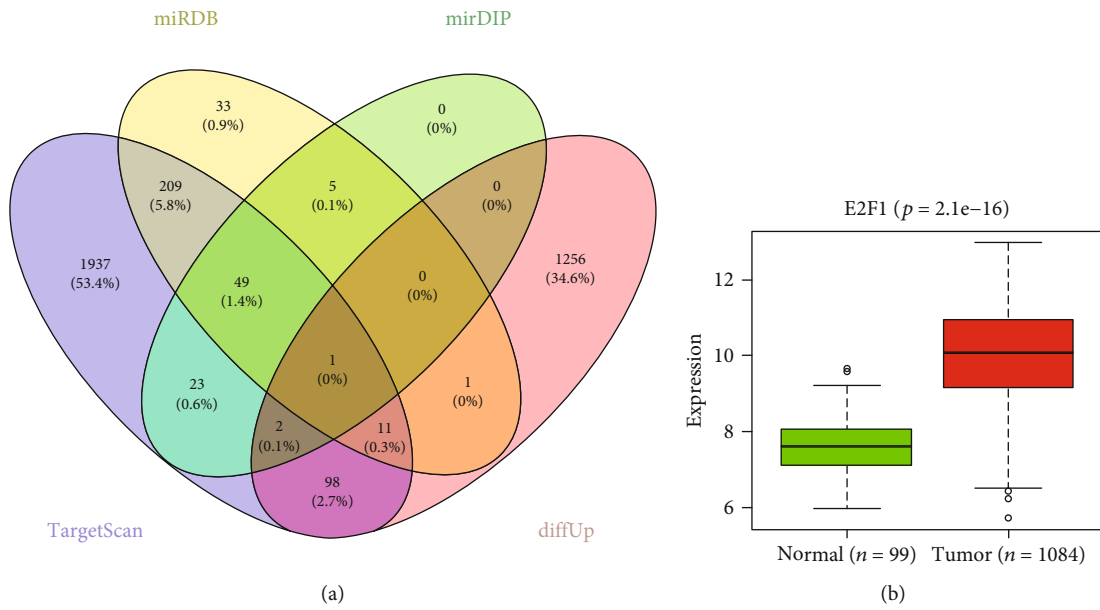


FIGURE 3: Continued.

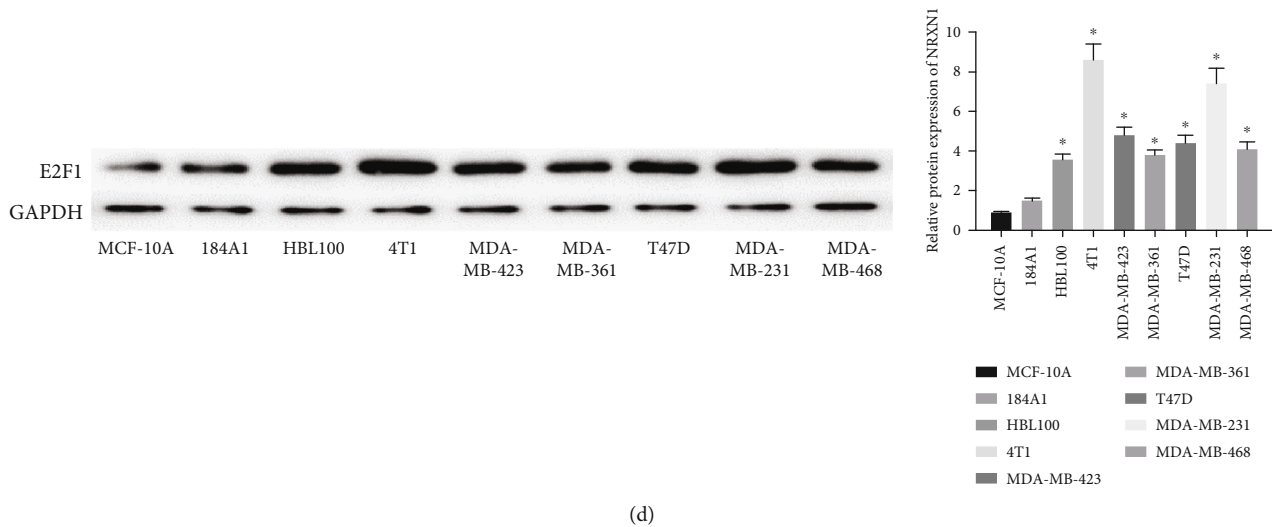


FIGURE 3: E2F1 is highly expressed in BC cells. (a) Venn diagram shows the intersection between the predicted target genes of miR-1258 and upregulated DE mRNAs. (b) Expression of E2F1 in the normal and tumor groups in the TCGA-BRCA dataset. (c, d) The mRNA and protein expressions of E2F1 in BC cell lines were detected by qRT-PCR and WB. All results were representative of 3 independent experiments, * $p < 0.05$.

3. Results

3.1. miR-1258 Is Poorly Expressed in BC Cells. A total of 74 DE miRNAs and 2,161 DE mRNAs were obtained by differential analysis between BC tumor and normal tissue samples (Figures 1(a) and 1(b)), and miR-1258 was found to be significantly lowly expressed in tumor tissue (Figure 1(c)). Therefore, qRT-PCR was used to detect the expression of miR-1258 in BC cell lines HBL100, 4T1, MDA-MB-435, MDA-MB-361, T47D, MDA-MB-231, MDA-MB-468, and immortalized mammary epithelial cell lines MCF-10A and 184A1 to verify the prediction by bioinformatics. It was observed that miR-1258 was downregulated in all BC cell lines relative to that in MCF-10A and 184A1 cell lines (Figure 1(d)), which was consistent with the bioinformatics result.

3.2. Overexpression of miR-1258 Affects Cell Cycle and Proliferation of BC Cells. Survival analysis showed that the low expression of miR-1258 in tumor tissue had a significant impact on prognosis, and the survival time of patients with low expression of miR-139 was significantly shorter than those with a high expression (Figure 2(a)). GSEA enrichment analysis exhibited that miR-1258 was closely related to cell cycle (Figure 2(b)). Therefore, the 4T1 cell line with the lowest miR-1258 expression was selected to further study the effect of miR-1258 on BC cell cycle and proliferation. We firstly verified the expression of miR-1258 in the cells transfected with miR-1258 mimic and NC mimic by qRT-PCR, and the result indicated that the content of miR-1258 in the cells transfected with miR-1258 mimic was significantly higher than that in the cells with NC mimic (Figure 2(c)).

After confirming that miR-1258 was overexpressed successfully in cells, MTT (Figure 2(d)) and colony formation (Figure 2(e)) assays were carried out and displayed that the overexpression of miR-1258 reduced the proliferation and

growth activity of BC cells. In apoptosis and cell cycle experiments, the overexpression of miR-1258 promoted the apoptosis of BC cells and blocked cells in the G0/G1 phase (Figures 2(f) and 2(g)). In summary, miR-1258 may act as a tumor suppressor to inhibit cell proliferation, thereby suppressing the progression of BC.

3.3. E2F1 Is Highly Expressed in BC Cells. For further study, we predicted the target genes of miR-1258 through TargetScan, miRDB, and miRDIIP databases and identified the target gene E2F1 with the binding sites of miR-1258 from the intersection of 1,369 upregulated DE mRNAs and predicted genes (Figure 3(a)). In the meantime, we found that E2F1 was dramatically highly expressed in BC tissue than in normal tissue (Figure 3(b)). Therefore, we further observed the mRNA and protein expressions of E2F1 in BC cell lines and immortalized breast epithelial cell lines through qRT-PCR and WB. The results exhibited that the mRNA and protein expression of E2F1 in BC cell lines were both higher than those in normal cell lines (Figures 3(c) and 3(d)), which was in agreement with the bioinformatics result.

3.4. Silencing E2F1 Inhibits BC Cell Cycle and Proliferation. GSEA enrichment analysis was also conducted on E2F1, and it was found that E2F1 was remarkably enriched in the cell cycle (Figure 4(a)). So, it was necessary to observe the effect of E2F1 on the proliferation and cell cycle of BC cells. At first, we successfully silenced E2F1 in BC cells (Figure 4(b)). Then, MTT (Figure 4(c)) and colony formation (Figure 4(d)) assays were performed, and it was observed that silencing E2F1 inhibited the proliferation and growth of BC cells. In cell apoptosis and cell cycle experiments, silencing E2F1 promoted cell apoptosis and induced cell cycle arrested in the G0/G1 phase (Figures 4(e) and 4(f)), which was in accordance with the results of bioinformatics.

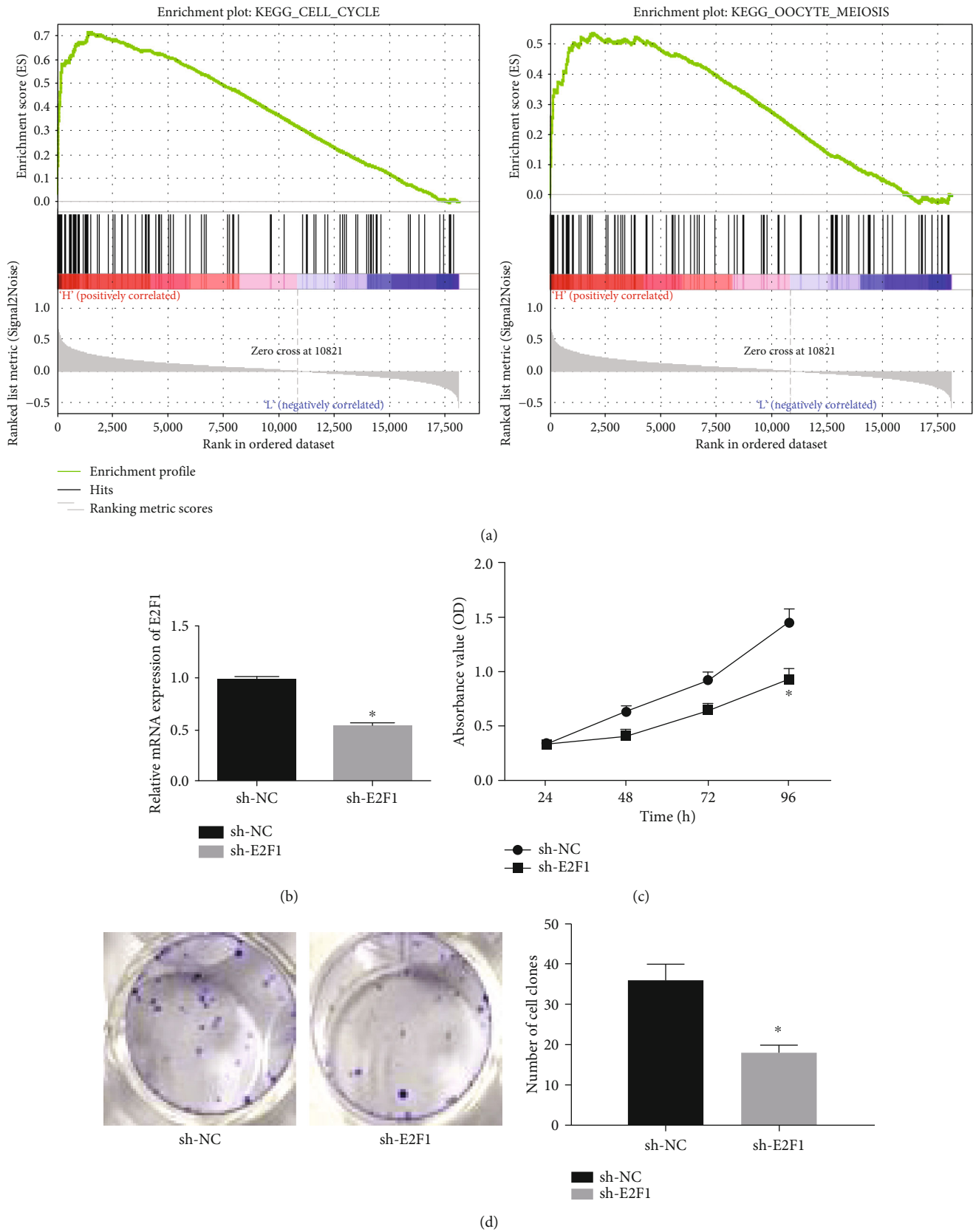


FIGURE 4: Continued.

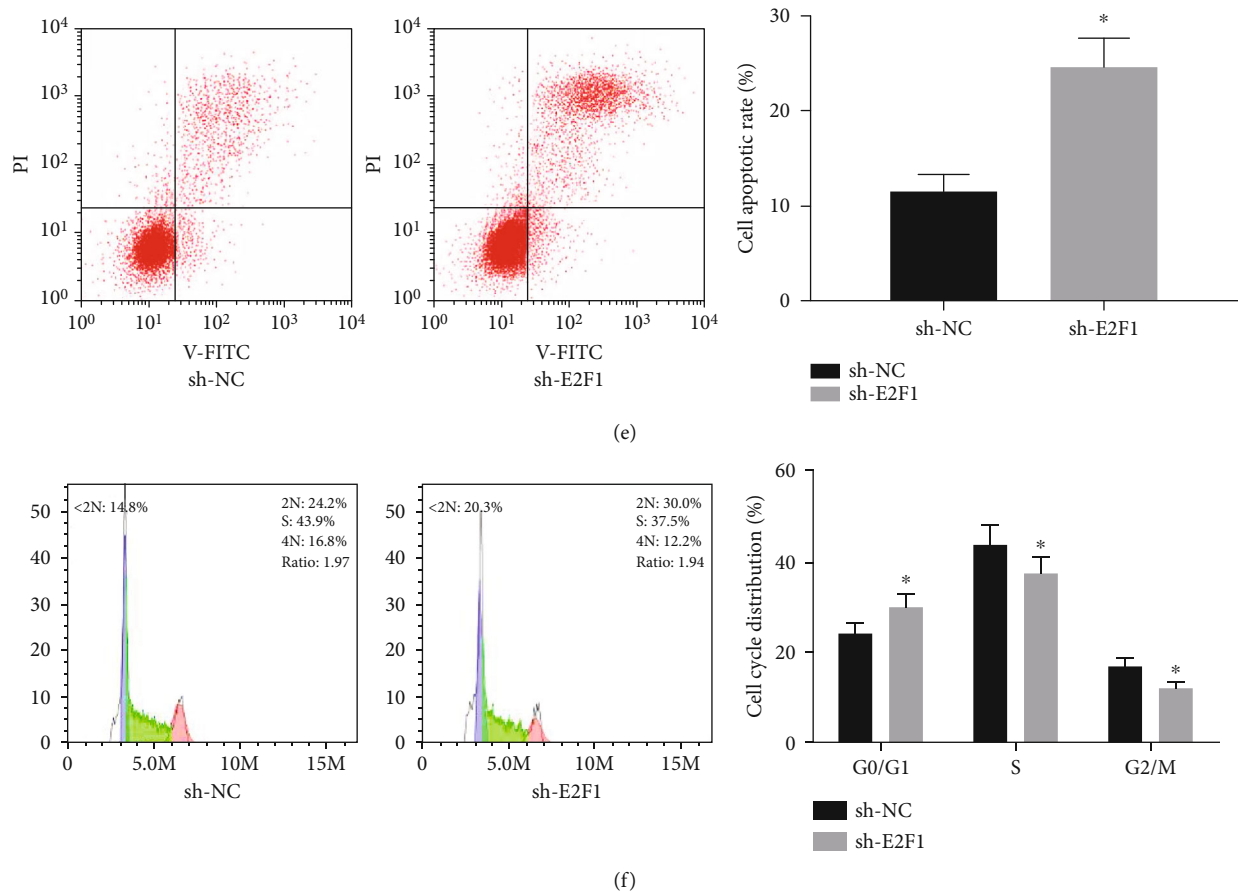


FIGURE 4: Silencing E2F1 inhibits BC cell cycle and proliferation. (a) GSEA enrichment analysis of E2F1. (b) The mRNA expression of E2F1 in BC cells was tested by qRT-PCR. (c–f) Effects of silencing E2F1 on cell (c) proliferation, (d) growth, (e) apoptosis, and (f) cell cycle were detected by MTT, colony formation assay, and flow cytometry. Each result was a representative of 3 experiments, * $p < 0.05$.

3.5. *There Is a Targeting Relationship between miR-1258 and E2F1.* Bioinformatics predicted that E2F1 may be the downstream target gene of miR-1258, and the above experiments showed that both of them had effects on the proliferation and cell cycle of BC cells. In order to prove the targeting relationship, qRT-PCR and WB were applied and it was displayed that overexpression of miR-1258 in BC cells significantly decreased the mRNA and protein expressions of E2F1, indicating that miR-1258 could affect the expression of E2F1 (Figures 5(a) and 5(b)). Then, we used dual-luciferase reporter assay to further determine the targeting relationship. The result suggested that compared with the cotransfection group of E2F1-WT and NC mimic, the luciferase activity of the cotransfection group of E2F1-WT and miR-1258 mimic was significantly lower ($P < 0.05$). Meanwhile, the luciferase activity of the cotransfection group of E2F1-MUT and miR-1258 mimic had no significant change relative to that of the group of E2F1-MUT and NC mimic ($p > 0.05$) (Figure 5(c)). The above findings collectively indicated a targeting relationship between miR-1258 and E2F1.

3.6. *miR-1258 Regulates BC Cell Proliferation, Apoptosis, and Cell Cycle by Targeting E2F1.* Through the dual-luciferase assay, we concluded that miR-1258 and E2F1 had a targeting relationship. In order to further confirm the result, we con-

ducted rescue experiments. Three groups were constructed: NC-mimic+oe-NC, miR-1258 mimic+oe-NC, and miR-1258 mimic+oe-E2F1. The mRNA and protein expressions of E2F1 in each group were detected through qRT-PCR and WB, after which MTT, colony formation, and flow cytometry were conducted to assess cell proliferation, colony formation, cell apoptosis, and cell cycle. The results showed that overexpression of E2F1 in BC cells could reverse the inhibiting effect of miR-1258 overexpression on BC cells (Figures 6(a)–6(f)).

4. Discussion

BC is a cancer with a high rate of morbidity and mortality, but its pathogenesis remains unclear and there have been no effective curative methods currently [11]. Many previous studies indicated that miRNAs can play an important role in the progression of cancer by directly targeting mRNAs [12, 13]. For instance, miRNAs can act as a tumor suppressor to further inhibit tumor progression through downregulating oncogene translation.

miR-1258 is believed to regulate the cell cycle of multiple tumor cells and inhibit their proliferation. For example, miR-1258 suppresses tumor progression via the GRB2/Ras/Erk pathway in non-small-cell lung cancer [14]. In osteosarcoma,

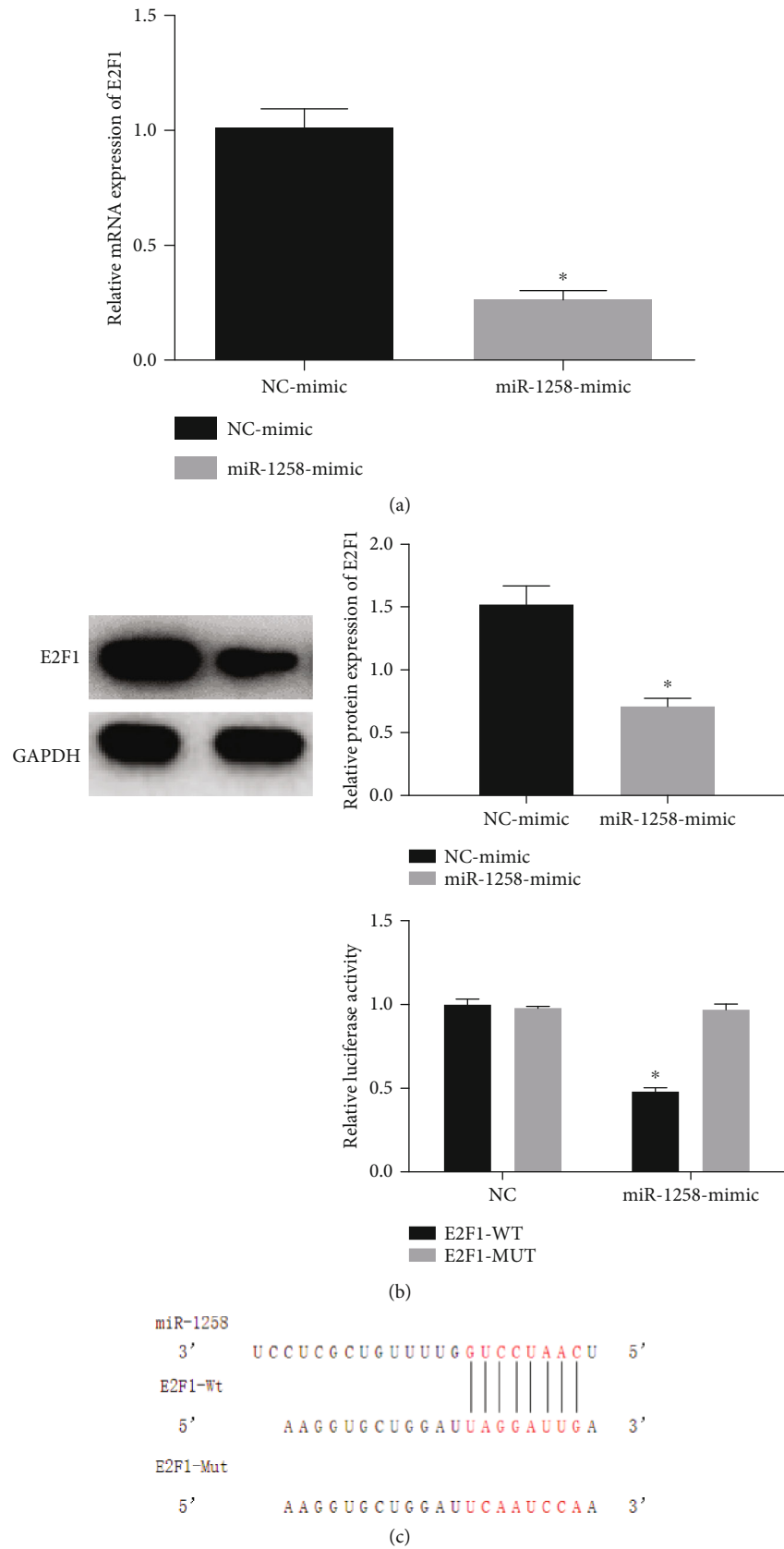
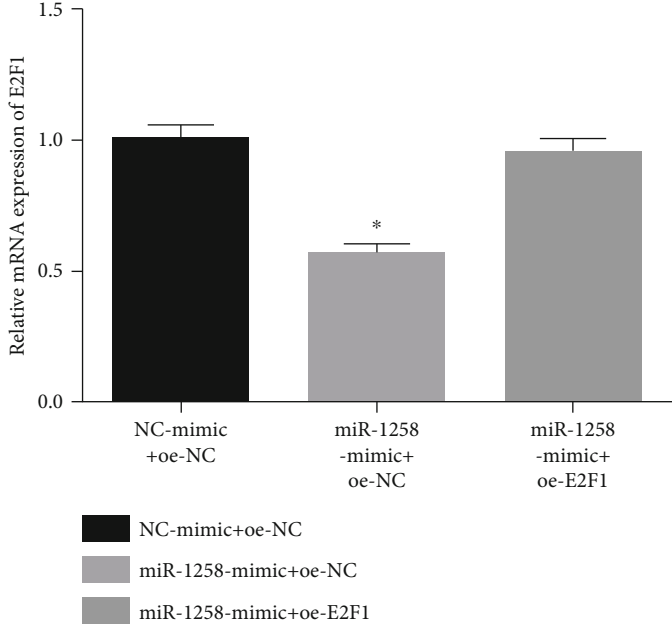
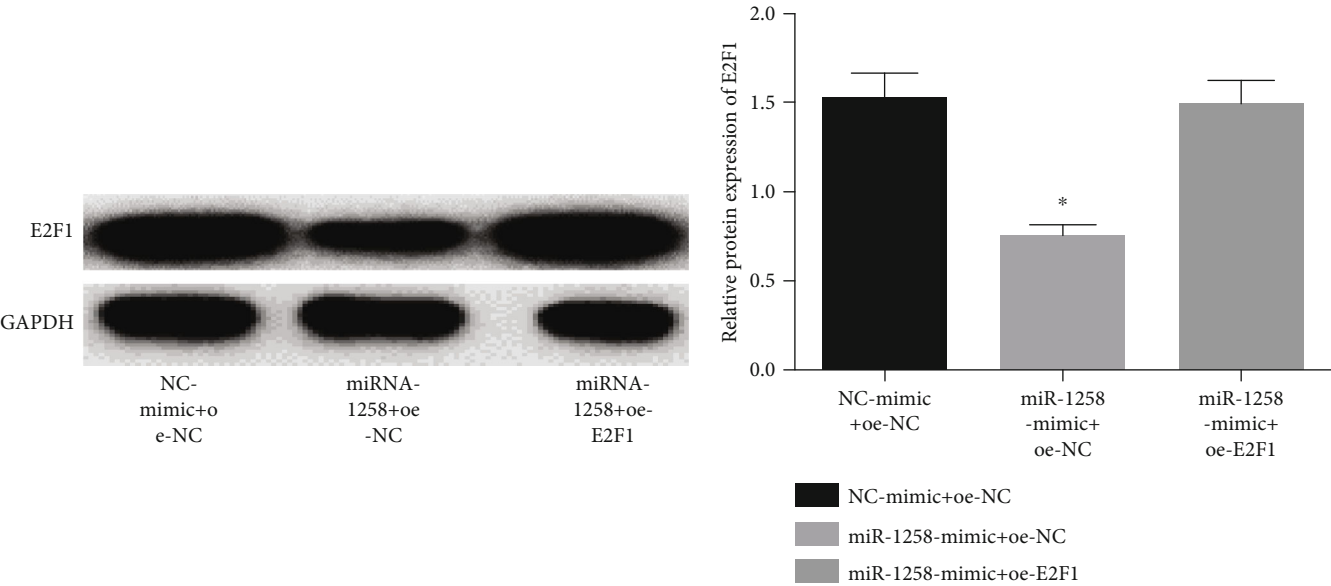


FIGURE 5: There is a targeting relationship between miR-1258 and E2F1. (a, b) The mRNA and protein expressions of E2F1 after overexpressing miR-1258 in BC cells. (c) The targeting relationship between miR-1258 and E2F1 was validated by dual-luciferase assay. At least 3 independent replicates were performed, * $p < 0.05$.

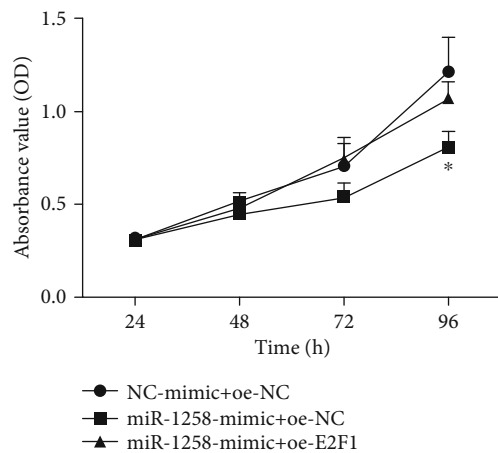


(a)

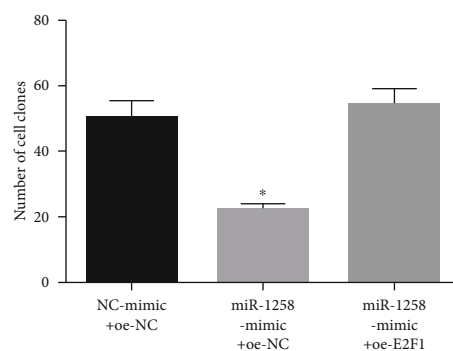
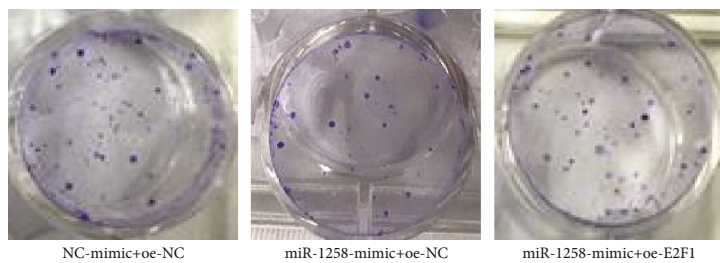


(b)

FIGURE 6: Continued.

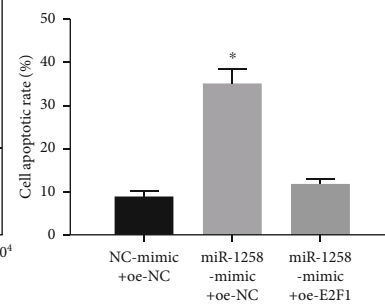
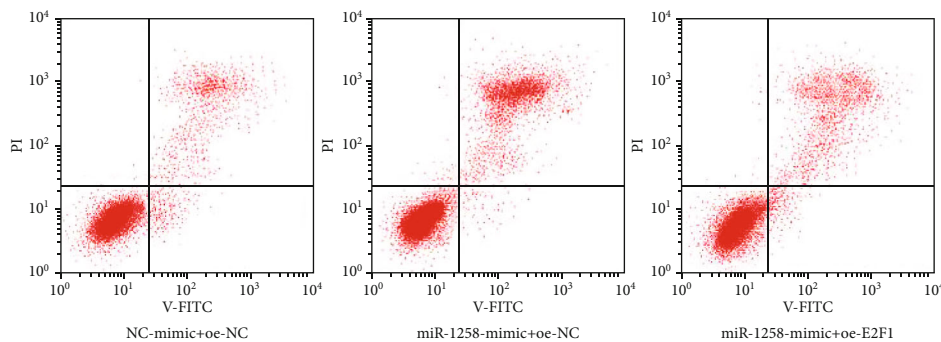


(c)



■ NC-mimic+oe-NC
 ■ miR-1258-mimic+oe-NC
 ■ miR-1258-mimic+oe-E2F1

(d)



(e)

FIGURE 6: Continued.

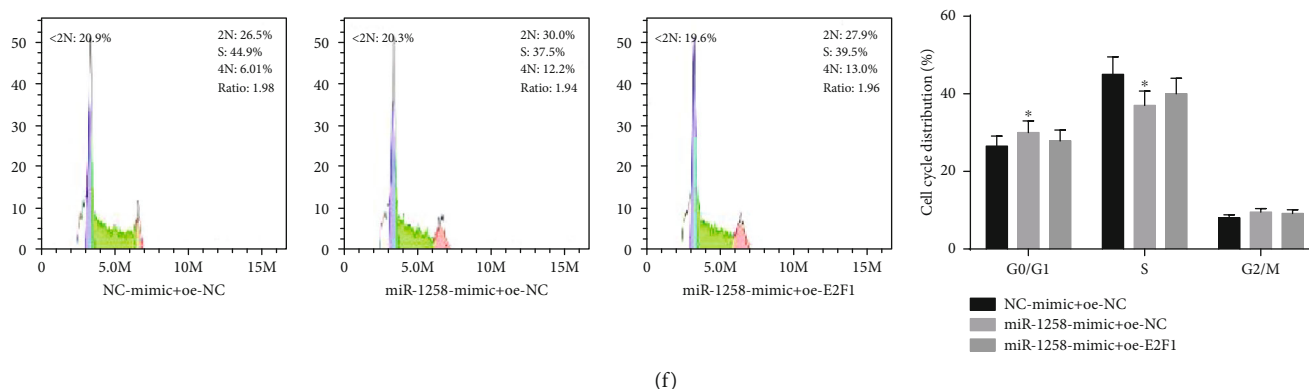


FIGURE 6: miR-1258 regulates the proliferation, apoptosis, and cell cycle of BC cells by targeting E2F1. BC cells were cotransfected with NC-mimic+oe-NC, miR-1258-mimic+oe-NC, and miR-1258-mimic+oe-E2F1. (a) The mRNA expression of E2F1 was tested by qRT-PCR. (b) The protein expression of E2F1 was tested by WB. (c–f) Effects of miR-1258/E2F1 on BC cell (c) proliferation, (d) growth, (e) apoptosis, and (f) cell cycle were detected by MTT, colony formation, and flow cytometry. The results were representative of at least 3 independent experiments, * $p < 0.05$.

miR-1258 inhibits cell proliferation and promotes cell cycle to be arrested in the G0/G1 phase through targeting AKT3 [15]. Recent studies have discovered that miR-1258 is also associated with BC, and it can inhibit the BC metastasis by targeting heparanase [16]. In this study, we found that the expression of miR-1258 was downregulated in multiple BC cell lines, compared with normal breast cells, and the upregulation of miR-1258 promoted the apoptosis of BC cells, indicating that miR-1258 was a tumor suppressor of BC.

The transcription factor E2F1 family consists of 8 different family members, respectively, which are E2F1 to E2F8. E2F1 as a transcription factor can regulate cell cycle and induce apoptosis of numerous cells. For instance, miR-136 inhibits the proliferation of cervical carcinoma cells by targeting E2F1 and promotes the apoptosis through the NF- κ B pathway [17]. CDCA5, transcribed by E2F1, potentiates the initiation of tumor by enhancing cell proliferation and inhibiting apoptosis via the AKT pathway in hepatocellular carcinoma [18]. Relevant literature has reported that E2F1 can promote the progression of BC. For example, E2F1 can drive BC metastasis by changing cell migration via regulating its target gene FGF13 [19]. The long noncoding RNA LINC00511 contributes to BC tumorigenesis by inducing the miR-185-3p/E2F1/Nanog axis [20]. However, the effect of E2F1 on BC cell cycle has not been studied yet. Hence, this study conducted research through relevant experiments and found that E2F1 could induce cell cycle arrest in THE G0/G1 phase and foster the apoptosis of BC cells.

Since both miR-1258 and E2F1 were observed to be able to regulate the progression of BC, we studied their correlation and observed that there is a targeting relationship between miR-1258 and E2F1. In summary, the study demonstrated that overexpression of miR-1258 inhibits BC cell proliferation and blocks cell cycle in the G0/G1 phase, while promoting cell apoptosis via downregulating E2F1. This study provides a basis for the potential of miR-1258 and E2F1 as new indicators for the prognosis and diagnosis of BC and also determines their targeting relationship, which are of profound significance for the study of BC.

Data Availability

The data used to support the findings of this study are included within the article. The data and materials in the current study are available from the corresponding author on reasonable request.

Additional Points

Highlights. The study confirms that miR-1258 is downregulated in breast cancer while E2F1 is upregulated. In breast cancer, miR-1258 can target and downregulate the expression of E2F1. The effects of miR-1258 and E2F1 on the proliferation and cell cycle of breast cancer cells are determined.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

The author contributed to the data analysis, drafted and revised the article, gave final approval of the version to be published, and agreed to be accountable for all aspects of the work.

References

- [1] J.-W. Lee, W. Guan, S. Han, D. K. Hong, L. S. Kim, and H. Kim, "MicroRNA-708-3p mediates metastasis and chemoresistance through inhibition of epithelial-to-mesenchymal transition in breast cancer," *Cancer Science*, vol. 109, no. 5, pp. 1404–1413, 2018.
- [2] K. Ikeda, K. Horie-Inoue, T. Ueno et al., "miR-378a-3p modulates tamoxifen sensitivity in breast cancer MCF-7 cells through targeting GOLT1A," *Scientific Reports*, vol. 5, no. 1, pp. 13170–13170, 2015.
- [3] M.-J. Wang, H. Zhang, J. Li, and H.-D. Zhao, "MicroRNA-98 inhibits the proliferation, invasion, migration and promotes

- apoptosis of breast cancer cells by binding to HMGA2," *Bioscience Reports*, vol. 38, no. 5, 2018.
- [4] F. Peng, T.-T. Li, K.-L. Wang et al., "H19/let-7/LIN28 reciprocal negative regulatory circuit promotes breast cancer stem cell maintenance," *Cell Death & Disease*, vol. 8, no. 1, pp. e2569–e2569, 2017.
- [5] M. Hu, M. Wang, H. Lu et al., "Loss of miR-1258 contributes to carcinogenesis and progression of liver cancer through targeting CDC28 protein kinase regulatory subunit 1B," *Oncotarget*, vol. 7, no. 28, pp. 43419–43431, 2016.
- [6] H. Zhang, S. Jiang, L. Guo, and X. Li, "MicroRNA-1258, regulated by c-Myb, inhibits growth and epithelial-to-mesenchymal transition phenotype via targeting SP1 in oral squamous cell carcinoma," *Journal of Cellular and Molecular Medicine*, vol. 23, no. 4, pp. 2813–2821, 2019.
- [7] J. Shi, P. Chen, J. Sun et al., "MicroRNA-1258: an invasion and metastasis regulator that targets heparanase in gastric cancer," *Oncology Letters*, vol. 13, no. 5, pp. 3739–3745, 2017.
- [8] D. Tang, Q. Zhang, S. Zhao et al., "The expression and clinical significance of microRNA-1258 and heparanase in human breast cancer," *Clinical Biochemistry*, vol. 46, no. 10-11, pp. 926–932, 2013.
- [9] A. K. Biswas and D. G. Johnson, "Transcriptional and non-transcriptional functions of E2F1 in response to DNA damage," *Cancer Research*, vol. 72, no. 1, pp. 13–17, 2012.
- [10] C. Stevens and N. B. La Thangue, "The emerging role of E2F-1 in the DNA damage response and checkpoint control," *DNA Repair (Amst)*, vol. 3, no. 8-9, pp. 1071–1079, 2004.
- [11] D. Froes Brandao, K. Strasser-Weippl, and P. E. Goss, "Prolactin and breast cancer: the need to avoid undertreatment of serious psychiatric illnesses in breast cancer patients: a review," *Cancer*, vol. 122, no. 2, pp. 184–188, 2016.
- [12] J. Li, S. Tan, R. Kooger, C. Zhang, and Y. Zhang, "MicroRNAs as novel biological targets for detection and regulation," *Chemical Society Reviews*, vol. 43, no. 2, pp. 506–517, 2014.
- [13] L. Paladini, L. Fabris, G. Bottai, C. Raschioni, G. A. Calin, and L. Santarpia, "Targeting microRNAs as key modulators of tumor immune response," *Journal of experimental & clinical cancer research*, vol. 35, no. 1, pp. 103–103, 2016.
- [14] W. Jiang, K. Wei, C. Pan et al., "MicroRNA-1258 suppresses tumour progression via GRB2/Ras/Erk pathway in non-small-cell lung cancer," *Cell Proliferation*, vol. 51, no. 6, pp. e12502–e12502, 2018.
- [15] W. Liu, Z. Zhou, Q. Zhang et al., "Overexpression of miR-1258 inhibits cell proliferation by targeting AKT3 in osteosarcoma," *Biochemical and Biophysical Research Communications*, vol. 510, no. 3, pp. 479–486, 2019.
- [16] L. Zhang, P. S. Sullivan, J. C. Goodman, P. H. Gunaratne, and D. Marchetti, "MicroRNA-1258 suppresses breast cancer brain metastasis by targeting heparanase," *Cancer Research*, vol. 71, no. 3, pp. 645–654, 2011.
- [17] H.-J. Lu, P.-Y. Jin, Y. Tang et al., "MicroRNA-136 inhibits proliferation and promotes apoptosis and radiosensitivity of cervical carcinoma through the NF- κ B pathway by targeting E2F1," *Life Sciences*, vol. 199, pp. 167–178, 2018.
- [18] H. Chen, J. Chen, L. Zhao et al., "CDCA5, transcribed by E2F1, promotes oncogenesis by enhancing cell proliferation and inhibiting apoptosis via the AKT pathway in hepatocellular carcinoma," *Journal of Cancer*, vol. 10, no. 8, pp. 1846–1854, 2019.
- [19] J. Wei, Q. Man, F. Guo et al., "Precise and systematic survey of the efficacy of multicomponent drugs against functional dyspepsia," *Scientific Reports*, vol. 9, no. 1, pp. 10713–10713, 2019.
- [20] G. Lu, Y. Li, Y. Ma et al., "Long noncoding RNA LINC00511 contributes to breast cancer tumorigenesis and stemness by inducing the miR-185-3p/E2F1/Nanog axis," *Journal of experimental & clinical cancer research*, vol. 37, no. 1, pp. 289–289, 2018.

Research Article

Identification of Differentially Expressed Genes and miRNAs Associated with Esophageal Squamous Cell Carcinoma by Integrated Analysis of Microarray Data

Lemeng Zhang , Jianhua Chen, Tianli Cheng, Hua Yang, Changqie Pan, and Haitao Li

Thoracic Medicine Department 1, Hunan Cancer Hospital, Changsha, Hunan Province, China

Correspondence should be addressed to Lemeng Zhang; zhanglemeng@hnca.org.cn

Received 18 April 2020; Accepted 10 June 2020; Published 2 July 2020

Guest Editor: Tao Huang

Copyright © 2020 Lemeng 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.

To identify candidate key genes and miRNAs associated with esophageal squamous cell carcinoma (ESCC) development and prognosis, the gene expression profiles and miRNA microarray data including GSE20347, GSE38129, GSE23400, and GSE55856 were downloaded from the Gene Expression Omnibus (GEO) database. Clinical and survival data were retrieved from The Cancer Genome Atlas (TCGA). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of differentially expressed genes (DEGs) was analyzed via DAVID, while the DEG-associated protein-protein interaction network (PPI) was constructed using the STRING database. Additionally, the miRNA target gene regulatory network and miRNA coregulatory network were constructed, using the Cytoscape software. Survival analysis and prognostic model construction were performed via the survival (version 2.42-6) and rbsurv R packages, respectively. The results showed a total of 2575, 2111, and 1205 DEGs, and 226 differentially expressed miRNAs (DEMs) were identified. Pathway enrichment analyses revealed that DEGs were mainly enriched in 36 pathways, such as the proteasome, p53, and beta-alanine metabolism pathways. Furthermore, 448 nodes and 1144 interactions were identified in the PPI network, with *MYC* having the highest random walk score. In addition, 7 DEMs in the microarray data, including miR-196a, miR-21, miR-205, miR-194, miR-103, miR-223, and miR-375, were found in the regulatory network. Moreover, several reported disease-related miRNAs, including miR-198a, miR-103, miR-223, miR-21, miR-194, and miR-375, were found to have common target genes with other DEMs. Survival analysis revealed that 85 DEMs were related to prognosis, among which hsa-miR-1248, hsa-miR-1291, hsa-miR-421, and hsa-miR-7-5p were used for a prognostic survival model. Taken together, this study revealed the important roles of DEGs and DEMs in ESCC development, as well as DEMs in the prognosis of ESCC. This will provide potential therapeutic targets and prognostic predictors for ESCC.

1. Introduction

Esophageal carcinoma (EC) remains the sixth leading cause of cancer-associated mortality worldwide, with approximately 442,000 new cases and 440,000 mortalities globally in 2014 [1, 2]. As the predominant subtype of EC, esophageal squamous cell carcinoma (ESCC) is frequently diagnosed in Eastern Asian countries, including China, where it accounts for 95% of all EC cases [3, 4]. A series of risk factors, such as tobacco smoking and consumption of alcohol and salted vegetables, are reported to be associated with the high incidence of ESCC [5]. Currently, despite advances in diagnosis and treatment of ESCC, its prognosis remains poor, with a

5-year overall survival rate of less than 40% [6, 7]. Therefore, identification of the molecular mechanisms underlying the progression and prognosis of ESCC is of utmost importance.

As a gene detection technique, gene profiling or gene chips have been widely used during the last decade for the comprehensive screening of differentially expressed genes (DEGs) [8]. With the development and application of gene chips, more and more data have been generated and stored in public databases, which will provide valuable clues for new research. Currently, an increasing number of studies have reported the identification of DEGs in ESCC. For example, Yuan et al. [9] analyzed the gene expression profile in the GSE20347 dataset, identifying a total of 538 DEGs. Likewise,

Xing and Liu [10] identified 1204 DEGs based on the GSE23400 dataset generated from ESCC and matched mucosa tissues. Furthermore, Hu et al. [11] focus on biallelic loss and its relation to mRNA expression based on the GSE38129 array data. Recently, differentially expressed miRNAs (DEMs) has been reported to be associated with differentiation, invasion, and metastasis of patients with ESCC [12]. Moreover, miRNA expression profiling analysis was also recently performed [13]. Jang et al. [14] identified prognostic markers for postoperative recurrence of ESCC by building an easy-to-use prognostic model with three small noncoding RNAs (sncRNAs) based on the GSE55856 dataset. However, the majority of these reports were based on a single dataset, which leads to the results being limited or inconsistent. Hence, the combination of bioinformatics methods and expression profiling techniques using different datasets may provide valuable information for the development of therapeutic strategies for patients with this disease.

In the present study, we obtained the original microarray data of the GSE20347, GSE38129, and GSE23400 datasets, as well as the miRNA microarray data of the GSE55856 dataset from the Gene Expression Omnibus (GEO). Clinical data and survival data were retrieved from The Cancer Genome Atlas (TCGA). Subsequently, the DEGs or DEMs were screened. Pathway enrichment analyses were performed, and protein-protein interaction (PPI) networks were created for the DEGs, in order to identify key genes and their biological function in ESCC. Additionally, the miRNA target gene regulatory network and miRNA coregulatory network were constructed to investigate the underlying functions of these miRNAs. As such, based on survival analysis of DEMs and univariate Cox analysis, a prognostic survival model based on the expression of different miRNAs was constructed.

2. Materials and Method

2.1. Microarray Data Collection and Preprocessing. Gene expression profiles from the GSE20347 (34 samples), GSE38129 (60 samples), and GSE23400 (106 samples) datasets between ESCC samples and matched normal controls were obtained from the GEO (<http://www.ncbi.nlm.nih.gov/geo/>) database using the Affymetrix Human Genome U133A 2.0 Array platform. In addition, the miRNA microarray GSE55856 dataset, which is composed of 216 samples (108 ESCC samples and 108 normal controls), was obtained using the Affymetrix Multispecies miRNA-2_0 Array.

For the preprocessing of the gene expression profile chip, the raw data of the GSE20347, GSE38129, and GSE23400 datasets were preprocessed using the R Affy package (version: 1.46.1) with a standardized RMA method [15]. The processing included background corrections, normalization, and calculation of gene expression. Afterwards, the probe ID was converted to a gene symbol, with probes that had no corresponding gene symbols being removed. As for the case where multiple probes correspond to the same gene symbol, we selected the mean value of the probes as the final gene expression value. The preprocessing of miRNA microarray data was done in a similar manner using the miRNA chip platform of Affy.

2.2. Identification of DEGs and DEMs. The limma package of R (version: 3.30.2) [16] was used to identify genes or miRNAs that were significantly differentially expressed between the tumor and normal tissues. $P < 0.05$ and \log_2 fold change (FC) ≥ 0.58 were selected as the cutoff values for statistically significant DEGs or miRNAs. Subsequently, we selected 3 groups of DEGs and then analyzed whether the genes were also significantly differentially expressed in the 3 datasets.

2.3. Prediction of DEGs Based on a Meta-analysis. By screening DEGs based on a meta-analysis, more reliable DEGs can be obtained due to the collection of multiple experimental datasets and enhancement of statistical ability. In order to integrate the DEGs that were combined in the three datasets, the MetaDE package of R (version: 1.0.5) was used [17]. Gene expression values were examined for heterogeneity with statistical parameters including τ^2 , Q value, and QP value. Criteria standards of $\tau^2 = 0$ and $QPval > 0.05$ were selected as the homogeneity test parameter. A P value of < 0.05 was the threshold for a significant difference in gene expression. Moreover, a heatmap was generated with the pheatmap [18] R package (version: 3.25).

2.4. Pathway Enrichment Analysis for DEGs. To investigate the biofunctions of DEGs, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed using DAVID (version: 6.8) [19]. $P < 0.05$ was selected as the cutoff criterion.

2.5. Identification of PPI Network and Key Genes. To better understand the interactions of the DEGs, the Search Tool for the Retrieval of Interacting Genes (STRING) database (version: 10) was employed to develop a DEG-encoded PPI [20], with a reliability threshold of > 0.9 . The prediction methods were derived from the neighborhood, gene fusion, cooccurrence, coexpression, experiments, databases, and text mining. The Cytoscape software (version: 3.2.1) was utilized to construct the PPI.

Next, the random walk algorithm was used to analyze important nodes in the PPI network. Briefly, the random walk was started at the seed node and the importance of each node was expressed by calculating the frequency of each node after the random walk between nodes in the network. The corresponding higher frequency genes may be candidate genes that have important physiological regulatory functions. The RWOAG package of R [21] was used to calculate the network node score.

2.6. Construction of the miRNA-Target Gene Regulatory Network. Mirwalk2 (<http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2>) [22] was used for the prediction of target genes regulated by miRNAs, and differentially expressed target genes were filtered by using the “validated target” module. Based on the data of differentially expressed miRNAs and DEGs, the opposite relationship pairs (upregulated miRNA-downregulated gene or downregulated miRNA-upregulated gene) were selected from the miRNA-target gene data. The regulatory network of miRNA-target genes was constructed using the Cytoscape software [23]. Meanwhile, we screened several miRNAs from the miR2Disease database

(<http://www.mir2disease.org/>) [24] which were reported to be related to ESCC.

2.7. Functional Analysis of miRNAs. In order to obtain information regarding the pathways associated with the identified miRNAs, we performed KEGG pathway analysis for differentially expressed target genes using the clusterProfiler package of R (version: 3.3.1). The enrichment significance P value was corrected with the BH method and a P value of less than 0.05 was considered to be significant.

2.8. Construction of miRNA Coregulatory Network. Based on the regulatory network of miRNAs and their target genes, miRNA pairs that regulate the same target genes were screened to construct the coregulatory network between miRNAs.

2.9. Survival Analysis of DEMs. Clinical and survival data from 251 patients with ESCC were retrieved from TCGA, which were downloaded from the database of University of California Santa Cruz (UCSC) Xena (<https://xenabrowser.net>). Moreover, miRNA-seq data was downloaded from http://gdac.broadinstitute.org/runs/stddata_2016_01_28/data/ESCA/20160128/.

In general, TCGA data directly downloaded cannot be directly analyzed. Therefore, in order to link different data, we need to match, select, and delete different data by screening samples. In this study, the DEMs obtained from the integration analysis of different GEO databases intersected with the miRNA-seq data filenames downloaded by TCGA using an R package [25]. Additionally, miRNAs with value = 0 in more than half of the total samples were removed from the existing intersection data.

The optimal miRNA cutoff was determined via the surv_cutpoint of survminer (version 0.4.3) of R package, with >optimal cutoff being considered high expression and <optimal cutoff being considered low expression. Survival analysis was conducted with the survival (version 2.42-6) R package, and P values < 0.05 were taken as the threshold. miRNAs with significant correlation to prognosis were selected and survival curves were made.

2.10. Univariate Cox Analysis and Prognostic Model Construction. Univariate survival Cox analysis was continued for miRNAs significantly correlated with survival, and miRNAs with P values < 0.05 were used for the construction of the prognostic model.

After univariate analysis, there were still many significant univariate factors, which were not conducive to inclusion in the prognostic model. Therefore, some dimensionality reduction methods were adopted to select the most important univariate factors to be included in the prognostic model for downstream analysis. In this study, the rbsurv R package was used to investigate the robustness of univariate survival analysis. Briefly, 3/4 samples were randomly selected as training data and the remaining 1/4 samples as validation data. Multivariate Cox analysis was carried out on the obtained models in the test training set and verification set, and risk scores for survival verification of the model were obtained.

Finally, the overall evaluation effect of the model on prognosis was checked.

3. Results

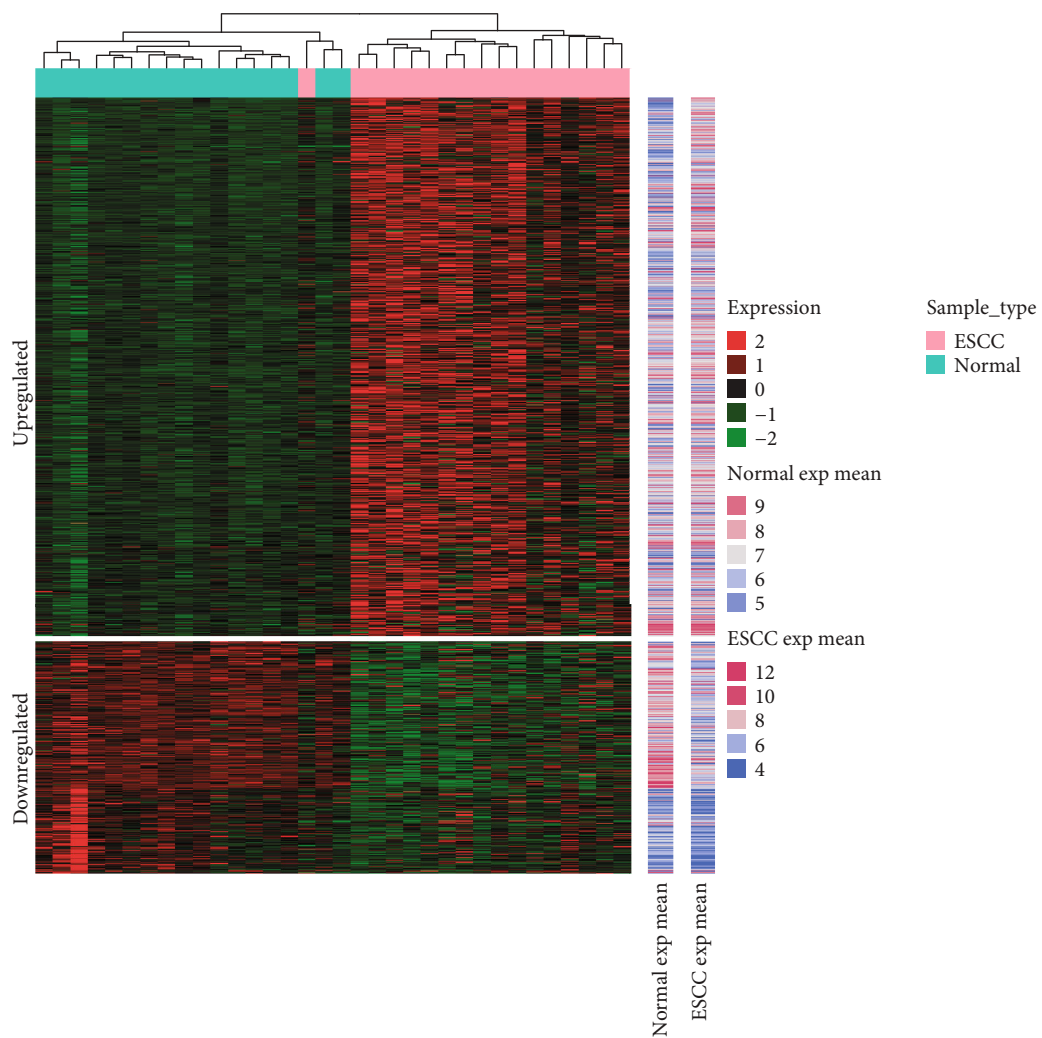
3.1. Identification of DEGs and DEMs. After data preprocessing, a total of 2575, 2111, and 1205 DEGs between ESCC and normal tissues were identified in the gene expression profile of the GSE20347, GSE38129, and GSE23400 datasets, respectively, based on the cutoff criteria. Moreover, 226 DEMs were identified in GSE55856, of which 190 were upregulated and 36 were downregulated.

3.2. Meta-analysis of DEGs and Hierarchical Clustering. Based on the meta-analysis, 1001 DEGs, including 700 upregulated genes and 301 downregulated genes were obtained. As shown in Figure 1, hierarchical clustering revealed that the DEGs obtained from the meta-analysis and DEMs could effectively cluster the samples from the GSE20347, GSE38129, GSE23400, and GSE55856 datasets, which suggests that the ESCC samples could easily be distinguished from the normal controls by analyzing the DEGs or miRNAs.

3.3. Functional Enrichment Analysis of DEGs Screened from the Meta-analysis. As illustrated in Figure 2, DEGs were classified into four functional categories, including pathways, biological process, cellular components, and molecular function. KEGG pathway analysis revealed that the upregulated DEGs were mainly enriched in 18 pathways, including DNA replication, cell cycle, proteasome, base excision repair, and the spliceosome signaling pathway (all, P < 0.05; Table 1). Similarly, 18 KEGG pathways were significantly enriched in downregulated DEGs, such as regulation of actin cytoskeleton, glycine, serine, and threonine metabolism, cGMP-PKG signaling pathway, valine, leucine, and isoleucine degradation and the arginine and proline metabolism signaling pathway (all, P < 0.05).

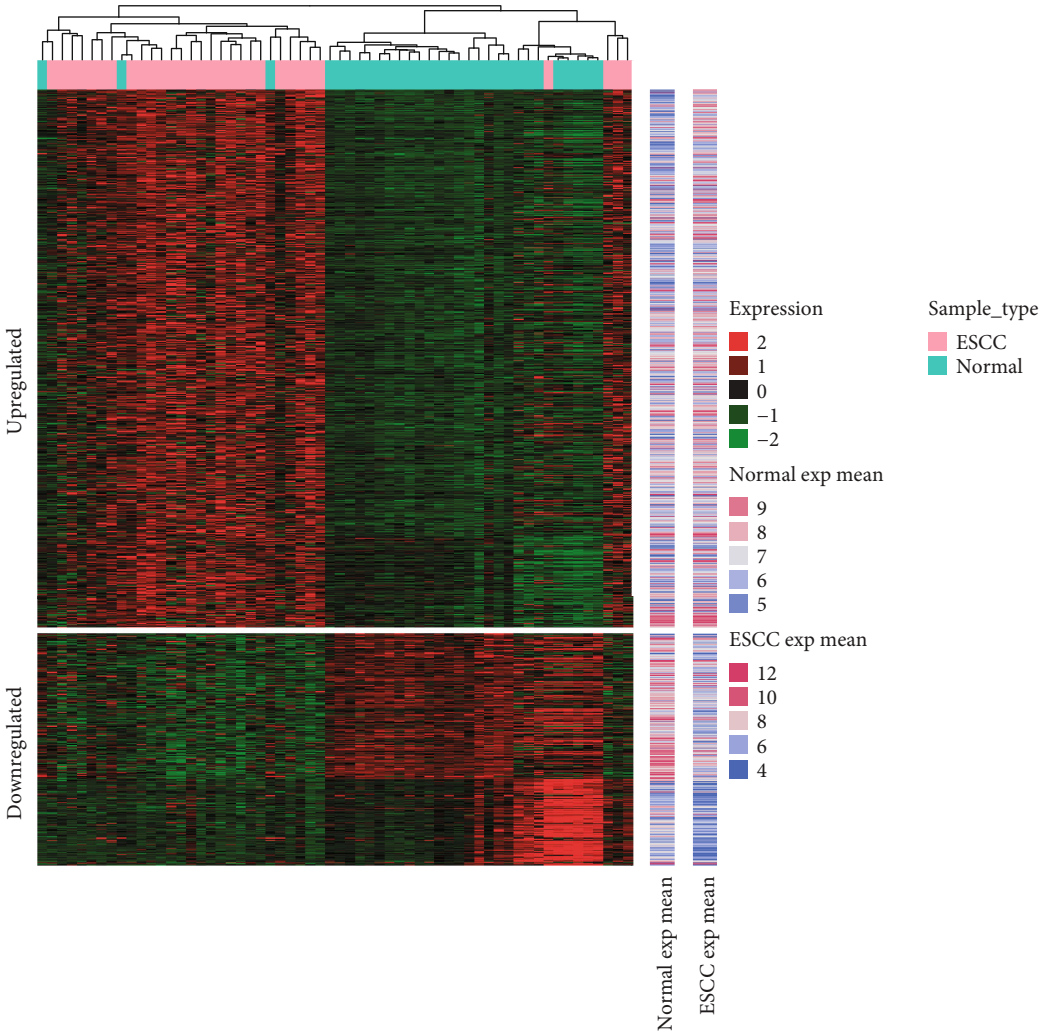
3.4. Identification of Key Candidate Genes and Pathways by DEGs PPI Network Analysis. As shown in Figure 3, a series of DEGs were filtered into the PPI network, which contained 448 nodes and 1144 interaction pairs. Among the nodes, the key candidate node genes were identified by filtering the random walk score. The top 20 nodes, including 17 upregulated and 3 downregulated genes were summarized in Table 2. Among these DEGs, the MYC protooncogene (MYC) had the highest score.

3.5. Regulatory Network of miRNA Target Genes. According to the screening principles of an upregulated miRNA-downregulated gene or downregulated miRNA-upregulated gene, we constructed the miRNA-target gene regulatory network. As shown in Figure 4, a total of 72 upregulated miRNAs which targeted 130 downregulated genes, as well as 19 downregulated miRNAs which targeted 133 upregulated genes were filtered in the network. Based on the miRNAs associated with ESCC obtained from miR2Disease (<http://www.mir2disease.org/>), 8 miRNAs, including miR-196a, miR-21, miR-205, miR-194, miR-103, miR-223, miR-203, and miR-375, were found to be significantly differentially

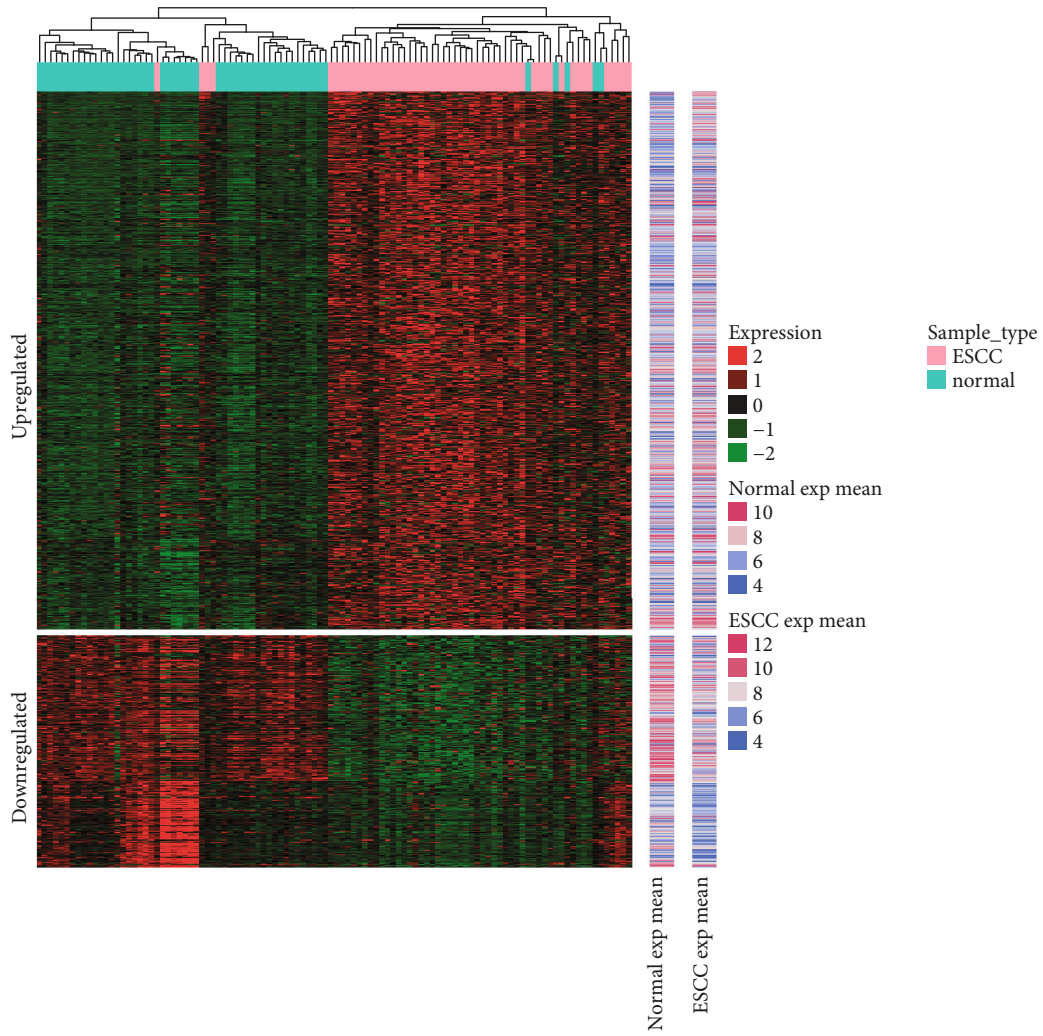


(a)

FIGURE 1: Continued.

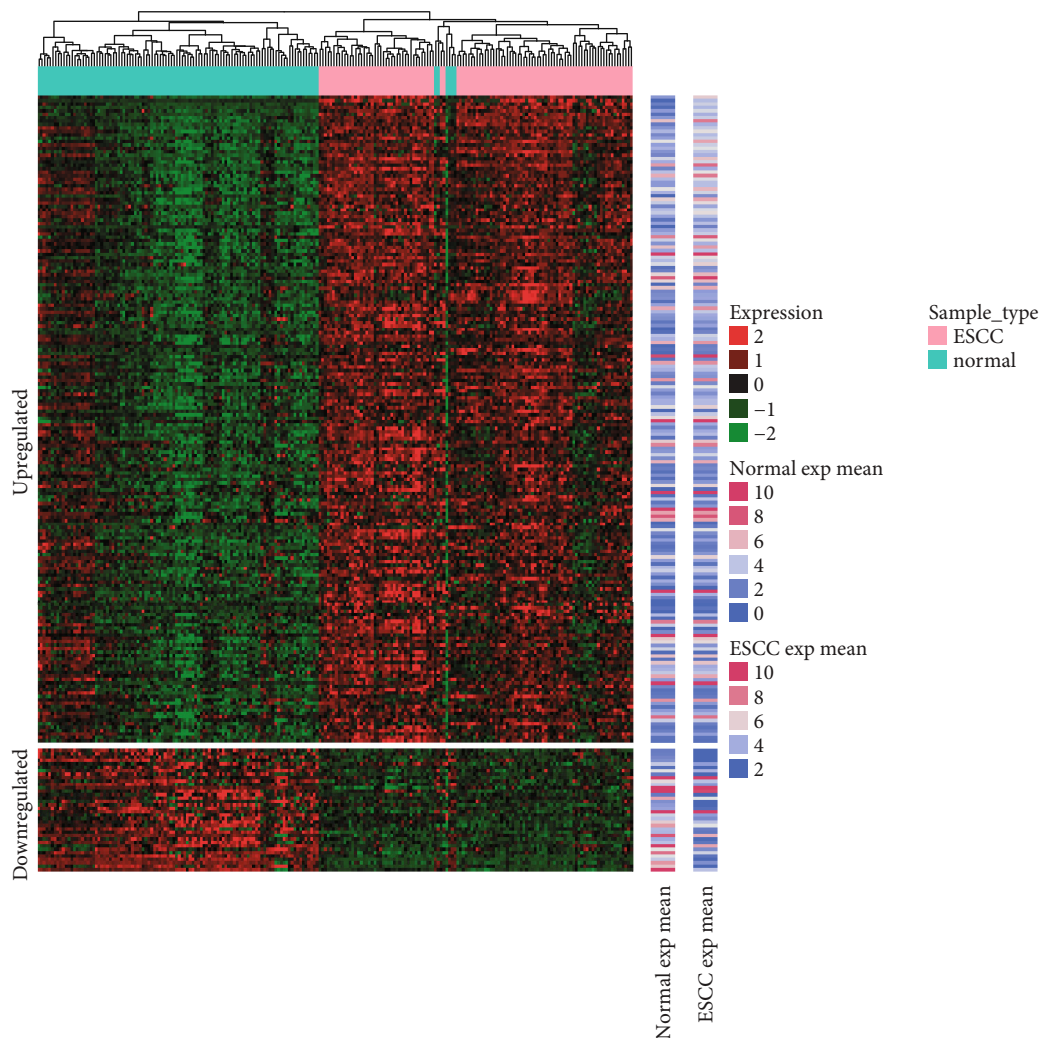


(b)
FIGURE 1: Continued.



(c)

FIGURE 1: Continued.



(d)

FIGURE 1: Heatmap clustering of the differentially expressed genes (DEGs) and miRNAs (DEMs) between ESCC and normal tissues samples in the GSE20347 (a), GSE38129 (b), GSE23400 (c), and GSE55856 (d) datasets. “Red” represents high relative expression and “green” represents a low relative expression.

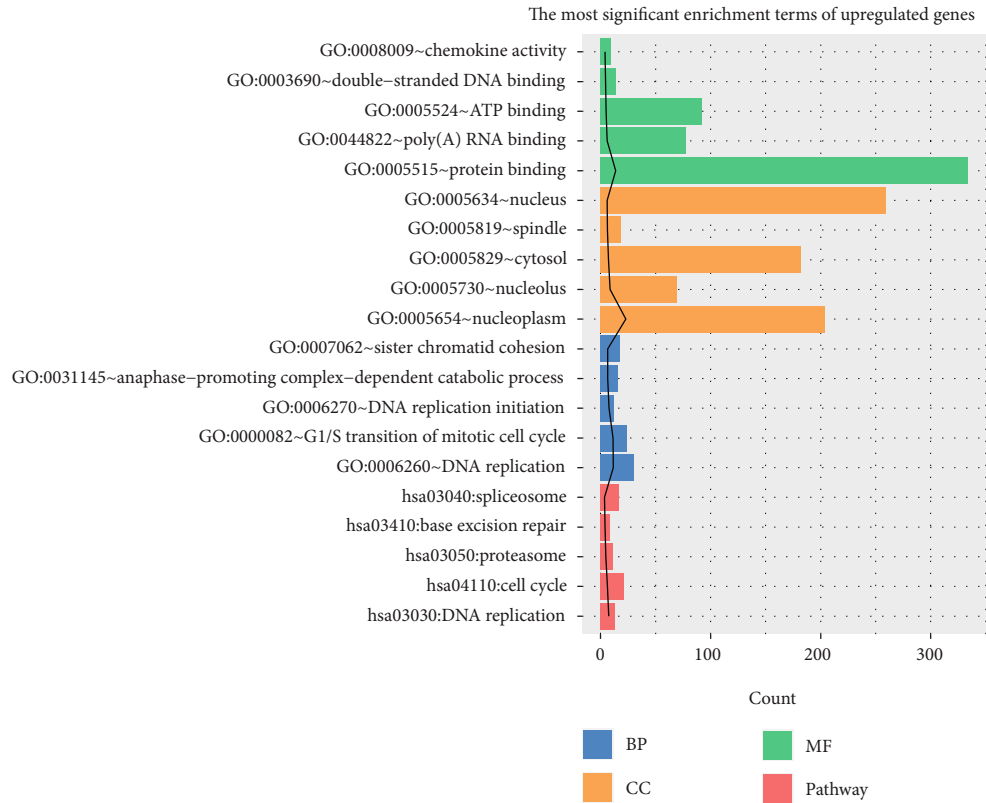
expressed in miRNA microarray datasets. Among these miRNAs, 7 miRNAs excluding miR-203 were found in the network. Moreover, the gene with the highest score in the PPI network, *MYC*, was coregulated by miR-125a-3p, miR-940, and miR-375, among which miR-375 has been reported to be related to ESCC.

3.6. Coregulatory Network between miRNAs. In order to construct the coregulatory network, the miRNAs that regulated the same target gene were identified via the miRNA-target gene regulatory network. As illustrated in Figure 5, several miRNAs which are reported to be disease related, such as hsa-miR-198a, hsa-miR-103, hsa-miR-223, hsa-miR-21, hsa-miR-194, and hsa-miR-375, had common target genes with other DEMs. These miRNAs have played an essential in proliferation, invasion, and metastasis of malignant disease, which is closely related to pathogenesis and prognosis.

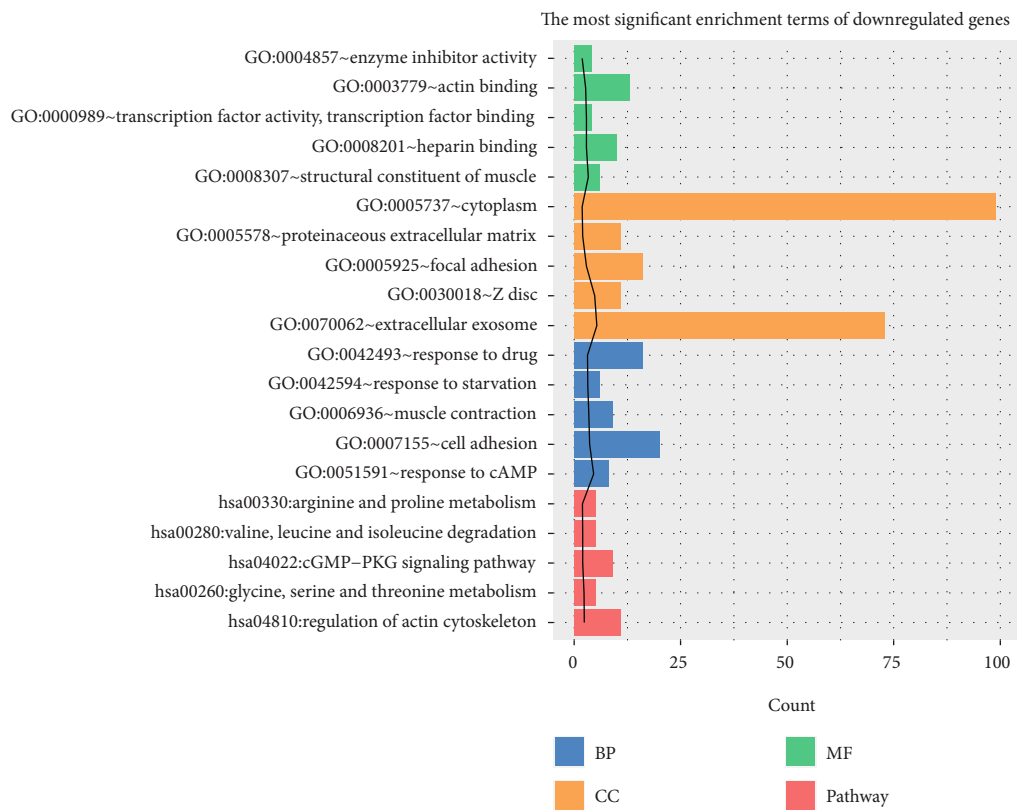
3.7. Survival Analysis of DEMs. After processing of the TCGA data described above, a total of 184 cancer samples and 174 differentially matched miRNAs were obtained. Survival analysis revealed that there were 91 DEMs significantly correlated with the outcome of ESCC patients (Supplementary Table 1).

3.8. Univariate Cox Analysis and Prognostic Model Construction. A total of 13 DEMs were obtained after the univariate Cox analysis. According to the model analysis, a prognostic survival model with 4 DEMs, including hsa-miR-1248, hsa-miR-1291, hsa-miR-421, and hsa-miR-7-5p was obtained. Among these miRNAs, hsa-miR-1248, hsa-miR-1291, and hsa-miR-421 were the DEMs in the GEO data. Moreover, hsa-miR-7-5p was concentrated as the precursor of hsa-miR-7 in the GEO data.

Multivariate Cox analysis was carried out for the following 4 DEMs, hsa-miR-1248, hsa-miR-1291, hsa-miR-421,



(a)



(b)

FIGURE 2: DEG functional enrichment analysis in ESCC. Upregulated (a) and downregulated (b) DEGs were enriched in four functional categories, including pathways, biological processes, cellular components, and molecular functions.

TABLE 1: Signaling pathway enrichment analysis of DEGs in ESCC.

| | Term | Count | P value |
|---------------------------------|---|----------|----------|
| Up | hsa03030:DNA replication | 13 | 3.51E-08 |
| | hsa04110:cell cycle | 21 | 8.02E-07 |
| | hsa03050:proteasome | 11 | 2.33E-05 |
| | hsa03410:base excision repair | 9 | 9.52E-05 |
| | hsa03040:spliceosome | 17 | 3.83E-04 |
| | hsa03430:mismatch repair | 7 | 4.78E-04 |
| | hsa00240:pyrimidine metabolism | 14 | 9.29E-04 |
| | hsa03420:nucleotide excision repair | 8 | 5.35E-03 |
| | hsa05222:small-cell lung cancer | 11 | 5.50E-03 |
| | hsa00480:glutathione metabolism | 8 | 8.42E-03 |
| | hsa05230:central carbon metabolism in cancer | 9 | 8.77E-03 |
| | hsa04062:chemokine signaling pathway | 17 | 1.18E-02 |
| | hsa03013:RNA transport | 16 | 1.29E-02 |
| | hsa04510:focal adhesion | 17 | 2.83E-02 |
| | hsa00230:purine metabolism | 15 | 3.25E-02 |
| | hsa04115:p53 signaling pathway | 8 | 3.39E-02 |
| | hsa04512:ECM-receptor interaction | 9 | 4.68E-02 |
| hsa04145:phagosome | 13 | 5.00E-02 | |
| Down | hsa04810:regulation of actin cytoskeleton | 11 | 4.23E-03 |
| | hsa00260:glycine, serine, and threonine metabolism | 5 | 4.86E-03 |
| | hsa04022:cGMP-PKG signaling pathway | 9 | 9.30E-03 |
| | hsa00280:valine, leucine and isoleucine degradation | 5 | 9.45E-03 |
| | hsa00330:arginine and proline metabolism | 5 | 1.17E-02 |
| | hsa00072:synthesis and degradation of ketone bodies | 3 | 1.29E-02 |
| | hsa00410:beta-alanine metabolism | 4 | 1.72E-02 |
| | hsa04710:circadian rhythm | 4 | 1.72E-02 |
| | hsa04270:vascular smooth muscle contraction | 7 | 1.91E-02 |
| | hsa04924:renin secretion | 5 | 2.68E-02 |
| | hsa04713:circadian entrainment | 6 | 2.68E-02 |
| | hsa04510:focal adhesion | 9 | 3.02E-02 |
| | hsa00982:drug metabolism-cytochrome P450 | 5 | 3.25E-02 |
| | hsa04610:complement and coagulation cascades | 5 | 3.40E-02 |
| | hsa00360:phenylalanine metabolism | 3 | 3.59E-02 |
| | hsa04020:calcium signaling pathway | 8 | 4.01E-02 |
| | hsa05146:amoebiasis | 6 | 4.02E-02 |
| hsa00071:fatty acid degradation | 4 | 4.30E-02 | |

and hsa-miR-7-5p, in the training set and validation set, respectively, and the regression coefficients were obtained (Table 3). Furthermore, the corresponding risk score was calculated for survival analysis and survival test. The threshold determination of the prognostic model was performed. The threshold of the cutoff point in the training set is 1.48 (Figure 6(a)) and in the validation set is 1.56 (Figure 6(b)), respectively. As illustrated in Figures 6(c) and 6(d), the survival analysis results of the risk score obtained by the prognostic model composed of the 4 miRNAs were appropriate in both the training and validation sets (both, $P < 0.01$).

4. Discussion

During the past few decades, an increasing number of studies have investigated the causes and potential mechanisms of ESCC tumorigenesis. However, the high incidence and mortality of ESCC worldwide still pose a challenge, as most studies only focus on a single genetic event [9, 26]. Gene profiling or microarray technologies have been widely used to predict potential targets for the treatment of different tumors. In our study, we downloaded three original microarray data and identified 2575, 2111, and 1205 DEGs, as well as 226 DEMs, between ESCC and normal tissues. Moreover, the meta-

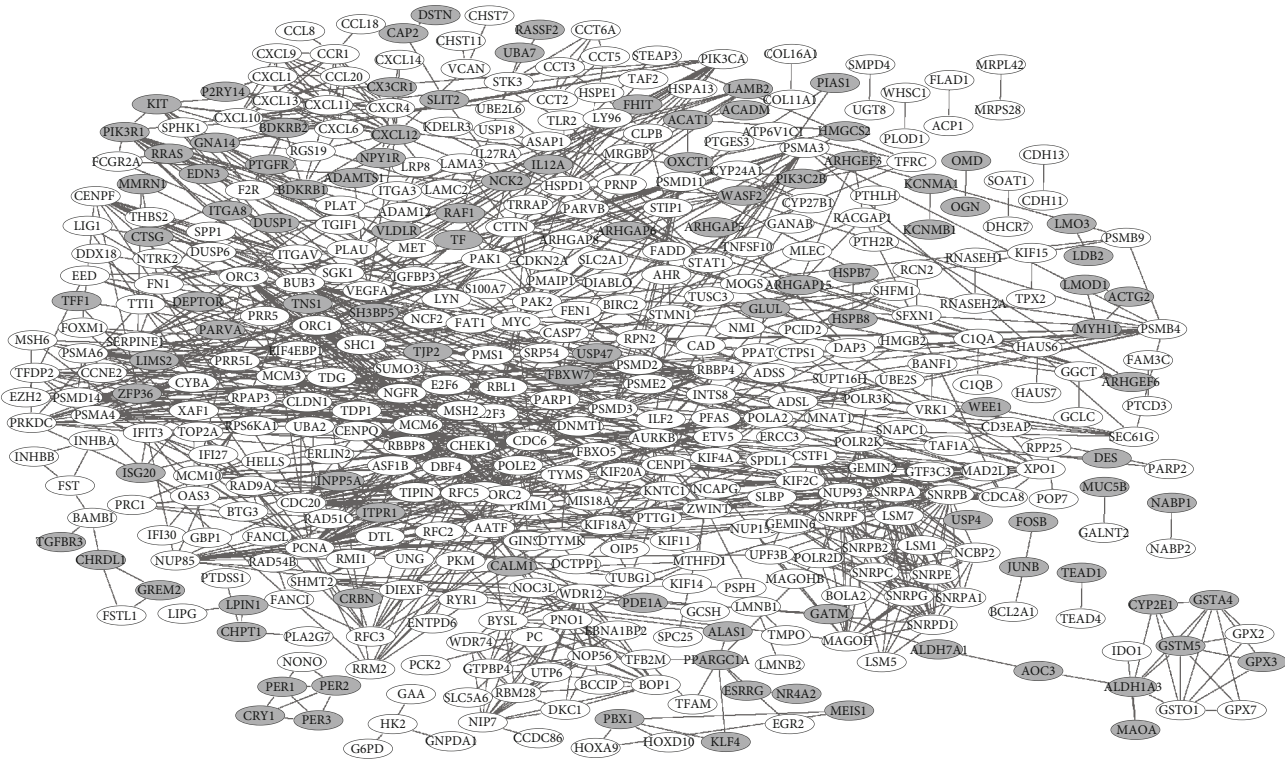


FIGURE 3: DEGs protein-protein interaction (PPI) network in ESCC. There were 448 nodes and 1144 interactions identified in the network. White represents upregulated genes and gray represents downregulated genes.

TABLE 2: The key genes (top 20) in the PPI network in ESCC.

| Node | Direction | Random walk score |
|----------|-----------|-------------------|
| MYC | Up | 0.00656632 |
| PCNA | Up | 0.005105377 |
| AURKB | Up | 0.005066962 |
| STAT1 | Up | 0.004433487 |
| CXCL12 | Down | 0.003986094 |
| POLR2K | Up | 0.003855735 |
| CDC20 | Up | 0.003688161 |
| PIK3CA | Up | 0.003613145 |
| CDC6 | Up | 0.003570703 |
| PRKDC | Up | 0.00355442 |
| SHMT2 | Up | 0.003539976 |
| CHEK1 | Up | 0.003533405 |
| PPARGC1A | Down | 0.003496619 |
| MAGOH | Up | 0.003355163 |
| SERPINE1 | Up | 0.003290157 |
| CCR1 | Up | 0.003277042 |
| LYN | Up | 0.003267606 |
| HK2 | Up | 0.003262363 |
| WDR12 | Up | 0.003251239 |
| PIK3R1 | Down | 0.003241252 |

analysis was used to further obtain more reliable DEGs, and 1001 genes that could be used for sample clustering of each dataset were identified. Furthermore, based on the clinical,

survival, and miRNA-seq data downloaded from TCGA, a total of 85 DEMs were found to be significantly associated with the outcome of ESCC patients. As such, a prognostic survival model composed of 4 DEMs, including hsa-miR-1248, hsa-miR-1291, hsa-miR-421, and hsa-miR-7-5p was constructed.

It has been known that tumorigenesis is a complex process that involves the interaction of various genes and signaling pathways [26, 27]. In ESCC, an increasing number of signaling pathways have been reported to play important roles in the progression of the disease [28–30]. Therefore, the analysis of pathways related to neoplasia could provide information regarding tumor development. In the present study, the KEGG pathway analysis revealed that both upregulated and downregulated DEGs were mainly enriched in 18 pathways, such as the DNA replication, cell cycle, p53, proteasome, BcGMP-PKG signaling pathway, valine, leucine, and isoleucine degradation, beta-alanine metabolism, and arginine and proline metabolism signaling pathways. Among these pathways, recent studies have shown a correlation between the proteasome pathway and ESCC carcinogenesis [31]. Zhou et al. [32] revealed that the cell growth and apoptosis of ESCC could be regulated via activation of the p53 signaling pathway. Moreover, the beta-alanine metabolism pathway was reported as a novel subpathway related to ESCC, by the cooperative use of different genes in different pathways [33]. Taken together, the detection of molecules related to these pathways may help predict the occurrence and progression of ESCC.

Furthermore, the PPI network of proteins encoded by the identified DEGs was constructed. It contained 448 nodes and

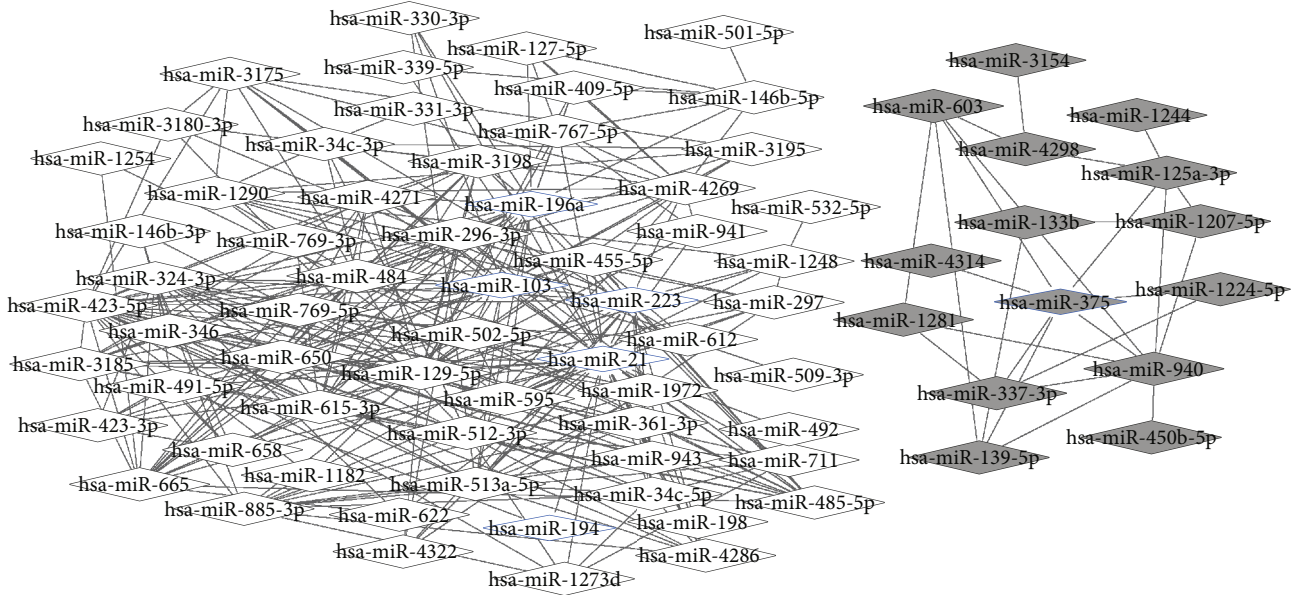


FIGURE 4: MiRNA-target gene regulatory network. There were 72 upregulated miRNAs and 130 downregulated target genes, as well as 19 downregulated miRNAs and 133 upregulated target genes in the network. Gray represents a downregulated expression, while white represents an upregulated expression. Ovals represent the differentially expressed target gene, and diamonds represents the DEMs (blue edge represents miRNAs reported to be disease related).

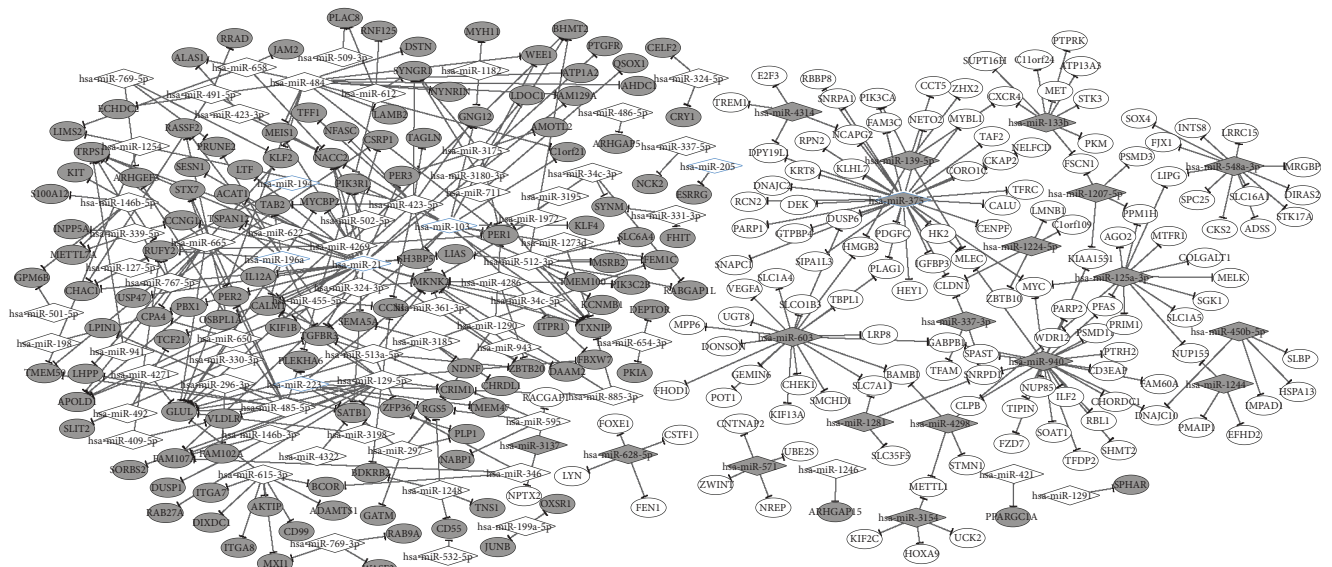


FIGURE 5: The coregulatory network between miRNAs. Several reported disease-related miRNAs shared common target genes with other DEMs. Gray represents a downregulated expression and white represents an upregulated expression. The blue thickened edge represents miRNAs reported to be disease related.

TABLE 3: The regression coefficients of 4 miRNAs.

| Training set | Validation set | | |
|--------------|----------------|--------------|---------|
| hsa-miR-1248 | 0.62396 | PANTR1 | 0.37498 |
| hsa-miR-1291 | 0.40231 | LINC01266 | 0.09693 |
| hsa-miR-421 | -0.02532 | FGF13-AS1 | 0.10180 |
| hsa-miR-7-5p | 0.19866 | TMEM132D-AS1 | 0.34272 |

1144 interactions. *MYC*, with the highest random walk score, could be regarded as the key gene in the PPI network. Indeed, accumulating evidence has demonstrated that *MYC* is an important factor in biological development and tumorigenesis. For instance, Kwon et al. [34] showed that *MYC* was overexpressed in ESCC patients and that its expression could predict better overall survival (OS) for patients. Likewise, Zhong et al. [35] revealed that *MYC* is involved in the tumorigenicity of ESCC by regulating the expression of the

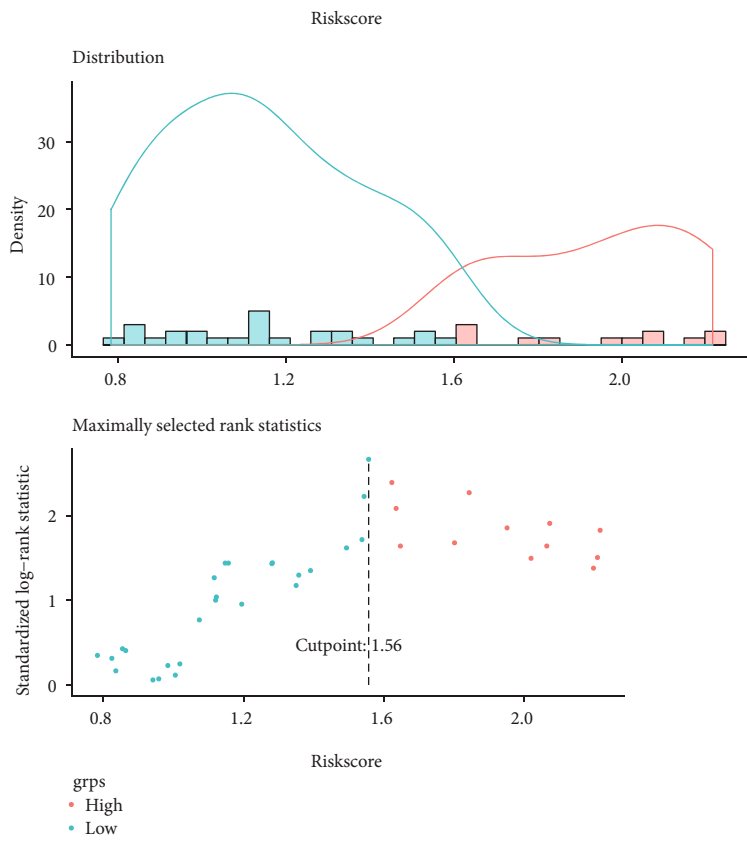
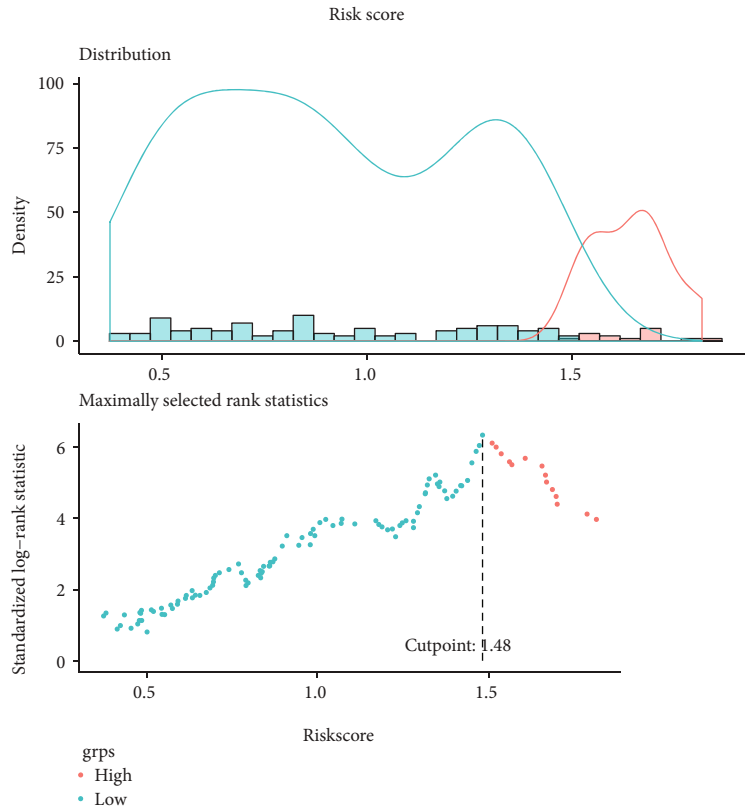


FIGURE 6: Continued.

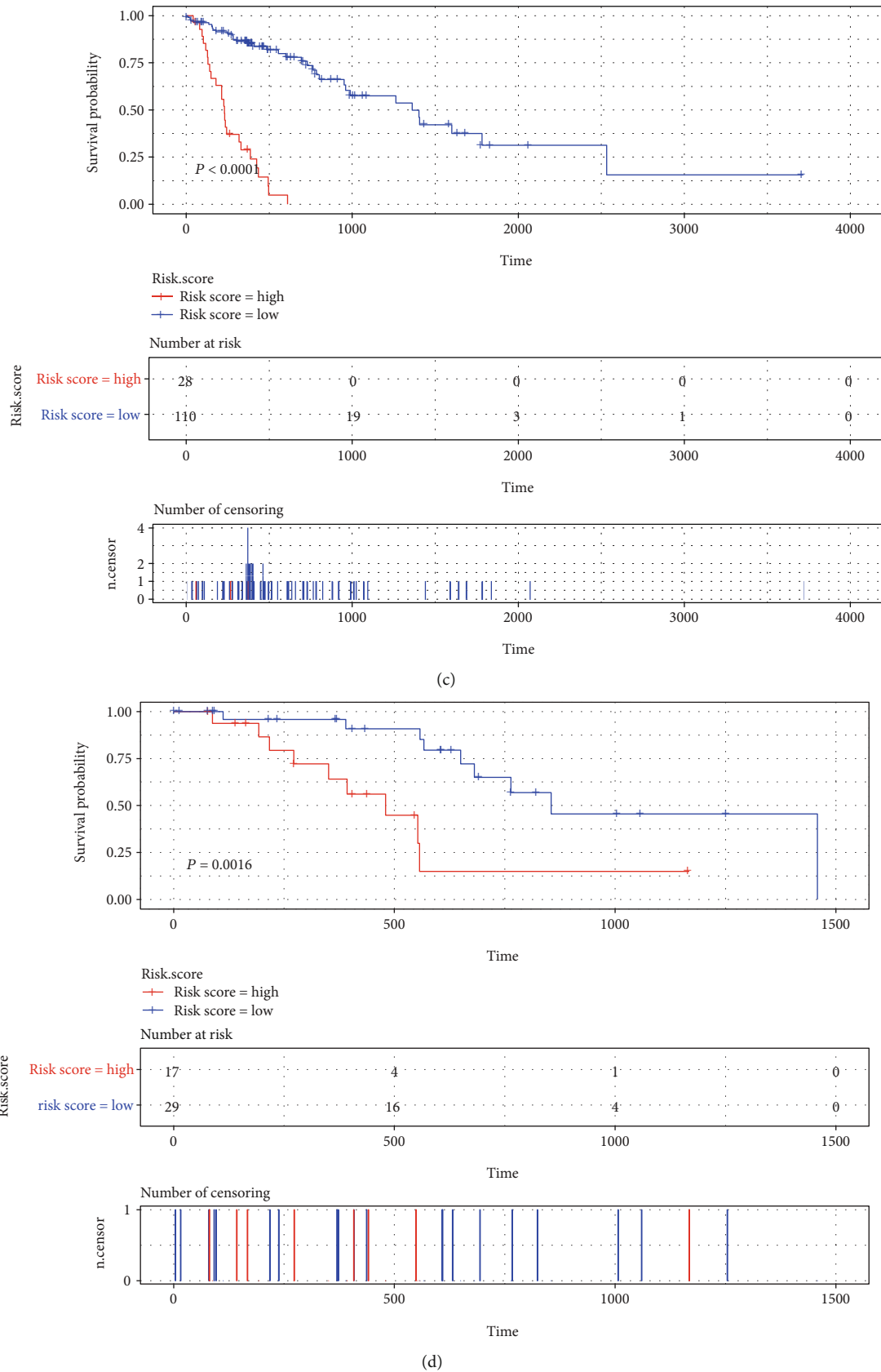


FIGURE 6: The threshold determination and survival test of the risk score of the prognostic model. The threshold of the cutoff point training set (a) and validation set (b). Survival results of the risk score obtained in the training set (c) and validation set (d) of the prognostic model composed of 4 miRNAs.

hydroxymethylglutaryl coenzyme A reductase (*HMGCR*). Therefore, further research on *MYC* will provide a basis for targeted therapy against ESCC.

miRNAs, which are endogenous, noncoding single-stranded RNAs, have been reported to play critical roles in various biological processes via binding to their target mRNAs [36, 37]. In the present study, a miRNA-target gene regulatory network using the identified DEMs was constructed, and seven miRNAs associated with ESCC, including miR-196a, miR-21, miR-205, miR-194, miR-103, miR-223, and miR-375 were identified in both the microarray data and the miRNA-target gene regulatory network. Among these miRNAs, miR-196a, miR-21, and miR-205 were previously found to be abnormally expressed in ESCC [38–40]. Notably, miR-125a-3p, miR-940, and miR-375 collectively regulated the expression of *MYC*, suggesting that these miRNAs might play a role in ESCC by regulating the *MYC* expression. Additionally, the miRNAs that regulated the same target gene were identified using the coregulatory network, and miR-198a, miR-103, miR-223, miR-21, miR-194, and miR-375 were reported to be miRNAs related to ESCC [41–44]. Survival analysis showed that a total of 85 DEMs were related to prognosis, among which hsa-miR-1248, hsa-miR-1291, hsa-miR-421, and hsa-miR-7-5p were used in a prognostic survival model. Indeed, miR-1248 has been reported to be involved in the microRNA signature model for the prediction of prognosis in patients with the Wilms tumors [45]. Moreover, miR-1291 and miR-421 are associated with OS in patients with lung adenocarcinoma [46, 47]. However, a prognostic survival model using these miRNAs has not been constructed for ESCC. The results of our study revealed that the tumorigenesis of ESCC may be the result of the coregulation of multiple miRNAs. Moreover, several miRNAs, such as hsa-miR-1248, hsa-miR-1291, hsa-miR-421, and hsa-miR-7-5p could be used to construct a prognostic survival model for the prediction of ESCC patient outcomes.

In conclusion, in the present study, we conducted a comprehensive bioinformatics analysis of DEGs, DEMs, and pathways, based on different datasets. As such, we identified DEGs such as *MYC*, miRNAs such as miR-125a-3p, miR-940, miR-375, miR-196a, miR-21, miR-205, miR-194, miR-103, miR-223, and miR-198a, and pathways such as the proteasome, p53, and beta-alanine metabolism pathways, which may be involved in ESCC development. Notably, survival analysis showed that 85 DEMs were related to prognosis, among which hsa-miR-1248, hsa-miR-1291, hsa-miR-421, and hsa-miR-7-5p were used to construct a prognostic survival model. Taken together, these findings have important clinical significance, as they can improve our understanding of the pathogenesis and molecular mechanisms of ESCC. Moreover, our results provide potential biomarkers for the prediction of ESCC prognosis. However, further studies are still needed to confirm the function of the identified genes.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Le-meng Zhang and Jianhua Chen contributed to the clinical data collection. Tianli Cheng, Hua Yang, Changqie Pan, Haitao Li, and Yongzhong Luo contributed to the data statistics analysis. Le-meng Zhang contributed to the study design, manuscript writing, and data analysis. All authors read and approved the final manuscript.

Acknowledgments

This study was supported by grants from the Hunan Provincial Key Research and Development Program for Social Development (2016SK2006), National Natural Science Foundation of Hunan Province (2019JJ50353), Natural Science Foundation of Hunan Province National Health Commission (B2019091), Natural Science Foundation of Changsha Science and Technology Bureau (kp1901084), Wu Jieping Medical Foundation (320.6750.18368), and Cancer Foundation of China (NCC2018B58).

Supplementary Materials

Supplementary Table 1: 85 miRNAs significantly correlated with prognosis. (*Supplementary Materials*)

References

- [1] H. Zeng, R. Zheng, S. Zhang et al., "Esophageal cancer statistics in China, 2011: estimates based on 177 cancer registries," *Thoracic Cancer*, vol. 7, no. 2, pp. 232–237, 2016.
- [2] J. Lagergren, "Oesophageal cancer in 2014: advances in curatively intended treatment," *Nature Reviews Gastroenterology & Hepatology*, vol. 12, no. 2, pp. 74–75, 2015.
- [3] P. Zhang, X. M. Li, X. K. Zhao et al., "Novel genetic locus at MHC region for esophageal squamous cell carcinoma in Chinese populations," *PLoS One*, vol. 12, no. 5, article e0177494, 2017.
- [4] W. Chen, R. Zheng, P. D. Baade et al., "Cancer statistics in China, 2015," *Ca: A Cancer Journal for Clinicians*, vol. 66, no. 2, pp. 115–132, 2016.
- [5] G. D. Tran, X. D. Sun, C. C. Abnet et al., "Prospective study of risk factors for esophageal and gastric cancers in the Linxian general population trial cohort in China," *International Journal of Cancer*, vol. 113, no. 3, pp. 456–463, 2005.
- [6] L. M. Best, M. Mughal, and K. S. Gurusamy, "Non-surgical versus surgical treatment for oesophageal cancer," *Cochrane Database of Systematic Reviews*, vol. 3, no. 3, article CD011498, 2016.
- [7] P. van Hagen, M. Hulshof, van Lanschot et al., "Preoperative chemoradiotherapy for esophageal or junctional cancer," *New England Journal of Medicine*, vol. 366, no. 22, pp. 2074–2087, 2012.
- [8] B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, D. L. Jr, and K. W. Kinzler, "Cancer genome landscapes," *Science*, vol. 339, no. 6127, pp. 1546–1558, 2013.

- [9] Y. Yuan, L. Xue, and H. Fan, "Screening of differentially expressed genes related to esophageal squamous cell carcinoma and functional analysis with DNA microarrays," *International Journal of Oncology*, vol. 44, no. 4, pp. 1163–1170, 2014.
- [10] J. Xing and C. Liu, "Identification of genes associated with histologic tumor grade of esophageal squamous cell carcinoma," *FEBS Open Bio*, vol. 7, no. 9, pp. 1246–1257, 2017.
- [11] N. Hu, C. Wang, R. J. Clifford et al., "Integrative genomics analysis of genes with biallelic loss and its relation to the expression of mRNA and micro-RNA in esophageal squamous cell carcinoma," *BMC Genomics*, vol. 16, pp. 1–11, 2015.
- [12] H. L. Fu, D. P. Wu, X. F. Wang et al., "Altered miRNA expression is associated with differentiation, invasion, and metastasis of esophageal squamous cell carcinoma (ESCC) in patients from Huaian, China," *Cell Biochemistry & Biophysics*, vol. 67, no. 2, pp. 657–668, 2013.
- [13] J. Slotta-Huspenina, E. Drecoll, M. Feith et al., "MicroRNA expression profiling for the prediction of resistance to neoadjuvant radiochemotherapy in squamous cell carcinoma of the esophagus," *Journal of Translational Medicine*, vol. 16, no. 1, p. 109, 2018.
- [14] H. J. Jang, H. S. Lee, B. M. Burt et al., "Integrated genomic analysis of recurrence-associated small non-coding RNAs in oesophageal cancer," *Gut*, vol. 66, no. 2, pp. 215–225, 2017.
- [15] L. Gautier, L. Cope, B. M. Bolstad, and R. A. Irizarry, "Affy—analysis of Affymetrix GeneChip data at the probe level," *Bioinformatics*, vol. 20, no. 3, pp. 307–315, 2004.
- [16] G. Smyth and G. K. Smyth, "Limma: linear models for microarray data," in *Bioinformatics and Computational Biology Solution Using R and Bioconductor*, Statistics for Biology and Health, R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit, Eds., Springer, New York, NY, 2013.
- [17] C. Qi, L. Hong, Z. Cheng, and Q. Yin, "Identification of metastasis-associated genes in colorectal cancer using metaDE and survival analysis," *Oncology Letters*, vol. 11, no. 1, pp. 568–574, 2016.
- [18] R. Kolde, *pheatmap: pretty heatmaps*, 2015.
- [19] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.
- [20] D. Szklarczyk, A. Franceschini, S. Wyder et al., "STRING v10: protein–protein interaction networks, integrated over the tree of life," *Nucleic Acids Research*, vol. 43, no. D1, pp. D447–D452, 2015.
- [21] S. Köhler, S. Bauer, D. Horn, and P. N. Robinson, "Walking the interactome for prioritization of candidate disease genes," *American Journal of Human Genetics*, vol. 82, no. 4, pp. 949–958, 2008.
- [22] H. Dweep and N. Gretz, "miRWalk2.0: a comprehensive atlas of microRNA–target interactions," *Nature Methods*, vol. 12, no. 8, p. 697, 2015.
- [23] Q. Jiang, Y. Wang, Y. Hao et al., "miR2Disease: a manually curated database for microRNA deregulation in human disease," *Nucleic Acids Research*, vol. 37, Supplement 1, pp. D98–D104, 2009.
- [24] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *Omics—a Journal of Integrative Biology*, vol. 16, no. 5, pp. 284–287, 2012.
- [25] R. C. Team, *R: A Language and Environment for Statistical Computing*, R foundation for statistical computing, 2013.
- [26] X. Wang, G. Li, Q. Luo, and C. Gan, "Identification of crucial genes associated with esophageal squamous cell carcinoma by gene expression profile analysis," *Oncology Letters*, vol. 15, pp. 8983–8990, 2018.
- [27] J. Long, Z. Liu, W. U. Kingda, X. U. Yuanhong, and G. E. Chunlin, "Gene expression profile analysis of pancreatic cancer based on microarray data," *Molecular Medicine Reports*, vol. 13, no. 5, pp. 3913–3919, 2016.
- [28] H. Xiaoyu, Y. Yiru, S. Shuisheng et al., "The mTOR pathway regulates PKM2 to affect glycolysis in esophageal squamous cell carcinoma," *Technology in Cancer Research & Treatment*, vol. 17, 2018.
- [29] Y. Tu, F. Tan, J. Zhou, and J. Pan, "Pristimerin targeting NF- κ B pathway inhibits proliferation, migration, and invasion in esophageal squamous cell carcinoma cells," *Cell Biochemistry & Function*, vol. 36, no. 4, pp. 228–240, 2018.
- [30] X. Li, D. Chen, M. Li, X. Gao, G. Shi, and H. Zhao, "The CADM2/Akt pathway is involved in the inhibitory effect of miR-21-5p downregulation on proliferation and apoptosis in esophageal squamous cell carcinoma cells," *Chemico-Biological Interactions*, vol. 288, pp. 76–82, 2018.
- [31] Q. Yang, W. Lin, Z. Liu et al., "RAP80 is an independent prognosis biomarker for the outcome of patients with esophageal squamous cell carcinoma," *Cell Death & Disease*, vol. 9, no. 2, p. 146, 2018.
- [32] P. Zhou, R. Zhang, Y. Wang et al., "Cepharanthine hydrochloride reverses the mdrl (P-glycoprotein)-mediated esophageal squamous cell carcinoma cell cisplatin resistance through JNK and p 53 signals," *Oncotarget*, vol. 8, no. 67, pp. 111144–111160, 2017.
- [33] C. Li, Q. Wang, J. Ma et al., "Integrative pathway analysis of genes and metabolites reveals metabolism abnormal subpathway regions and modules in esophageal squamous cell carcinoma," *Molecules*, vol. 22, no. 10, p. 1599, 2017.
- [34] D. Kwon, Y. Y. Ji, B. Keam, Y. T. Kim, and Y. K. Jeon, "Prognostic implications of FGFR1 and MYC status in esophageal squamous cell carcinoma," *World Journal of Gastroenterology*, vol. 22, no. 44, pp. 9803–9812, 2016.
- [35] C. Zhong, L. Fan, F. Yao, J. Shi, W. Fang, and H. Zhao, "HMGR is necessary for the tumorigenicity of esophageal squamous cell carcinoma and is regulated by Myc," *Tumour Biology*, vol. 35, no. 5, pp. 4123–4129, 2014.
- [36] G. C. Shukla, J. Singh, and S. Barik, "MicroRNAs: processing, maturation, target recognition and regulatory functions," *Molecular & Cellular Pharmacology*, vol. 3, p. 83, 2011.
- [37] D. P. Bartel, "MicroRNAs: target recognition and regulatory functions," *Cell*, vol. 136, no. 2, pp. 215–233, 2009.
- [38] M. Fendereski, M. F. Zia, M. Shafiee, F. Safari, M. H. Saneie, and M. Tavassoli, "MicroRNA-196a as a potential diagnostic biomarker for esophageal squamous cell carcinoma," *Cancer Investigation*, vol. 35, no. 2, pp. 78–84, 2017.
- [39] K. Wang, D. Chen, Y. Meng, J. Xu, and Q. Zhang, "Clinical evaluation of 4 types of microRNA in serum as biomarkers of esophageal squamous cell carcinoma," *Oncology Letters*, vol. 16, pp. 1196–1204, 2018.
- [40] H. Xu, Y. Yao, F. Meng et al., "Predictive value of serum miR-10b, miR-29c, and miR-205 as promising biomarkers in esophageal squamous cell carcinoma screening," *Medicine*, vol. 94, no. 44, article e1558, 2015.

- [41] Z. Huang, L. Zhang, D. Zhu, X. Shan, and X. Zhou, "A novel serum microRNA signature to screen esophageal squamous cell carcinoma," *Cancer Medicine*, vol. 6, no. 1, pp. 109–119, 2017.
- [42] F. Li, J.-H. Lv, L. Liang et al., "Downregulation of microRNA-21 inhibited radiation-resistance of esophageal squamous cell carcinoma," *Cancer Cell International*, vol. 18, no. 1, p. 39, 2018.
- [43] C. Wu, C. Wang, X. Guan et al., "Diagnostic and prognostic implications of a serum miRNA panel in oesophageal squamous cell carcinoma," *PLoS One*, vol. 9, no. 3, article e92292, 2014.
- [44] J. Yi, L. Jin, J. Chen et al., "MiR-375 suppresses invasion and metastasis by direct targeting of SHOX2 in esophageal squamous cell carcinoma," *Acta biochimica et biophysica Sinica*, vol. 49, no. 2, pp. 159–169, 2017.
- [45] Y. Gong, B. Zou, J. Chen et al., "Potential five-microRNA signature model for the prediction of prognosis in patients with Wilms tumor," *Medical Science Monitor*, vol. 25, pp. 5435–5444, 2019.
- [46] S. Y. Sathipati and S.-Y. Ho, "Identifying the miRNA signature associated with survival time in patients with lung adenocarcinoma using miRNA expression profiles," *Scientific Reports*, vol. 7, no. 1, p. 7507, 2017.
- [47] Y. Li, X. Cui, Y. Li, T. Zhang, and S. Li, "Upregulated expression of miR-421 is associated with poor prognosis in non-small-cell lung cancer," *Cancer Management and Research*, vol. Volume 10, pp. 2627–2633, 2018.

Research Article

Prognostic Value of Complement Component 2 and Its Correlation with Immune Infiltrates in Hepatocellular Carcinoma

Gang Ning , Yan-Lin Huang , Li-Min Zhen, Wen-Xiong Xu , Xue-Jun Li , Li-Na Wu, Ying Liu, Chan Xie , and Liang Peng 

Department of Infectious Diseases, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

Correspondence should be addressed to Chan Xie; happyxiechan@hotmail.com and Liang Peng; pliang@mail.sysu.edu.cn

Received 21 March 2020; Revised 20 May 2020; Accepted 25 May 2020; Published 15 June 2020

Guest Editor: Tao Huang

Copyright © 2020 Gang Ning 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. Single nucleotide polymorphism (SNP) of complement component 2 (C2) has been found to be significantly associated with hepatocellular carcinoma (HCC). However, little is known about the role and mechanism of C2 in HCC. In the present study, we aimed to explore the prognostic value of C2 and its correlation with tumor-infiltrating immune cells in HCC. **Materials and Methods.** mRNA expression was downloaded from TCGA (365 HCC patients and 50 healthy controls), GSE14520 (220 HCC patients and 220 adjacent normal tissues), and ICGC HCC (232 HCC patients) cohorts. Unpaired Student's *t*-tests or ANOVA tests were used to evaluate differences of C2 expression. Univariate and multivariate analyses were used to analyze the prognostic value of C2. CIBERSORT was used to calculate the proportion of 22 kinds of tumor-infiltrating immune cells. **Results.** Significantly lower C2 expression was found at HCC compared to healthy controls, and C2 was associated with TNM stages. Higher C2 expression was significantly associated with better prognosis, and multivariate analysis showed that C2 was also an independent factor for the prognosis of HCC. Moreover, elevated CD4 T cells were found at HCC patients with higher C2 expression while the higher proportion of macrophage M0 cells was found in HCC patients with lower C2 expression. KEGG analysis showed that “cell cycle,” “AMPK signaling pathway,” and “PPAR signaling pathway” were enriched in HCC patients with higher C2 expression. **Conclusion.** C2 is a prognostic factor for HCC and may be used as a therapeutic target for future treatment of HCC.

1. Introduction

As one of the most common cancers and a leading cause of cancer-related death worldwide, hepatocellular carcinoma (HCC) continues to be a tremendous public burden on society [1]. Up to 50% of HCC patients are estimated to develop recurrence after resection partly due to limited amenable curative treatment options and rapid development of resistance [2]. A growing number of studies have indicated that tumor microenvironment (TME) plays important roles in almost every key aspect of HCC tumorigenesis, such as tumor initiation, progression, immune invasion, metastasis, recurrence, and resistance to therapy [3–5]. Therefore, understanding the interactions of stromal cells with cancer cells will help to develop effective strategies to conquer resistance and improve the therapeutic effect for HCC [6].

The complement system is a fundamental branch of innate immunity and could rapidly respond to invading pathogens by promoting cell lysis [7]. Remarkably, studies carried out over the last decade have shed new light on complement activation in the TME, which contributes to tumor-promoting and tumor-suppressing immune responses [8]. For example, complement component 7 (C7) and complement factor H (CFH) are found to be necessary for maintaining stemness of HCC cells as silence of C7 and CFH inhibits tumor-sphere formation and promotes cell differentiation while overexpression of them elevates stemness factor expression and cell growth in vivo [9]. C3 is required for the local and systemic immune responses against tumor in mice with G422 gliomas generated by photodynamic therapy (PDT), because knockout of C3 reduces the infiltration of immune cells and release of TNF- α and IFN- γ , which

indicates a crucial role played by C3 in mediating antitumor immunity [10]. Interestingly, bilateral effect of complement component 5a (C5a) on tumor progression has also been observed. High levels of C5a are related to tumorigenesis accompanied by reduced IFN- γ -producing CD8 and CD4 cells, while a low level of C5a is associated with decreased tumor burden with increased IFN- γ -producing CD4 and CD8 T cells in mice with lymphoma [11]. These studies indicate entirely different roles played by each complement component on cancer development.

Complement component 2 (C2) is an important part of the complement system, and single nucleotide polymorphism (SNP) of C2 has been found to be significantly associated with HCC [12, 13]. For example, C2 SNP rs9267665 is associated with the risk of developing HCC while rs10947223 shows protective effects against HCC, indicating an important role played by C2 in HCC. However, up to now, little is known about the role and mechanism of C2 in HCC, so in the present study, we aimed to explore the prognostic value of C2 and its correlation with tumor-infiltrating immune cells of TME in HCC patients.

2. Materials and Methods

2.1. Ethics Statement. All the data analyzed in the present study were attained from The Cancer Genome Atlas (TCGA) dataset, Gene Expression Omnibus (GEO) dataset, and the International Cancer Genome Consortium (ICGC) dataset, and informed consents had been gained from each patient before our study.

2.2. Acquisition of mRNA Expression and Corresponding Clinical-Pathological Parameters from TCGA, GEO, and ICGC. In the present study, a total of three cohorts, including the TCGA HCC cohort, GSE14520 HCC cohort, and ICGC HCC cohort, were employed. mRNA expression and corresponding clinical-pathological parameters of the TCGA HCC cohort were downloaded from TCGA (<https://cancergenome.nih.gov/>). In the TCGA HCC cohort, clinical-pathological parameters of 377 HCC patients, including gender, age, histologic grades, cirrhosis, TNM stage, status, and time of overall survival (OS), were attained. Meanwhile, mRNA expression of 374 HCC patients and 50 healthy controls was also downloaded. The GSE14520 HCC cohort was downloaded from GEO (<https://www.ncbi.nlm.nih.gov/geo/>). In the GSE14520 HCC cohort, clinical-pathological parameters of 220 HCC patients, including gender, age, tumor size, cirrhosis, TNM staging, status, and time of OS, and mRNA expression of 220 HCC patients and their corresponding adjacent normal tissues were available. mRNA expression and corresponding clinical-pathological parameters of the ICGC HCC cohort were got from the ICGC portal (<https://dcc.icgc.org/projects/LIRI-JP>). In the ICGC HCC cohort, mRNA expression and clinical-pathological parameters of 232 HCC patients, including gender, age, TNM staging, status, and time of OS, were available. Basic demographic characteristics of TCGA, GSE14520, and ICGC HCC cohorts are summarized in Table 1.

2.3. CIBORSORT. An online tool, CIBERSORT, was used to calculate the proportion of 22 kinds of tumor-infiltrating immune cells with transcriptomic data (<https://cibersort.stanford.edu>). In CIBERSORT, relative fractions of 22 kinds of immune cells were deconvolved from the transcriptional expression of tumor samples basing on a referenced signature matrix by linear support vector regression.

2.4. Kyoto Encyclopedia of Genes and Genomes (KEGG) Analysis. The underlying mechanism of C2 in hepatocarcinogenesis was analyzed by KEGG analysis in the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (<https://david.ncifcrf.gov/summary.jsp>). First, HCC patients of TCGA or GSE14520 or ICGC HCC cohort were divided into high C2 and low C2 groups according to the median C2 expression. Then, differentially expressed genes (DEGs) between the two groups were found with a cut-off value of $p < 0.05$. Finally, KEGG pathways enriched by these DEGs were identified in DAVID, and a cut-off value of $p < 0.05$ was considered as statistically significant.

2.5. Statistical Analysis. GraphPad Prism 6 (GraphPad Software, La Jolla, CA, USA) was used to carry out the statistical analyses. Data were presented as the median. Unpaired Student's t -tests or ANOVA tests were performed to compare the difference of C2 expression between HCC patients and healthy controls/adjacent normal tissues or among HCC patients with different histologic grades or TNM stages. Univariate and multivariate Cox regression analyses were carried out to analyze the prognostic value of C2 expression, and Kaplan-Meier analysis with a two-sided log-rank test was also performed to compare the OS of HCC patients with high or low C2 expression. Additional statistical analysis was performed with STAMP [14]. $p < 0.05$ was considered as statistically significant.

3. Results

3.1. C2 Expression between HCC Patients and Healthy Controls/Adjacent Normal Tissues. We first analyzed the C2 expression between HCC patients and healthy controls/adjacent normal tissues of TCGA and GSE14520 HCC cohorts. As is shown in Figure 1, in the TCGA cohort, significantly lower C2 expression was found at HCC patients compared to healthy controls ($p < 0.01$, Figure 1(a)). Similar results were also found in the GSE14520 cohort, in which significantly lower C2 expression was also found at HCC tissues compared to adjacent normal tissues ($p < 0.001$, Figure 1(b)). In short, the above results suggested that C2 expression of HCC was lower than that of healthy controls.

3.2. Association of C2 Expression with Clinical-Pathological Parameters of HCC Patients. After reduced C2 expression was found in HCC patients, we next analyzed the association of C2 expression with clinical-pathological parameters. As is shown in Figure 2, in the TCGA cohort, C2 expression was remarkably correlated with TNM stages. HCC patients who were in more advanced TNM stages expressed lower C2 expression ($p = 0.02$, Figure 2(a)). Besides, HCC patients who were in more decreased differentiation tended to express

TABLE 1: Basic characteristics of HCC patients from the TCGA, GSE14520, and ICGC HCC cohorts.

| Variables | TCGA cohort (N = 377) | GSE14520 cohort (N = 220) | ICGC cohort (N = 232) |
|--|-----------------------|---------------------------|-----------------------|
| Gender (male/female) | 255/122 | 190/30 | 171/61 |
| Age (years, median) | 60 (16-90) | 50 (21-77) | 69 (31-89) |
| Cirrhosis (yes/no/NA) | 81/137/159 | 202/18 | NA |
| Histologic grade (G1/G2/G3/G4/NA) | 55/180/124/13/5 | NA | NA |
| TNM stage (I/II/III/IV/NA) | 175/87/86/5/24 | 93/77/48/-/2 | 36/106/71/76 |
| Main tumor size (>/≤5 cm) | NA | 80/140 | NA |
| Overall survival status (alive/dead) | 245/132 | 136/84 | 189/43 |
| Overall survival time (months, median) | 19.67 (0-120.73) | 51.65 (1.8-67.4) | 26.0 (0.3-72) |

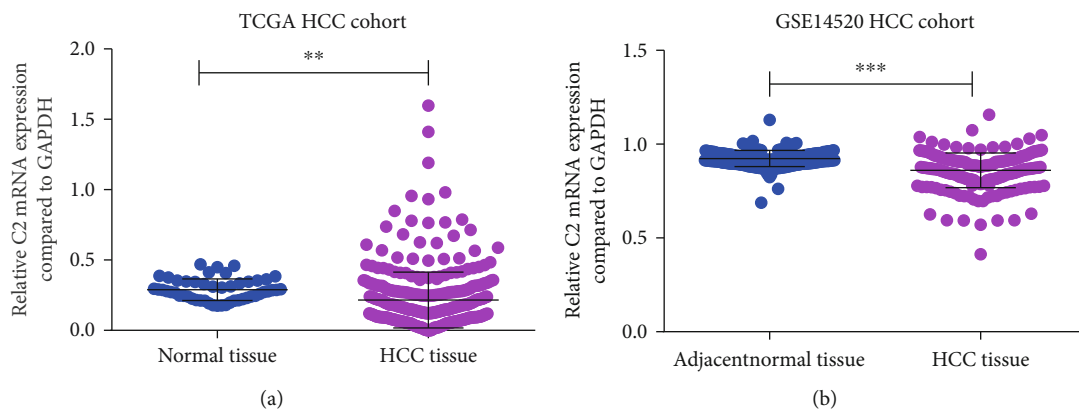


FIGURE 1: C2 expression between HCC patients and healthy controls/adjacent normal tissues. Expression of C2 in HCC patients and healthy controls of the TCGA cohort (a); expression of C2 in HCC patients and adjacent normal tissues of the GSE14520 cohort (b).

lower C2 expression too, but the difference was not significant ($p = 0.18$, Figure 2(b)). Similar results were also found at ICGC and GSE14520 HCC cohorts. In the ICGC cohort, HCC patients who were in more advanced TNM stages expressed lower C2 expression ($p = 0.004$, Figure 2(c)). In the GSE14520 cohort, HCC patients who were in more advanced TNM stages expressed lower C2 expression ($p < 0.0001$, Figure 2(d)). Moreover, C2 expression of HCC patients with tumor size ≥ 5 cm was lower than that of HCC patients with tumor size ≤ 5 cm ($p = 0.002$, Figure 2(e)). Taken together, C2 expression was significantly associated with TNM stages in HCC patients, and the higher the TNM stages, the lower the C2 expression.

3.3. Prognostic Value of C2 in HCC Patients. Having found that C2 expression was significantly associated with clinical-pathological parameters in HCC patients, we next analyzed the prognostic value of C2 in HCC patients. As is shown in Figure 3, C2 expression was significantly associated with the prognosis of HCC patients. In the TCGA cohort, univariate Cox analysis showed that higher C2 expression was significantly associated with better OS of HCC patients (HR = 0.66, 95% CI: 0.47-0.93, $p = 0.02$, Figure 3(a)). Moreover, multivariate analysis also showed that C2 expression was independently associated with OS of HCC patients after adjusting for gender, age, cirrhosis, histologic grade, and TNM stage (HR = 0.68, 95% CI: 0.47-0.99, $p = 0.04$, Table 2). In the GSE14520 cohort, univariate Cox indicated

that higher C2 expression was also significantly associated with better OS of HCC patients (HR = 0.56, 95% CI: 0.36-0.87, $p = 0.01$, Figure 3(b)), but multivariate analysis showed that C2 expression was not independently associated with OS of HCC patients after adjusting for gender, age, cirrhosis, main tumor size, and TNM stage (HR = 0.73, 95% CI: 0.47-1.16, $p = 0.18$, Table 3). In the ICGC cohort, univariate Cox suggested that higher C2 expression was significantly related with better OS (HR = 0.46, 95% CI: 0.25-0.86, $p = 0.02$, Figure 3(c)), and multivariate analysis also showed that C2 expression was independently associated with OS of HCC patients after adjusting for gender, age, and TNM stage (HR = 0.73, 95% CI: 0.27-0.96, $p = 0.04$, Table 4). In a word, C2 may play an important role in HCC suppression.

3.4. Correlation of C2 Expression with Tumor-Infiltrating Immune Cells in HCC Patients. Increasing studies had proved an immunoregulatory effect of complement on TME, and interactions between complement and tumor-infiltrating immune cells contribute to the development and progression of many kinds of cancers [8]. So, we further analyzed the correlation of C2 expression with tumor-infiltrating immune cells in HCC patients. As is shown in Figure 4, C2 expression was significantly associated with tumor-infiltrating immune cells in HCC patients. In the TCGA cohort, significantly higher proportions of resting CD4 memory T cells and macrophage M1 cells were found in HCC patients with higher C2 expression, while significantly higher proportions

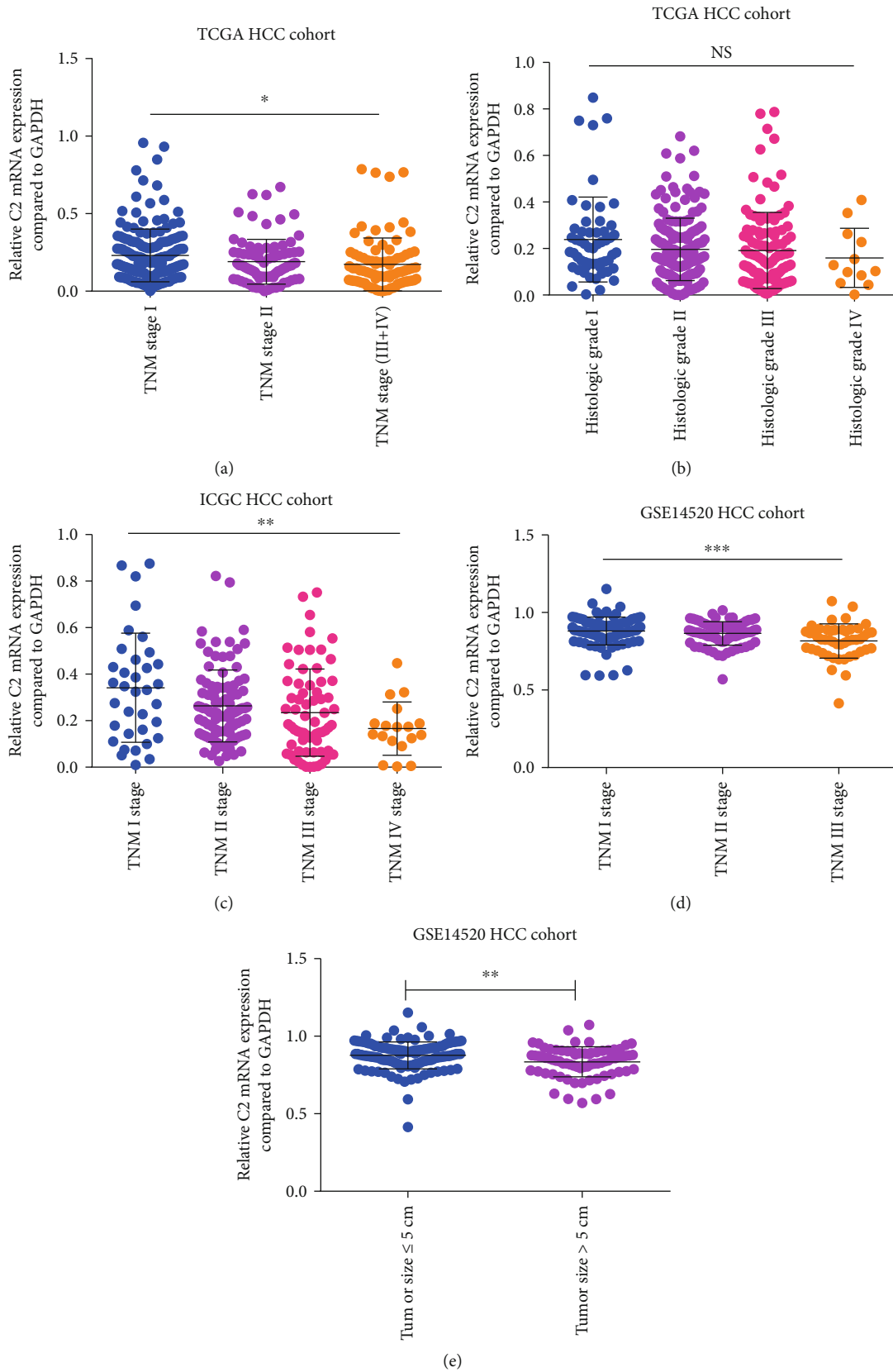


FIGURE 2: Association of C2 expression of HCC patients with clinical-pathological parameters. Association of C2 expression with TNM stages and histologic grades of TCGA cohort (a, b); association of C2 expression with TNM stages of ICGC cohort (c); association of C2 expression with TNM stages and main tumor size of cohort (d, e).

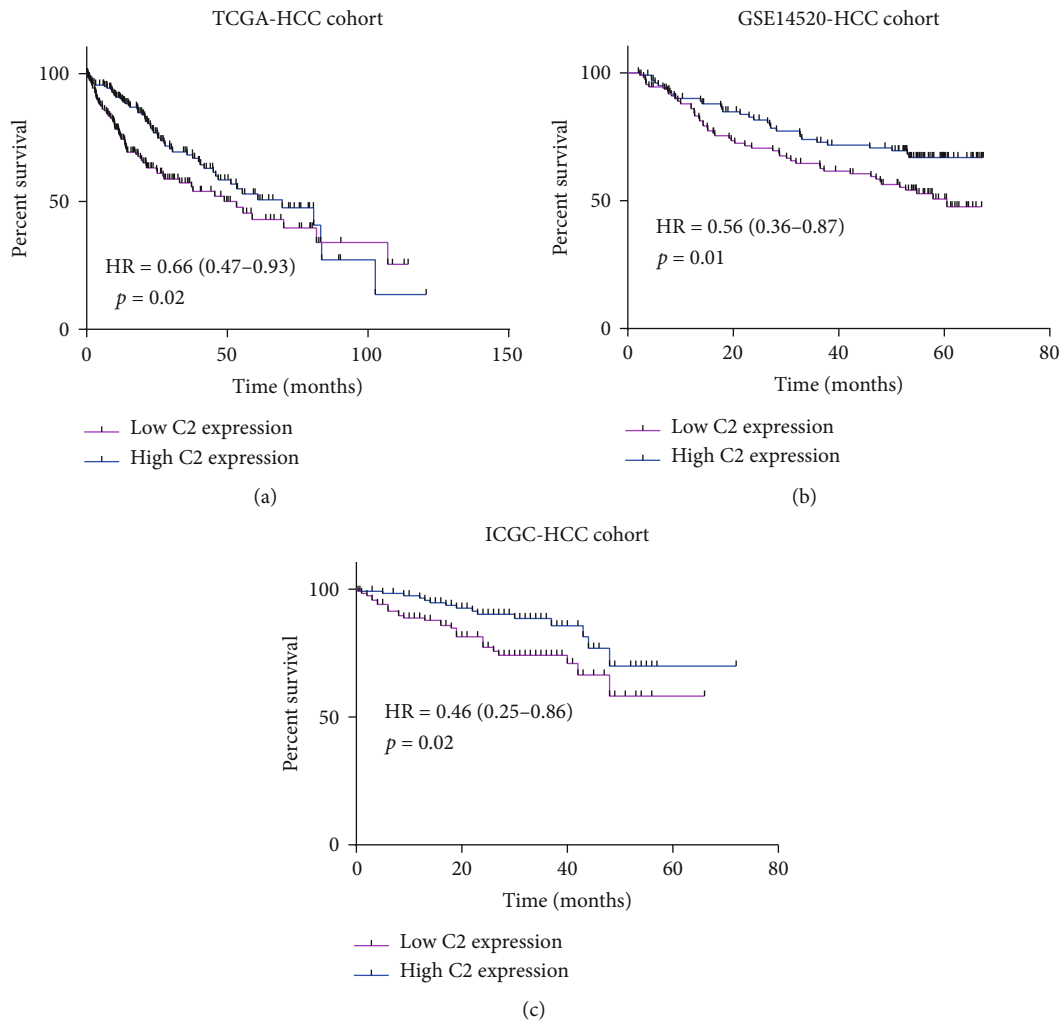


FIGURE 3: Prognostic value of C2 in HCC patients. Kaplan-Meier analysis of overall survival time of HCC patients with high C2 and low C2 expression in the TCGA cohort (a); Kaplan-Meier analysis of overall survival time of HCC patients with high C2 and low C2 expression in the GSE14520 cohort (b); Kaplan-Meier analysis of overall survival time of HCC patients with high C2 and low C2 expression in the ICGC cohort (c).

TABLE 2: Univariate and multivariate analyses of C2 expression for overall survival in HCC patients of the TCGA cohort.

| Variables | Univariate analysis | | | Multivariate analysis | | |
|------------------------------------|---------------------|-----------|---------|-----------------------|-----------|---------|
| | Hazard ratio | 95% CI | p value | Hazard ratio | 95% CI | p value |
| Gender (male vs. female) | 0.82 | 0.57-1.16 | 0.26 | | | |
| Age (>60 vs. ≤60) | 1.25 | 0.88-1.77 | 0.21 | | | |
| Cirrhosis (yes vs. no) | 0.83 | 0.48-1.42 | 0.49 | | | |
| Histologic grade (G3+G4 vs. G1+G2) | 1.12 | 0.78-1.61 | 0.54 | | | |
| TNM stage (III+IV vs. I+II) | 2.43 | 1.69-3.55 | 0.000* | 2.38 | 1.64-3.46 | 0.000* |
| C2 (high vs. low) | 0.66 | 0.47-0.93 | 0.02* | 0.68 | 0.47-0.99 | 0.04* |

of macrophage M0 cells, activated CD4 memory cells, and plasma cells were found at HCC patients with lower C2 expression (all $p < 0.05$, Figure 4(a)). In the GSE14520 cohort, significantly higher proportions of resting mast cells, follicular helper T cells, and resting CD4 memory T cells were found in HCC patients with higher C2 expression, while significantly higher proportions of macrophage M0 cells, acti-

vated mast cells, and plasma cells were found at HCC patients with lower C2 expression (all $p < 0.05$, Figure 4(b)). In the ICGC cohort, significantly higher proportions of resting CD4 memory T cells were found in HCC patients with higher C2 expression, while significantly higher proportions of macrophage M0 cells, activated CD4 memory T cells, and Treg cells were found at HCC patients with lower C2

TABLE 3: Univariate and multivariate analysis of C2 expression for overall survival in HCC patients of the GSE14520 cohort.

| Variables | Univariate analysis | | | Multivariate analysis | | |
|-----------------------------------|---------------------|------------|----------------|-----------------------|------------|----------------|
| | Hazard ratio | 95% CI | <i>p</i> value | Hazard ratio | 95% CI | <i>p</i> value |
| Gender (male vs. female) | 1.68 | 0.81-3.49 | 0.16 | | | |
| Age (>50 vs. ≤50) | 1.0 | 0.65-1.53 | 0.99 | | | |
| Cirrhosis (yes vs. no) | 4.57 | 1.13-18.62 | 0.03* | 3.66 | 0.89-14.98 | 0.07 |
| Main tumor size (>5 cm vs. ≤5 cm) | 1.97 | 1.28-3.30 | 0.002* | 1.18 | 0.69-2.01 | 0.56 |
| TNM stage (III+IV vs. I+II) | 3.43 | 2.20-6.11 | 0.000* | 2.38 | 1.64-3.46 | 0.000* |
| C2 (high vs. low) | 0.56 | 0.36-0.87 | 0.01* | 0.73 | 0.47-1.16 | 0.18 |

TABLE 4: Univariate and multivariate analyses of C2 expression for overall survival in HCC patients of the ICGC cohort.

| Variables | Univariate analysis | | | Multivariate analysis | | |
|-----------------------------|---------------------|-----------|----------------|-----------------------|-----------|----------------|
| | Hazard ratio | 95% CI | <i>p</i> value | Hazard ratio | 95% CI | <i>p</i> value |
| Gender (male vs. female) | 0.51 | 0.28-0.97 | 0.04* | 0.43 | 0.23-0.81 | 0.01* |
| Age (>70 vs. ≤70) | 1.06 | 0.58-1.94 | 0.84 | | | |
| TNM stage (III+IV vs. I+II) | 2.38 | 1.30-4.36 | 0.004* | 2.59 | 1.39-4.80 | 0.003* |
| C2 (high vs. low) | 0.46 | 0.25-0.86 | 0.02* | 0.73 | 0.27-0.96 | 0.04* |

expression (all $p < 0.05$, Figure 4(c)). Taken together, C2 expression was significantly associated with CD4 memory T cells and macrophage M0 cells in HCC patients from all these three HCC cohorts.

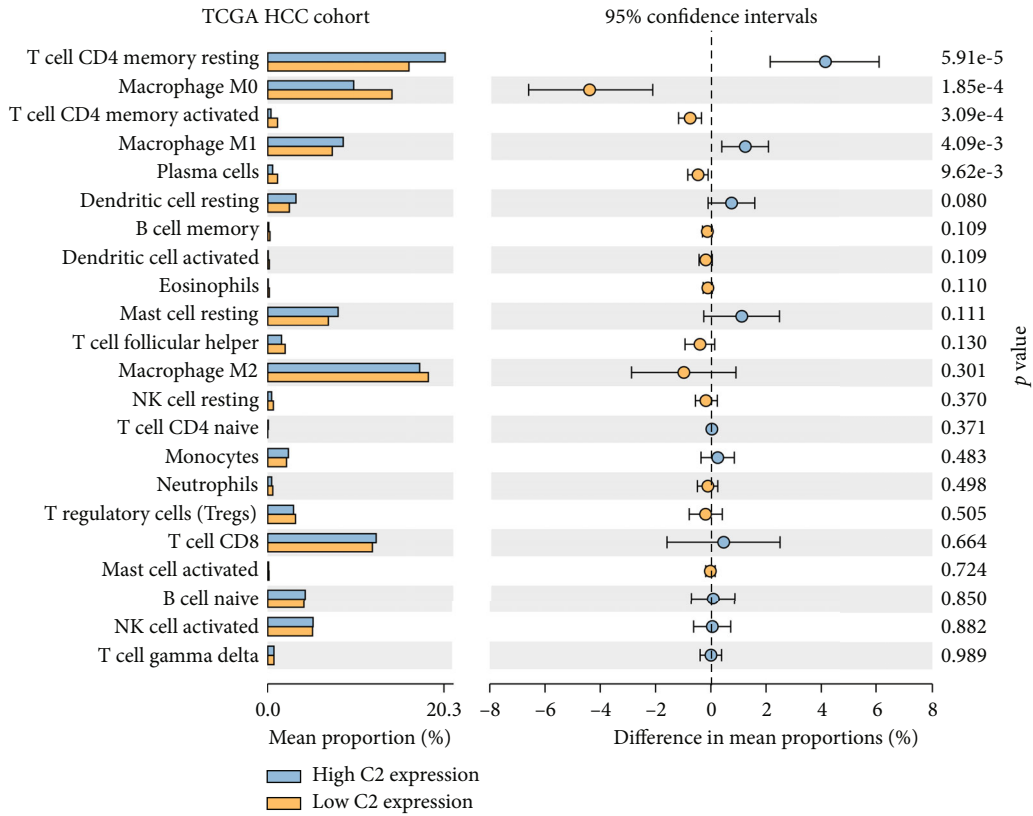
3.5. Molecular Mechanisms of C2 in HCC. KEGG analysis was performed to explore the underlying biological mechanism by which C2 influenced the prognosis of HCC. As is shown in Figure 5, in the TCGA cohort, KEGG pathways, such as “cell cycle,” “ubiquitin-mediated proteolysis,” “complement and coagulation cascades,” “spliceosome,” and “RNA transport,” were most significantly enriched in HCC patients with higher C2 expression compared to HCC patients with lower C2 expression (all $p < 0.001$, Figure 5(a)). In the GSE14520 cohort, KEGG pathways, such as “complement and coagulation cascades,” “carbon metabolism,” “biosynthesis of amino acids,” “peroxisome,” and “ribosome,” were most significantly enriched in HCC patients with higher C2 expression (all $p < 0.001$, Figure 5(b)). In the ICGC cohort, KEGG pathways, such as “ribosome,” “cell cycle,” “complement and coagulation cascades,” “spliceosome,” and “RNA transport,” were most significantly enriched in HCC patients with higher C2 expression (all $p < 0.001$, Figure 5(c)). In total, as is shown in Table 5, 17 significant KEGG pathways, such as “cell cycle,” “complement and coagulation cascades,” “AMPK signaling pathway,” and “PPAR signaling pathway,” overlapped in the three HCC cohorts, indicating that C2 may influence the prognosis of HCC by regulation of these 17 overlapped KEGG pathways.

4. Discussion

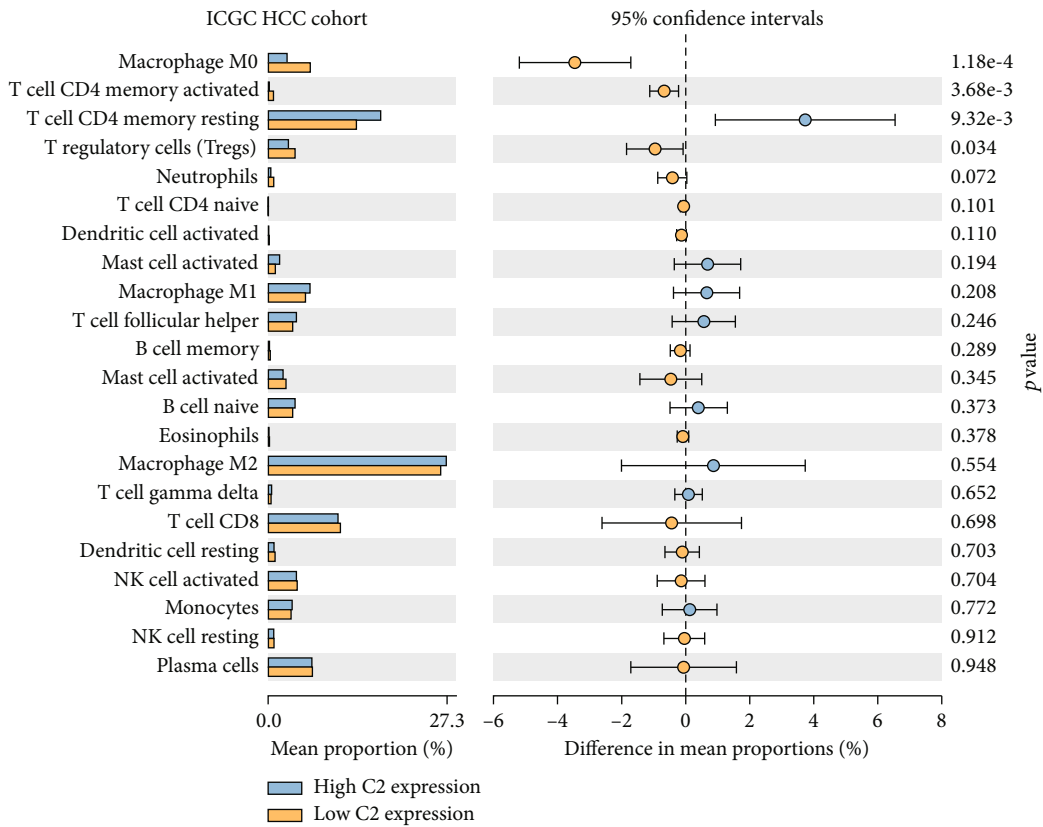
HCC is one of the most malignant kinds of cancer worldwide, and TME has been found to play important roles of

recurrence and resistance to therapy of HCC [8, 15]. Increasing studies have proved an immunoregulatory effect of complement on TME, and interactions between complement and TME contribute to the development and progression of many kinds of cancers [8]. C2 is an important part of the complement system, and SNP of C2 has been found to be significantly associated with HCC [12, 13]. In the present study, we found that C2 was associated with the prognosis of HCC and explored the relevant underlying mechanism.

Previously, Imamura et al. have found that higher expression of the C5a receptor from breast cancer tissues is significantly associated with clinical parameters, such as larger tumor size, advanced histologic grade, lymph node metastasis, higher TNM stages, and poorer prognosis [16]. Lin et al. have also observed that higher tissue C3 expression is significantly associated with better prognosis of non-small-cell lung cancer (NSCLC) patients, indicating an important role played by C3 in NSCLC suppression [17]. Similarly, in our study, lower expression of C2 was found at HCC patients, and C2 expression was significantly associated with TNM stages and better OS. Besides, C2 expression was independently associated with OS of HCC patients in TCGA and ICGC HCC cohorts, but it was not in the GSE14520 cohort. The reason for the difference in the predictive value of C2 expression in different HCC cohorts may be that some pathological parameters among the three HCC cohorts were different. For example, in the TCGA HCC cohort, cirrhosis was not associated with OS, but in the GSE14520 HCC cohort, cirrhosis and main tumor size were significantly associated with OS. Moreover, in the TCGA and GSE1420 HCC cohorts, gender was not associated with OS, but in the ICGC HCC cohort, gender was associated with OS. The difference of cirrhosis, main tumor size, and gender among the three HCC cohorts may result in the difference of factors used



(a)



(b)

FIGURE 4: Continued.

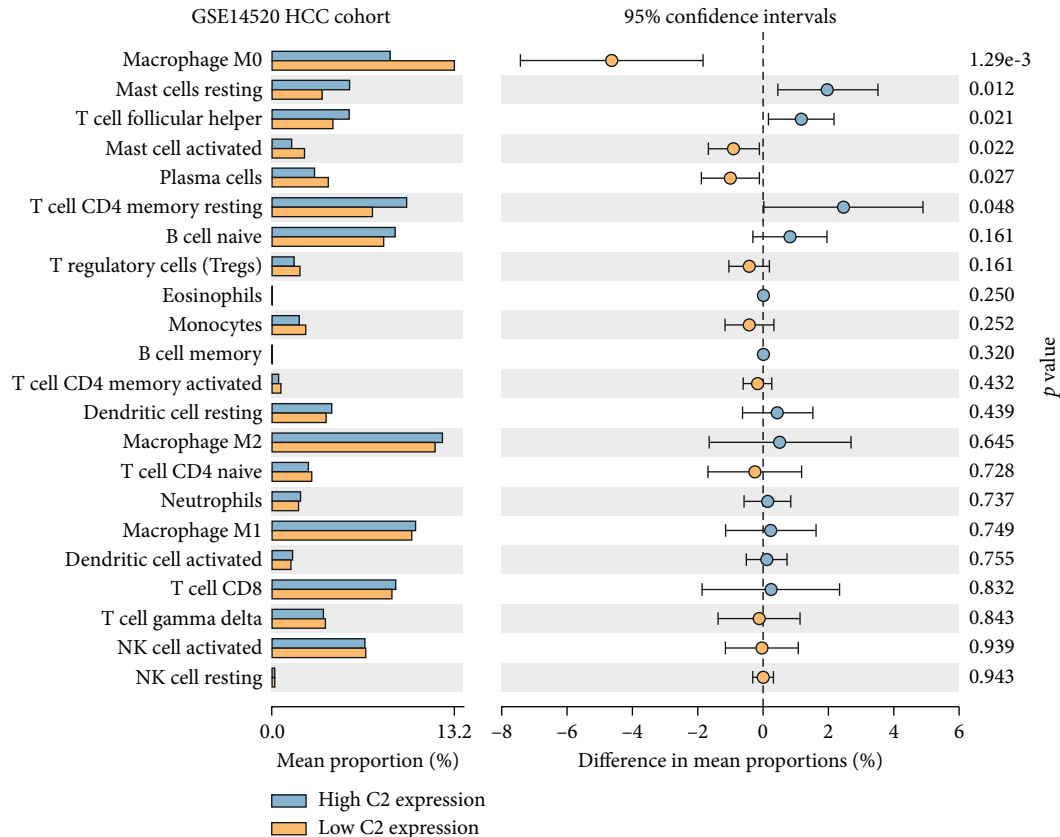


FIGURE 4: Correlation of C2 expression with tumor-infiltrating immune cells in HCC patients of the TCGA cohort (a), GSE14520 cohort (b), and ICGC cohort (c).

for multivariate analysis, thus leading to the difference of the predictive value of C2 expression for OS. In a word, C2 expression was associated with OS of HCC patients, but it still needed external and multicenter prospective cohorts with large sample sizes to validate whether C2 expression could be an independent prognostic factor for OS.

A series of studies have proved that the complement system takes part in the regulation of TME, and interaction of complement with tumor-infiltrating immune cells plays an important role in the development and progression of many kinds of cancers [8]. Tumor cell-derived C5a can recruit and differentiate myeloid-derived suppressor cells (MDSCs) in TME, which could protect tumor cells against the immune system and immunotherapy and promote tumor progression by inhibiting T cell responses and promoting the generation of Tregs [18]. Moreover, blockade of C5aR can reduce the number of MDSCs [19]. Lin et al. have also showed that higher tissue C3 expression was positively correlated with higher numbers of tumor-infiltrating CD4 T cells and CD8 T cells, which may also contribute to tumor suppression by C3, as higher C3 expression predicts better prognosis in NSCLC patients [17]. In the present study, elevated CD4 T cells were found at HCC patients with higher C2 expression while a higher proportion of macrophage M0 cells was found in HCC patients with lower C2 expression in all three HCC cohorts. Garnelo et al. have showed that the degree of infil-

trated T cells and B cells of tumor tissues significantly relates to the improved prognosis of HCC patients [20]. CD4+ T cells can improve antitumor immune responses by producing cytokines which are important for the activation of CD8+ T cells and B cells. Imai et al. have found that CD4+ T cells are important for the formation and maintenance of polyfunctionality of cytotoxic CD8+ T cells, which is a key determinant of the success of immunological control of tumor growth [21]. Moreover, the important role of virus control and antitumor immunity played by CD4+ T cell-mediated cytotoxicity is being increasingly recognized. Fu et al. have found that circulating and tumor-infiltrating CD4+ cytotoxic T cells decrease in HCC patients with advanced stages, and loss of CD4+ cytotoxic T cells is associated with a high mortality rate and reduced survival time [22]. Hsiao et al. have observed that the higher number of macrophage M0 cells is significantly associated with poorer prognosis of HCC patients [23]. Macrophages can be recruited into HCC tissue to become tumor-associated macrophages (TAMs) by upregulation of HMGB1 and then take part in the cancer progression and metastasis [24]. TAMs locate in the stroma of HCC tissue and are polarized toward the M2 phenotype. A lot of studies have showed that TAMs can promote tumor proliferation, angiogenesis, invasion, and metastasis [25]. For example, Yeung et al. have showed that M2 macrophages are associated with a poor prognosis

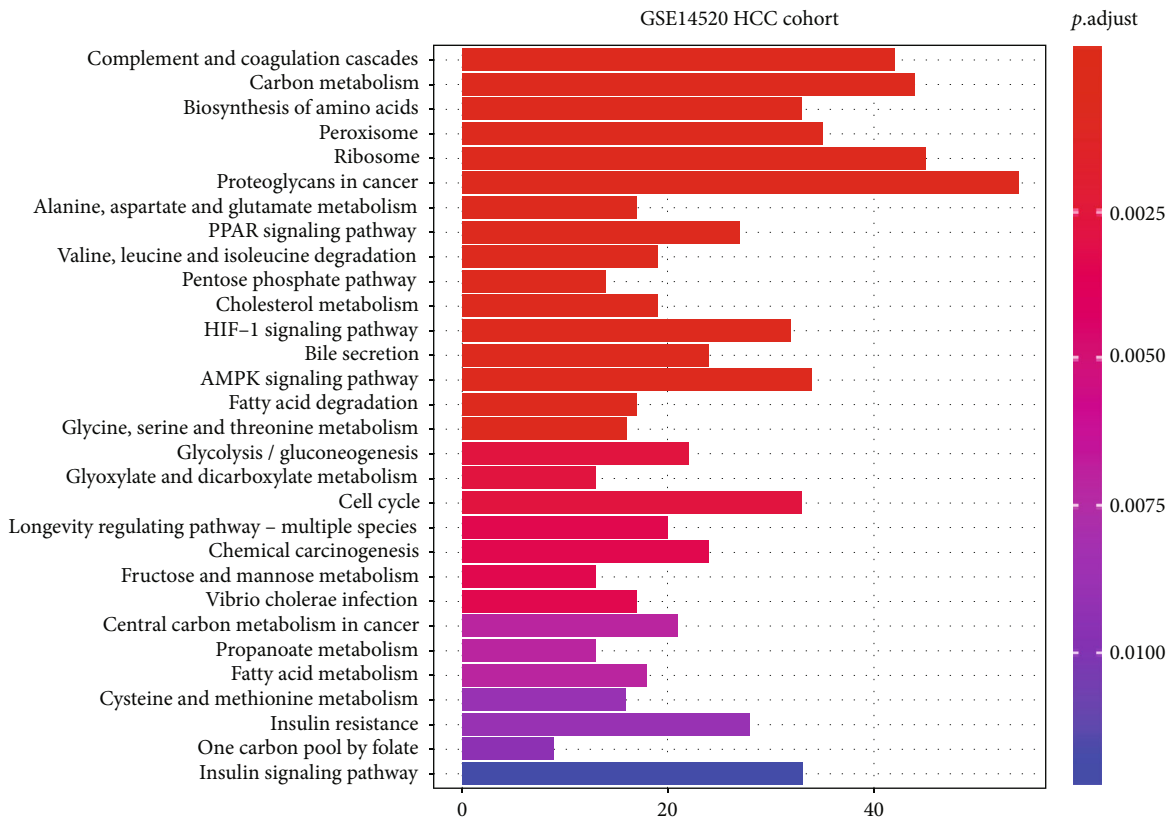
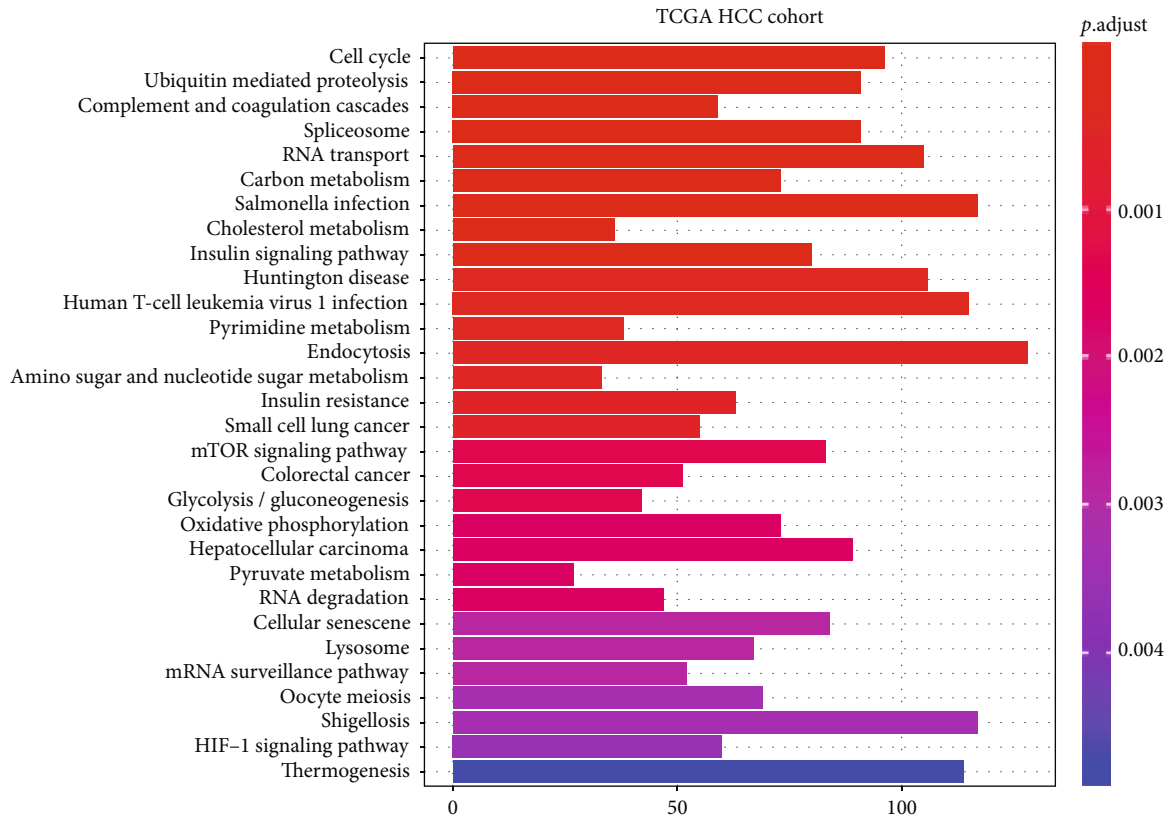


FIGURE 5: Continued.

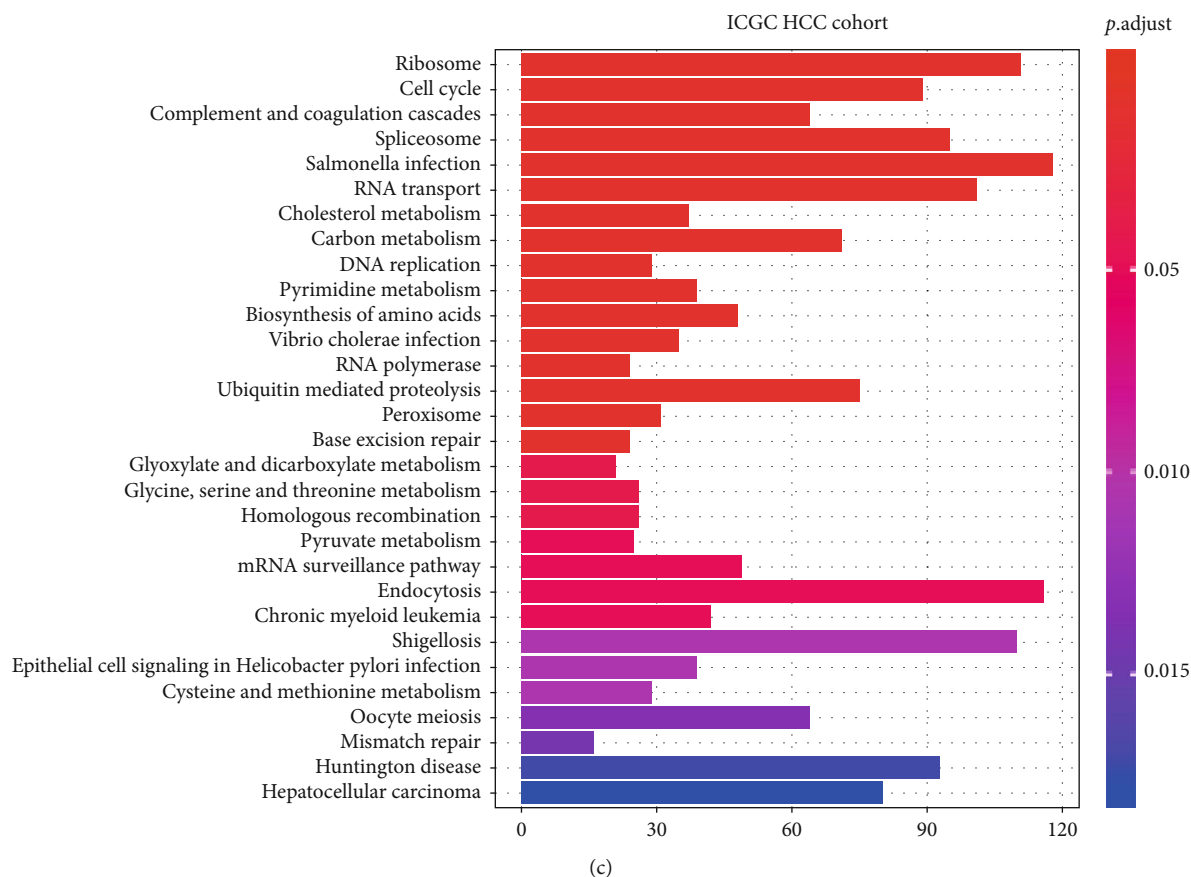


FIGURE 5: The most significant 30 KEGG pathways enriched in HCC patients with higher C2 expression compared to HCC patients with lower C2 expression of the TCGA cohort (a), GSE14520 cohort (b), and ICGC cohort (c).

TABLE 5: Overlap of significant KEGG pathways enriched in HCC patients with higher C2 expression from the TCGA, GSE14520, and ICGC HCC cohorts.

| ID | Description | TCGA cohort | | GSE14520 cohort | | ICGC cohort | |
|----------|---|-------------|-----------------|-----------------|-----------------|-------------|-----------------|
| | | Count | <i>p.adjust</i> | Count | <i>p.adjust</i> | Count | <i>p.adjust</i> |
| hsa04110 | Cell cycle | 96 | 0.000 | 33 | 0.003 | 89 | 0.000 |
| hsa04610 | Complement and coagulation cascades | 59 | 0.000 | 42 | 0.000 | 64 | 0.000 |
| hsa04152 | AMPK signaling pathway | 61 | 0.023 | 34 | 0.001 | 58 | 0.031 |
| hsa03320 | PPAR signaling pathway | 42 | 0.014 | 27 | 0.000 | 40 | 0.025 |
| hsa04910 | Insulin signaling pathway | 80 | 0.000 | 28 | 0.008 | 66 | 0.026 |
| hsa01200 | Carbon metabolism | 73 | 0.000 | 44 | 0.000 | 71 | 0.000 |
| hsa00630 | Glyoxylate and dicarboxylate metabolism | 20 | 0.012 | 13 | 0.002 | 21 | 0.004 |
| hsa00620 | Pyruvate metabolism | 27 | 0.001 | 13 | 0.017 | 25 | 0.006 |
| hsa01230 | Biosynthesis of amino acids | 42 | 0.011 | 33 | 0.000 | 48 | 0.000 |
| hsa00410 | Beta-alanine metabolism | 19 | 0.039 | 11 | 0.021 | 19 | 0.030 |
| hsa00270 | Cysteine and methionine metabolism | 29 | 0.017 | 16 | 0.008 | 29 | 0.012 |
| hsa05110 | Vibrio cholerae infection | 31 | 0.006 | 17 | 0.004 | 35 | 0.000 |
| hsa00010 | Glycolysis/gluconeogenesis | 42 | 0.001 | 22 | 0.002 | 36 | 0.029 |
| hsa04979 | Cholesterol metabolism | 36 | 0.000 | 19 | 0.001 | 37 | 0.000 |
| hsa00280 | Valine, leucine, and isoleucine degradation | 27 | 0.042 | 19 | 0.000 | 27 | 0.029 |
| hsa00640 | Propanoate metabolism | 21 | 0.027 | 54 | 0.000 | 20 | 0.041 |
| hsa00260 | Glycine, serine, and threonine metabolism | 23 | 0.048 | 16 | 0.001 | 26 | 0.004 |

in HCC by promoting tumor growth and invasiveness through CCL22-induced epithelial-mesenchymal transition (EMT) [26]. Besides, tumor-suppressive M1 macrophages were found to be enriched in HCC patients with higher C2 expression in the TCGA cohort, which may contribute to better prognosis of HCC patients as studies have showed that M1 macrophages are involved in killing pathogens and tumor cells by producing large amounts of proinflammatory cytokines and expressing MHC molecules [27]. Moreover, the higher proportion of Tregs was found at HCC patients with lower C2 expression in the ICGC cohort, which may also contribute to the unfavorable prognosis of HCC patients as Treg cells are immunosuppressive cells and could promote the occurrence and development of HCC by inhibiting the function of T cells [28]. Taken together, the interaction of C2 with tumor-infiltrating immune cells may influence the prognosis of HCC.

In addition, to analyze the association of C2 expression with tumor-infiltrating immune cells, we also explore the underlying signaling pathways exploited by C2 to influence prognosis by KEGG analysis. Signaling pathways, such as “AMPK signaling pathway” and “PPAR signaling pathway,” were found to be significantly enriched in HCC patients with higher C2 expression from all three HCC cohorts, which have been reported to take part in the development and progression of HCC. For example, Han et al. have observed that hispidulin could suppress the growth and metastasis of HCC through AMPK signaling-mediated PPAR γ activation both in vitro and in vivo [29]. Tuo et al. have also found that phosphoenolpyruvate carboxykinase 1 (PCK1), one of the key enzymes of gluconeogenesis, could inhibit the progression of cell cycle and proliferation of hepatoma cells via the AMPK/p27^{Kip1} axis [30]. Basing on these studies, we speculated that C2 may influence the prognosis of HCC by regulation of the AMPK signaling pathway and/or PPAR signaling pathway.

Several limitations of the present study should also be noted. First, complement is mainly synthesized in the liver and then secreted into blood, but we do not analyze the difference of serum C2 between HCC and healthy controls and the prognostic value of serum C2 in HCC. Second, C2 expression is found to be associated with tumor-infiltrating cells, such as CD4 T cells and macrophage M0 cells, but we do not explore how C2 regulate these immune cells to influence the prognosis of HCC patients. Finally, we do not validate the KEGG pathways enriched in HCC patients with higher C2 expression in in vitro studies.

In conclusion, higher C2 expression is associated with better prognosis of HCC, and C2 may influence the prognosis of HCC by interaction with CD4 T cells and macrophage M0 cells and regulation of pathways, such as the AMPK signaling pathway and PPAR signaling pathway.

Abbreviations

| | |
|------|--------------------------|
| HCC: | Hepatocellular carcinoma |
| TME: | Tumor microenvironment |
| C7: | Complement component 7 |
| CFH: | Complement factor H |
| PDT: | Photodynamic therapy |

| | |
|--------|--|
| C5a: | Complement component 5a |
| C2: | Complement component 2 |
| SNP: | Single nucleotide polymorphism |
| TCGA: | The Cancer Genome Atlas |
| GEO: | Gene Expression Omnibus |
| ICGC: | International Cancer Genome Consortium |
| OS: | Overall survival |
| KEGG: | Kyoto Encyclopedia of Genes and Genomes |
| DAVID: | Database for Annotation, Visualization, and Integrated Discovery |
| DEGs: | Differentially expressed genes |
| NSCLC: | Non-small-cell lung cancer |
| MDSCs: | Myeloid-derived suppressor cells |
| EMT: | Epithelial-mesenchymal transition |
| MBL: | Mannan binding lectin |
| PCK1: | Phosphoenolpyruvate carboxykinase 1. |

Data Availability

All the data used in the present study could be downloaded from TCGA (<https://cancergenome.nih.gov/>), GEO (<https://www.ncbi.nlm.nih.gov/geo/>), and ICGC portal (<https://dcc.icgc.org/projects/LIRI-JP>).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Gang Ning, Yan-Lin Huang, and Li-Min Zhen have contributed equally to this work.

Acknowledgments

This study was supported by grants from the Natural Science Foundation of China (Nos. 81472259, 81570539, and 81873572), National Major Science and Technology Project for the Prevention and Treatment of AIDS and Viral Hepatitis (2018ZX10302205-002 and 2018ZX10302204), Natural Science Foundation of Guangdong Province (2014A030313042), Guangzhou Science and Technology Plan Project (201904010442), Guangdong Provincial Science and Technology Projects (2015A020212007), Sun Yat-sen University Clinical Research 5010 Program (2018009), Research Project on Degree and Postgraduate Education Reform in Guangdong Province (2018JGXM04), Young Teacher Training Program of Sun Yat-sen University (16ykpy40), Open Project of Key Lab of Tropical Disease Control (Sun Yat-sen University), Ministry of Education (2019kfkt07), and Ten Five-Year Plan Fund of the Third Affiliated Hospital of Sun Yat-sen University (SW201701).


References

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2019,” *CA: a cancer journal for clinicians*, vol. 69, no. 1, pp. 7–34, 2018.
- [2] P. Tabrizian, G. Jibara, B. Shrager, M. Schwartz, and S. Roayaie, “Recurrence of hepatocellular cancer after

- resection: patterns, treatments, and prognosis," *Annals of surgery.*, vol. 261, no. 5, pp. 947–955, 2015.
- [3] D. F. Quail and J. A. Joyce, "Microenvironmental regulation of tumor progression and metastasis," *Nature medicine.*, vol. 19, no. 11, pp. 1423–1437, 2013.
- [4] D. Hanahan and L. M. Coussens, "Accessories to the crime: functions of cells recruited to the tumor microenvironment," *Cancer Cell*, vol. 21, no. 3, pp. 309–322, 2012.
- [5] M. V. Novikova, N. V. Khromova, and P. B. Kopnin, "Components of the hepatocellular carcinoma microenvironment and their role in tumor progression," *Biochemistry Biokhimiia.*, vol. 82, no. 8, pp. 861–873, 2017.
- [6] D. Moris, E. W. Beal, J. Chakedis et al., "Role of exosomes in treatment of hepatocellular carcinoma," *Surgical oncology.*, vol. 26, no. 3, pp. 219–228, 2017.
- [7] D. Ricklin, G. Hajishengallis, K. Yang, and J. D. Lambris, "Complement: a key system for immune surveillance and homeostasis," *Nature Immunology*, vol. 11, no. 9, pp. 785–797, 2010.
- [8] E. S. Reis, D. C. Mastellos, D. Ricklin, A. Mantovani, and J. D. Lambris, "Complement in cancer: untangling an intricate relationship," *Nature reviews Immunology.*, vol. 18, no. 1, pp. 5–18, 2018.
- [9] H. S. Seol, S. E. Lee, J. S. Song et al., "Complement proteins C7 and CFH control the stemness of liver cancer cells via LSF-1," *Cancer letters.*, vol. 372, no. 1, pp. 24–35, 2016.
- [10] F. Li, Y. Cheng, J. Lu, R. Hu, Q. Wan, and H. Feng, "Photodynamic therapy boosts anti-glioma immunity in mice: a dependence on the activities of T cells and complement C3," *Journal of Cellular Biochemistry*, vol. 112, no. 10, pp. 3035–3043, 2011.
- [11] L. Gunn, C. Ding, M. Liu et al., "Opposing roles for complement component C5a in tumor progression and the tumor microenvironment," *Journal of Immunology*, vol. 189, no. 6, pp. 2985–2994, 2012.
- [12] S. Namgoong, J. G. Shin, H. S. Cheong et al., "Genetic association of complement component 2 variants with chronic hepatitis B in a Korean population," *Liver international: official journal of the International Association for the Study of the Liver.*, vol. 38, no. 9, pp. 1576–1582, 2018.
- [13] R. J. Clifford, J. Zhang, D. M. Meerzaman et al., "Genetic variations at loci involved in the immune response are risk factors for hepatocellular carcinoma," *Hepatology*, vol. 52, no. 6, pp. 2034–2043, 2010.
- [14] D. H. Parks, G. W. Tyson, P. Hugenholtz, and R. G. Beiko, "STAMP: statistical analysis of taxonomic and functional profiles," *Bioinformatics*, vol. 30, no. 21, pp. 3123–3124, 2014.
- [15] Z. Yin, C. Dong, K. Jiang et al., "Heterogeneity of cancer-associated fibroblasts and roles in the progression, prognosis, and therapy of hepatocellular carcinoma," *Journal of hematology & oncology.*, vol. 12, no. 1, p. 101, 2019.
- [16] T. Imamura, M. Yamamoto-Ibusuki, A. Sueta et al., "Influence of the C5a-C5a receptor system on breast cancer progression and patient prognosis," *Breast Cancer*, vol. 23, no. 6, pp. 876–885, 2016.
- [17] K. Lin, S. He, L. He et al., "Complement component 3 is a prognostic factor of nonsmall cell lung cancer," *Molecular medicine reports.*, vol. 10, no. 2, pp. 811–817, 2014.
- [18] R. J. Tesi, "MDSC; the most important cell you have never heard of," *Trends in pharmacological sciences.*, vol. 40, no. 1, pp. 4–7, 2019.
- [19] L. Corrales, D. Ajona, S. Rafail et al., "Anaphylatoxin C5a creates a favorable microenvironment for lung cancer progression," *Journal of Immunology*, vol. 189, no. 9, pp. 4674–4683, 2012.
- [20] M. Garnelo, A. Tan, Z. Her et al., "Interaction between tumour-infiltrating B cells and T cells controls the progression of hepatocellular carcinoma," *Gut*, vol. 66, no. 2, pp. 342–351, 2017.
- [21] N. Imai, I. Tawara, M. Yamane, D. Muraoka, H. Shiku, and H. Ikeda, "CD4⁺ T cells support polyfunctionality of cytotoxic CD8⁺ T cells with memory potential in immunological control of tumor," *Cancer Science*, 2020.
- [22] J. Fu, Z. Zhang, L. Zhou et al., "Impairment of CD4⁺ cytotoxic T cells predicts poor survival and high recurrence rates in patients with hepatocellular carcinoma," *Hepatology*, vol. 58, no. 1, pp. 139–149, 2013.
- [23] Y.-W. Hsiao, L.-T. Chiu, C.-H. Chen, W.-L. Shih, and T.-P. Lu, "Tumor-infiltrating leukocyte composition and prognostic power in hepatitis B- and hepatitis C-related hepatocellular carcinomas," *Genes*, vol. 10, no. 8, p. 630, 2019.
- [24] R. Ostuni, F. Kratochvill, P. J. Murray, and G. Natoli, "Macrophages and cancer: from mechanisms to therapeutic implications," *Trends in immunology.*, vol. 36, no. 4, pp. 229–239, 2015.
- [25] Z. Tian, X. Hou, W. Liu, Z. Han, and L. Wei, "Macrophages and hepatocellular carcinoma," *Cell & Bioscience*, vol. 9, no. 1, 2019.
- [26] O. W. H. Yeung, C.-M. Lo, C.-C. Ling et al., "Alternatively activated (M2) macrophages promote tumour growth and invasiveness in hepatocellular carcinoma," *Journal of hepatology*, vol. 62, no. 3, pp. 607–616, 2015.
- [27] C. Lamagna, M. Aurrand-Lions, and B. A. Imhof, "Dual role of macrophages in tumor growth and angiogenesis," *Journal of leukocyte biology.*, vol. 80, no. 4, pp. 705–713, 2006.
- [28] H. Q. Zhao, W. M. Li, Z. Q. Lu, and Y. M. Yao, "Roles of Tregs in development of hepatocellular carcinoma: a meta-analysis," *World journal of gastroenterology.*, vol. 20, no. 24, pp. 7971–7978, 2014.
- [29] M. Han, H. Gao, P. Ju et al., "Hispidulin inhibits hepatocellular carcinoma growth and metastasis through AMPK and ERK signaling mediated activation of PPAR γ ," *Biomedicine & Pharmacotherapy*, vol. 103, pp. 272–283, 2018.
- [30] L. Tuo, J. Xiang, X. Pan et al., "PCK1 negatively regulates cell cycle progression and hepatoma cell proliferation via the AMPK/p27^{Kip1} axis," *Journal of experimental & clinical cancer research: CR.*, vol. 38, no. 1, p. 50, 2019.

Research Article

Profiles of Immune Infiltration and Prognostic Immunoscore in Lung Adenocarcinoma

Yanyan Li,¹ Liping Tao,² and Weiyang Cai^{1,2} 

¹Department of Ultrasound, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, Zhejiang Province, China

²Department of Gastroenterology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang Province, China

Correspondence should be addressed to Weiyang Cai; caiweiyang@sjtu.edu.cn

Received 26 March 2020; Accepted 14 May 2020; Published 9 June 2020

Guest Editor: Tao Huang

Copyright © 2020 Yanyan Li 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.

Lung tissue is abundant with immune cells that form a powerful first defense against exotic particles and microbes. The malignant phenotype of lung adenocarcinoma (LUAD) is defined not only by intrinsic tumor cells but also by tumor-infiltrating immune cells (TIICs) recruited to the immune microenvironment. Understanding more about the immune microenvironment of LUAD could function in sorting out patients more likely with high risk and benefit from immunotherapy. Twenty-two types of TIICs were estimated based on large public LUAD cohorts from the TCGA and GEO datasets using the CIBERSORT algorithm. Then principal component analysis (PCA), meta-analysis, and single-sample gene set enrichment analysis (ssGSEA) were used to measure and evaluate the specific immune responses and relative mechanisms. Moreover, an immunoscore model based on the percent of immune cells was constructed via the univariate and multivariate Cox regression models, which provided an in-depth overview of the LUAD immune microenvironment and shed light on the immune regulatory mechanism. The differential expression genes (DEGs) were acquired based on the immunoscore model, and prognostic immune-related genes were further identified. GSEA and the protein-protein interaction network (PPI) further revealed that these genes were mostly enriched in many immune-related biological processes. It is hoped that this immune landscape could provide a more accurate understanding for LUAD development and tumor immune therapy.

1. Introduction

Lung adenocarcinoma (LUAD), a dominating subtype of non-small-cell lung cancer, is acknowledged and regarded as one of the most leading cause of cancer-related mortality worldwide, with a five-year survival rate of only 22.1% [1]. The primary reason accounting for this disappointing prognosis is that the majority of LUAD patients are diagnosed at advanced stage III or IV and the uneven efficiency of chemotherapy selection [2]. The admitted first-line treatment for NSCLC without driver mutations is cytotoxic chemotherapy, immune checkpoint inhibitors, or a combination of both modalities.

The lung surrounding contacting with with the outer world is abundant with immune cells that construct a powerful defense against noxious particles and microbes. Various immune cells, IFN response, immune checkpoints, HLA,

cytokine, inflammation factor, and adoptive cell transplantation have been proven to make a great contribution in the progression of lung cancer. Immunotherapy, such as programmed cell death protein 1 (PD-1)/programmed death ligand 1 (PD-L1) checkpoint inhibitors and CTLA4, concentrates on revitalizing immunologic cells to release molecular components to defend against cancer cell in the tumor microenvironment. Two anti-PD-1 agents, pembrolizumab and nivolumab, and two anti-PD-L1 agents, atezolizumab and durvalumab, have been authorized by FDA for the treatment of lung tumor. It has been studied that pembrolizumab and nivolumab exhibited a surprising antitumor activity in advanced NSCLC with better overall survival (OS) and progression-free survival (PFS) than traditional second-line chemotherapy [3, 4]. Tumor PD-L1 expression is the only identified clinical biomarker to screen out patients most likely to respond to immunotherapy. Patients with high PD

– L1 expression [TPS] $\geq 50\%$ (tumor proportion score) and no EGFR or ALK genomic mutant are suitable for anti-PD-L1 agents. Disappointedly, the response rates for acceptable patients to PD-L1 have been variable. Only a proportion of high PD-L1 LUAD patients effectively respond to immunotherapy and gained a satisfactory clinical benefit [5], whereas some metastatic NSCLC patients given pembrolizumab were shown to have obvious longer PFS and OS regardless of PD-L1 expression [6]. More seriously, an increase of immunotherapy does take a leap of the probability of side effects without survival benefit in many cases. There are many reasons contributing to this phenomenon; anti-PD-L1 agents' therapeutic effects are limited by many biological characters, such as intratumoral heterogeneity and temporal change expression. At present, the clinical used flow cytometry and immunohistochemistry could not precisely meter the expression level of PD-L1, with low accuracy and credibility. In addition, some researches provided that there existed some intrinsic resistant mechanisms in the tumor immune environment for immunotherapy, such as improving the expression of immunosuppressive cells and checkpoint molecules such as tumor-associated macrophages (TAMs), T follicular helper (Tfh) and regulatory T cells (Tregs), and absence of antigenic proteins and antigenic presentation [7, 8]. The disorder of immune cells and molecules then generate the insensitivity of T cells, reprogram the phenotype of macrophage, alternate the traditional immune response signaling pathways, and engender T cell exclusion, which shut the door for the powerful immune therapy.

The ultimate aim is to identify distinguished LUAD patients that would benefit from immunotherapy, ensure optimal clinical response, minimize immune related adverse events, and decrease treatment costs. Therefore, it is warranted to propose other predictive immune therapy markers, not merely PD-L1, for the effectiveness of LUAD immunotherapy. Many ongoing clinical trials have been investigated focusing on the identification of predictive and prognostic biomarkers for immunotherapy. For example, peripheral blood inflammatory parameters have been shown to be correlated with poor prognosis and lower response rate to immune therapy in NSCLC [9, 10]; high tumor mutational burden level was associated with greater expression of neoantigens, which fosters anticancer immune response [10]; high LDH seems to negatively correlate with cytotoxic T lymphocyte activation and impairs aerobic glycolysis, which participated in the prediction of immune therapy [11]. However, the efficiency of these single indicators is still low and could not convince the public. They ignored the essence property of tumor immune microenvironment but concentrated on the index only reflecting - one side face of immunological indicators, impossibly whole. Thus, it is necessary to lucubrate the complexity of the LUAD microenvironment and identify subclasses of the tumor immune microenvironment existing in the LUAD tissues, which is beneficial for predicting and administering corresponding immunotherapeutics.

A comprehensive analysis of the tumor immune microenvironment in LUAD is imperatively required. In this study, we systematically described the constitutive pattern of the immune cell proportions and how it influenced progression

of tumor development. Moreover, we constructed an immunoscore signature model to predict 1-, 3-, and 5-year overall survival for patients, with the hope to select patients for adjuvant immunotherapy and guide the development of new treatment options.

2. Materials and Methods

2.1. Data Acquisition. This study utilized data from a public database. We retrospectively selected the LUAD gene expression and its clinicopathological data from the TCGA (<https://cancergenome.nih.gov/>) and GEO datasets (<https://www.ncbi.nlm.nih.gov/geo/>). We searched the key words “lung cancer” and “Homo sapiens”. All matrices with LUAD gene expression data (containing at least 20 samples) were considered eligible, with no specific exclusion criteria. All candidate series were assessed by two independent reviewers. A total of 371 normal and 990 LUAD cases were eligible for subsequent research. Raw microarray data Affymetrix were downloaded and normalized using the limma package. The platform profiles of the Affymetrix matrix were listed in table S1. The relevant clinical data from TCGA were retrieved and organized manually. The concrete working algorithm was demonstrated in Figure 1.

2.2. Inference of Infiltrating Immune Cells. Normalized gene expression data were used to calculate the proportions of 22 types of infiltrating immune cells via the CIBERSORT algorithm. CIBERSORT is a gene expression-based deconvolution algorithm, which infers cell type proportions in data from bulk tumor samples of mixed cell types using support vector regression [12]. CIBERSORT derives P values for the deconvolution for each sample using the Monte Carlo sampling, providing the confidence of the outcomes. At a threshold of $P \leq 0.050$, the results of the inferred fractions of immune cell populations produced were considered precise and accurate [13]. Then, patients with CIBERSORT $P \leq 0.050$ were considered eligible for the following analysis.

2.3. Systematic Meta-Analysis. Meta-analysis was referred as the standard method, offering an average impact estimate of the heterogeneity of effects across a series of results. In this study, we take advantage of the meta-analysis to layout the expression of particular immune cell infiltrate. Continuous outcomes were estimated as a standard mean difference (SMD) with a 95% confidence interval (CI). Setting $P < 0.05$ as the cut-off, we deeply explored the composition of TIICs in the LUAD to implement more convincing conclusions.

2.4. Evaluation of Tumor Immune Reaction Score. We obtained a series of genes in the immune-relative pathways from KEGG and published articles [13], then applied the single-sample gene set enrichment analysis (ssGSEA) to quantify the score by the GSVA and GSEABase packages [14]. The definition of each immune term was listed in the table S2. The correlation of the composition of the TIICs and the immune reaction were calculated by the Pearson correlation and showed by heat map.

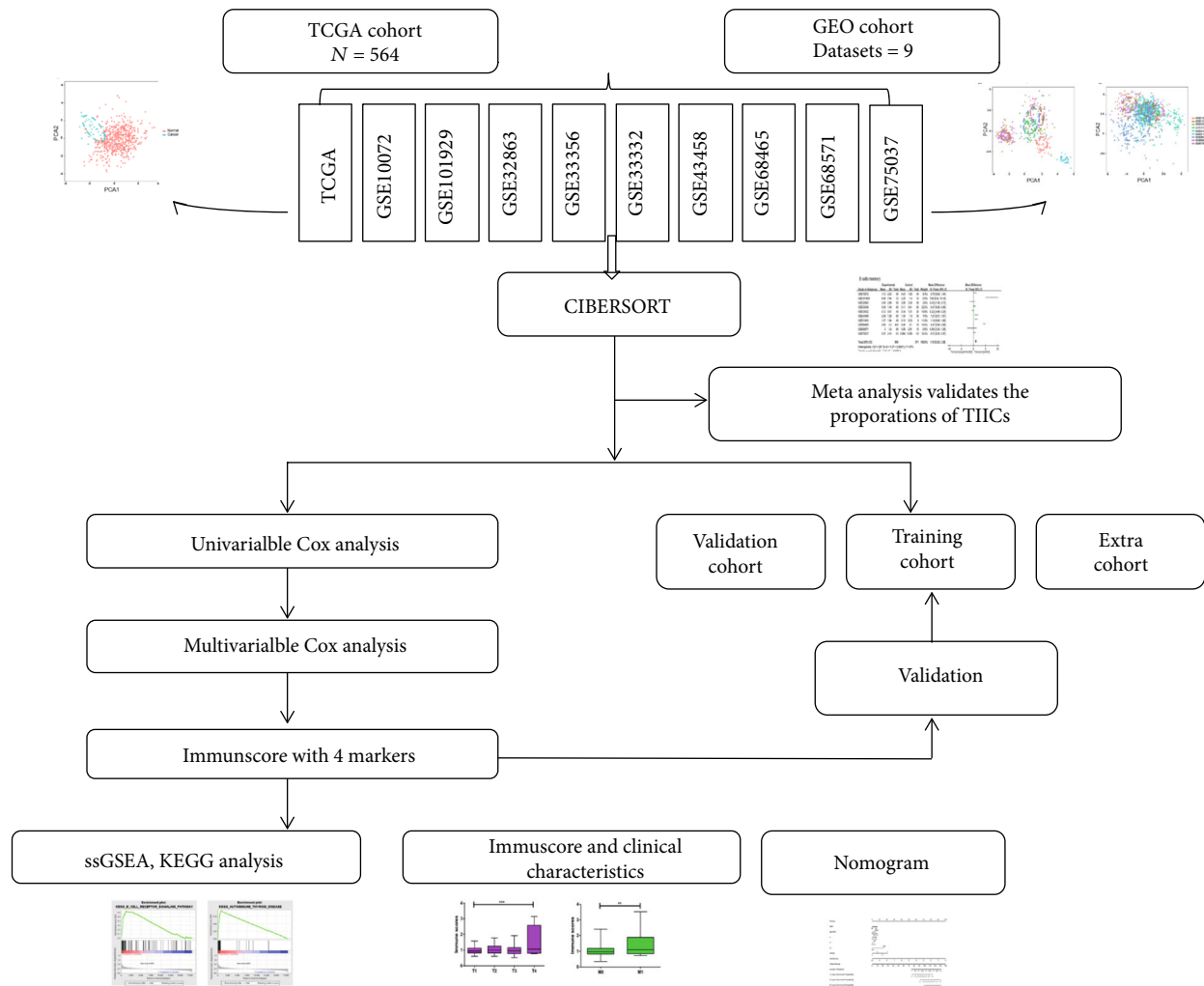


FIGURE 1: Study flowchart detailing the flow of sample collection and analysis.

2.5. Independence of the Prognostic Immunoscore for LUAD Survival Prediction. Only patients with CIBERSORT $P \leq 0.050$ were selected for further survival analyses. Survival package was used to perform univariate Cox proportional hazard regression analysis to filter out prognostic immune cell, and then, multivariate Cox regression analysis was used to construct a powerful predictive immunoscore model. The optimal cut-off value was evaluated by the survminer package. With the median score established as the cut-off line, these LUAD patients were divided into the high- and low-immunoscore groups. The results of Cox regression analysis were showed by forest plot. A nomogram was generated via the survival, rms, and ggplot2 packages.

2.6. Random Grouping Method. The LUAD patients were randomly divided into the training and validation groups in a ratio of 6:4 using the stratified randomization method, which generated random values from a normal distribution with specified mean (0) and standard deviation (1) values and ordered them from high to low. These training and validation groups were both used to validate the precision of the immune model.

2.7. GSEA and PPI Analysis. We used the GSEA program to derive the enrichment scores of each immune-related term by calculated immunoscore [15]. Normalized gene expression data arranged by immunoscore were analyzed via the limma package to identify the immune DEGs, with $P \leq 0.05$ and $|\log_{2}FC| \geq 1.5$ as the cut-off criterion. The connectivity degree of each node of the network was calculated by the STRING database and reconstructed via Cytoscape software.

2.8. Statistical Analysis. Continuous variables were exhibited for means, medians, range, and standard deviation (SD) and compared using an independent t -test or Wilcoxon test; categorical variables were compared between two groups by means of the Chi-squared test. Spearman's correlation coefficient was used for variable correlation. The associations of immune cell infiltrates and corresponding overall survival (OS) were analyzed by log-rank survival test, and the results were shown in the forest plot. To identify prognostic immune cells, the Cox proportional hazard regression model was employed. All statistical tests were two sided and $P < 0.050$ was considered statistically significant. Statistical analyses were conducted using R software and Stats.

3. Results

3.1. The Immune Landscape of the Microenvironment in LUAD. After applying data filter criteria, 564 annotated lung adenocarcinoma samples with immune cell fraction were available for further analyses. We systematically described the immune microenvironment pattern of LUAD tissues from the TCGA cohort. As shown in Figures 2(a) and 2(b), the proportions of TIIC displayed distinct group bias clustering and expression pattern between normal and cancer tissues. The proportions of B cell memory, plasma, T cell CD4 memory activated, T cell follicular helper, T cell regulatory, and macrophage M1 were significantly improved, whereas the proportions of T cell CD4 memory resting, NK cell resting, monocytes, macrophage M2, and mast cell resting were downregulated (Figure 2(c)).

To confirm the accuracy of the results of this study, the researchers inferred its accuracy in other independent LUAD datasets both containing rental tumor and adjacent normal specimens. Figure S1 summarized the compromise of TIIC subpopulations of each normal (Figure S1A) and LUAD tissues (Figure S1B) and CIBERSORT *P* value (Figure S1C). Although these LUAD cohorts were obtained from different platforms (Table S1) and variable signature matrix influences the accuracy of inferred TIIC constitutes, they did not show evident cohort bias both in normal and cancer tissues. We summarized each selected matrix and then eliminate the interaction between the components of the original data by PCA analysis (Figures 2(d) and 2(e), Figure S1). Then, precisely offering an average impact estimate of the heterogeneity of effects was performed to validate the proportions of each TIIC. Obviously, B cell memory (95% CI, 0.92-1.29; $P < 0.01$, $I^2 = 97\%$), plasma (95% CI, 3.58-4.67; $P < 0.01$, $I^2 = 91\%$), T cell CD4 memory activated (95% CI, 0.68-1.17; $P < 0.01$, $I^2 = 91\%$), T cell follicular helpers (95% CI, 1.42-2.06; $P < 0.01$, $I^2 = 96\%$), Treg (95% CI, 0.27-0.44; $P < 0.01$, $I^2 = 93\%$), and macrophage M1 (95% CI, 3.47-4.19; $P < 0.01$, $I^2 = 76\%$) exhibited a decreasing tendency, whereas T cell CD4 memory resting (95% CI, -2.75--1.49; $P < 0.01$, $I^2 = 92\%$), NK cell resting (95% CI, -1.16--0.57; $P < 0.01$, $I^2 = 94\%$), monocytes (95% CI, -3.05--2.58; $P < 0.01$, $I^2 = 98\%$), macrophage M2 (95% CI, -3.09--1.75; $P < 0.01$, $I^2 = 92\%$), Mast cell resting (95% CI, -4.11--3.71; $P < 0.01$, $I^2 = 100\%$), and eosinophils (95% CI, -0.21--0.06; $P < 0.01$, $I^2 = 94\%$) exhibited an increasing tendency (Figure 3). Together, we provided an in-depth overview of the LUAD TIIC subpopulation, which was tightly bounded with LUAD development and immune therapy.

3.2. TIIC Subpopulation Closely Correlated with Immune Signatures in LUAD. To explore the correlation between various tumor immunity cell activities, we found that different TIIC subpopulations were weakly to moderately correlated, specially for T cell CD4 memory activated and T cell CD8 (Figure 4(a)). Next, we obtained a set of genes of the relative immune system from KEGG and published articles, and then used ssGSEA to quantify important immune signatures,

including HLA expression, T cell cosimulation, inflammation-promoting mechanism, PD-L1 reaction, type II IFN response, type I IFN response, check point reaction, T cell coinhibition, parainflammation, and CCR. As shown in Figure 4(b), plasma, macrophage M0, and macrophage M2 were negatively correlated with immune signatures, whether macrophage M1 and T cell CD8 in mice. Moreover, we deeply explored the concrete association between TIIC checkpoint response (Figure 4(c)), PD-L1 reaction (Figure 4(d)), and inflammation-promoting mechanism (Figure 4(e)).

3.3. Establishment of Immunoscore for LUAD Patients. Considering the important role of the composition of the TIICs in the prognosis, we further explored their clinical significance. The unadjusted HRs and 95% confidence intervals for the median proportion of TIIC subsets were depicted in Figure 5(a). Macrophage M1 and dendritic cell resting were significantly associated with LUAD survival. Then, to identify a more predictive model with the best explanatory and informative efficacy, 4 subgroup immune cells were further selected to build an immunoscore model. We yielded the immune score for each tumor sample based on its immune proportion profiles: Formula = (the percentage of NK cell activated $\times 0.066$) + (the percentage of macrophage M1 $\times 0.035$) - (the percentage of dendritic cell resting $\times 0.028$) + (the percentage of dendritic cells activated $\times 0.088$). As shown in Figure 5(b), patients were divided into a high-immune and a low-immune group based on the median immunoscore. The overall survival of patients with a high-immune score was worse than that of those with a low-risk score ($P = 0.814 - e04$, Figure 5(c)). This immunoscore also had strong predictive power for T and M stages (Figure 5(d)).

To further investigate the prognostic value of the immune signature model, univariable and multivariable Cox regression analyses were performed considering immune score, sex, T stage, N stage, M stage, and pathological stage. The results of univariable and multivariable analyses of the above clinicopathological variables were presented in Figures 5(e) and 5(f). In conclusion, the immunoscore was a significantly independent prognostic factor for LUAD patients.

3.4. Variation in Prognostic Effect of Immunoscore in Intra- and Extracohort. To confirm that the proposed immunoscore model has a powerful prognostic value in different populations, the formula was applied to the validation cohort and also to the other cohort. We randomly divided total patients in the primary and validation cohort. Consistent with the findings in the total TCGA cohort, patients in the high-immunoscore group had a significantly lower overall survival rate than those in the low-immunoscore group in both the validation cohort (HR = 2.57, 95% CI: 1.537-4.301) and the training cohort (HR = 1.419, 95% CI: 1.152-1.844) (Table 1). In line with these findings, the immunoscore model was also validated to be independently associated with survival in extracohort GSE101929 and GSE6857 (Figure S2, only this two cohorts possessed clinical information in the above meta-analysis data).

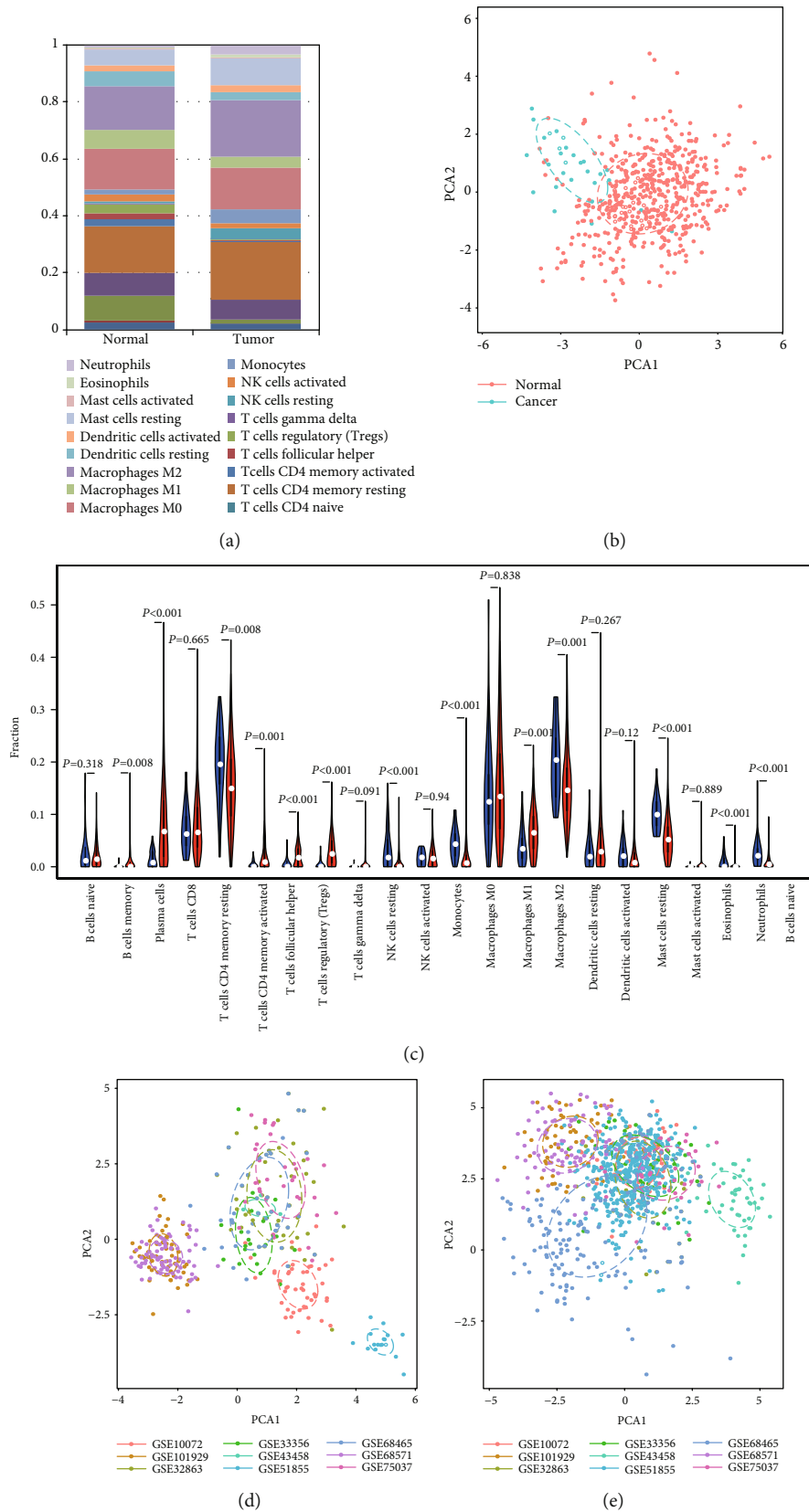
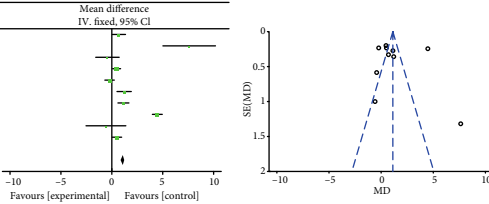


FIGURE 2: Landscape of microenvironment TIIC composition in LUAD. (a) The composition of TIICs of normal and RCC tissues. (b) The proportions of immune cells from normal and LUAD displayed distinct group bias clustering and individual differences in the TCGA dataset. (c) Violin plot of the proportions of the TIIC subpopulation (blue represents normal tissue, red represents LUAD). (d) PCA analysis of the subgroup of normal samples from the GEO dataset. (e) PCA analysis of the subgroup of LUAD samples from the GEO dataset.

B cells memory

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI | Mean difference IV, fixed, 95% CI |
|-------------------|--------------|------|-------|---------|-------|-------|--------|--------------------------------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | | |
| GSE10072 | 1.13 | 2.23 | 58 | 0.43 | 1.05 | 49 | 8.3% | 0.70 [0.06, 1.34] | |
| GSE101929 | 9.85 | 7.44 | 33 | 2.25 | 1.4 | 33 | 0.5% | 7.60 [5.02, 10.18] | |
| GSE32863 | 2.46 | 2.98 | 58 | 2.89 | 3.34 | 58 | 2.6% | -0.43 [-1.58, 0.72] | |
| GSE33356 | 0.58 | 1.49 | 60 | 0.11 | 0.41 | 60 | 22.5% | 0.47 [0.08, 0.86] | |
| GSE33532 | 0.12 | 0.47 | 40 | 0.34 | 1.01 | 20 | 15.8% | -0.22 [-0.69, 0.25] | |
| GSE43458 | 2.56 | 1.28 | 80 | 1.35 | 1.8 | 30 | 7.0% | 1.21 [0.51, 1.91] | |
| GSE18855 | 1.27 | 1.84 | 49 | 0.13 | 0.25 | 9 | 11.8% | 1.14 [0.60, 1.68] | |
| GSE68465 | 4.93 | 5.2 | 443 | 0.46 | 0.1 | 19 | 14.5% | 4.47 [3.98, 4.96] | |
| GSE68571 | 3 | 3.6 | 86 | 3.58 | 2.91 | 10 | 0.9% | -0.58 [-2.54, 1.38] | |
| GSE75037 | 0.57 | 2.14 | 83 | 0.064 | 0.056 | 83 | 16.2% | 0.51 [0.05, 0.97] | |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | 1.10 [0.92, 1.29] | |

Heterogeneity: Chi² = 267.18, df = 9 (< 0.00001); I² = 97%
Test for overall effect: Z = 11.64 (P < 0.00001)

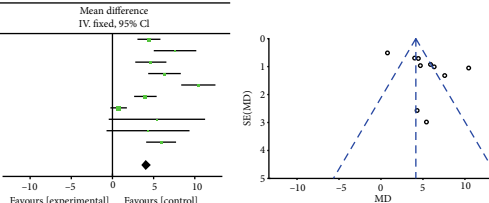


(a)

Plasma

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI | Mean difference IV, fixed, 95% CI |
|-------------------|--------------|------|-------|---------|------|-------|--------|--------------------------------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | | |
| GSE10072 | 7.43 | 4.64 | 58 | 3 | 2.58 | 49 | 15.4% | 4.43 [3.03, 5.83] | |
| GSE101929 | 9.85 | 7.44 | 33 | 2.25 | 1.4 | 33 | 4.5% | 7.60 [5.02, 10.18] | |
| GSE32863 | 5.78 | 6.91 | 58 | 1.12 | 2.49 | 58 | 8.4% | 4.66 [2.77, 6.55] | |
| GSE33356 | 10.48 | 7.37 | 60 | 4.16 | 2.61 | 60 | 7.7% | 6.32 [4.54, 8.30] | |
| GSE33532 | 14.17 | 6.41 | 40 | 3.74 | 1.4 | 20 | 6.9% | 10.43 [8.35, 12.51] | |
| GSE43458 | 8.52 | 2.54 | 80 | 4.52 | 3.49 | 30 | 16.0% | 4.00 [2.63, 5.37] | |
| GSE18855 | 2.08 | 1.73 | 49 | 1.33 | 1.34 | 9 | 29.9% | 0.75 [-0.25, 1.75] | |
| GSE68465 | 30.15 | 3.44 | 443 | 2.475 | 1.3 | 19 | 0.9% | 5.40 [-0.45, 11.25] | |
| GSE68571 | 10.88 | 8.56 | 86 | 6.56 | 7.59 | 10 | 1.2% | 4.32 [-0.72, 9.36] | |
| GSE75037 | 7.26 | 7.85 | 83 | 1.35 | 3.1 | 83 | 9.1% | 5.91 [4.09, 7.73] | |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | 4.12 [3.58, 4.67] | |

Heterogeneity: Chi² = 95.15, df = 9 (< 0.00001); I² = 91%
Test for overall effect: Z = 14.76 (P < 0.00001)

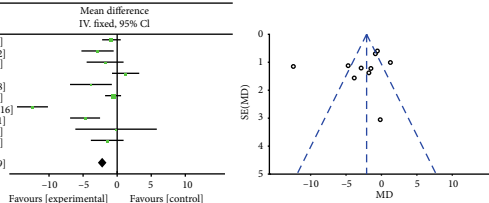


(b)

T cells CD4 memory resting

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI | Mean difference IV, fixed, 95% CI |
|-------------------|--------------|------|-------|---------|------|-------|--------|--------------------------------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | | |
| GSE10072 | 2.33 | 3.42 | 58 | 3.17 | 3.72 | 49 | 21.2% | -0.84 [-2.20, 0.52] | |
| GSE101929 | 9.04 | 4.81 | 33 | 11.92 | 4.99 | 33 | 7.0% | -2.88 [-5.24, -0.52] | |
| GSE32863 | 15.64 | 8.32 | 58 | 17.41 | 6.45 | 58 | 5.4% | -1.77 [-4.48, 0.94] | |
| GSE33356 | 6.31 | 6.13 | 60 | 5.07 | 4.97 | 60 | 9.9% | 1.24 [-0.76, 3.24] | |
| GSE33532 | 5.15 | 5.41 | 40 | 9 | 5.86 | 20 | 4.2% | -2.85 [-6.92, -0.78] | |
| GSE43458 | 4.56 | 3.48 | 80 | 5.16 | 2.49 | 30 | 28.6% | -0.60 [-1.77, 0.57] | |
| GSE18855 | 5.12 | 4.72 | 49 | 17.55 | 2.83 | 9 | 7.6% | -12.43 [-14.70, -10.16] | |
| GSE68465 | 0.023 | 0.27 | 443 | 4.73 | 4.89 | 19 | 8.1% | -4.71 [-6.91, -2.51] | |
| GSE68571 | 24.9 | 8.48 | 86 | 25.06 | 9.2 | 10 | 1.1% | -0.16 [-6.14, 5.82] | |
| GSE75037 | 16.99 | 9.26 | 83 | 18.45 | 6.23 | 83 | 6.8% | -1.46 [-3.86, 0.94] | |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | -2.12 [-2.75, -1.49] | |

Heterogeneity: Chi² = 102.53, df = 9 (< 0.00001); I² = 91%
Test for overall effect: Z = 7.38 (P < 0.00001)

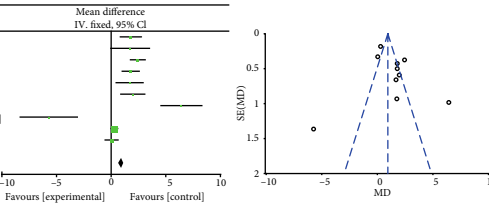


(c)

T cells CD4 memory activated

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI | Mean difference IV, fixed, 95% CI |
|-------------------|--------------|------|-------|---------|------|-------|--------|--------------------------------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | | |
| GSE10072 | 7.4 | 2.86 | 58 | 5.6 | 2.31 | 49 | 6.3% | 1.80 [0.82, 2.78] | |
| GSE101929 | 5.54 | 3.87 | 33 | 3.8 | 3.7 | 33 | 1.8% | 1.74 [-0.09, 3.57] | |
| GSE32863 | 2.63 | 2.81 | 58 | 0.19 | 0.49 | 58 | 11.2% | 2.44 [1.71, 3.17] | |
| GSE33356 | 4 | 2.72 | 60 | 2.21 | 1.89 | 60 | 8.6% | 1.79 [0.95, 2.63] | |
| GSE33532 | 4.33 | 3 | 40 | 2.64 | 2.05 | 20 | 3.6% | 1.69 [0.40, 2.98] | |
| GSE43458 | 8.46 | 3.49 | 80 | 6.48 | 2.43 | 30 | 4.5% | 1.98 [0.82, 3.14] | |
| GSE18855 | 13.3 | 5.2 | 49 | 6.87 | 1.94 | 9 | 1.6% | 6.43 [4.50, 8.36] | |
| GSE68465 | 10.48 | 2.81 | 443 | 10.53 | 9.92 | 19 | 0.8% | -5.71 [-8.38, -3.04] | |
| GSE68571 | 0.53 | 1.06 | 86 | 0.23 | 0.45 | 10 | 47.2% | 0.30 [-0.06, 0.66] | |
| GSE75037 | 1.31 | 2.48 | 83 | 1.27 | 1.73 | 83 | 14.3% | 0.04 [-0.61, 0.69] | |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | 0.93 [0.68, 1.17] | |

Heterogeneity: Chi² = 102.53, df = 9 (< 0.00001); I² = 91%
Test for overall effect: Z = 7.38 (P < 0.00001)

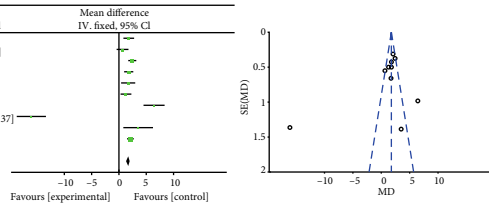


(d)

T cells follicular helpers

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI | Mean difference IV, fixed, 95% CI |
|-------------------|--------------|------|-------|---------|------|-------|--------|--------------------------------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | | |
| GSE10072 | 7.4 | 2.86 | 58 | 5.6 | 2.31 | 49 | 10.5% | 1.80 [0.82, 2.78] | |
| GSE101929 | 3.17 | 2.16 | 33 | 2.51 | 2.33 | 33 | 8.5% | 0.66 [-0.42, 1.74] | |
| GSE32863 | 2.63 | 2.81 | 58 | 0.19 | 0.49 | 58 | 18.6% | 2.44 [1.71, 3.17] | |
| GSE33356 | 4 | 2.72 | 60 | 2.21 | 1.89 | 60 | 14.3% | 1.79 [0.95, 2.63] | |
| GSE33532 | 4.33 | 3 | 40 | 2.64 | 2.05 | 20 | 6.0% | 1.69 [0.40, 2.98] | |
| GSE43458 | 3.48 | 2.49 | 80 | 2.19 | 2.3 | 30 | 10.3% | 1.29 [0.30, 2.28] | |
| GSE18855 | 13.3 | 5.2 | 49 | 6.87 | 1.94 | 9 | 2.7% | 6.43 [4.50, 8.36] | |
| GSE68465 | 4.82 | 2.81 | 443 | 20.86 | 5.92 | 19 | 1.4% | -16.04 [-18.71, -13.37] | |
| GSE68571 | 7.61 | 5.63 | 86 | 4.09 | 3.94 | 10 | 1.4% | 3.52 [0.80, 6.24] | |
| GSE75037 | 2.2 | 2.83 | 83 | 0.099 | 0.49 | 83 | 26.3% | 2.10 [1.48, 2.72] | |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | 1.74 [1.42, 2.06] | |

Heterogeneity: Chi² = 203.52, df = 9 (< 0.00001); I² = 96%
Test for overall effect: Z = 10.75 (P < 0.00001)

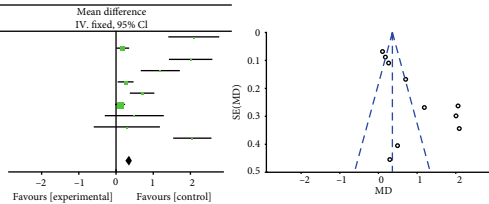


(e)

Treg

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI | Mean difference IV, fixed, 95% CI |
|-------------------|--------------|------|-------|---------|--------|-------|--------|--------------------------------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | | |
| GSE10072 | 4.13 | 2.07 | 58 | 2.04 | 1.47 | 49 | 1.6% | 2.09 [1.42, 2.76] | |
| GSE101929 | 0.18 | 0.51 | 33 | 0.0021 | 0.011 | 33 | 24.6% | 0.18 [0.00, 0.35] | |
| GSE32863 | 2.08 | 2.26 | 58 | 0.075 | 0.19 | 58 | 2.2% | 2.00 [1.42, 2.59] | |
| GSE33356 | 1.37 | 1.92 | 60 | 0.18 | 0.79 | 60 | 2.7% | 1.19 [0.66, 1.72] | |
| GSE33532 | 0.26 | 0.69 | 40 | 0.002 | 0.0046 | 20 | 16.3% | 0.26 [0.04, 0.47] | |
| GSE43458 | 1.04 | 1.25 | 80 | 0.34 | 0.51 | 30 | 6.9% | 0.70 [0.37, 1.03] | |
| GSE18855 | 0.11 | 0.47 | 49 | 0.0014 | 0.046 | 9 | 40.8% | 0.11 [-0.03, 0.24] | |
| GSE68465 | 1.72 | 1.85 | 443 | 1.23 | 1.72 | 19 | 1.2% | 0.49 [-0.30, 1.28] | |
| GSE68571 | 0.72 | 1.25 | 86 | 0.43 | 1.37 | 10 | 0.9% | 0.29 [-0.60, 1.18] | |
| GSE75037 | 2.18 | 2.35 | 83 | 0.13 | 0.43 | 83 | 2.8% | 2.05 [1.54, 2.56] | |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | 0.35 [0.27, 0.44] | |

Heterogeneity: Chi² = 129.62, df = 9 (< 0.00001); I² = 93%
Test for overall effect: Z = 8.06 (P < 0.00001)



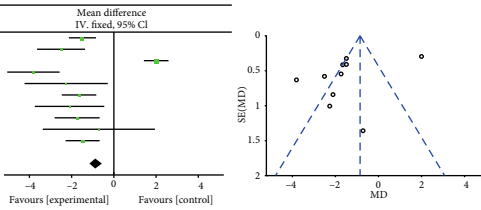
(f)

FIGURE 3: Continued.

NK cells resting

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI |
|-----------------------|--------------|------|------------|---------|------|------------|---------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | |
| GSE10072 | 0.59 | 1.27 | 58 | 2.08 | 1.97 | 49 | 21.2% | -1.49 [-2.13, -0.85] |
| GSE101929 | 0.5 | 1.06 | 33 | 3 | 3.16 | 33 | 6.7% | -2.50 [-3.64, -1.36] |
| GSE32863 | 2.08 | 2.26 | 58 | 0.075 | 0.19 | 58 | 25.6% | 2.00 [1.42, 2.59] |
| GSE33356 | 1.49 | 2.26 | 60 | 5.3 | 4.35 | 60 | 5.7% | -3.81 [-5.05, -2.57] |
| GSE33532 | 1.87 | 3.04 | 40 | 4.13 | 3.95 | 20 | 2.2% | -2.26 [-4.23, -0.29] |
| GSE43458 | 1.84 | 2.83 | 80 | 3.49 | 1.49 | 30 | 13.0% | -1.65 [-2.47, -0.83] |
| GSE51855 | 2.16 | 1.26 | 49 | 4.26 | 2.46 | 9 | 3.2% | -2.10 [-3.75, -0.45] |
| GSE68465 | 1.62 | 2.12 | 443 | 3.36 | 2.33 | 19 | 7.7% | -1.74 [-2.81, -0.67] |
| GSE68571 | 3.92 | 3.44 | 86 | 4.63 | 4.12 | 10 | 1.2% | -0.71 [-3.37, -1.95] |
| GSE75037 | 1.07 | 1.78 | 83 | 2.55 | 3.3 | 83 | 13.4% | -1.48 [-2.29, -0.67] |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | -0.86 [-1.16, -0.57] |

Heterogeneity: $\chi^2 = 138.59$, $df = 9$ (< 0.00001); $I^2 = 94\%$
 Test for overall effect: $Z = 5.72$ ($P < 0.00001$)

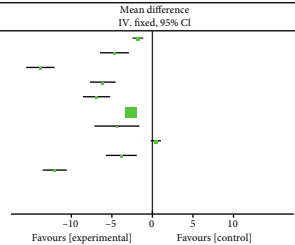


(g)

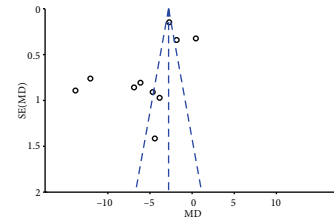
Monocytes

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI |
|-----------------------|--------------|------|------------|---------|------|------------|---------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | |
| GSE10072 | 0.99 | 1.48 | 58 | 2.79 | 1.97 | 49 | 12.2% | -1.80 [-2.47, -1.13] |
| GSE101929 | 0.68 | 2.25 | 33 | 5.36 | 4.7 | 33 | 1.7% | -4.68 [-6.46, -2.90] |
| GSE32863 | 3.53 | 3.55 | 58 | 17.35 | 5.8 | 58 | 1.8% | -13.82 [-15.57, -12.07] |
| GSE33356 | 1.59 | 2.8 | 60 | 7.7 | 5.6 | 60 | 2.2% | -6.11 [-7.69, -4.53] |
| GSE33532 | 1.23 | 1.82 | 40 | 8.12 | 3.62 | 20 | 1.9% | -6.89 [-8.59, -5.21] |
| GSE43458 | 0.11 | 0.37 | 80 | 2.81 | 0.8 | 30 | 61.9% | -2.70 [-3.00, -2.40] |
| GSE51855 | 1.25 | 1.49 | 49 | 5.64 | 4.19 | 9 | 0.7% | -4.39 [-7.16, -1.62] |
| GSE68465 | 1.11 | 1.88 | 443 | 0.69 | 1.36 | 19 | 13.5% | -0.42 [-0.22, -1.06] |
| GSE68571 | 1.76 | 2.44 | 86 | 5.59 | 2.96 | 10 | 1.5% | -3.83 [-5.74, -1.92] |
| GSE75037 | 3.09 | 3.74 | 83 | 15.14 | 5.82 | 83 | 2.5% | -12.05 [-13.54, -10.56] |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | -2.82 [-3.05, -2.58] |

Heterogeneity: $\chi^2 = 454.23$, $df = 9$ (< 0.00001); $I^2 = 98\%$
 Test for overall effect: $Z = 23.58$ ($P < 0.00001$)



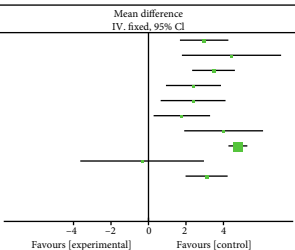
(h)



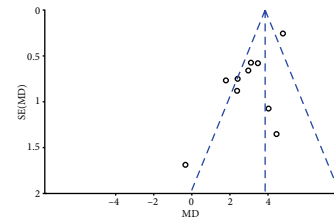
Macrophages M1

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI |
|-----------------------|--------------|------|------------|---------|------|------------|---------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | |
| GSE10072 | 7.47 | 4 | 58 | 4.5 | 2.79 | 49 | 7.7% | 2.97 [1.68, 4.26] |
| GSE101929 | 10.61 | 5.57 | 33 | 6.18 | 5.4 | 33 | 1.8% | 4.43 [1.78, 7.08] |
| GSE32863 | 5.18 | 4.02 | 58 | 1.71 | 1.86 | 58 | 9.9% | 3.47 [2.33, 4.61] |
| GSE33356 | 7.52 | 4.4 | 60 | 5.12 | 3.78 | 60 | 6.0% | 2.40 [0.93, 3.87] |
| GSE33532 | 6.96 | 3.79 | 40 | 4.58 | 2.9 | 20 | 4.3% | 2.38 [0.65, 4.11] |
| GSE43458 | 5.94 | 4.49 | 80 | 4.16 | 3.19 | 30 | 5.7% | 1.78 [0.27, 3.29] |
| GSE51855 | 7.66 | 3.36 | 49 | 3.65 | 2.88 | 9 | 2.9% | 4.01 [1.91, 5.11] |
| GSE68465 | 6.87 | 3.3 | 443 | 2.1 | 0.89 | 19 | 50.4% | 4.77 [4.27, 5.27] |
| GSE68571 | 2.49 | 2.92 | 86 | 2.83 | 5.23 | 10 | 1.2% | -0.34 [-3.64, 2.96] |
| GSE75037 | 5.07 | 4.67 | 83 | 1.97 | 2.33 | 83 | 10.2% | 3.10 [1.98, 4.22] |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | 3.83 [3.47, 4.19] |

Heterogeneity: $\chi^2 = 36.85$, $df = 9$ (< 0.00001); $I^2 = 76\%$
 Test for overall effect: $Z = 20.96$ ($P < 0.00001$)



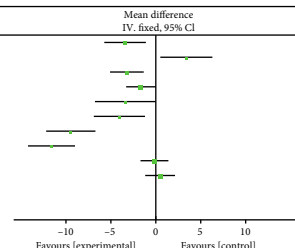
(i)



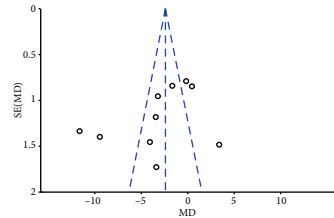
Macrophages M2

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI |
|-----------------------|--------------|------|------------|---------|------|------------|---------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | |
| GSE10072 | 18.91 | 5.28 | 58 | 22.33 | 6.69 | 49 | 8.4% | -3.42 [-5.73, -1.11] |
| GSE101929 | 10.62 | 5.07 | 33 | 7.22 | 6.82 | 33 | 5.3% | -3.40 [0.50, 6.30] |
| GSE32863 | 9.82 | 5.07 | 58 | 13.04 | 5.24 | 58 | 12.8% | -3.22 [-5.10, -1.34] |
| GSE33356 | 10.52 | 3.25 | 60 | 12.17 | 5.65 | 60 | 16.5% | -1.65 [-3.30, -0.00] |
| GSE33532 | 14.8 | 5.37 | 40 | 18.19 | 6.72 | 20 | 3.9% | -3.39 [-6.77, -0.01] |
| GSE43458 | 10.49 | 8.19 | 80 | 14.54 | 6.19 | 30 | 5.5% | -4.05 [-6.90, -1.20] |
| GSE51855 | 14.03 | 5.71 | 49 | 23.49 | 3.41 | 9 | 6.0% | -9.46 [-12.20, -6.72] |
| GSE68465 | 0.11 | 0.31 | 443 | 11.72 | 5.82 | 19 | 6.6% | -11.61 [-14.23, -8.99] |
| GSE68571 | 1.4 | 2.2 | 86 | 1.55 | 2.39 | 10 | 18.6% | -0.15 [-1.70, 1.40] |
| GSE75037 | 10.84 | 5.22 | 83 | 10.36 | 5.68 | 83 | 16.3% | 0.48 [-1.18, 1.14] |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | -2.42 [-3.09, -1.75] |

Heterogeneity: $\chi^2 = 111.93$, $df = 9$ (< 0.00001); $I^2 = 92\%$
 Test for overall effect: $Z = 7.08$ ($P < 0.00001$)



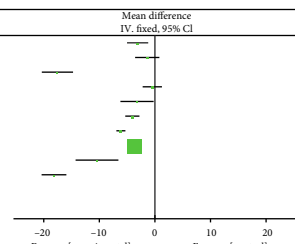
(j)



Mast cells resting

| Study or subgroup | Experimental | | | Control | | | Weight | Mean difference IV, fixed, 95% CI |
|-----------------------|--------------|------|------------|---------|------|------------|---------------|--------------------------------------|
| | Mean | SD | Total | Mean | SD | Total | | |
| GSE10072 | 2.75 | 3.43 | 58 | 5.86 | 6.03 | 49 | 1.1% | -3.11 [-5.02, -1.20] |
| GSE101929 | 2.83 | 3.68 | 33 | 4.21 | 5.23 | 33 | 0.8% | -1.28 [-3.56, 0.80] |
| GSE32863 | 11.98 | 6.95 | 58 | 29.52 | 8.43 | 58 | 0.5% | -17.54 [-20.35, -14.73] |
| GSE33356 | 4.76 | 4.94 | 60 | 5.23 | 4.76 | 60 | 1.3% | -0.47 [-2.21, 1.27] |
| GSE33532 | 5.3 | 5.11 | 40 | 8.52 | 5.78 | 20 | 0.5% | -3.22 [-6.21, -0.23] |
| GSE43458 | 2.15 | 2.6 | 80 | 6.21 | 3.19 | 30 | 2.5% | -4.06 [-5.34, -2.78] |
| GSE51855 | 0.45 | 1.56 | 49 | 6.57 | 1.05 | 9 | 6.1% | -6.12 [-6.93, -5.31] |
| GSE68465 | 0.61 | 0.94 | 443 | 4.22 | 0.44 | 19 | 86.1% | -3.61 [-3.83, -3.39] |
| GSE68571 | 5.9 | 6.18 | 86 | 16.32 | 5.85 | 10 | 0.3% | -10.42 [-14.27, -6.57] |
| GSE75037 | 11.16 | 6.65 | 83 | 29.32 | 7.95 | 83 | 0.8% | -18.16 [-20.39, -15.93] |
| Total (95% CI) | | | 990 | | | 371 | 100.0% | -3.91 [-4.11, -3.71] |

Heterogeneity: $\chi^2 = 315.06$, $df = 9$ (< 0.00001); $I^2 = 97\%$
 Test for overall effect: $Z = 38.21$ ($P < 0.00001$)



(k)

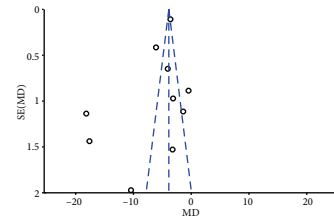


FIGURE 3: Continued.

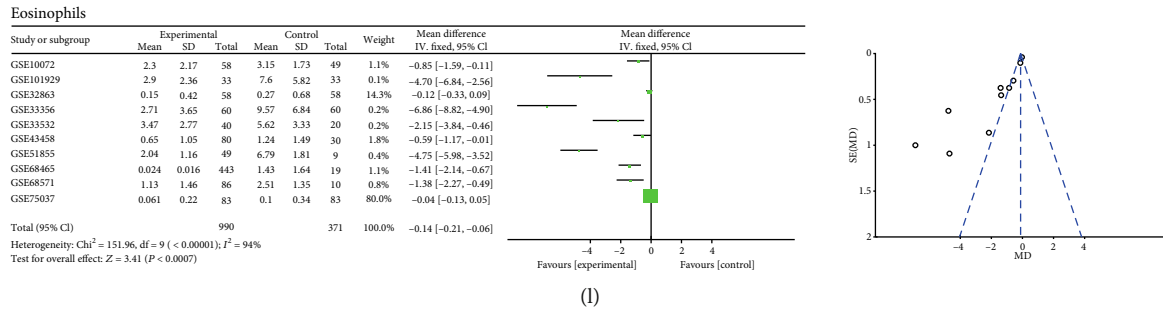


FIGURE 3: Meta-analysis verified each discrepant TIIC composition in LUAD.

3.5. Prognostic Nomograms for Prediction of LUAD Patients' Overall Survival. To develop a practical method for clinicians to predict the LUAD survival probability, we constructed a prognostic nomogram that integrated the immunoscore, age, sex, T stage, N stage, M stage, and pathological stage. These variables from the Cox proportional analysis were all considered. As shown in Figure 6, the immunoscore contributed the most risk points (ranged 0 to 10), whereas the other clinical factors contributed much less (ranged 0 to 20). In general, the immunoscore was an independent risk predictor for the overall survival for LUAD patients.

3.6. Functional Analysis of Immune-Relative Genes. We compared the gene expression profiles between the high and low immunoscores and identified the KEGG pathways that were enriched in the cancerous immune activity by GSEA. It suggested that the high-immunoscore group was significantly enriched in B cell receptor signaling pathway, autoimmune disease, primary immunodeficiency, and T cell receptor signaling pathway (Figure S3). These results confirmed the elevated immune activity in the progression of LUAD.

To investigate the potentially altered molecular mechanisms, immune-related genes were selected by comparing the high- and low-immunoscore groups. Totally, 24 genes were upregulated and 24 genes downregulated. The relationship and function of immune-related genes were revealed using the PPI network, and there exist three significant immune relative modules (Figure 7). In the most significant one (left), 95 edges involving 30 nodes were formed in the network (Table S3). ALB, APDC3, LPA, APDC3, and CRP were hub genes and remarkable for having many connections with others.

4. Discussion

In many studies, the important impact of immune cells in determining the progression of the LUAD pathogenesis has been proven. However, owing to technical limitations, relative studies have only heavily relied on limited repertoire of immune phenotypic markers. The defects of single-agent immunotherapy are restricted to a small number of patients, with a response rate of 11-21% [6]. Immunotherapy concentrates on the activation of immunologic molecular components to defend against the cancer cell in the tumor microenvironment. Thus, it is important to deeply explore the influence of the tumor microenvironment on LUAD

immunotherapy, which is the key to overcoming de novo resistance. Compared with traditional single-factor predictors, a comprehensive systematic assessment of the immune system is urgently needed to better understand the development of LUAD.

In our study, LUAD gene expression data from the TCGA database was analyzed based on CIBERSORT method, we found that the proportions of TIICs varied between the normal and cancer tissues. CD4 memory resting, macrophage M0, and macrophages M2 ranked the top three proportions of TIICs in LUAD. CD4 memory resting T cells can differentiate into a variety of helper T cells and regulatory T cells, which play an important role in adaptive immune response, participating in humoral immunity and regulation of cellular immunity [16]. Meanwhile, tumor-associated macrophages (TAM) have been proved to be the most important immune cells in the tumor stroma and played an important role in the occurrence and development of malignant tumors, accounting for more than 50% of the total number of immune cells in the tumor stroma [17, 18].

Further research showed that B cell memory, plasma cells, T cell CD4 memory activated, T cell follicular helper, Tregs, and macrophage M1 were significantly increased in LUAD compared with the adjacent normal tissues. In contrast, T cell CD4 memory resting, NK cell resting, monocyte, macrophage M2, mast cell resting, eosinophils, and neutrophils were decreased in LUAD compared with normal tissues. Then, the proportions of TIIC were validated in GEO the database, and the trend was consistent with our previous results. The varying degree of infiltration of some immune cells between the tumor and adjacent normal tissues suggested that these cells played an important role in the development of LUAD. Macrophages are the main immune-infiltrating cells in tumors, the key cell types that link inflammation and cancer [19]. The major states of macrophages, macrophage M1 and macrophage M2, differently infiltrated between the tumor and adjacent normal tissue. Macrophages have a series of continuous functional states, and M1-type and M2-type macrophages are the two extremes of this continuous state [20]. Macrophage M1 activates the production cytokine, recruited the proimmunostimulating leukocytes TME, which ultimately leads to the phagocytosis of tumor cells [21]. However, macrophage M2 was reported to have an opposite function in tumor; it promotes the development of tumor via the breakdown of basement membrane, recruitment of leukocytes recruitment, angiogenesis, and immune

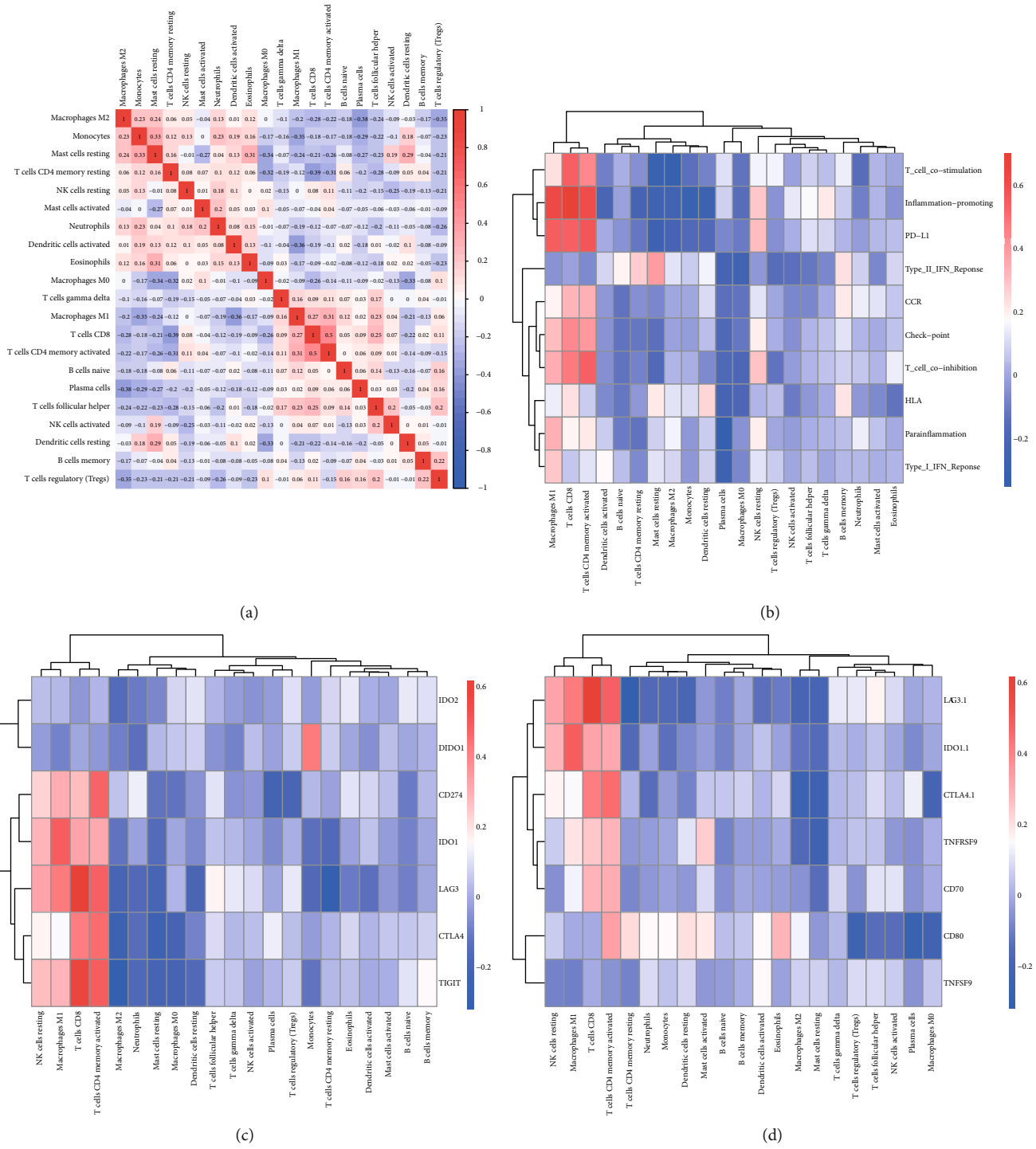


FIGURE 4: Continued.

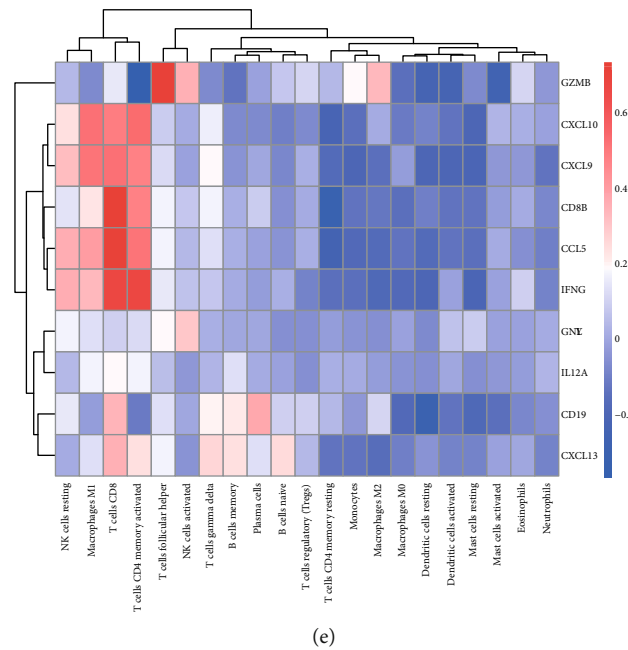


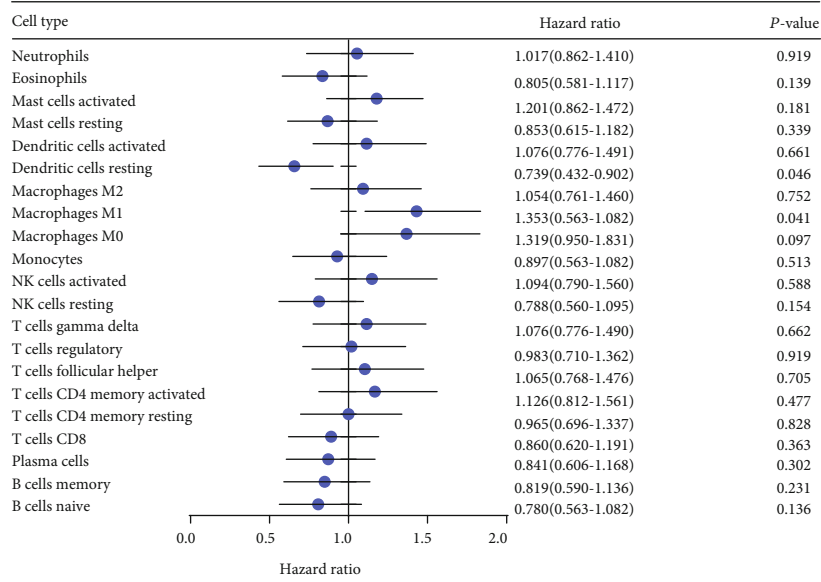
FIGURE 4: TIIC subpopulation was closely correlated with immune signatures. (a) The Spearman correlation matrix of all 22 immune cell proportions in the TCGA cohort. (b) The Spearman correlation between the proportions of infiltrating immune cells and specific immune signatures. (c) Correlation matrix of all 22 immune cell proportions and checkpoint response; (d). Correlation matrix of all 22 immune cell proportions and PD-L1 activity. (e) Correlation matrix of all 22 immune cell proportions and inflammation-promoting mechanisms.

suppression [22, 23]. Interestingly, previous studies showed that the increased level of macrophage M1 level was associated with better prognosis of tumor [24, 25], while the increase of macrophage M2 indicates that the prognosis of tumor patients is poor [26]. Studies have demonstrated that cytotoxic activity in the environment of the tumor would be increased through reversion of macrophages from M2 to M1 phenotypes [27]. Thus, scientists proposed that reprogramming the macrophage from M2 to M1 phenotypes in different ways increases the antitumor activity of macrophages and provides new ideas for the treatment of tumors. Some studies have been carried out. For example, in a tail metastasis model established by breast cancer, the activation of NF- κ B in macrophages lead to a reversion from macrophage M2 to macrophages M1 and ended up with an increased apoptosis of tumor cells and a reduction of lung metastases [28]. In addition, more studies reported that plasma cells play a positive role in antitumor immunity, indicating a positive prognostic effect in human cancers [29]. The number of follicular T helper cells increased significantly in MLD patients, and the proportion of follicular T helper cells decreased in effective treatment patients, which could be used as an indicator of efficacy [30]. Among the 13 immune cells, other than those discussed above, some are poorly investigated in LUAD. As for their biological function in LAUD, we believe that further investigations are urgently needed.

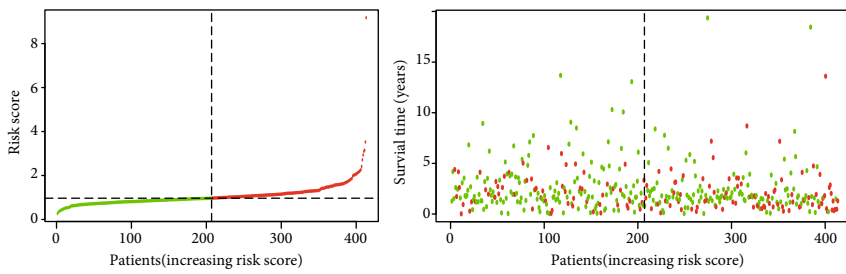
The interaction of programmed cell death ligand-1 (PD-L1) interaction with programmed cell death protein-1 (PD-1) provides an important target for blockade-based immunotherapy in LUAD [3, 31, 32]. Thus, based on our researches, we would like to explore the association between TIICs and PD-L1 reaction was further explored. The results showed

that the expression of PD-L1 was positively associated with macrophage M1, T cell CD8, and T cell CD4 memory activated. The association between macrophages and anti-PD-1 therapy has been reported before. Previous studies of anti-PD-1 therapy in osteosarcoma (OS) have shown that anti-PD-1 therapy can cause phenotypic metastasis of macrophages from M2 to M1, resulting in regression of OS pulmonary metastasis. Macrophage depletion significantly reduced the efficacy of anti-PD1, confirming their role in anti-PD-1 against OS pulmonary metastasis [33]. Another study showed that the expression of PD-L1 negatively correlates with phagocytic ability against tumor cells, and the blockade of PD-1-PD-L1 increased the phagocytosis ability of macrophage and reduced the progression of tumor, as well as prolonged survival time of mice in mouse models of cancer, indicating that anti-PD-1 therapy may function through a direct effect on macrophages [34]. In conclusion, the above results suggested that macrophage cells played a key role in the anti-PD-1 therapy of LUAD and can be a potential target of immunity therapy of LUAD in the future.

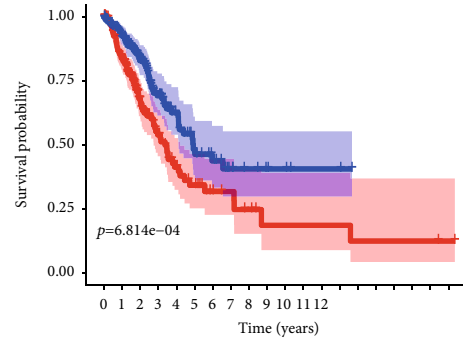
The result of univariate analysis showed that not overall immune cell infiltration but a certain TIIC subpopulation was related to the prognosis of LUAD. In this way, a Cox regression model was constructed based on the estimated fractions of signature immune cells. In recent years, several models based on immunoscore have been published to quantify the immune contexture and provide a statistically robust parameter for prognosis in patients with various types of solid tumor. Our study showed that patients with high immune scores have shorter survival time than patients with low immune scores, and this result was validated in a testing dataset and two independent databases from the GEO



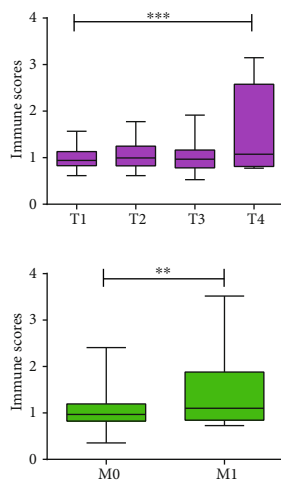
(a)



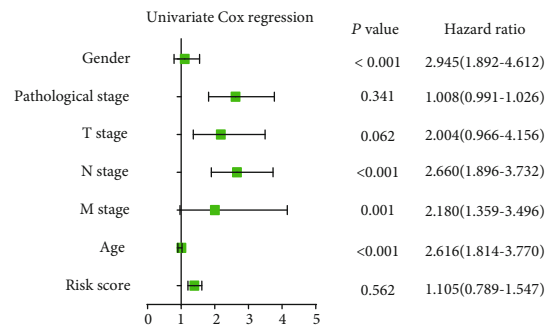
(b)



(c)



(d)



(e)

FIGURE 5: Continued.

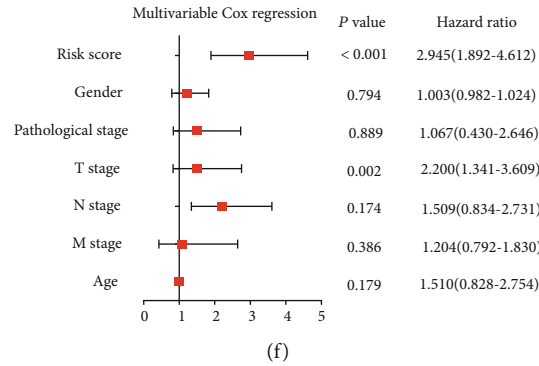


FIGURE 5: Construction of the immunoscore model. (a). Forest plots showed the association between each immune cell subsets and overall survival in the total TCGA dataset. Unadjusted hazard ratios are shown with 95 percent confidence intervals. (b). The distributions of the immune score and survival status for each LUAD patients. (c). Kaplan-Meier curves for the high- and low-immune score subgroup. (d) Association between the TIICs and clinicopathological features. Infiltrating immune cell function in distinguishing T and M stages. (e) Forest plot summary of the univariable analyses of overall survival of the immune score model. (f). Forest plot summary of the multivariable analyses of overall survival of the immune score model.

TABLE 1: Validation of the immunoscore results in the training and validation cohort in the TCGA database.

| Variable | P value | Training cohort | | Validation cohort | |
|--------------|---------|-----------------|-----------------------|-------------------|----------------------|
| | | P value | Hazard ratio (95% CI) | P value | Hazard ratio(95% CI) |
| M | 0.897 | | | 0.810 | |
| M0 | 0.698 | | 0.753 (0.180-3.149) | 0.541 | 0.642 (0.155-2.663) |
| M1 | 0.647 | | 0.723 (0.180-2.898) | 0.622 | 0.699 (0.168-2.901) |
| N | 0.621 | | | 0.759 | |
| N0 | 0.505 | | 0.617 (0.149-2.554) | 0.766 | 0.710 (0.074-6.828) |
| N1-3 | 0.329 | | 0.581 (0.196-1.727) | 0.457 | 0.693 (0.264-1.822) |
| T | 0.241 | | | 0.813 | |
| T1 | 0.152 | | 0.285 (0.051-1.586) | 0.314 | 7.388 (0.203-14.299) |
| T2 | 0.459 | | 0.590 (0.146-2.385) | 0.806 | 0.842 (0.214-3.317) |
| T3 | 0.606 | | 0.842 (0.438-1.620) | 0.990 | 0.922 (0.266-3.700) |
| T4 | 0.800 | | 1.724 (0.279-5.236) | 0.796 | 1.208 (0.288-5.067) |
| Stage | 0.144 | | | 0.785 | |
| Stage1 | 0.133 | | 0.408 (0.126-1.315) | 0.326 | 0.520 (0.141-1.921) |
| Stage2 | 0.023 | | 1.263 (1.016-1.832) | 0.389 | 0.555 (0.145-2.119) |
| Stage3-4 | 0.418 | | 0.802 (0.470-1.368) | 0.711 | 0.877 (0.438-1.757) |
| Gender | 0.648 | | | 0.782 | 0.682 (0.216-1.243) |
| Man | 0.369 | | 0.781 (0.455-1.340) | 0.619 | 0.856 (0.463-1.582) |
| Woman | 0.364 | | 0.787 (0.469-1.032) | 0.546 | 1.207 (0.655-2.223) |
| Immune score | 0.029 | | 2.570(1.537-4.301) | 0.001 | 1.419(1.152-1.844) |
| | 0.062 | | 1.016(0.969-1.021) | 0.325 | 1.027(0.988-1.036) |

database. Further studies showed that the immune score was significantly related to the T stage and M stage. TNM staging system was limited in clinical practice, due to the ignorance of age, gender, or clinical stage of patients [35]. We believe our immunoscore signature is a great complement to TNM staging system. Thus, a nomogram based on age, gender, T stage, N stage, M stage, tumor stage, and immunoscore signature was constructed, with which we can accurately predict the prognosis of LUAD patients.

The potential mechanism for the differences in immune score was explored through GSEA. The results showed that

a series of pathways associated with immune responses were significantly enriched in patients with high immune score. Moreover, DEGs of the high- and low-immune score group were identified and sent to PPI network. The results show that ALB, APOC3, LPA, AGT, and SSX1 were the hub gene of the module of immune-related cells. Totally, the above results suggested that immune responses play a critical role in the development, diagnosis, and treatment of LUAD and strong attention should be paid to it.

Despite the remarkable results of our research, we should also recognize the limitations. Firstly, our model

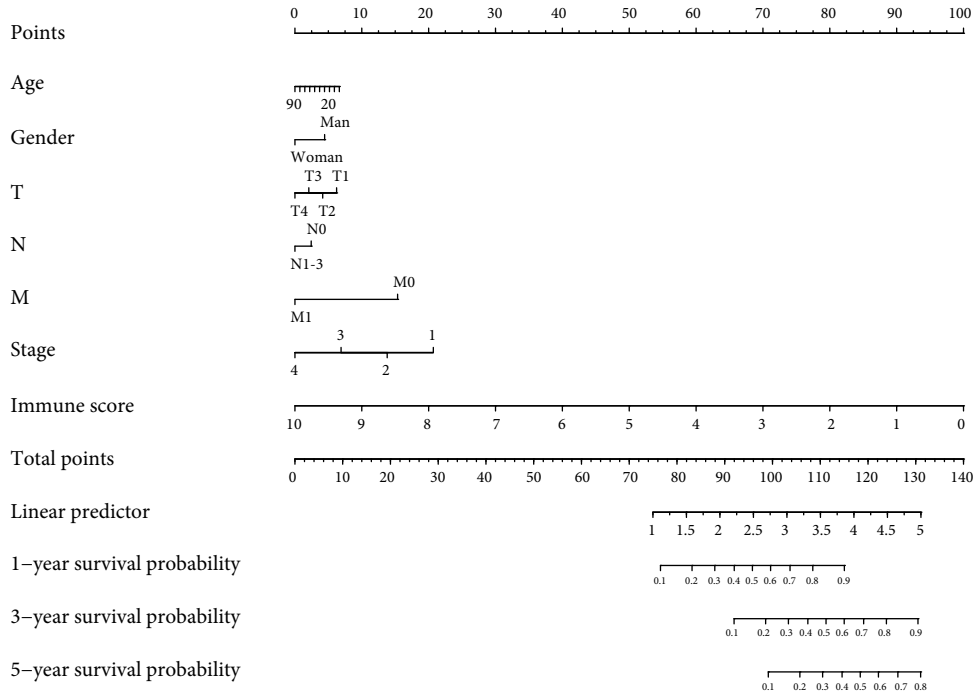


FIGURE 6: The nomogram to predict the 1-, 3-, and 5-year overall survival rates of LUAD patients.

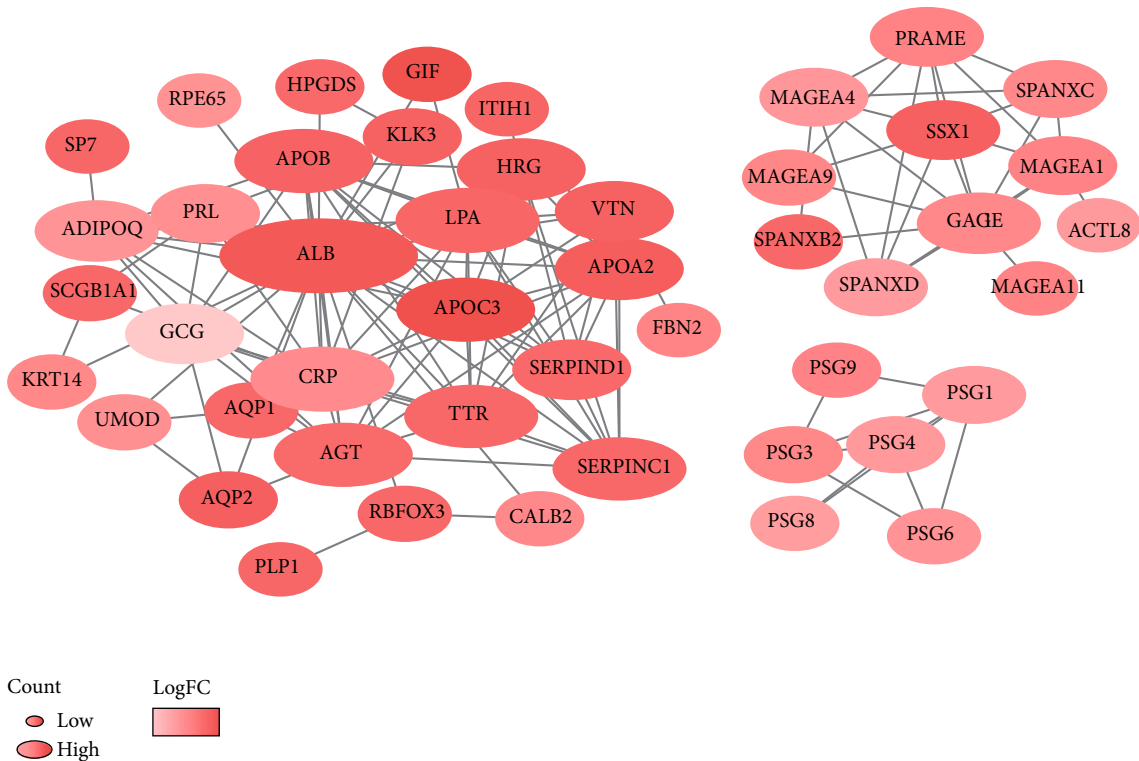


FIGURE 7: PPI network of immune prognostic genes. The color of a node in the PPI network reflects the log(FC) value of gene expression; the size of node indicated the number of interacting proteins with the designated protein.

was built based on the TCGA database and was validated in the GEO dataset, and multicenter clinical data should be collected to validate our model. Secondly, the level of tumor-associated immune cells was estimated using

CIBERSORT method, and future studies should be needed to confirm the findings.

In summary, we provided an in-depth review of TIIC subsets in LUAD, which were closely related to the

development of LUAD and immunotherapy. The immune score signature was constructed based on four survival-related immune markers. In addition, a nomogram was constructed by combining the characteristic immune score signature and meaningful clinical factor, providing an accurate model for predicting the prognosis of patients with LUAD. However, experimental researches on the mechanism between immune cells and LUAD are still needed in the future.

Data Availability

The data are available in the TCGA and GEO datasets.

Conflicts of Interest

The authors have declared that no conflict of interest exists.

Authors' Contributions

Weiyang Cai conceived and designed the experiments. Liping Tao performed the data collection. Weiyang Cai and Yanyan Li analyzed the data and wrote the manuscript. Yanyan Li and Liping Tao contributed equally to this work.

Acknowledgments

This work was supported by the Scientific Research Foundation of the First Affiliated Hospital of Wenzhou Medical University (Grant No. FHY2019002).

Supplementary Materials

Supplementary Figure 1: summary of GEO inferred GEO TIIC composition. (A) Summary of GEO chips concrete immune cell subset proportions of normal lung tissues. (B) Summary of GEO chips concrete immune cell subset proportions of LUAD tissues. (C) The proportion of samples with different *P* value threshold between studies. Supplementary Figure 2: validation of the immunoscore results in other cohorts from the GEO database. Kaplan-Meier survival curves showed that this immunoscore was associated with overall survival in GSE101929 and GSE68571. Supplementary Figure 3: GSEA differentiates the molecular mechanism between the high- and low-immune score group. Table S1: the detailed information of GEO chips. Table S2: the specific markers of immune signatures for ssGSEA. Table S3: the specific construction of PPI interaction. (*Supplementary Materials*)

References

- [1] H. C. Kim, C. Y. Jung, D. G. Cho et al., "Clinical characteristics and prognostic factors of lung cancer in Korea: a pilot study of data from the Korean Nationwide Lung Cancer Registry," *Tuberculosis and Respiratory Diseases*, vol. 82, no. 2, pp. 118–125, 2019.
- [2] H. S. Park, K. B. C. Society, B. J. Chae et al., "Effect of Axillary Lymph Node Dissection after Sentinel Lymph Node Biopsy on Overall Survival in Patients with T1 or T2 Node-positive Breast Cancer: Report from the Korean Breast Cancer Society," *Annals of Surgical Oncology*, vol. 21, no. 4, pp. 1231–1236, 2014.
- [3] H. Borghaei, L. Paz-Ares, L. Horn et al., "Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer," *New England Journal of Medicine*, vol. 373, no. 17, pp. 1627–1639, 2015.
- [4] R. S. Herbst, P. Baas, D. W. Kim et al., "Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial," *Lancet*, vol. 387, no. 10027, pp. 1540–1550, 2016.
- [5] N. Hanna, F. A. Shepherd, F. V. Fossella et al., "Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy," *Journal of Clinical Oncology*, vol. 22, no. 9, pp. 1589–1597, 2004.
- [6] E. S. Nakasone, H. A. Askautrud, T. Kees et al., "Imaging tumor-stroma interactions during chemotherapy reveals contributions of the microenvironment to resistance," *Cancer Cell*, vol. 21, no. 4, pp. 488–503, 2012.
- [7] P. Sharma, S. Hu-Lieskovan, J. A. Wargo, and A. Ribas, "Primary, adaptive, and acquired resistance to cancer immunotherapy," *Cell*, vol. 168, no. 4, pp. 707–723, 2017.
- [8] J. M. Sun, "Response to Costantini et al. "comments on increased response rates to salvage chemotherapy administered after PD-1/PD-L1 inhibitors in patients with non-small cell lung cancer"," *Journal of Thoracic Oncology*, vol. 13, no. 4, pp. e56–e57, 2018.
- [9] A. J. Templeton, M. G. McNamara, B. Šeruga et al., "Prognostic role of neutrophil-to-lymphocyte ratio in solid tumors: a systematic review and meta-analysis," *JNCI: Journal of the National Cancer Institute*, vol. 106, no. 6, 2014.
- [10] M. T. Agullo-Ortuno, O. Gomez-Martin, S. Ponce et al., "Blood predictive biomarkers for patients with non-small-cell lung cancer associated with clinical response to nivolumab," *Clinical Lung Cancer*, vol. 21, no. 1, pp. 75–85, 2020.
- [11] N. L. Syn, M. W. L. Teng, T. S. K. Mok, and R. A. Soo, "De-novo and acquired resistance to immune checkpoint targeting," *The Lancet Oncology*, vol. 18, no. 12, pp. e731–e741, 2017.
- [12] 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.
- [13] 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.
- [14] M. Foroutan, D. D. Bhuva, R. Lyu, K. Horan, J. Cursons, and M. J. Davis, "Single sample scoring of molecular phenotypes," *BMC Bioinformatics*, vol. 19, no. 1, p. 404, 2018.
- [15] 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.
- [16] A. E. Oja, B. Piet, D. van der Zwan et al., "Functional heterogeneity of CD4(+) tumor-infiltrating lymphocytes with a resident memory phenotype in NSCLC," *Frontiers in Immunology*, vol. 9, p. 2654, 2018.
- [17] C. Ngambenjawong, H. H. Gustafson, and S. H. Pun, "Progress in tumor-associated macrophage (TAM)-targeted

- therapeutics,” *Advanced Drug Delivery Reviews*, vol. 114, pp. 206–221, 2017.
- [18] M. Yang, D. McKay, J. W. Pollard, and C. E. Lewis, “Diverse functions of macrophages in different tumor microenvironments,” *Cancer Research*, vol. 78, no. 19, pp. 5492–5503, 2018.
- [19] E. M. Conway, L. A. Pikor, S. H. Y. Kung et al., “Macrophages, inflammation, and lung cancer,” *American Journal of Respiratory and Critical Care Medicine*, vol. 193, no. 2, pp. 116–130, 2016.
- [20] R. Tamura, T. Tanaka, Y. Yamamoto, Y. Akasaki, and H. Sasaki, “Dual role of macrophage in tumor immunity,” *Immunotherapy*, vol. 10, no. 10, pp. 899–909, 2018.
- [21] R. Evans and P. Alexander, “Cooperation of immune lymphoid cells with macrophages in tumour immunity,” *Nature*, vol. 228, no. 5272, pp. 620–622, 1970.
- [22] H. W. Wang and J. A. Joyce, “Alternative activation of tumor-associated macrophages by IL-4: priming for protumoral functions,” *Cell Cycle*, vol. 9, no. 24, pp. 4824–4835, 2010.
- [23] D. F. Quail and J. A. Joyce, “Microenvironmental regulation of tumor progression and metastasis,” *Nature Medicine*, vol. 19, no. 11, pp. 1423–1437, 2013.
- [24] J. Ma, L. Liu, G. Che, N. Yu, F. Dai, and Z. You, “The M1 form of tumor-associated macrophages in non-small cell lung cancer is positively associated with survival time,” *BMC Cancer*, vol. 10, no. 1, 2010.
- [25] J. Mei, Z. Xiao, C. Guo et al., “Prognostic impact of tumor-associated macrophage infiltration in non-small cell lung cancer: a systemic review and meta-analysis,” *Oncotarget*, vol. 7, no. 23, pp. 34217–34228, 2016.
- [26] Z. Y. Yuan, R. Z. Luo, R. J. Peng, S. S. Wang, and C. Xue, “High infiltration of tumor-associated macrophages in triple-negative breast cancer is associated with a higher risk of distant metastasis,” *OncoTargets and Therapy*, vol. 7, pp. 1475–1480, 2014.
- [27] A. Shapouri-Moghaddam, S. Mohammadian, H. Vazini et al., “Macrophage plasticity, polarization, and function in health and disease,” *Journal of Cellular Physiology*, vol. 233, no. 9, pp. 6425–6440, 2018.
- [28] L. Connelly, W. Barham, H. M. Onishko et al., “NF-kappaB activation within macrophages leads to an anti-tumor phenotype in a mammary tumor lung metastasis model,” *Breast Cancer Research*, vol. 13, no. 4, p. R83, 2011.
- [29] M. C. A. Wouters and B. H. Nelson, “Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer,” *Clinical Cancer Research*, vol. 24, no. 24, pp. 6125–6135, 2018.
- [30] D. M. Zhou, Y. X. Xu, L. Y. Zhang et al., “The role of follicular T helper cells in patients with malignant lymphoid disease,” *Hematology*, vol. 22, no. 7, pp. 412–418, 2017.
- [31] X. Meng, Y. Liu, J. Zhang, F. Teng, L. Xing, and J. Yu, “PD-1/PD-L1 checkpoint blockades in non-small cell lung cancer: new development and challenges,” *Cancer Letters*, vol. 405, pp. 29–37, 2017.
- [32] J. Remon, N. Chaput, and D. Planchard, “Predictive biomarkers for programmed death-1/programmed death ligand immune checkpoint inhibitors in nonsmall cell lung cancer,” *Current Opinion in Oncology*, vol. 28, no. 2, pp. 122–129, 2016.
- [33] P. Dhupkar, N. Gordon, J. Stewart, and E. S. Kleinerman, “Anti-PD-1 therapy redirects macrophages from an M2 to an M1 phenotype inducing regression of OS lung metastases,” *Cancer Medicine*, vol. 7, no. 6, pp. 2654–2664, 2018.
- [34] S. R. Gordon, R. L. Maute, B. W. Dulken et al., “PD-1 expression by tumour-associated macrophages inhibits phagocytosis and tumour immunity,” *Nature*, vol. 545, no. 7655, pp. 495–499, 2017.
- [35] Q. Song, J. Shang, Z. Yang et al., “Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma,” *Journal of Translational Medicine*, vol. 17, no. 1, p. 70, 2019.

Research Article

Effect of Different Expression of Immune-Related lncRNA on Colon Adenocarcinoma and Its Relation to Prognosis

Meiwei Mu,¹ Yi Tang,² Zheng Yang ,¹ Yuling Qiu,¹ Xiaohong Li ,¹ Wuning Mo ,¹ and Qisheng Su ¹

¹Department of Clinical Laboratory, First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region, China

²College of Basic Medicine, Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region, China

Correspondence should be addressed to Qisheng Su; 1753359993@qq.com

Received 9 March 2020; Accepted 15 May 2020; Published 6 June 2020

Guest Editor: Tao Huang

Copyright © 2020 Meiwei Mu 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.

Objective. To explore the expression of immune-related lncRNAs in colon adenocarcinoma and find out the effect on how these lncRNAs influence the development and prognosis of colon adenocarcinoma. **Method.** Transcriptome data of colon adenocarcinoma from The Cancer Genome Atlas (TCGA) were downloaded, and gene sets “IMMUNE RESPONSE” and “IMMUNE SYSTEM PROCESS” were sought from the Molecular Signatures Database (MSigDB). The expression of immune-related genes was extracted that were immune-related mRNAs. Then, the immune-related lncRNAs were sought out by utilizing of the above data. Clinical traits were combined with immune-related lncRNAs, so that prognostic-related lncRNAs were identified by Cox regression. Multivariate Cox regression was built to calculate risk scores. Relationships between clinical traits and immune-related lncRNAs were also calculated. **Result.** A total of 480 colorectal adenocarcinoma patients and 41 normal control patients’ transcriptome sequencing data of tissue samples were obtained from TCGA database. 918 immune-related lncRNAs were screened. Cox regression showed that 34 immune-related lncRNAs were associated with colon adenocarcinoma prognosis. Seven lncRNAs were independent risk factors. **Conclusion.** This study revealed that some lncRNAs can affect the development and prognosis of colon adenocarcinoma. It may provide new theory evidence of molecular mechanism for the future research and molecular targeted therapy of colon adenocarcinoma.

1. Background

Colon adenocarcinoma (COAD) is the third most common cancer in the gastrointestinal tract. Most COAD are sporadic and have no genetic or family history. Though the research of the molecular mechanism and pathway has been making some progress, there is no perfect explanation on the origin, development, and metastasis of COAD [1–3]. Long noncoding RNA (lncRNA) is a class of RNAs with a length of more than 200 bases, lacking complete open reading frames. Studies have shown that lncRNA can regulate gene expression at multiple levels, such as transcription, translation, regulation of protein modification, and RNA proteins [4]. lncRNAs influence cellular physiology and pathology, and they are involved in the tumor development [5]. Some former researches found a number of lncRNAs

involved in the development of many diseases: SCHLAPI promotes proliferation and metastasis of prostate cancer by targeting miR-198 and promoting the MAPK1 pathway [6]. And some diseases, such as melanoma, have satisfactory results which are presented in the treatment [7]. In COAD, CASC15 demonstrated to have promoted growth and metastasis through the activation of the Wnt/ β -catenin signaling pathway in a miR-4310/LGR5-dependent manner [8]. And ROR1-AS1 level in COAD tissues was remarkably higher than that in normal tissues [9].

In this study, we attempted to explore the effect of lncRNAs, regulating immune-related processes, on the survival of colorectal adenocarcinoma patients. Based on these lncRNAs, we constructed a model to predict the risk of death of COAD patients. We found out that the relationship between these immune-related lncRNAs and prognosis of

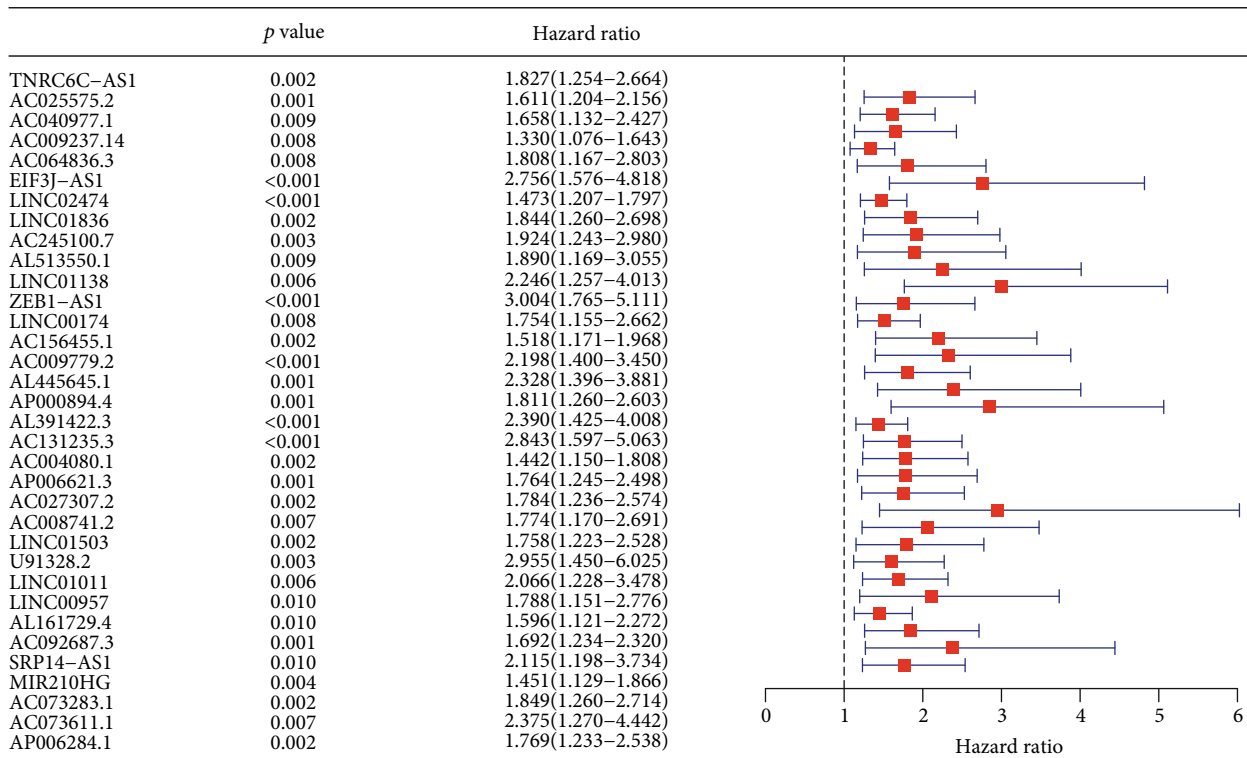


FIGURE 1: Univariate Cox regression showed that 34 immune-related lncRNAs were risk factors for COAD patients (hazard ratio > 1).

COAD are of great importance and may be helpful in building a new way to treat COAD.

2. Materials and Methods

2.1. Acquisition of Colon Adenocarcinoma Data from TCGA. COAD transcriptome expression data in the format of Fragments Per Kilobase Million (FPKM) of The Cancer Genome Atlas (TCGA) were downloaded from Genomic Data Commons (GDC, <https://gdc.cancer.gov/>), which are open to the public. Clinical traits were also downloaded. Among these, there were 480 colorectal adenocarcinoma patients and 41 normal control patients.

2.2. Acquisition of Immune-Related Genes from MSigDB. In order to extract immune-related lncRNAs, two gene sets “IMMUNE RESPONSE” and “IMMUNE SYSTEM PROCESS” were downloaded from the Molecular Signatures Database (MSigDB, <https://www.gsea-msigdb.org/gsea/msigdb/>). Then, immune-related mRNA expression was extracted from the above two gene sets.

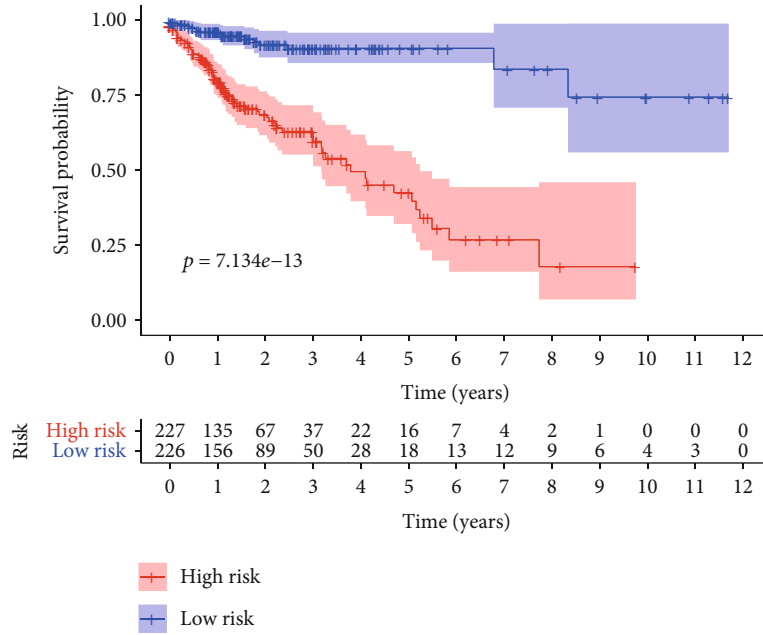
2.3. Immune-Associated lncRNAs. Through Pearson’s method, correlations between the mRNAs and lncRNAs were calculated. lncRNAs with the absolute value of correlation coefficient > 0.4 and *p* < 0.001 were considered as immune-related lncRNAs for further analysis.

2.4. Cox Regression. Immune-related lncRNA expression matrix was merged with clinical traits. Univariate Cox regression was performed on the above data to extract immune-related lncRNAs which are associated with survival time,

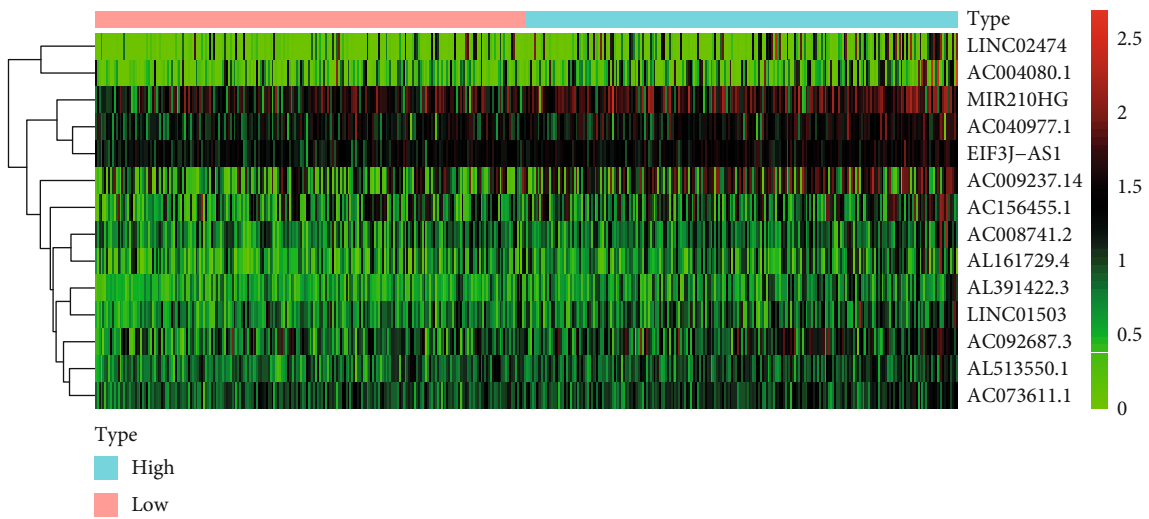
TABLE 1: The prognosis model of COAD patients was constructed with immune-related lncRNA based on multivariate Cox regression.

| ID | coef | HR |
|-------------|----------|----------|
| AC040977.1 | 0.487058 | 1.627521 |
| AC009237.14 | 0.372021 | 1.450663 |
| EIF3J-AS1 | 0.630542 | 1.878628 |
| LINC02474 | 0.450276 | 1.568745 |
| AL513550.1 | 0.605145 | 1.831517 |
| AC156455.1 | 0.356569 | 1.42842 |
| AL391422.3 | 0.550559 | 1.734222 |
| AC004080.1 | 0.348324 | 1.416691 |
| AC008741.2 | -1.07714 | 0.340569 |
| LINC01503 | 0.614045 | 1.847891 |
| AL161729.4 | 0.744444 | 2.105272 |
| AC092687.3 | 0.305246 | 1.356959 |
| MIR210HG | 0.513543 | 1.671202 |
| AC073611.1 | 0.523188 | 1.687398 |

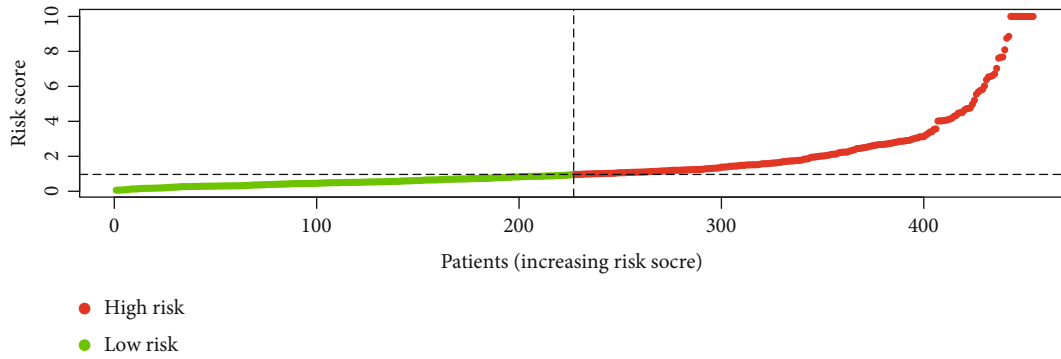
and *p* < 0.01 was used as threshold. All these lncRNAs were divided into two categories: high-risk lncRNAs with HR > 1 and low-risk lncRNAs with HR < 1. A stepwise multivariate Cox regression was conducted. The risk score of each patient was calculated according to the multivariate Cox regression. The formula can be expressed as follows: risk score = $\beta_{\text{gene1}} \times \text{Expression}_{\text{gene1}} + \beta_{\text{gene2}} \times \text{Expression}_{\text{gene2}} + \beta_{\text{gene3}} \times \text{Expression}_{\text{gene3}} + \dots + \beta_{\text{genen}} \times \text{Expression}_{\text{genen}}$. Patients were divided into two groups according to the median value



(a)



(b)



(c)

FIGURE 2: Continued.

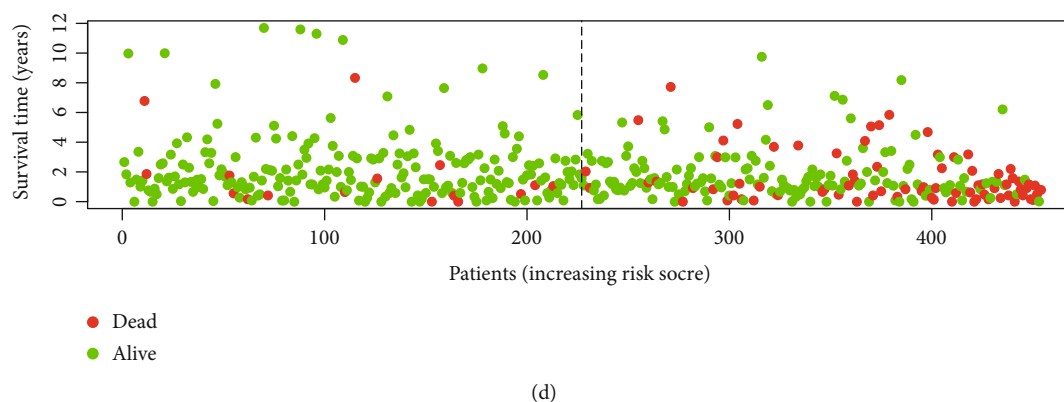


FIGURE 2: The survival curve shows that patients in the high-risk score group have a shorter survival time than those in the low-risk score group (a). The heat map of prognosis-related lncRNAs in COAD (b). Risk curves were drawn based on the data of 480 patients. Take the risk scores as the vertical coordinate and the patient data as the horizontal coordinate. We can see the risk scores were increasing, in which the low-risk curve was in green whereas the high-risk curve was in red (c). Green dots mean people who were still alive, and red dots mean people who were already dead. As the risk scores were increasing, the deaths were added; the red dots were concentrated in the high-risk area. And in the low-risk area, where the green dots were concentrated, patients are almost still alive (d).

of risk scores. If the risk score was higher than the median, it was assigned to the high-risk group; otherwise, it was assigned to the low-risk group. The survival curve and risk curve were drawn based on risk scores and survival information.

2.5. Independent Prognostic Analysis. Comparing clinical traits and risk scores with survival status and survival time, the risk score was determined whether it can be used as an independent prognostic factor. Single factor and multifactorial analysis were performed to judge if these factors can be used as independent prognostic factors. The receiver operator characteristic (ROC) curve was drawn to determine whether the result was reliable.

2.6. Clinical Relevance Analysis. A box chart was drawn to analyze whether clinical traits were associated with lncRNAs in the risk prediction model.

3. Result

3.1. Prognosis-Related Immune lncRNA. 329 immune-related mRNAs were obtained. 918 immune-related lncRNAs were obtained through Pearson's correlation. The expression of immune-related lncRNAs were extracted and fused with clinical survival data. 34 immune-related lncRNAs were associated with the survival of COAD ($p < 0.01$). Among them, all of these prognosis-related lncRNAs were classified as high risk factors (Figure 1).

3.2. Cox Regression. Based on the above prognosis-related immune lncRNAs, the Cox regression model was built (Table 1). Next, all risk scores were calculated according to the Cox regression model. All COAD patients were divided by the median value of risk scores: high-risk group and low-risk group. According to the two groups, survival curves were drawn. The chart showed that the survival probability of the low-risk group was higher than that of the high-risk group. The chart below confirmed this conclusion, because patients' survival time in the high-risk group was evidently

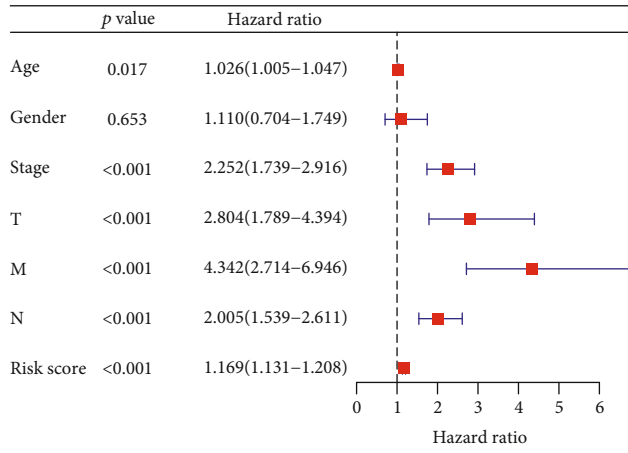
lower than that in the low-risk group. The number of deaths in the high-risk group increased faster than that in the low-risk group. By 11th year, only the patients in the low-risk group still survived (Figure 2).

3.3. Independent Prognostic Analysis. In univariate Cox regression, clinical stage, T stage, M stage, N stage, and risk score were thought to be associated with survival time in COAD patients ($p < 0.001$). Among them, age and risk score were independent prognostic factors in multivariate Cox regression ($p < 0.001$), but gender, clinical stage, T stage, M stage, and N stage were not. The ROC was drawn, and the area under the curve (AUC) was calculated to confirm the reliability of this conclusion. The area under the curve of gender was 0.444; all of the other factors including the risk score, age, clinical stage, T stage, M stage, and N stage were over 0.500. Among them, the AUC of risk score was 0.776 (Figure 3).

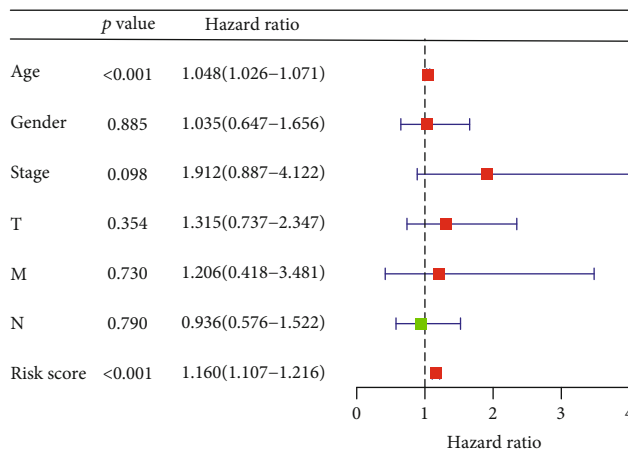
3.4. Clinical Relevance Analysis. A box chart was drawn to analyze whether lncRNAs are associated with clinical traits. The following results were obtained: T stage was associated with AC004080.1, AC040977.1, and AC073611.1; M stage was associated with AC004080.1, AC008741.2, AC009237.14, AC092687.3, AL161729.4, 29.4AL391422.3, and EIF3J-AS1 3; N stage was associated with AC004080.1, AC008741.2, AC009237.14, AC092687.3 AL391422.3, EIF3J-AS1, and LINC02474 4; and clinical stage was associated with AC008741.2, AC009237.14, AC092687.3, AC156455.1, AL391422, EIF3J-AS, and LINC0247 (Figure 4).

4. Discussion

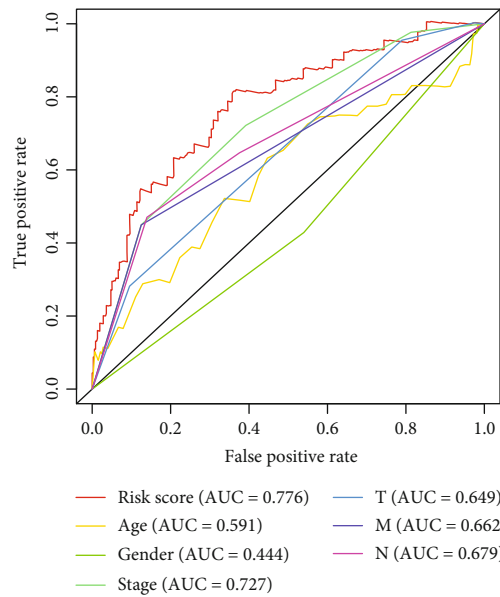
The usual treatment of COAD is surgical treatment, which still has the insufficiency. In recent years, emerging immunotherapy methods such as PD-1/PD-L1 inhibitor which targeted at immune checkpoints have developed rapidly in the treatment of tumor and achieved a certain curative effect [10, 11]. These indicate that immune intervention has a



(a)

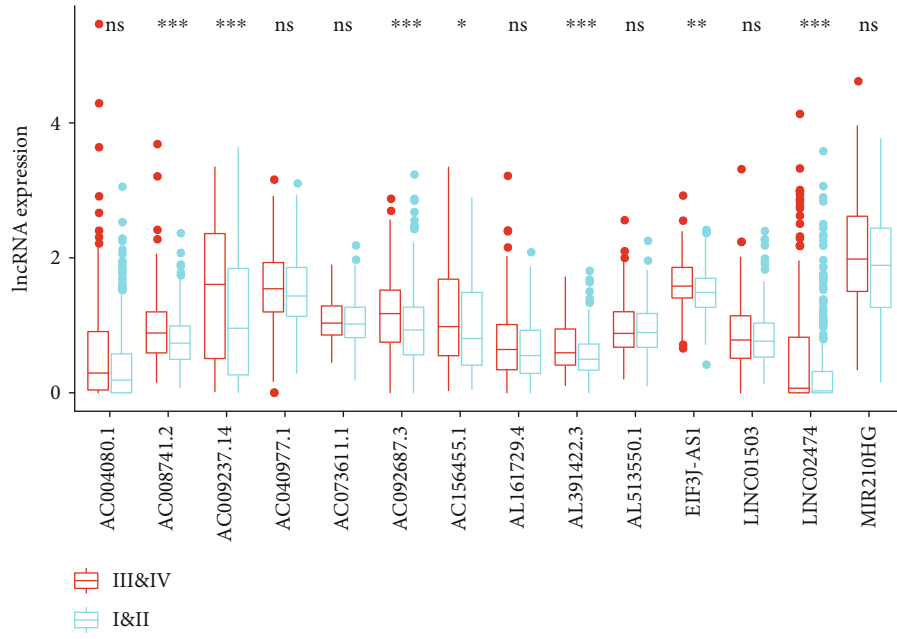


(b)

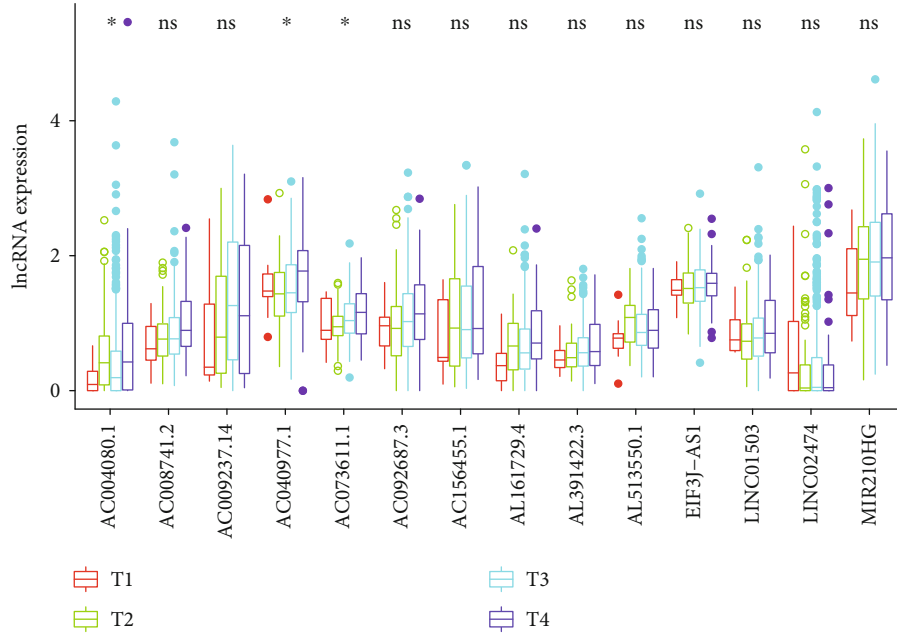


(c)

FIGURE 3: Univariate Cox regression showed that age, clinical stage, T stage, N stage, M stage, and risk score were all correlated with the prognosis of COAD patients (a). Multivariate Cox regression showed that only the risk score was an independent risk factor for COAD (b), and the ROC curve showed that the risk score had the largest area under the curve (c).



(a)



(b)

FIGURE 4: Continued.

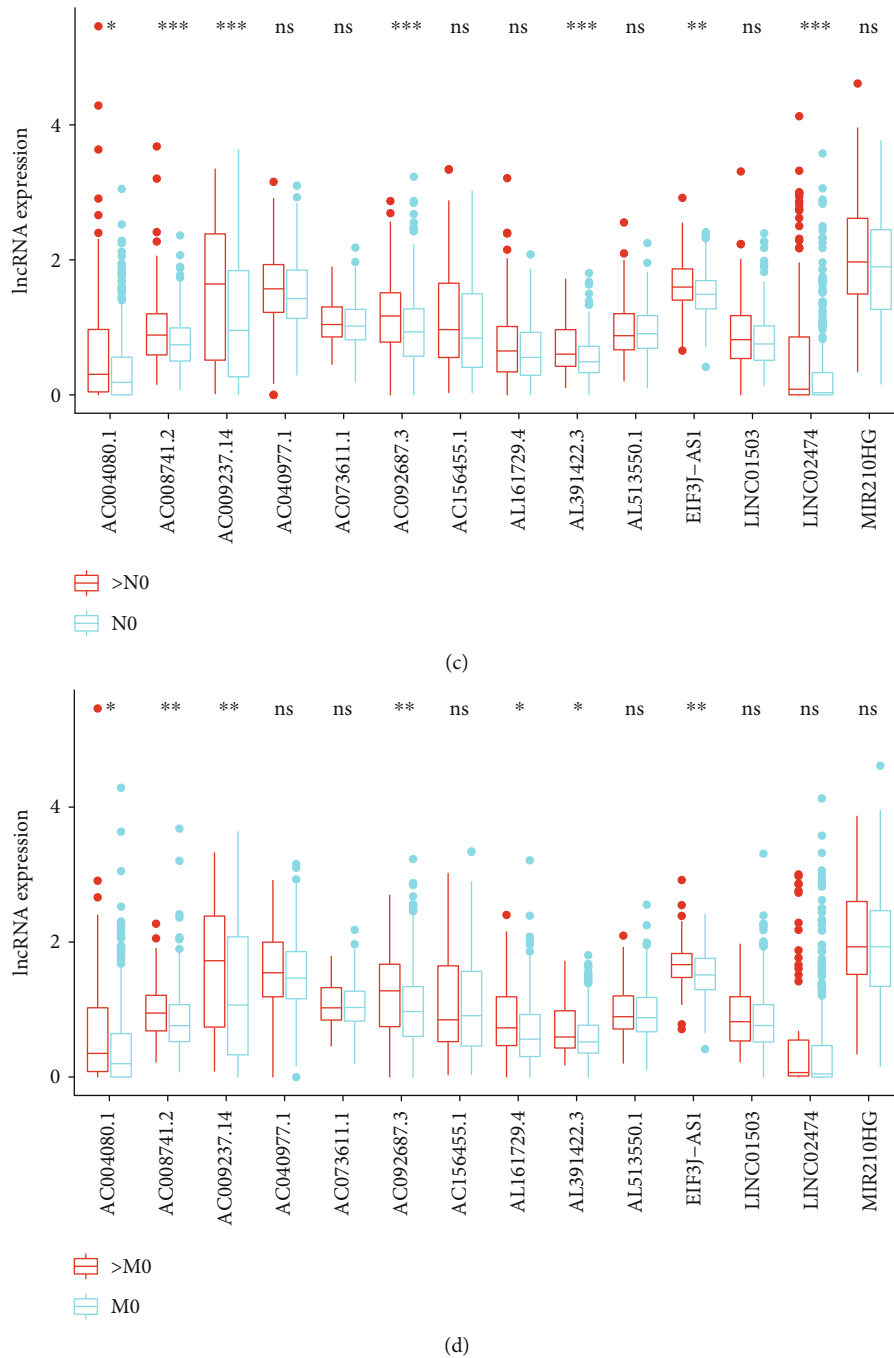


FIGURE 4: Multiple immune-related lncRNAs were associated with clinical traits (* means they had relevance, and the more they had, the greater the correlation; “ns” means they had no relevance). (a) Relationship between immune-related lncRNAs and clinical stages. (b) Relationship between immune-related lncRNAs and T stages. (c) Relationship between immune-related lncRNAs and N stages. (d) Relationship between immune-related lncRNAs and M stages.

broad application prospect in tumor therapy. However, there is much room for improvement in the management of colon cancer at present, and the prognosis of patients with advanced COAD is still very limited by the existing treatments. Therefore, new strategies for the treatment and management of COAD are necessary. Based on these current situations, we attempted to analyze the transcriptome sequencing data of COAD in TCGA, a public database. Because lncRNAs play a wide range of regulatory roles in

various biological processes of the organism, we focused our research on lncRNAs and tried to correlate with immunity. We speculate that immune-related lncRNAs may be involved in important regulation in the immune process of the body against COAD.

In this study, we extracted the expression amount of lncRNAs regulating immune gene expression in COAD patients. The results of univariate Cox analysis showed that 34 immune-related lncRNAs were significantly associated

with the survival time of patients with COAD, which is consistent with our view that lncRNAs may be involved in the regulation of the body's antitumor immune process. The results of multivariate Cox regression showed that the formula containing 14 immune-related lncRNA expressions could well distinguish patients with good prognosis from those with poor prognosis. This index has advantages over widely used clinical stages, T stage, N stage, M stage, and age. We believe that this formula can improve the management of patients with COAD.

Among 14 immune-related lncRNAs, EIF3J-AS1 was reported to promote COAD cell proliferation and inhibit apoptosis [12]. Our study found that EIF3J-AS1 was associated with the expression of immune genes in COAD, and its high expression predicted shorter survival time. LINC01503 and LINC02474 were confirmed to be associated with the survival of COAD in bioinformatics-based studies [13, 14], and LINC01503 could regulate the expression of miR4492/FOXK1 signaling which promotes proliferation and invasion of cells of COAD [15]. MIR210HG has been experimentally demonstrated to promote proliferation and invasion of tumor cells in hepatocellular carcinoma, non-small-cell lung cancer, breast cancer, cervical cancer, osteosarcoma, and other tumors. Its high expression is associated with poor prognosis of cancer patients [16–20]. He et al. and Ruan et al. found that they also predicted poor prognosis in COAD [21, 22], which is consistent with the results of our study, suggesting that our prediction of their survival in COAD is reliable. The remaining lncRNAs have not been studied in COAD to our knowledge, and we found that they are also associated with the survival of patients with COAD. Their association with clinical stage, T stage, M stage, and N stage also suggests that they may contribute to the malignant development of COAD. Their roles and mechanisms in tumors are still unknown, and we look forward to future experimental studies on their roles in COAD that will show us their mechanisms in COAD and their role in tumor immune response, which we believe will be helpful and complementary to tumor immunotherapy.

5. Conclusion

This study identified 329 immune-related lncRNAs in colon adenocarcinoma. And 34 prognosis-related lncRNAs were obtained by univariate Cox regression. The risk score calculated by 14 immune-related lncRNAs was identified as an independent prognostic factor. Compared with clinical stages, T stages, N stages, M stages, and age, the risk score is the best in predicting poor prognosis of patients. The identification of prognostic immune-related lncRNAs and the discovery of the relationship between these lncRNAs and prognosis of COAD may provide a new direction for future research and provide a new target for a molecular targeted therapy of COAD.

Data Availability

The data used in this study are from open public databases, and how to obtain them has been explained in the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Y. Zheng, C. Yang, S. Tong et al., "Genetic variation of long non-coding RNA TINCR contribute to the susceptibility and progression of colorectal cancer," *Oncotarget*, vol. 8, no. 20, pp. 33536–33543, 2017.
- [2] C. Hanisch, J. Sharbati, B. Kutz-Lohroff, O. Huber, R. Einspanier, and S. Sharbati, "TFF3-dependent resistance of human colorectal adenocarcinoma cells HT-29/B6 to apoptosis is mediated by miR-491-5p regulation of lncRNA PRINS," *Cell Death Discovery*, vol. 3, no. 1, p. 16106, 2017.
- [3] P. Han, J.-w. Li, B.-m. Zhang et al., "The lncRNA CRNDE promotes colorectal cancer cell proliferation and chemoresistance via miR-181a-5p-mediated regulation of Wnt/ β -catenin signaling," *Molecular Cancer*, vol. 16, no. 1, p. 9, 2017.
- [4] W.-X. Peng, P. Koirala, and Y.-Y. Mo, "lncRNA-mediated regulation of cell signaling in cancer," *Oncogene*, vol. 36, no. 41, pp. 5661–5667, 2017.
- [5] F. Ferrè, A. Colantoni, and M. Helmer-Citterich, "Revealing protein-lncRNA interaction," *Briefings in Bioinformatics*, vol. 17, no. 1, pp. 106–116, 2016.
- [6] Y. Li, H. Luo, N. Xiao, J. Duan, Z. Wang, and S. Wang, "Long noncoding RNA SCHLAP1 accelerates the proliferation and metastasis of prostate cancer via targeting miR-198 and promoting the MAPK1 pathway," *Oncology Research*, vol. 26, no. 1, pp. 131–143, 2018.
- [7] M. A. Curran, W. Montalvo, H. Yagita, and J. P. Allison, "PD-1 and CTLA-4 combination blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 9, pp. 4275–4280, 2010.
- [8] N. Jing, T. Huang, H. Guo et al., "lncRNA CASC15 promotes colon cancer cell proliferation and metastasis by regulating the miR-4310/LGR5/Wnt/ β -catenin signaling pathway," *Molecular Medicine Reports*, vol. 18, no. 2, pp. 2269–2276, 2018.
- [9] X.-Y. Wang, X. Jian, B.-Q. Sun, X.-S. Ge, F.-J. Huang, and Y.-Q. Chen, "lncRNA ROR1-AS1 promotes colon cancer cell proliferation by suppressing the expression of DUSP5/CDKN1A," *European Review for Medical and Pharmacological Sciences*, vol. 24, no. 3, pp. 1116–1125, 2020.
- [10] A. S. Jakola, K. S. Myrnes, R. Kloster et al., "Comparison of a strategy favoring early surgical resection vs a strategy favoring watchful waiting in low-grade gliomas," *Journal of the American Medical Association*, vol. 308, no. 18, pp. 1881–1888, 2012.
- [11] M. K. Aghi, B. V. Nahed, A. E. Sloan, T. C. Ryken, S. N. Kalkanis, and J. J. Olson, "The role of surgery in the management of patients with diffuse low grade glioma," *Journal of Neuro-Oncology*, vol. 125, no. 3, pp. 503–530, 2015.
- [12] D. Liu, H. Zhang, J. Cong et al., "H3K27 acetylation-induced lncRNA EIF3J-AS1 improved proliferation and impeded apoptosis of colorectal cancer through miR-3163/YAP1 axis," *Journal of Cellular Biochemistry*, vol. 121, no. 2, pp. 1923–1933, 2019.
- [13] X. Wang, J. Zhou, M. Xu et al., "A 15-lncRNA signature predicts survival and functions as a ceRNA in patients with colorectal cancer," *Cancer Management and Research*, vol. Volume 10, pp. 5799–5806, 2018.

- [14] H. Yang, H.-C. Lin, H. Liu et al., "A 6 lncRNA-Based Risk Scoresystem for Predicting the Recurrence of Adenocarcinoma Patients," *Frontiers in Oncology*, vol. 10, p. 81, 2020.
- [15] S.-. R. Lu, Q. Li, J.-. L. Lu, C. Liu, X. Xu, and J.-. Z. Li, "Long non-coding RNA LINC01503 promotes colorectal cancer cell proliferation and invasion by regulating miR-4492/FOXK1 signaling," *Experimental and Therapeutic Medicine*, vol. 16, no. 6, pp. 4879–4885, 2018.
- [16] A.-H. Wang, C.-H. Jin, G.-Y. Cui et al., "MIR210HG promotes cell proliferation and invasion by regulating miR-503-5p/TRAF4 axis in cervical cancer," *Aging*, vol. 12, no. 4, pp. 3205–3217, 2020.
- [17] X.-Y. Li, L.-Y. Zhou, H. Luo et al., "The long noncoding RNA MIR210HG promotes tumor metastasis by acting as a ceRNA of miR-1226-3p to regulate mucin-1c expression in invasive breast cancer," *Aging*, vol. 11, no. 15, pp. 5646–5665, 2019.
- [18] X. Kang, F. Kong, K. Huang et al., "LncRNA MIR210HG promotes proliferation and invasion of non-small cell lung cancer by upregulating methylation of CACNA2D2 promoter via binding to DNMT1," *OncoTargets and Therapy*, vol. Volume 12, pp. 3779–3790, 2019.
- [19] Y. Wang, W. Li, X. Chen, Y. Li, P. Wen, and F. Xu, "MIR210HG predicts poor prognosis and functions as an oncogenic lncRNA in hepatocellular carcinoma," *Biomedicine & Pharmacotherapy*, vol. 111, pp. 1297–1301, 2019.
- [20] J. Li, Q.-M. Wu, X.-Q. Wang, and C.-Q. Zhang, "Long noncoding RNA miR210HG sponges miR-503 to facilitate osteosarcoma cell invasion and metastasis," *DNA and Cell Biology*, vol. 36, no. 12, pp. 1117–1125, 2017.
- [21] Z. He, J. Dang, A. Song, X. Cui, Z. Ma, and Z. Zhang, "Identification of LINC01234 and MIR210HG as novel prognostic signature for colorectal adenocarcinoma," *Journal of Cellular Physiology*, vol. 234, no. 5, pp. 6769–6777, 2019.
- [22] Z. Ruan, Z. Xu, Z. Li, and Y. Lv, "Integral analyses of survival-related long non-coding RNA MIR210HG and its prognostic role in colon cancer," *Oncology Letters*, vol. 18, no. 2, pp. 1107–1116, 2019.