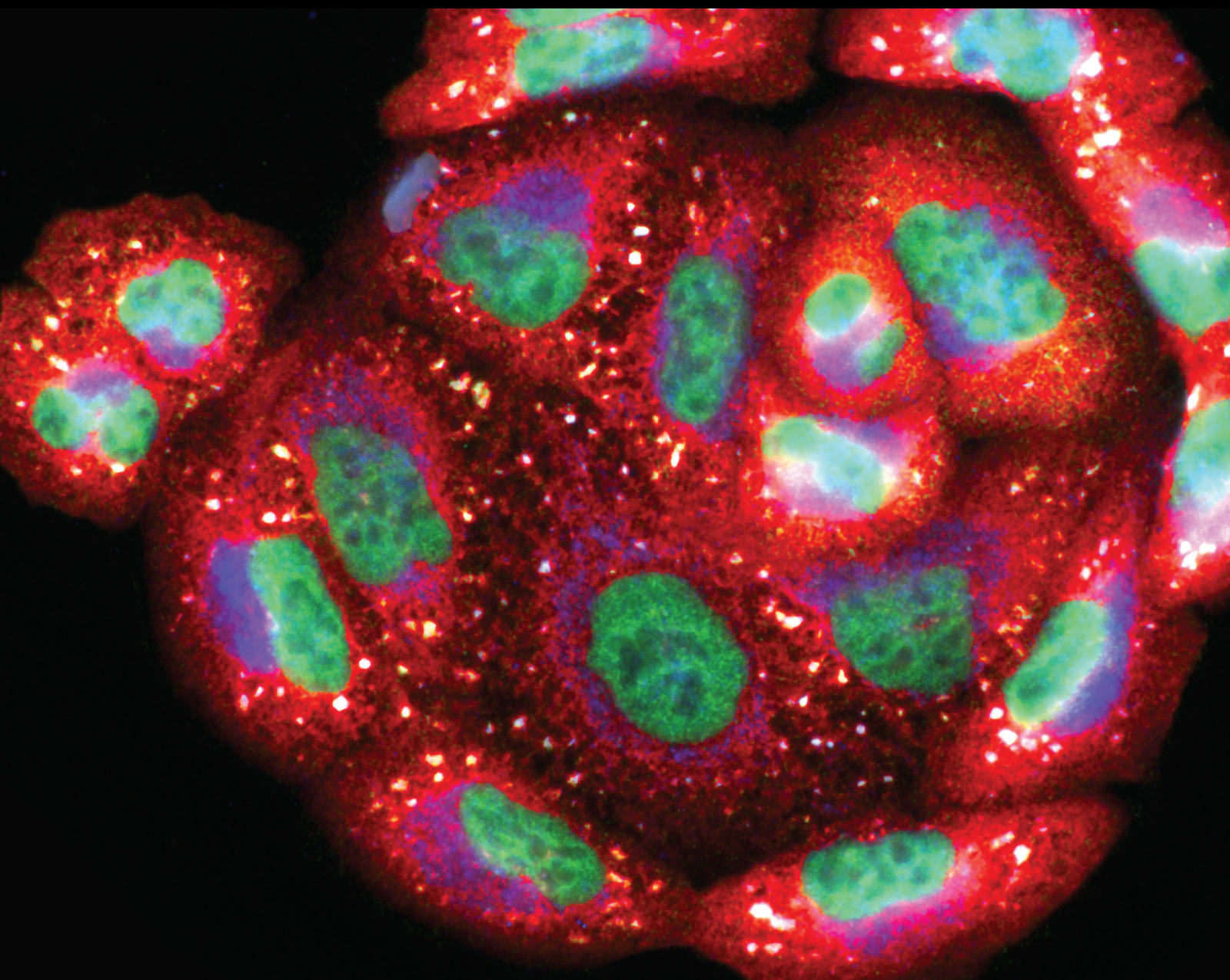# Translational Medicine and Redox Biology in the Diagnosis and Treatment of Tumors

Lead Guest Editor: Pei Jiang
Guest Editors: Liang Hu, Qian Yuan, and Wenjie Shi

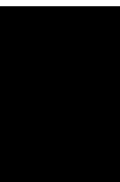# Translational Medicine and Redox Biology in the Diagnosis and Treatment of Tumors

# Translational Medicine and Redox Biology in the Diagnosis and Treatment of Tumors

Lead Guest Editor: Pei Jiang
Guest Editors: Liang Hu, Qian Yuan, and Wenjie Shi

Tim Hofer, Norway
John D. Horowitz, Australia
Silvana Hrelia, Italy
Dragan Hrncic, Serbia
Zebo Huang, China
Zhao Huang, China
Tarique Hussain, Pakistan
Stephan Immenschuh, Germany
Norsharina Ismail, Malaysia
Franco J. L, Brazil
Sedat Kacar, USA
Andleeb Khan, Saudi Arabia
Kum Kum Khanna, Australia
Neelam Khaper, Canada
Ramoji Kosuru, USA
Demetrios Kouretas, Greece
Andrey V. Kozlov, Austria
Chan-Yen Kuo, Taiwan
Gaocai Li, China
Guoping Li, USA
Jin-Long Li, China
Qiangqiang Li, China
Xin-Feng Li, China
Jialiang Liang, China
Adam Lightfoot, United Kingdom
Christopher Horst Lillig, Germany
Paloma B. Liton, USA
Ana Lloret, Spain
Lorenzo Loffredo, Italy
Camilo López-Alarcón, Chile
Daniel Lopez-Malo, Spain
Massimo Lucarini, Italy
Hai-Chun Ma, China
Nageswara Madamanchi, USA
Kenneth Maiese, USA
Marco Malaguti, Italy
Steven McAnulty, USA
Antonio Desmond McCarthy, Argentina
Sonia Medina-Escudero, Spain
Pedro Mena, Italy
Víctor M. Mendoza-Núñez, Mexico
Lidija Milkovic, Croatia
Alexandra Miller, USA
Sara Missaglia, Italy

Premysl Mladenka, Czech Republic
Sandra Moreno, Italy
Trevor A. Mori, Australia
Fabiana Morroni, Italy
Ange Mouithys-Mickalad, Belgium
Iordanis Mourouzis, Greece
Ryoji Nagai, Japan
Amit Kumar Nayak, India
Abderrahim Nemmar, United Arab
Emirates
Xing Niu, China
Cristina Nocella, Italy
Susana Novella, Spain
Hassan Obied, Australia
Pál Pacher, USA
Pasquale Pagliaro, Italy
Dilipkumar Pal, India
Valentina Pallottini, Italy
Swapnil Pandey, USA
Mayur Parmar, USA
Vassilis Paschalis, Greece
Keshav Raj Paudel, Australia
Ilaria Peluso, Italy
Tiziana Persichini, Italy
Shazib Pervaiz, Singapore
Abdul Rehman Phull, Republic of Korea
Vincent Pialoux, France
Alessandro Poggi, Italy
Zsolt Radak, Hungary
Dario C. Ramirez, Argentina
Erika Ramos-Tovar, Mexico
Sid D. Ray, USA
Muneeb Rehman, Saudi Arabia
Hamid Reza Rezvani, France
Alessandra Ricelli, Italy
Francisco J. Romero, Spain
Joan Roselló-Catafau, Spain
Subhadeep Roy, India
Josep V. Rubert, The Netherlands
Sumbal Saba, Brazil
Kunihiro Sakuma, Japan
Gabriele Saretzki, United Kingdom
Luciano Saso, Italy
Nadja Schroder, Brazil

Anwen Shao, China
Iman Sherif, Egypt
Salah A Sheweita, Saudi Arabia
Xiaolei Shi, China
Manjari Singh, India
Giulia Sita, Italy
Ramachandran Srinivasan, India
Adrian Sturza, Romania
Kuo-hui Su, United Kingdom
Eisa Tahmasbpour Marzouni, Iran
Hailiang Tang, China
Carla Tatone, Italy
Shane Thomas, Australia
Carlo Gabriele Tocchetti, Italy
Angela Trovato Salinaro, Italy
Rosa Tundis, Italy
Kai Wang, China
Min-qi Wang, China
Natalie Ward, Australia
Grzegorz Wegrzyn, Poland
Philip Wenzel, Germany
Guangzhen Wu, China
Jianbo Xiao, Spain
Qiongming Xu, China
Liang-Jun Yan, USA
Guillermo Zalba, Spain
Jia Zhang, China
Junmin Zhang, China
Junli Zhao, USA
Chen-he Zhou, China
Yong Zhou, China
Mario Zoratti, Italy

# Contents

*Research Article*

# Machine Learning Predicts the Oxidative Stress Subtypes Provide an Innovative Insight into Colorectal Cancer

**Haitao Zhong,[1] Le Yang,[1] Qingshang Zeng,[2] Weidong Chen,[1] Haibo Zhao,[1] Linlin Wu [ID],[3] Lei Qin [ID],[1] and Qing-Qing Yu [ID][1]**

[1]*Jining First People's Hospital, Jining Medical University, Jining 272000, China*
[2]*Shanghai Tianyou Hospital, Tongji University, Shanghai 200333, China*
[3]*Department of Oncology, Tengzhou Central People's Hospital Affiliated to Jining Medical College, Tengzhou 277500, China*

Correspondence should be addressed to Linlin Wu; linlinliff@163.com, Lei Qin; qinlei007@126.com,
and Qing-Qing Yu; yuqingqing_lucky@163.com

So far, it has been reached the academic consensus that the molecular subtypes are via genomic heterogeneity and immune infiltration patterns. Considering that oxidative stress (OS) is involved in tumorigenesis and prognosis prediction, we propose an innovative classification of colorectal cancer- (CRC-) OS subtypes. We obtain three datasets from The Cancer Genome Atlas Program (TCGA) and Gene Expression Omnibus (GEO) online databases. 1399 OS-related genes were selected from the GeneCards database. We remove the batch effect before conducting differentially expressed genes (DEGs) analyses between normal and tumor samples. Nonnegative matrix factorization (NMF) was used to perform an unsupervised cluster. Lasso regression and Cox regression were used to construct the signature model. DEGs, robust rank aggregation, and protein-protein interaction networks were used to select hub genes, and then use hub genes to predict OS subtypes by random forest algorithms. NMF identifies two OS-related subtypes of CRC patients. Eight OS-related gene signatures were built to predict the outcome of patients, based on the DEGs between two subtypes. A total of 61 DEGs overlap each dataset, and the RRA analysis shows that 17 genes are important in these three datasets, and 15 genes are shared genes between the two methods. PPI network suggests that five hub genes are confirmed, they are SPP1, SERPINE1, CAV1, PDGFRB, and PLAU. These five hub genes could predict the OS-related subtype of CRC accurately with AUC equal to 0.771. In our study, we identify two OS-related subtypes, which will provide an innovative insight into colorectal cancer.

## 1. Introduction

Nowadays, with an estimated 1,800,000 new cases and 900,000 deaths annually [1, 2], colorectal cancer (CRC) become the third most common cancer and the second leading cause of cancer death [3]. Despite rapid development in the diagnosis and treatments of CRC, the mortality remains high, especially in advanced stage at first diagnosis [4]. Therefore, the lack of biomarkers for early screening and prognosis prediction is still an urgent clinical problem to improve the treatments efficacy and reduce the cases mortality of CRC.

With the numerous studies on hallmarks of cancer, the characteristics of genomic variation in CRC have unique clonal, stromal, and immune dependencies [5]. So far, a molecular classification of CRC has been reached the academic consensus than the four molecular subtype groups via the current best description of the genomic heterogeneity [6]. In addition to the transcriptomic subtypes of CRC, the expression profile analysis of CRC showed that the immune infiltration patterns with different immune-tolerant microenvironment resulted in different effects of special immunotherapy [7, 8]. However, it is more notable that to maintain the high proliferation rate tumor cells, it demands high ROS concentrations, which the regulation of oxidative stress (OS) includes oxidative metabolism, for example, the conversion of the glycolytic pathway into the pentose phosphate pathway [9, 10]. And the prognosis of radiotherapy and
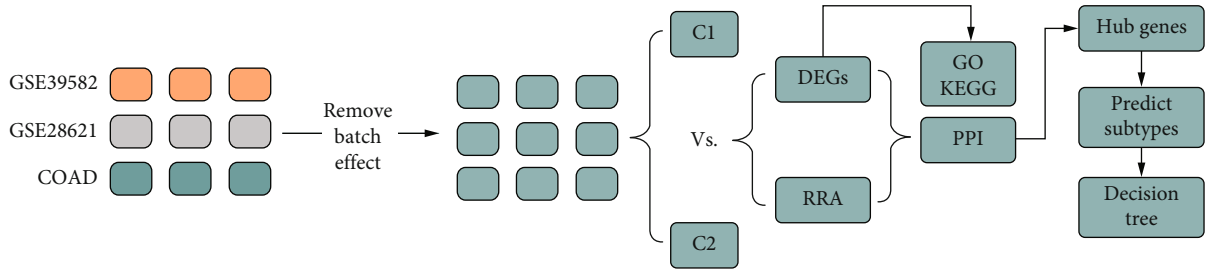
FIGURE 1: Workflow of the study. Three datasets (COAD, GSE39583, and GSE 28621) were obtained from TCGA and GEO including 1399 OS-related genes selected from the GeneCards database. Before DEGs analyses, the batch effect was removed. And nonnegative matrix factorization (NMF) was used to perform an unsupervised cluster. Hub genes were selected by Lasso regression and Cox regression to construct the signature model. DEGs, robust rank aggregation, protein-protein interaction networks were used to select hub genes to predict OS subtypes by random forest algorithms.

chemotherapy treatments is influenced through OS modulation indicating that the OS is also highly significant for cancer therapy [11–14]. Therefore, we propose an innovative classification of CRC-OS subtypes, which has never been studied in-depth.

Machine learning is a computer technology applied artificial intelligence which has a widespread application in improving medical research and clinical decision in clinic such as diagnostics, precision medicine, and clinical trials of cancer. In our study, we aimed at identifying potential OS molecular subtypes and predicting CRC outcomes through the gene expression profile in multiple datasets. By utilizing the nonnegative matrix factorization (NMF) clustering algorithm, 350 OS-related differentially expressed genes (DEGs) were distinctly classified into two molecular subtypes (named C1 and C2) in three CRC cohorts. Among them, C1 was associated with a better prognosis. Moreover, based on the intersection of C1 and C2 DEGs, we established a novel OS-related prognostic signature by multivariate Cox regression model and validated its significant prognostic values for CRC patients. Finally, we also explored the hub genes for predicting OS subtypes in CRC.

## 2. Materials and Methods

*2.1. Data Obtain.* We obtain datasets from three individual datasets, colon adenocarcinoma (COAD) from The Cancer Genome Atlas Program (TCGA), including 41 normal samples and 473 tumor samples; GSE39582, including 19 normal samples and 443 tumor samples; GSE29621, including 65 tumor samples from Gene Expression Omnibus (GEO), respectively. Gene expression profile and clinic information of the above samples were downloaded by R package TCGA biolinks [15]. Oxidative stress- (OS-) related genes were checked by the GeneCards database (https://www.genecards.org/) and 1399 OS-related genes were selected for future analysis.

*2.2. Batch Effect Correction.* Because the above datasets resource from different individual databases, we use R package sva to reduce the batch effect between samples. We merge these three datasets and calculate the overlap genes, then use the combat method to remove the batch effect.

*2.3. OS-Related Genes' Different Expressions.* We extract 1399 OS-related genes from the above merge expression matrix which has been removed batch effect and then use the limma package to conduct differential expression analysis between normal and tumor samples. Absolute value of LogFC was set to more than 1, and the $p$ value was set to less than 0.05.

*2.4. Nonnegative Matrix Factorization Identify OS-Related Subtype.* Nonnegative matrix factorization (NMF) is an excellent unsupervised learning algorithm that could identify the probable subtype between large samples. Here, we also use this method and aim to identify candidate OS-related subtypes in CRC samples. We set the rank from 2 to 10 using a method called brunet, and other algorithm parameters were set as default.

*2.5. Subtypes Validate and Survival Analysis.* After identifying candidate subtypes, we also use heat map and principal component analysis (PCA) to validate the results of the typing results. Survival analysis is also used to compare different subgroups. The survival difference was calculated by a log-rank test.

*2.6. Different Expression Genes between Subtypes.* To explore the potential mechanisms between subtypes and build a useful prediction model, we use the limma package to conduct differential expression analysis between groups. Here, we set the absolute value of LogFC as more than 0.5, and the $p$ value also was set to less than 0.05.

*2.7. Lasso Regression and Cox Regression to Identify Model Genes.* A useful machine learning method, Lasso regression and Cox regression model were used to dimensionalize the data. We use the batch univariate Cox regression model to obtain the prognosis OS-related genes, and then we input the above results into Lasso regression and calculate the minimum value of lambda to get the important genes of the model. Next, the important genes of Lasso regression results will perform the last step, and multivariate Cox regression will conduct this procedure to select the candidate model genes.
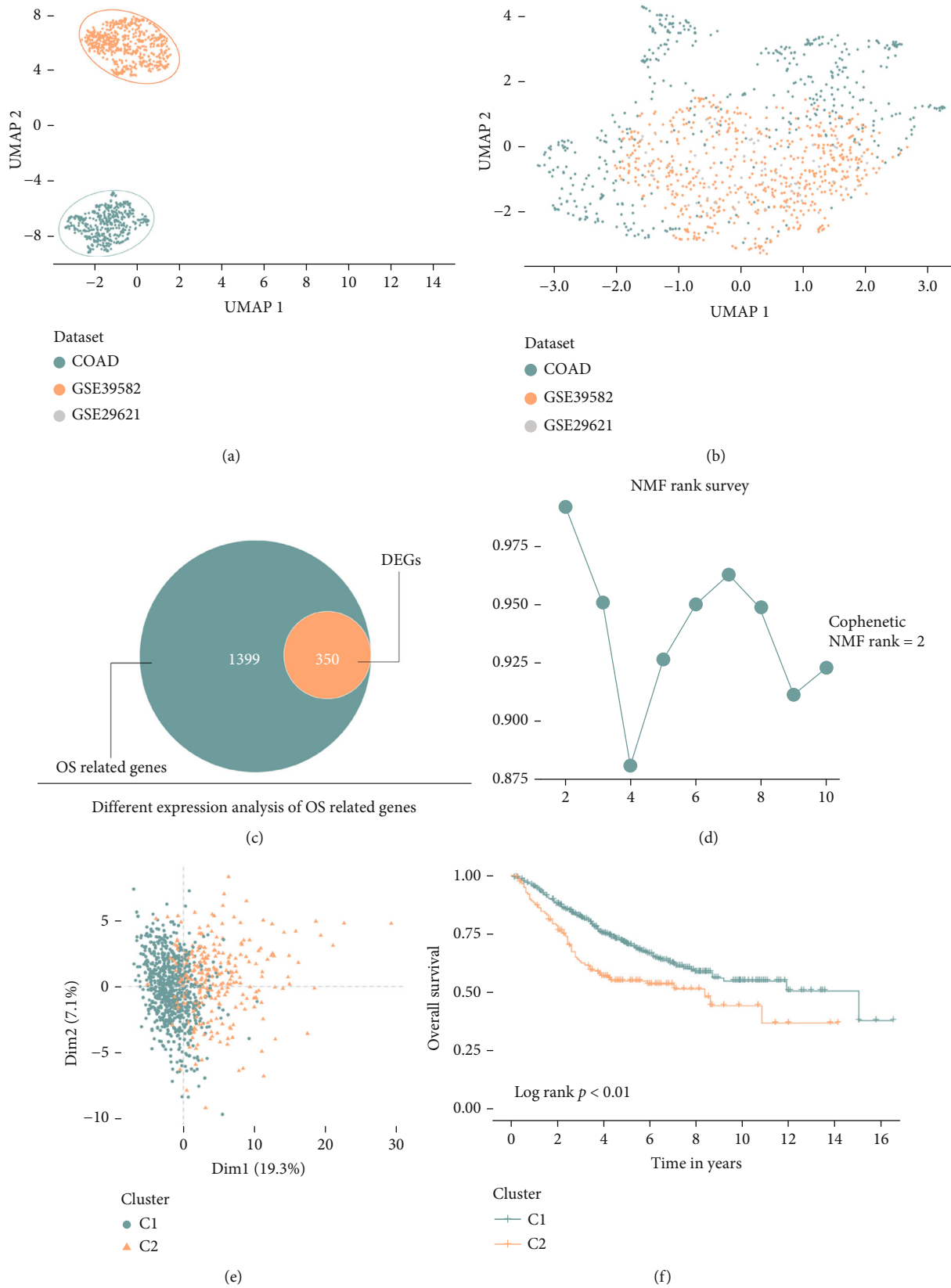
FIGURE 2: NMF cluster samples into two oxidative stress subtypes. Before batch correction, three datasets are scattered (a). After batch correction, the component data is evenly distributed (b). The OS-related DEGs are selected by OS-related genes and DEGs (c). $K = 2$ was best cut off by NMF analysis (d). PCA results show that two subtypes are grouped distinctly (e). Survival analysis demonstrates C2 patients with a poor prognosis, compared with C1 patients (f).
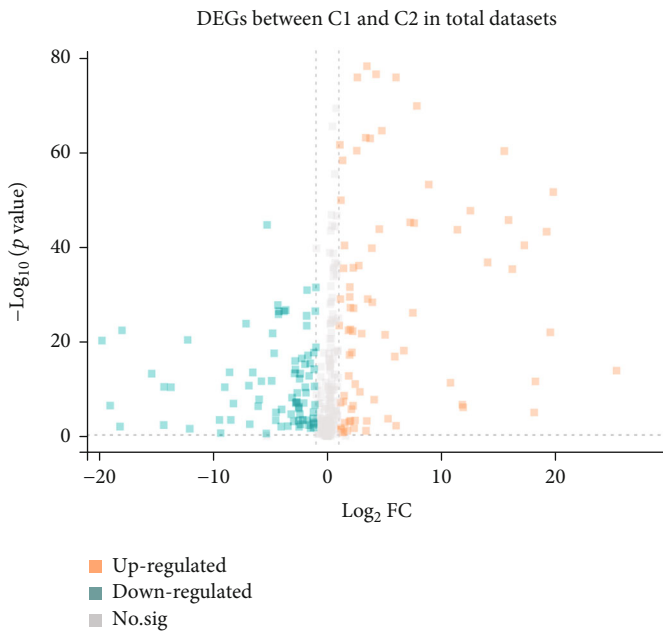
(a)
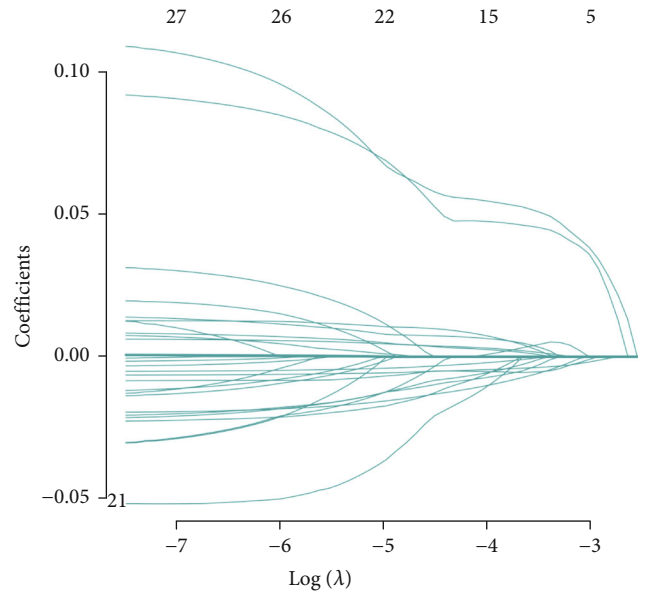


(b)
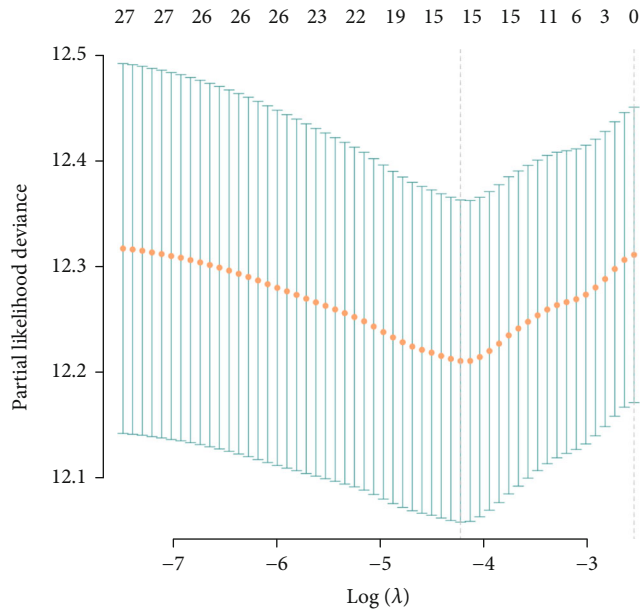


(c)



(d)

Figure 3: Continued.

(e)



(f)
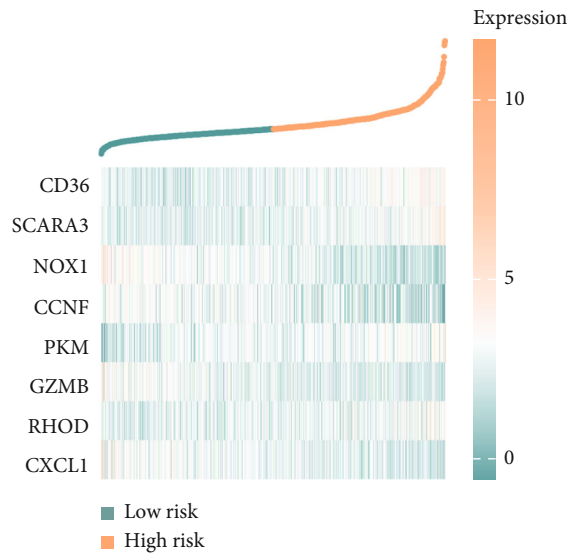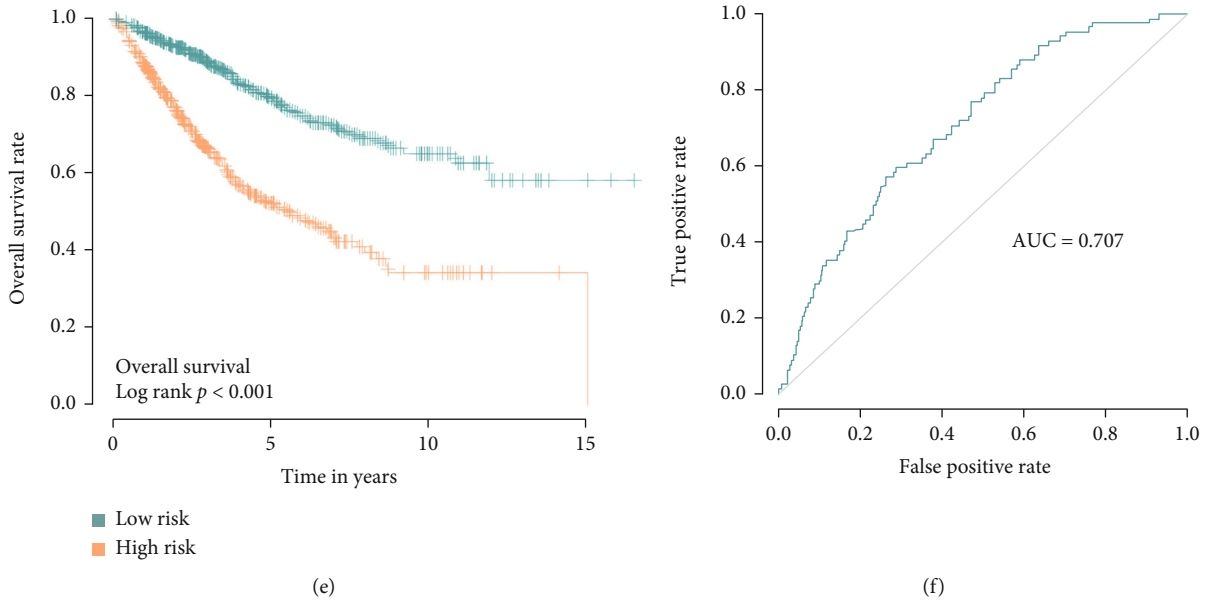
FIGURE 3: 8-gene signature model to predict CRC patients' outcome. DEGs between C1 and C2 in total datasets (a), Lasso regression discover 15 genes that are important (b&c), after multivariate Cox and eight model genes are selected and eight genes signature are built (d). In this model, patients with low-risk score always mean a better outcome, when compared with high-risk score patients (e), and the ROC demonstrates that this model has a good predictive ability for patients' prognosis (f).

TABLE 1: Eight OS-related prognostic genes for signature model by multivariate Cox regression model.

| Gene | Coef | HR | HR.95L | HR.95H | $p$ value |
|---|---|---|---|---|---|
| CD36 | 0.070893982 | 1.073467413 | 1.008614396 | 1.142490423 | 0.025764595 |
| SCARA3 | 0.061080412 | 1.062984388 | 0.992720726 | 1.138221233 | 0.080019034 |
| NOX1 | -0.007467179 | 0.992560631 | 0.987524046 | 0.997622904 | 0.004016322 |
| CCNF | -0.079307037 | 0.923756253 | 0.855402943 | 0.997571521 | 0.043181104 |
| PKM | 0.005245057 | 1.005258837 | 1.002050265 | 1.008477682 | 0.001301438 |
| GZMB | -0.019493712 | 0.980695062 | 0.964974318 | 0.996671918 | 0.018065092 |
| RHOD | 0.016390642 | 1.016525706 | 1.000464261 | 1.032845001 | 0.043686726 |
| CXCL1 | -0.005312535 | 0.994701551 | 0.98998481 | 0.999440765 | 0.028478596 |

*2.8. A Signature Model Constructs to Predict the Overall Survival of CRC Patients.* The multivariate regression model has selected candidate genes, which will be constructed as a signature. This signature was constructed in two steps, first, each candidate gene coefficient needs extract from multivariate regression results, and second, we calculate the risk score of each patient according to the following formula. Risk score = expression × gene A + expression × gene B + expression × gene C. After calculating all patients' risk scores, they will be divided into two groups, according to the median value. A survival curve will be used to demonstrate the differences between groups, and receiver operating characteristic (ROC) was used to evaluate the signature prediction ability.

*2.9. DEGs Combine Robust Rank Aggregation to Identify Important DEGs.* We calculate the different genes in each dataset and merge three DEGs results, then, we also use robust rank aggregation (RRA) method to select the important DEGs, and finally, we combine DEGs overlap results and RRA results to get the final overlap genes which are considered as important genes with different expression between subtypes. To explore the candidate mechanism of DEGs between subtypes, Kobas (http://kobas.cbi.pku.edu.cn/) was used to perform the enrichment analysis including Gene Oncology and KEGG analysis.

*2.10. Protein-Protein Interaction Networks Analysis and Hub Gene Screened.* We use the STRING database (https://cn.string-db.org/), which is an online search for known protein interactions to conduct protein-protein interaction networks (PPI) to show the internal interaction in important genes and use MCODE plugins, which are resources from Cytoscape. To identify the hub protein network, genes located in this network were considered hub genes.

*2.11. Hub Gene Predicts OS Subtypes by Random Forest Algorithms.* We perform hub genes to predict the OS

(a)



GSE29621
- ■ Up-regulated
- ■ Down-regulated
- ■ No.sig

(b)



GSE39582
- ■ Up-regulated
- ■ Down-regulated
- ■ No.sig

(c)



TCGA_COAD
- ■ Up-regulated
- ■ Down-regulated
- ■ No.sig

(d)



(e)

FIGURE 4: Continued.

(f)

FIGURE 4: DEGs and RRA analysis to select important genes. The workflow of hub genes to predict subtypes (a), and different expression genes in each datasets (b–d) merge different DEGs (e) and use RRA to select hub genes (f).

## 3. Results

### 3.1. OS-Related Different Expressions of Genes.
This study workflow is shown in Figure 1, and a total of 1041 samples were enrolled in our study, including 60 normal samples and 981 tumor samples, respectively. 1399 OS-related genes are selected from GeneCards, and we performed the remove batch effect before extracted this OS-related gene from 981 tumor samples. Before batch removal, we could find that samples are distributed in three different spaces (Figure 2(a)), and after removing the batch effect, all samples are distributed on average (Figure 2(b)).

### 3.2. Two OS-Related Subtypes of CRC Patients and DEGs.
Some OS-related genes were expressed differently between normal and tumor tissues in CRC patients, so we conduct DEGs analysis to filter the above genes. These results show that 350 DEGs are selected, including 204 upregulated genes and 146 downregulated genes, respectively (Figure 2(c)).

NMF conducts unsupervised clustering by these 350 DEGs. When rank = 2, the clustering result is best, and samples will be divided into two groups (Figure 2(d) and Supplement Figure 1). The PCA results also demonstrate the above conclusion (Figure 2(e)). Survival analysis results show that when compared with cluster 1, cluster 2 patients will obtain a poor prognosis (log-rank $p < 0.01$) (Figure 2(f)).

### 3.3. Eight OS-Related Gene Signatures Predict the Outcome of Patients.
DEGs between two subtypes are selected and input into batch univariate Cox regression to screen prognosis genes. To identify more accurate prognosis-related genes, we use strict criteria as a $p$ value less than 0.01. Finally, we obtain 27 prognosis-related genes (Supplement Table 1). These 27 genes for future study to continue to perform dimensionality reduction by Lasso regression (LR). The LR analysis results suggest that when lambda obtains the minimum value, 15 important genes are screened (Figures 3(b) and 3(c)). Next step, multivariate Cox regression will analyze these important genes and confirm the final model genes, then, we use the multivariate Cox coefficient and gene expression value to build the final 8 OS-related gene signatures to predict the survival status of CRC patients (Table 1). Each patient will obtain one risk score after inputting 8 genes expression into the model, and the risk curve and genes expression heat map is shown in Figure 3(d). Patients with a risk score of more than the median value will be defined as high-risk groups while others will become low-risk groups. Survival analysis results show that the low-risk group always means a better outcome, while the high-risk group with a poor overall survival rate (log-rank $p < 0.001$) (Figure 3(e)).

subtype of CRC by random forest (RF) algorithms, which are included in the caret package. The detailed steps are listed here. The dataset of standardization, which has removed the batch effect, will be divided into two random datasets: one account for 70% as a training set and another is 30%, as a test set. Then, we use hub genes to predict the OS subtype of CRC in the training set and validate the prediction ability of these hub genes in the test set. Finally, we visualize the decision tree of the model. In this step, the most important is that all input gene expression needs to conduct min-max normalization, which will transfer gene expression values from 0 to 1.

FIGURE 5: Five hub genes predict two oxidative stress subtypes by random forest. The merge cogenes of DEGs and RRA results (a) and to perform PPI network analysis to select hub genes (b). Five hub genes are identified (c) and could predict patients subtype with high accuracy by random forest algorithm (d–f).
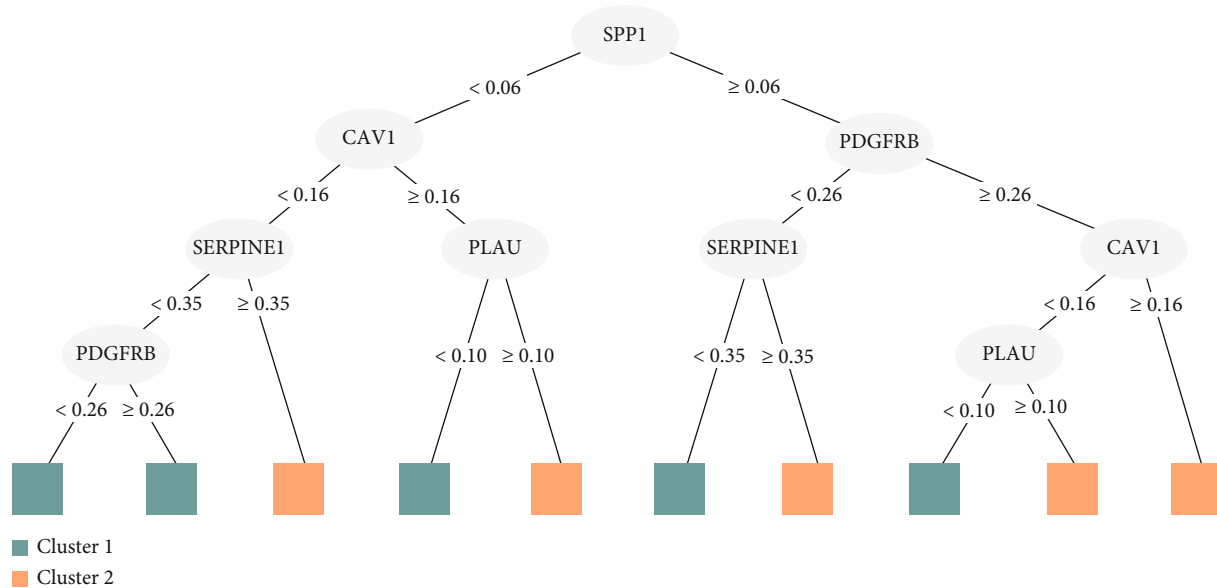
FIGURE 6: Decision tree of five hub genes prediction model. The decision tree of predict procedure, when we input patients genes expression value, the model will follow the value cut off to make decision, and step by step, at the end, return a subtype results C1 or C2 to assistant clinical decision.

Furthermore, The ROC results, AUC = 0.707, demonstrate that these 8 OS-related gene signatures also have better predictive performance (Figure 3(f)).

### 3.4. Different Expressions of Genes between Two Subtypes in each Dataset.

The workflow of how to select candidate hub genes, which will be performed to predict OS-related subtypes of CRC patients, is shown in Figure 3(a). In this workflow, we could find that DEGs analysis is the main idea to select hub genes (Figure 4(a)). So, in the next step, we conduct this process, and the different expressions analysis between C1 and C2 in each dataset show that a total of 77 DEGs in GSE29621, including 55 upregulated DEGs and 22 downregulated DEGs; while 205 DEGs in GSE39582, including 111 upregulated DEGs and 94 downregulated DEGs; 169 DEGs in COAD, including 90 upregulated DEGs and 79 downregulated DEGs, respectively (Figures 4(b)–4(d)). In addition, Kobas results show that these DEGs are enrichment in cellular response to oxidative stress, positive regulation of inflammatory response, HIF-1 signaling pathway, and metabolic pathways (Supplement Table 2 and 3).

### 3.5. Fifteen Important DEGs in Different DEGs Expression Profile.

We merge the above results of different expressions of genes in each dataset, and the results show that 61 genes overlap these three datasets (Figure 4(e)), and the RRA analysis shows that 17 genes are important genes in these three datasets (Figure 4(f)). Merge 61 genes and RRA results, we could find that 15 genes are shared genes (Figure 5(a)).

### 3.6. Five Hub Genes Identify.

We input fifteen important DEGs into the string database to construct the PPI network, and this network shows that excluding GPX3, AOC3, DES, and TPM1, the other 11 proteins interact closely (Figure 5(b)). In addition, we further apply MCODE plugins

in these 11 proteins, and the hub protein network was extracted. The five hub genes that construct this hub network also were identified. They are SPP1, SERPINE1, CAV1, PDGFRB, and PLAU (Figure 5(c)).

### 3.7. Prediction Model of OS-Related Subtype of CRC.

Five hub genes are used to build a prediction model of OS-related subtype of CRC by random forest, and the best mtry value is 2 while the number of trees is 200 in the training set when the model obtains robust predictability (Figure 5(d)). Figure 5(e) shows the importance of five hub genes. In addition, the high effective ability of the prediction model is also demonstrated by test data. The model has a high AUC value, 0.771, in the test group (Figure 5(f)). We also show the decision tree of the model (Figure 6). According to this decision tree, clinic physicians could evaluate patients' OS-related subtypes by judging five hub genes expression, step by step.

## 4. Discussion

The regulation of OS is an important factor in tumor development. OS not only induces the formation of tumors by abnormal cell proliferation [16] but also promotes further tumor development by altering the metabolism of tumors [17, 18]. Studies have also shown that targeting the antioxidant capacity of tumor cells can have a positive impact on cancer therapy [19]. OS is associated with colorectal carcinogenesis and has been identified as an important risk factor for colorectal adenoma in several studies [20–23]. So, focusing on the role of OS-related genes in CRC is necessary to promote diagnostic gene screening and therapeutic strategies for CRC. Moreover, the small size of the dataset will reduce the accuracy of the predictions and the robustness of the subtypes. Given these, we utilized 3 datasets from the TCGA and GEO databases to identify robust subtypes

of CRC for better understanding the underlying molecular pathogenesis of CRC. NMF algorithm has been widely applied to reveal various cancer subtypes through clustering tumor samples [24]. In our study, based on OS-related DEGs, we successfully classified the CRC sample into two subtypes (C1 and C2) by using the NMF algorithm. The results of the PCA analysis revealed that our classification was robust. Then, survival analysis results indicated that C1 subtype had a better prognosis compared to C2 subtype.

It is well known that the prognosis of cancer affects clinical decision-making. Recent clinical guidelines have emphasized the importance of using multigene tests to select patients who should receive adjuvant therapy [25]. Multiple genes of the tests are called cancer signatures, which are crucial for cancer prognosis. Considering the impact of OS and OS subtypes on clinical outcomes in CRC patients, therefore, univariate Cox regression, Lasso Cox regression, and multivariate Cox regression analyses were conducted to construct the optimal OS-related prognostic signature based on DEGs between C1 and C2 subtypes. In this 8-gene prognostic model, as expected, patients in the low-risk group had a better overall survival rate. At the same time, our results showed that this signature had good predictive accuracy in predicting the overall survival of CRC patients. Overall, the prognostic model we constructed may be useful for clinical treatment and decision-making in CRC.

Through machine learning algorithms and the PPI network, we successfully identified five OS-related hub genes, including SPP1, PDGFRB, SERPINE1, CAV1, and PLAU. All five hub genes play a significant role in tumor progression, invasion, and metastasis. Among these genes, secreted phosphoprotein 1 (SPP1, also known as osteopontin) is a secreted glycophosphoprotein, which can be secreted by a variety of cells, including macrophages and endothelial cells [26]. Previous works have demonstrated that SPP1 is overexpressed in various cancers (such as nonsmall cell lung cancer [27] and ovarian cancer [28]) and involved in the progression and metastasis of cancer. In colorectal cancer, SPP1 expression was significantly upregulated, and it promoted CRC metastasis by activating the epithelial-mesenchymal-transition pathway [29]. Platelet-derived growth factor receptor type $\beta$ (PDGFRB, also called PDGFR$\beta$) has been identified as a causal gene for idiopathic basal ganglia calcification [30]. Meanwhile, it is also correlated with CRC invasion and metastasis, for example, excessive PDGFR$\beta$ signaling leads to oversecretion of THBS4 and proliferative colorectal tumor development [31]. As for the serpin peptidase inhibitor, clade E, member 1 (SERPINE1, also called), it is expressed in many cancer cell and regulates cancer growth, invasion, and angiogenesis [32]. In the CRC cell line, the study has demonstrated that SERPINE1 expression is increased and related to tumor invasiveness and aggressiveness [33]. And, PAI-1 is regarded as a biomarker of poor prognosis in various human cancers and a possible therapeutic target for some cancers [34]. Caveolin-1 (CAV1), an oncogenic membrane protein related to endocytosis, extracellular matrix organization, cholesterol distribution, cell migration, and signaling has been linked to several cancers [35, 36]. For instance, Yang et al. [37] have reported that

overexpression of CAV1 markedly inhibits the proliferation, migration, and invasive potential of CRC cells, possibly by reducing phosphorylation of epidermal growth factor receptor activation. Lastly, plasminogen activator (PLAU) is also associated with the complex phenotype of human cancer, and its upregulation promotes metastatic cancers [38]. In a CRC study, downregulation of PLAU expression inhibits CRC cell proliferation and progression [39].

Although we have stratified OS molecular subtypes, built OS-related prognostic model and identified hub genes in CRC for the first time, there are some shortcomings in the current study. First, since we used different platform data and multiple CRC tissue samples, the effect of batch correction may not be completely eliminated. Second, the OS-related prognostic signature constructed from public datasets of the TCGA and GEO databases should be validated for their prognostic value through large-scale prospective studies. Finally, the function of hub genes and their mechanisms affecting CRC development also need to be further elucidated.

## 5. Conclusion

Our study successfully utilizes multiple datasets to stratify CRC samples into two novel OS subtypes, which can provide new insights into the molecular features of CRC. The OS-related prognostic gene signatures can serve as a powerful tool for overall survival prediction and treatment guidance in CRC patients. Additionally, we identified five key genes (SPP1, PDGFRB, SERPINE1, CAV1, and PLAU) as potential biomarkers for predicting the OS subtype and diagnosis of CRC. In general, these findings might enhance our understanding of the molecular pathogenesis of CRC and contribute to identifying new candidate biomarkers for CRC.

## Data Availability

The datasets were obtained from the TCGA (colon adenocarcinoma: COAD) and GEO (GSE39582 and GSE29621) databases, respectively. Oxidative stress (OS) related genes were checked by the GeneCards database (https://www.genecards.org/).

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Authors' Contributions

Haitao Zhong, Le Yang, Qingshang Zeng, Linlin Wu, Lei Qin, and Qing-Qing Yu contributed equally to this work.

## Acknowledgments

## Supplementary Materials

*Supplementary 1.* Supplement Table 1: batch univariate Cox regression discovers twenty-seven prognosis genes in DEGs between two subtypes.

*Supplementary 2.* Supplement Table 2: Gene Oncology enrichment results of DEGs between C1 and C2 cluster.

*Supplementary 3.* Supplement Table 3: KEGG enrichment results of DEGs between C1 and C2 cluster.

*Supplementary 4.* Supplement Figure 1: the heat map for two OS-related subtypes.

## References

[1] H. Brenner, M. Kloor, and C. P. Pox, "Colorectal cancer," *Lancet*, vol. 383, no. 9927, pp. 1490–1502, 2014.

[2] N. Keum and E. Giovannucci, "Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies," *Nature Reviews. Gastroenterology & Hepatology*, vol. 16, no. 12, pp. 713–732, 2019.

[3] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022," *CA: a Cancer Journal for Clinicians*, vol. 72, no. 1, pp. 7–33, 2022.

[4] D. Hanahan, "Hallmarks of cancer: new dimensions," *Cancer Discovery*, vol. 12, no. 1, pp. 31–46, 2022.

[5] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[6] J. Guinney, R. Dienstmann, X. Wang et al., "The consensus molecular subtypes of colorectal cancer," *Nature Medicine*, vol. 21, no. 11, pp. 1350–1356, 2015.

[7] N. J. Llosa, M. Cruise, A. Tam et al., "The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints," *Cancer Discovery*, vol. 5, no. 1, pp. 43–51, 2015.

[8] Z. Gatalica, C. Snyder, T. Maney et al., "Programmed cell death 1 (PD-1) and its ligand (PD-L1) in common cancers and their correlation with molecular cancer type," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 23, no. 12, pp. 2965–2970, 2014.

[9] A. Bar-Even, A. Flamholz, E. Noor, and R. Milo, "Rethinking glycolysis: on the biochemical logic of metabolic pathways," *Nature Chemical Biology*, vol. 8, no. 6, pp. 509–517, 2012.

[10] S. Matoba, J. G. Kang, W. D. Patino et al., "p53 regulates mitochondrial respiration," *Science*, vol. 312, no. 5780, pp. 1650–1653, 2006.

[11] S. Afzal, S. A. Jensen, J. B. Sørensen, T. Henriksen, A. Weimann, and H. E. Poulsen, "Oxidative damage to guanine nucleosides following combination chemotherapy with 5-fluorouracil and oxaliplatin," *Cancer Chemotherapy and Pharmacology*, vol. 69, no. 2, pp. 301–307, 2012.

[12] M. Diehn, R. W. Cho, N. A. Lobo et al., "Association of reactive oxygen species levels and radioresistance in cancer stem cells," *Nature*, vol. 458, no. 7239, pp. 780–783, 2009.

[13] J. Kozak, K. Jonak, and R. Maciejewski, "The function of miR-200 family in oxidative stress response evoked in cancer chemotherapy and radiotherapy," *Biomedicine & Pharmacotherapy*, vol. 125, article 110037, 2020.

[14] P. Ghosh, C. Vidal, S. Dey, and L. Zhang, "Mitochondria targeting as an effective strategy for cancer therapy," *International Journal of Molecular Sciences*, vol. 21, no. 9, p. 3363, 2020.

[15] A. Colaprico, T. C. Silva, C. Olsen et al., "TCGAbiolinks: an R/bioconductor package for integrative analysis of TCGA data," *Nucleic Acids Research*, vol. 44, no. 8, article e71, 2016.

[16] P. L. de Sá Junior, D. A. D. Câmara, A. S. Porcacchia et al., "The roles of ROS in cancer heterogeneity and therapy," *Oxidative Medicine and Cellular Longevity*, vol. 2017, Article ID 2467940, 12 pages, 2017.

[17] N. H. Kim, Y. H. Cha, J. Lee et al., "Snail reprograms glucose metabolism by repressing phosphofructokinase PFKP allowing cancer cell survival under metabolic stress," *Nature Communications*, vol. 8, no. 1, article 14374, 2017.

[18] P. K. Kopinski, L. N. Singh, S. Zhang, M. T. Lott, and D. C. Wallace, "Mitochondrial DNA variation and cancer," *Nature Reviews Cancer*, vol. 21, no. 7, pp. 431–445, 2021.

[19] C. Gorrini, I. S. Harris, and T. W. Mak, "Modulation of oxidative stress as an anticancer strategy," *Nature Reviews Drug Discovery*, vol. 12, no. 12, pp. 931–947, 2013.

[20] D. Basak, M. N. Uddin, and J. Hancock, "The role of oxidative stress and its counteractive utility in colorectal cancer (CRC)," *Cancers*, vol. 12, no. 11, p. 3336, 2020.

[21] S. Loft and H. E. Poulsen, "Cancer risk and oxidative DNA damage in man," *Journal of Molecular Medicine (Berlin, Germany)*, vol. 74, no. 6, pp. 297–312, 1996.

[22] E. C. Cheung and K. H. Vousden, "The role of ROS in tumour development and progression," *Nature Reviews Cancer*, vol. 22, no. 5, pp. 280–297, 2022.

[23] A. Tasdogan, J. M. Ubellacker, and S. J. Morrison, "Redox regulation in cancer cells during metastasis," *Cancer Discovery*, vol. 11, no. 11, pp. 2682–2692, 2021.

[24] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[25] M. J. Duffy, N. Harbeck, M. Nap et al., "Clinical use of biomarkers in breast cancer: updated guidelines from the European group on tumor markers (EGTM)," *European Journal of Cancer*, vol. 75, pp. 284–298, 2017.

[26] M. Ahmed, R. Behera, G. Chakraborty et al., "Osteopontin: a potentially important therapeutic target in cancer," *Expert Opinion on Therapeutic Targets*, vol. 15, no. 9, pp. 1113–1126, 2011.

[27] X. Wang, F. Zhang, X. Yang et al., "Secreted phosphoprotein 1 (SPP1) contributes to second-generation EGFR tyrosine kinase inhibitor resistance in non-small cell lung cancer," *Oncology Research*, vol. 27, no. 8, pp. 871–877, 2019.

[28] B. Zeng, M. Zhou, H. Wu, and Z. Xiong, "SPP1 promotes ovarian cancer progression via integrin $\beta$1/FAK/AKT signaling pathway," *Oncotargets and Therapy*, vol. 11, pp. 1333–1343, 2018.

[29] C. Xu, L. Sun, C. Jiang et al., "SPP1, analyzed by bioinformatics methods, promotes the metastasis in colorectal cancer by activating EMT pathway," *Biomedicine & Pharmacotherapy*, vol. 91, pp. 1167–1177, 2017.

[30] G. Nicolas, C. Pottier, D. Maltete et al., "Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification," *Neurology*, vol. 80, no. 2, pp. 181–187, 2013.

[31] M. S. Kim, H. S. Choi, M. Wu et al., "Potential role of PDGFR$\beta$-associated THBS4 in colorectal cancer development," *Cancers*, vol. 12, no. 9, p. 2533, 2020.

[32] B. McMahon and H. C. Kwaan, "The plasminogen activator system and cancer," *Pathophysiology of Haemostasis and Thrombosis*, vol. 36, no. 3-4, pp. 184–194, 2008.

[33] G. Mazzoccoli, V. Pazienza, A. Panza et al., "ARNTL2 and SERPINE1: potential biomarkers for tumor aggressiveness in colorectal cancer," *Journal of Cancer Research and Clinical Oncology*, vol. 138, no. 3, pp. 501–511, 2012.

[34] P. A. Andreasen, "PAI-1 - a potential therapeutic target in cancer," *Current Drug Targets*, vol. 8, no. 9, pp. 1030–1041, 2007.

[35] Z. C. Nwosu, M. P. Ebert, S. Dooley, and C. Meyer, "Caveolin-1 in the regulation of cell metabolism: a cancer perspective," *Molecular Cancer*, vol. 15, no. 1, p. 71, 2016.

[36] J. Ketteler and D. Klein, "Caveolin-1, cancer and therapy resistance," *International Journal of Cancer*, vol. 143, no. 9, pp. 2092–2104, 2018.

[37] J. Yang, T. Zhu, R. Zhao et al., "Caveolin-1 inhibits proliferation, migration, and invasion of human colorectal cancer cells by suppressing phosphorylation of epidermal growth factor receptor," *Medical Science Monitor*, vol. 24, pp. 332–341, 2018.

[38] M. Sudol, "From Rous sarcoma virus to plasminogen activator, Src oncogene and cancer management," *Oncogene*, vol. 30, no. 27, pp. 3003–3010, 2011.

[39] M. Lin, Z. Zhang, M. Gao, H. Yu, H. Sheng, and J. Huang, "MicroRNA-193a-3p suppresses the colorectal cancer cell proliferation and progression through downregulating the PLAU expression," *Cancer Management and Research*, vol. 11, pp. 5353–5363, 2019.

*Research Article*

# Necroptosis-Related Modification Patterns Depict the Tumor Microenvironment, Redox Stress Landscape, and Prognosis of Ovarian Cancer

**Rui Geng** [ID],[1] **Zihang Zhong** [ID],[1] **Senmiao Ni,**[1] **Wen Liu,**[1] **Zhiqiang He,**[1] **Shilin Gan,**[1] **Qinghao Huang,**[1] **Hao Yu** [ID],[1] **Jianling Bai** [ID],[1] **and Jinhui Liu** [ID][2]

[1]*Department of Biostatistics, School of Public Health, Nanjing Medical University, 101 Longmian Avenue, Jiangning District, Nanjing 211166, China*
[2]*Department of Gynecology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029 Jiangsu, China*

Correspondence should be addressed to Hao Yu; haoyu@njmu.edu.cn, Jianling Bai; baijianling@njmu.edu.cn, and Jinhui Liu; jinhuiliu@njmu.edu.cn

Necroptosis is one of programmed cell death discovered recently, which involves in tumorigenesis, cancer metastasis, and immune reaction. We studied the necroptosis-related genes (NRGs) in ovarian cancer (OV) tissues using data from public databases, which separated into two NRGclusters. Patients in cluster A would have severe clinical characteristics, poor prognosis, and worse tumor microenvironment infiltration characteristics. The NRG score was achieved through the Cox analysis, along with a construction of a prognostic model. People with lower risk score would have better prognosis, lower expression of redox related genes, higher immunogenicity, and better effect on immunotherapy. In addition, the NRG score was closely related to cancer stem cell index, copy number variations, tumor mutation load, and chemosensitivity. We built a nomogram to enhance clinical application of the signature. These outcomes can help use know the function of NRGs in OV and provide new ideas for evaluating clinical outcome and developing more effective treatment protocols.

## 1. Introduction

Ovarian cancer (OV) is a major gynecological malignancy around the world, and its mortality ranks first among gynecological malignancies [1]. Worldwide, the number of new OV cases was 313,959 and the number of deaths was 207,252 in 2020 [2]. Although the treatment of OV has made many progresses recently, the prognosis of OV is still not well [3]. More than 60% patients were in an advanced stage when diagnosed [4, 5]. Through timely diagnosis and appropriate treatment, the mortality of advanced stage and recurrence rate of OV can be reduced to great extent [6]. So, it is needed to found new diagnostic and therapeutic methods. With development for the branch of the cancer cell immune recognition and immune regulate molecules, immunotherapy has become a research hotspot recently [7]. Developing new biomarkers, identifying therapeutic targets, predicting

therapeutic effects, and screening potential immunotherapeutic drugs offer new orientation for the remedy of OV and may prolong the survival of patients [8].

Necroptosis is defined as a regulated type of necrosis whose morphology is similar to necrosis, such as cell swelling and rupture, regulated by certain signal pathways like apoptosis [9]. Necroptosis is crucial to cancer. For one thing, necroptosis can trigger adaptive immune response and impede tumor progression [10]. Meanwhile, the inflammatory response may also help the occurrence and development of cancer, and necroptosis can produce immunosuppressive tumor microenvironment (TME), which may contribute to cancer progression [11, 12]. Until now, several chemotherapeutics, natural compounds, and classical necroptosis inducers have been proved to disappear tumor cells through necroptosis. For instance, characteristics such as etoposide, 5-FU, and cisplatin may lead necroptosis of cancer cells [13]. Consequently, the

Altered in 89 (20.41%) of 436 samples.



(a)

Figure 1: Continued.

(b)



(c)



(d)

Figure 1: Heredity and transcriptional changes of necroptosis-related genes (NRGs) in OV. (a) In TCGA-OV population, 89 patients had gene mutations. (b) CNV frequency happened in NRG. (c) Location of CNV changes on 23 chromosomes in NRG. (d) Gene expression of NRGs in normal and tumor tissues. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

(a)



(b)



(c)
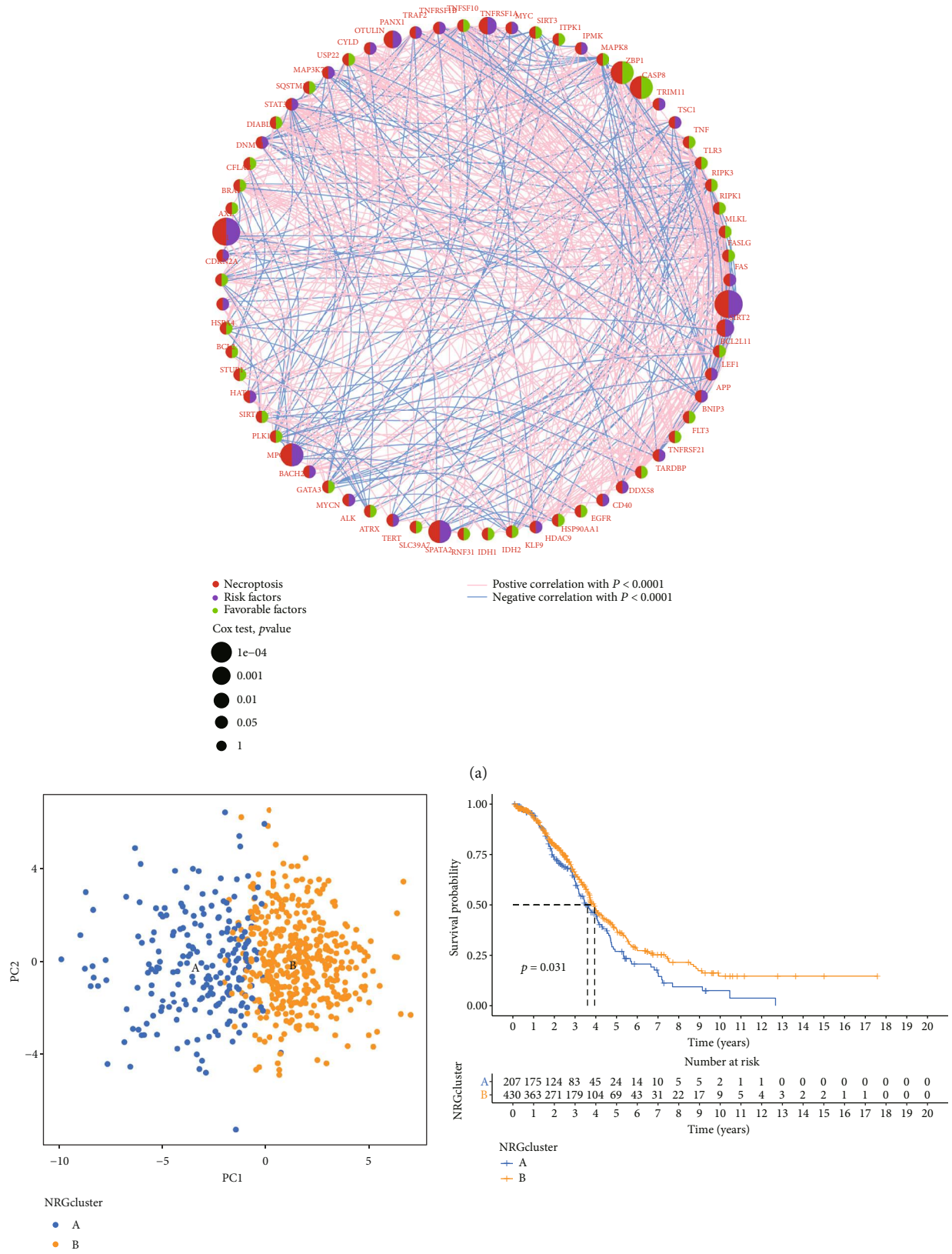
FIGURE 2: Continued.
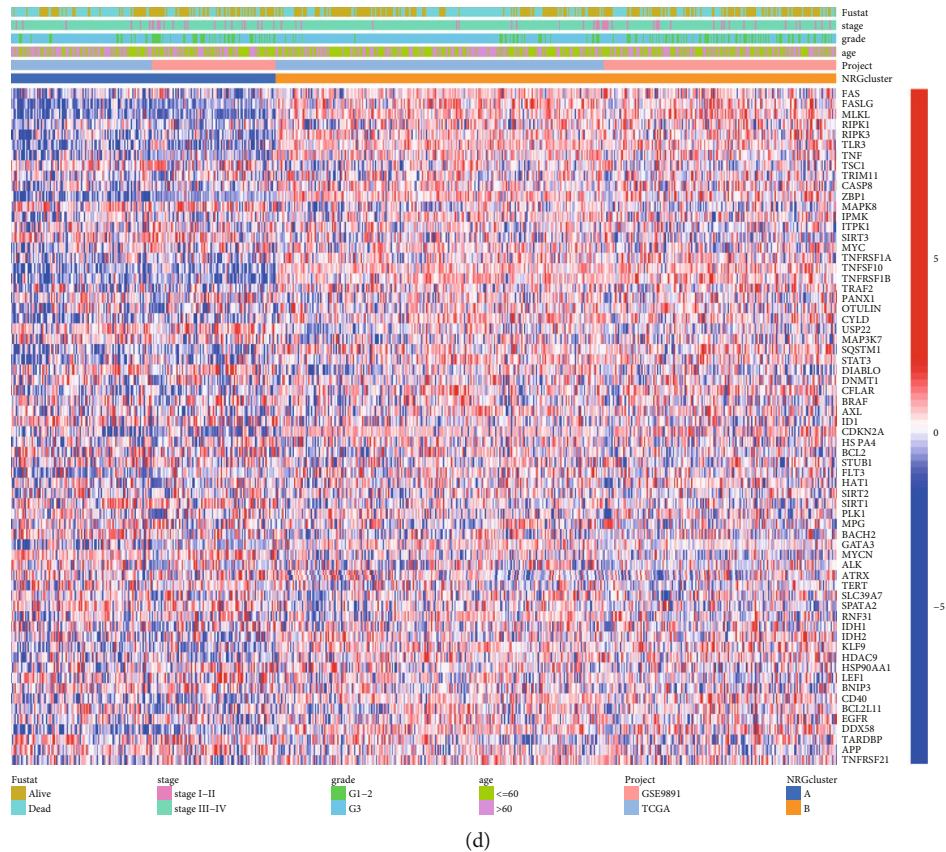
Figure 2: Clinical and biological factors of two clusters defined by clustering analysis. (a) Interrelationships among NRGs in OV. (b) PCA scatter plot reflecting the distinction between NRGclusters. (c) Survival probability of NRGcluster A and NRGcluster B. (d) Difference of clinical factors expression levels of NRGs between NRGclusters.

selection of necroptosis-related genes and using them to build predict signatures were promising methods to forecast the prognosis OV patients.

Except malignant transformed cells, tumors consisted of normal cells, like fibroblasts, muscle cells, and inflammatory immune cells, which make up with TME together [14]. The interaction between tumor cells and TME influences the treatment effect of cancer [15]. In the early stage, tumors are infiltrated by various adaptive and innate immune cells, which conduct tumor-promoting and antitumor functions [16]. For example, higher infiltration level of CD8 T cells is usually related to better prognosis [17], while the macrophages M2 is supposed to be a poor prognostic marker [18]. In fact, immunotherapy has become one of the most hopeful methods in oncology. Clarifying the status of infiltrating immune cells in TME and understanding their number and function may contribute to formulate strategies to enhance the response rate of immunotherapy. Necroptosis is becoming a new target of cancer immunotherapy. Necroptosis in tumor cells regulates TME and antitumor immunity, which will be particularly helpful to the treatment of immune desert tumors [19, 20]. Necroptosis has different influence on tumor progression in the light of tumor cell types and TME [21]. However, the mechanisms are still unclear [22].

Redox reaction is a part of normal cell metabolism. If the redox homeostasis is damaged, cell death may be induced [23]. Increasing oxidative stress by increasing reactive oxygen species level or decreasing cell antioxidant capacity is a promising anticancer way, and it takes part in the mechanism of many chemotherapy drugs that have been used in clinical application [24]. More and more evidence shows that the redox modification participated in the regulation of some cell death modes, like necrosis and apoptosis. In addition, thiol redox switches involve in regulating the crosstalk between apoptotic and necrotic forms of cell death [25]. Mitochondrial peroxidase has a upregulated expression in different tumor types, including OV [26].

Our study calculated the expression profile of necroptosis-related genes (NRGs) and downloaded the immune pattern in OV by using two computational algorithms. In terms of NRG expression level, OV patients were separated into two independent subgroups. Then, the patients were divided into three gene clusters according to the differentially expressed genes (DEGs) between NRGclusters. A prediction signature was further built to predict the prognosis, so as to realize accurate identification and therapeutic measures of individuals.

## 2. Materials and Methods

*2.1. Data Acquisition.* Genetic and clinical profiles of OV were downloaded from the cancer genome atlas (TCGA)

(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)

Figure 3: Continued.

(i)



(j)



(k)



(l)
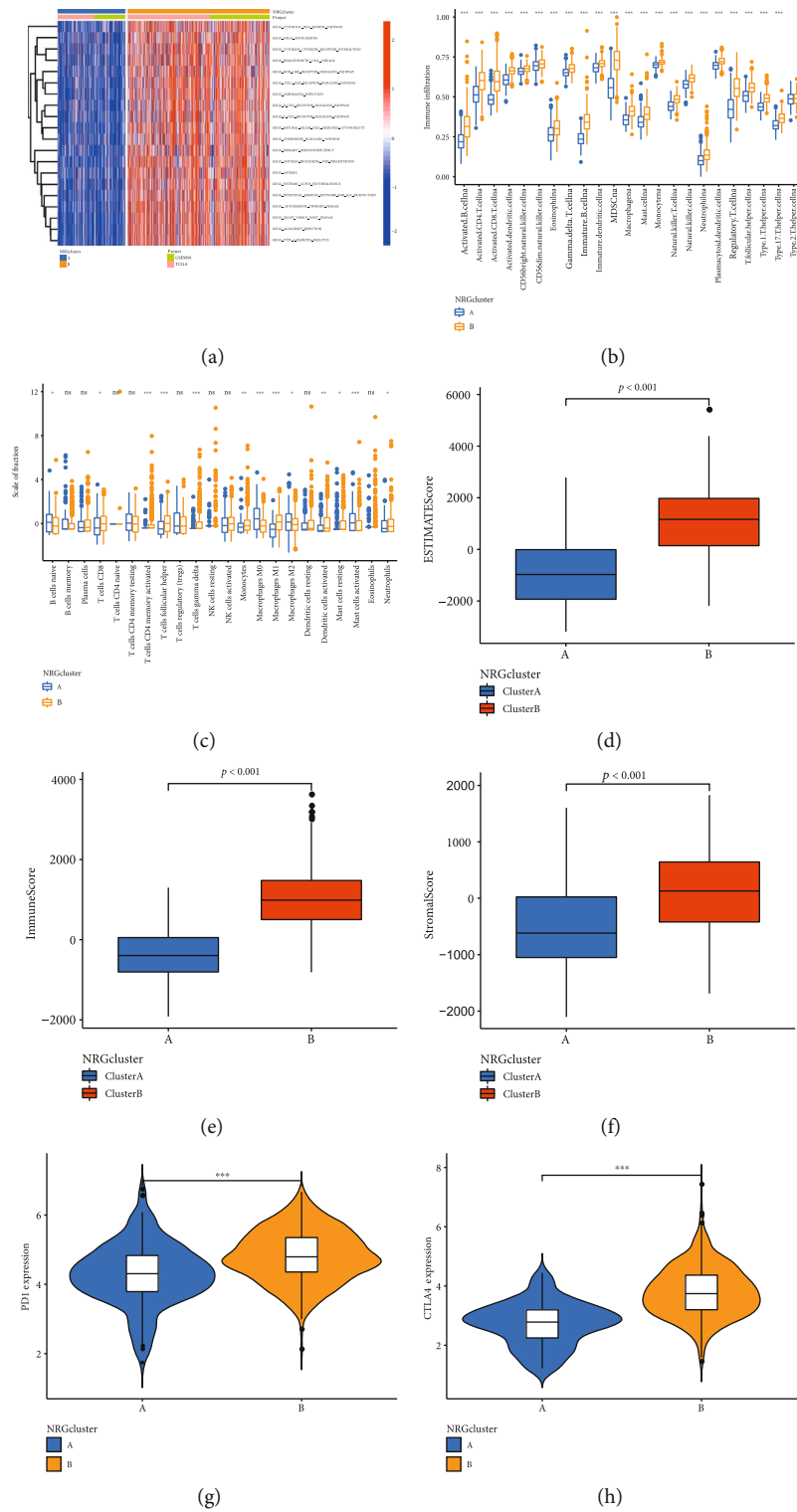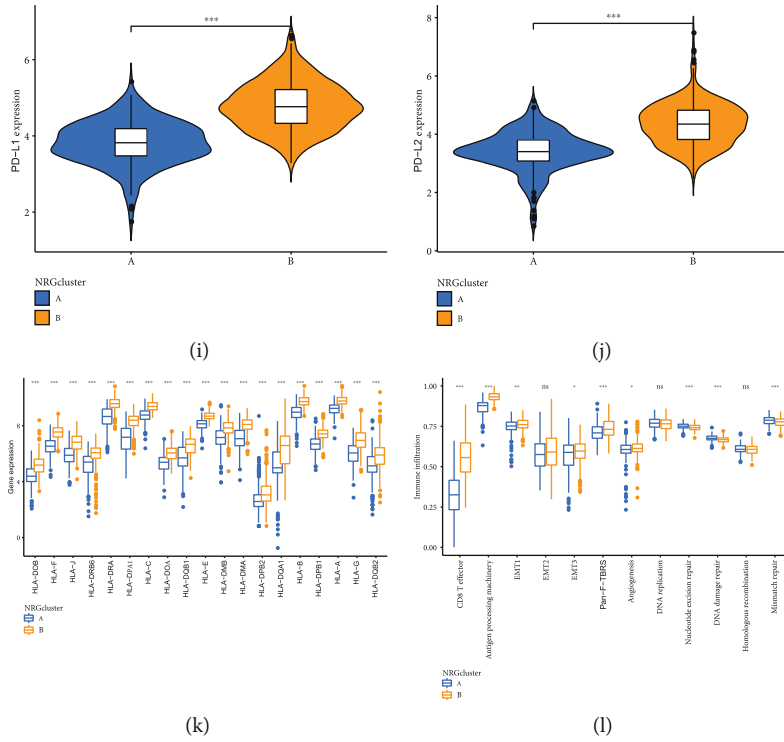
FIGURE 3: Tumor mutation load of the two NRGclusters. (a) Gene set enrichment analysis of NRGclusters. (b, c) Immune infiltration levels of two NRGclusters. (d) The stromal score, (e) immune score, and (f) estimated score of the two NRGclusters were compared. (g–j) Immune checkpoint expression of NRGclusters. (k) HLA expression of two NRGclusters. (l) Compare the scores of biological pathways between the two NRGclusters. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

and gene expression omnibus (GEO). This study used 836 samples from two cohorts, TCGA-OV and GSE9891. Table S1 shows clinical features of individuals involved in the research. We combine TCGA-OV and GSE9891 datasets and use the "ComBat" algorithm to correct the batch effect. In order to reduce statistical bias, samples without overall survival (OS) value and without follow-up data were screened out from the study. OV patients with relevant characteristics (age, grade, and stage) and survival data were used for further analysis. GSE9891 was employed as an external set to validate.

*2.2. Consensus Clustering Analysis.* 67 NRGs were obtained from published articles [27]. In accordance with gene expression data, we performed consumes clustering via R packages "ConsumusClusterPlus" [28]. To calculate the differences of NRGs in pathways, the gene set variation analysis (GSVA) was executed through a marker gene set (C2. Cp.kegg. V7.2) from Molecular Signatures Database.

*2.3. Gene Clusters Identification on the Basis of DEGs.* Firstly, the "limma" package was conducted to screen DEGs between gene clusters. The univariate Cox regression analysis was conducted on DEGs to screen DEGs related to OV survival. Secondly, OV patients were classified according to DEGs by using consistent clustering algorithm, and the patients were separated into three different subgroups.

*2.4. Build Prognostic NRG Score Related to Necroptosis.* After integrating the transcriptome and clinical data, we deleted individuals without prognostic data. All volunteers were randomly separated into training ($n = 319$) and testing subtypes ($n = 318$), and then the information of the training set was used to build the NRG score which related to OV patients. Based on prognostic genes associated with necroptosis, the least absolute shrinkage and selection operator (LASSO) regression algorithm was utilized to avoid over fitting through the "glmnet" R package. The multivariate Cox analysis identified key genes to establish predict model on the base of the training set. The formula is as follows: risk score = $\Sigma$ (coefi × expi), where coefi and expi, respectively, mean the coefficient and express level of each gene. In accordance with median risk score, samples were divided into two different subgroups.

*2.5. The Difference of Clinical Features Patients and Stratified Analysis.* The score of individuals with different clinical features was compared by box diagram and scatter diagram. Hierarchical analysis was employed to assess the differences in OS between different subgroups using the Kaplan–Meier curve achieved by the "survminer" R package, to determine that the model still has the ability to predict under different clinicopathological characteristics.

*2.6. Enrichment Analysis.* Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were applied to enrich NRG-related processes. KEGG is a dataset usually used to discover significantly altered pathways enriched in

(a)



(b)



(c)



(d)



(e)



(f)

Figure 4: Continued.

(g)



(h)



(i)



(j)



(k)



(l)



(m)



(n)

Figure 4: Continued.

(o)                                                                     (p)
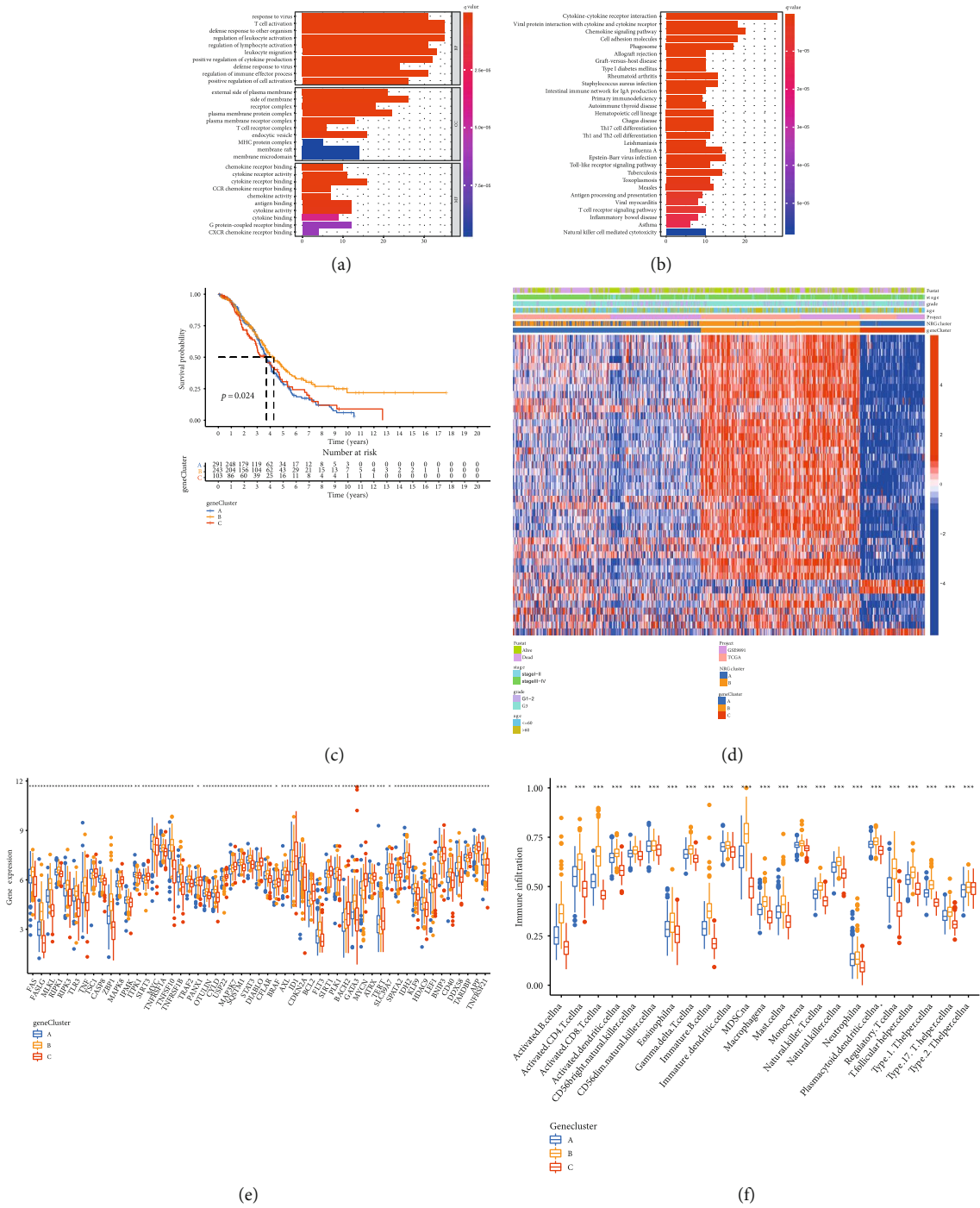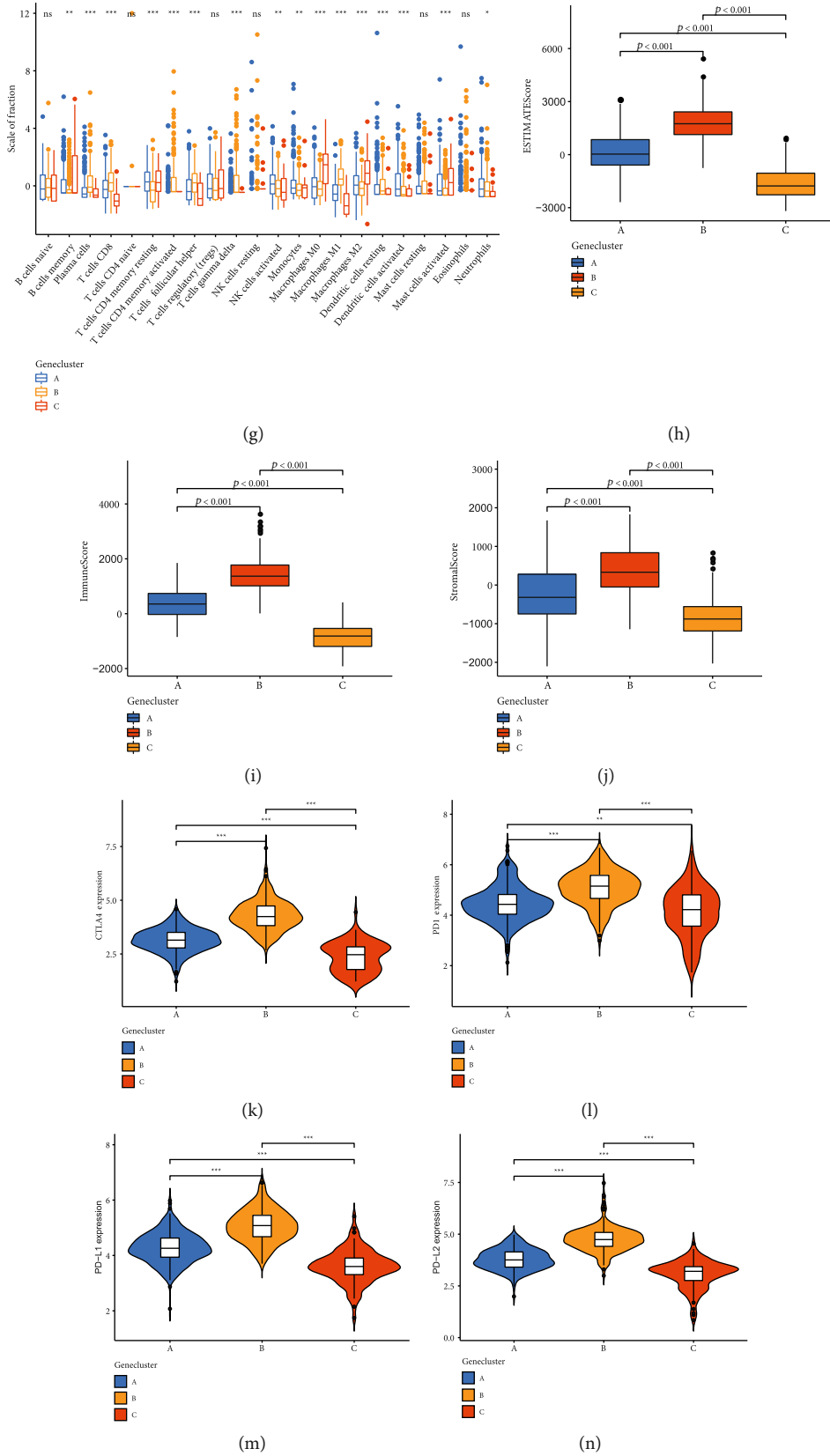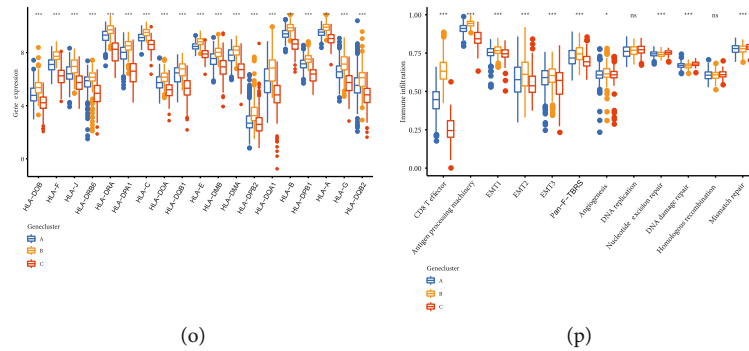
Figure 4: Identify gene subtypes in the light of DEGs (a) GO and (b) KEGG enrichment analysis of two gene clusters. (c) Survival probability of gene subtypes. (d) Clinicopathological features of three gene subtypes. (e) The expression of NRGs of gene clusters. (f, g) Immune infiltration of gene clusters. (h) Stromal score, (i) immune score, and (j) estimated score of three gene clusters were compared. (k–n) Expression of immune checkpoints. (o) HLA expression level of gene clusters. (p) Compare the score of biological pathways between the three gene clusters. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

gene list. GO and KEGG were carried out through the bioinformatics platform [29]. By aggregating gene into gene sets, the gene set enrichment analysis (GSEA) offer rich scores, which allows users to have an in-depth understanding of how biological processes are influenced [30].

### 2.7. Assessment of Immune Infiltration Level.
TME is widely involved in the tumorigenesis and tumor progression. ESTI-MATE can predict the TME status in the light of relevant biomarkers expression in immune and stromal cells which conducted by the R package "estimate" [31, 32]. The single-sample gene set enrichment analysis (ssGSEA) can quantitatively estimate immune cell components from complex gene expression data by using the R package "GSVA" [33, 34]. CIBERSORT can quantify the abundance of TIICs in risk groups (http://cibersort.stanford.edu/).

### 2.8. The Difference of Biological Processes between NRGclusters.
Rosenberg et al. defined a set of genes related to specific biological pathways, such as epithelial mesenchymal transition markers, DNA damage repair, and CD8 T-effector signature [35].

### 2.9. Phenotypes of RNAss Differentiation.
In cancer stem cells (CSCs), mRNA expression-based RNAss is a variable to describe the similarity between tumors and stem cells [36, 37]. The higher the score was, the stronger the degree of stemness and the lower differentiation degree. RNAss scores were achieved from the xena (https://xenabrowser.net/datapages/).

### 2.10. Predict the Effect of Immunotherapy.
Immunophenoscore (IPS) has been verified to predict patient response to immunotherapy [38], which can be achieved from The Cancer Immune Atlas (TCIA) (https://tcia.at/home). TMB can screen individuals who could benefit more from immunotherapy [39]. The burden of gain or loss of copy number variations (CNV) was calculated by gene pattern (https://cloud .genepattern.org) [40].

### 2.11. Analysis of Drug Sensitivity.
To assess the therapeutic efficacy of chemotherapeutics on OV patients, the half maximum inhibitory concentration (IC50) of chemotherapeutics

was achieved by the "prrophetic" R package [38]. We achieved the data of gene expression and drug sensitivity from CellMiner to calculate the correlation between some commonly used drugs and 8 genes.

### 2.12. Set Up a Nomograph.
A nomogram can evaluate the OS through the "rms" package [41], where each factors were given a score, then added up them, and achieved a final score [42].

Hosmer-Lemeshow was applied to testify whether the predicted results were consistent with the fact [42]. The predictive ability of the model was explored through the $C$-index and area under curve (AUC) [43, 44]. $C$-index can be calculated by restricted mean survival (RMS). The capacity of nomogram was calculated by $C$-index and AUC, ranging from 0.5 to 1.0 [45].

### 2.13. Statistical Analysis.
R version 4.1.0 was applied to analyses in this research. $P < 0.05$ was considered significant.

## 3. Results

### 3.1. Genetic and Transcriptional Changes of NRGs in OV.
The analysis flow chart is performed in Figure S1. Figure 1(a) presents the summary outcome of the incidence of somatic mutations in 67NRGs.There were 89 mutations that occurred in 436 samples (20.41%) where ATRX and ALK had highest mutation frequency (2%). The CNV of MYC and TNFSF10 increased significantly, while the CNV of TARDBP, TNFRSF21, HDAC9, AXL, TLR3, and CYLD were decreased (Figure 1(b)). The location of CNV of 67NRGs on chromosome is exhibited in Figure 1(c). We also compared the discrepancy of genes in control and OV samples (Figure 1(d)). Among the 67 genes, 38 genes had different expression levels, and the corresponding OS was also different (Figure S2). Genetic and transcriptional levels of NRGs between OV and control tissues were different which means that NRGs have a significant role in the progression of OV.

### 3.2. Discrimination of NRGclusters in OV.
The state of NRG interactions and regulator connections OV populations are

(a)



(b)



(c)



(d)



(e)



(f)
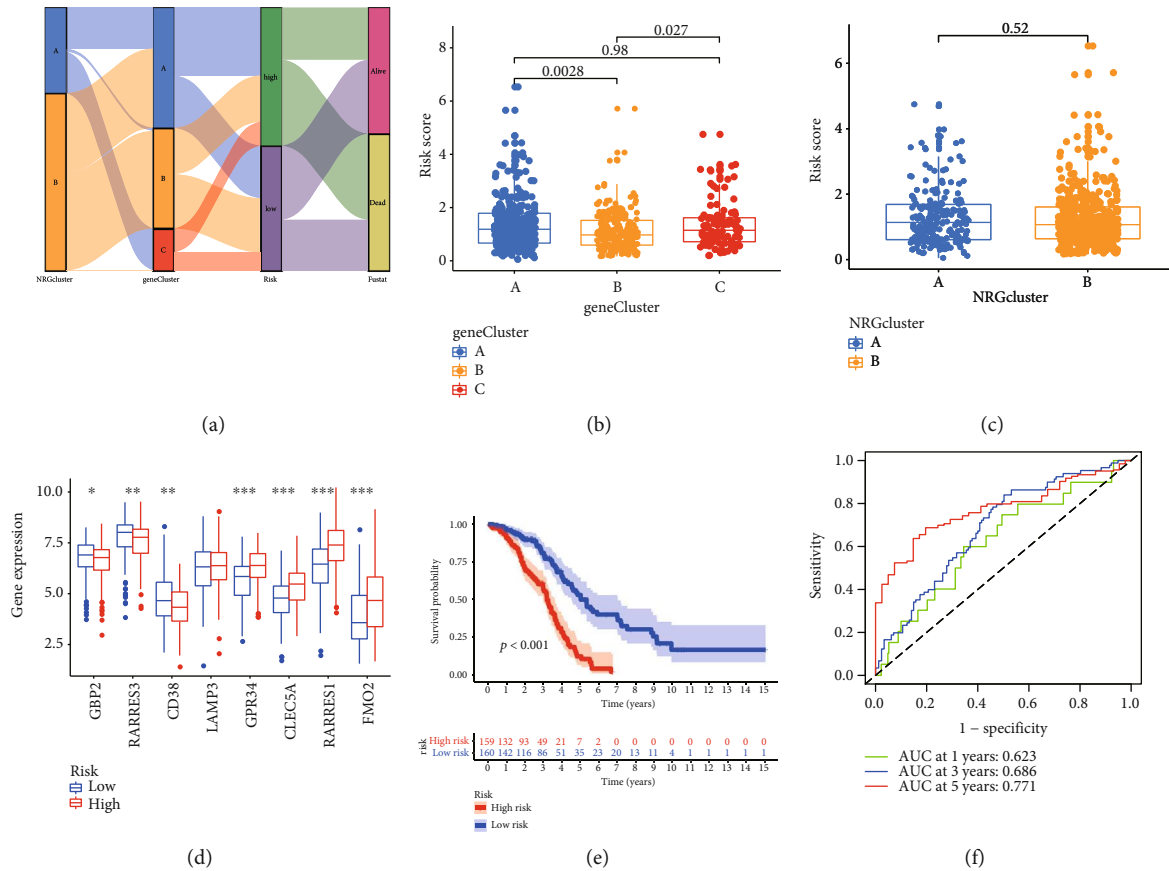
FIGURE 5: Construct NRG score based on the training set. (a) Distribution of groups with different classification criteria. Difference of NRG score between (b) gene clusters and (c) NRGclusters. (d) Expression of the 8 NRGs between risk groups. (e) The Kaplan–Meier analysis shows survival probability. (f) Assess sensitivity and specificity NRG score prediction in 1, 3, and 5 years. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

presented in Figure 2(a). For further analysis of the express characteristics of NRG in OV, we classified OV patients from $k = 1$ to $k = 9$ (Figure S3). PCA showed discrepancies in necroptosis transcription between the NRGclusters (Figure 2(b)). The Kaplan–Meier curve implied that NRGcluster A had higher survival probability than patients in NRGcluster B (Figure 2(c)). Furthermore, there were discrepancies in NRG expression and clinicopathological features among different OV subtypes (Figure 2(d)). Compared with NRGcluster B, patients in NRGcluster A had older age, more advanced stage and grade, and worse survival status.

3.3. Characteristics of Different Subtypes of TME. As performed in Figure 3(a), some immune activation-related processes like B cell receptor signaling pathway were enriched in NRGcluster B, which indicate immune activation (Figure 3(a)). The immune infiltration scores of NRGcluster A and NRGcluster B were compared, showing a great difference. The immune infiltration level in NRGcluster A was lower than that in NRGcluster B. Innate and adaptive immune cells were enriched in NRGcluster B (Figure 3(b)). Then, we calculated the association between two RNA modified subtypes and 22 TIICs. The proportion of immune cells was higher in NRGcluster B which means that NRGcluster B may related to immune activation (Figure 3(c)). Therefore,

we estimated the TME score of the NRGclusters (Figures 3(d)–3(f)). A higher estimate score represents a higher fraction of stromal and immune cells. The outcomes indicated that patients in NRGcluster B had a higher TME score. We noticed that the two NRGclusters had different immune infiltrations. Features of NRGcluster A were similar to the definition of "cold" tumors, which has less invasive immune cells and may achieve less benefit from immune therapy, while NRGcluster B is roughly similar to "hot" tumors. Regarding the express levels of immune checkpoints in two NRGclusters, we noticed that PD1 (Figure 3(g)), CTLA4 (Figure 3(h)), PDL1 (Figure 3(i)), and PD-L2 (Figure 3(j)) had a high expression level in NRGcluster B. The HLA expression of the NRGclusters was also different (Figure 3(k)). We then found that some immune-related pathways were more prominent in NRGcluster B (Figure 3(l)).

3.4. Identification of Gene Cluster Based on DEGs. We screened out DEGs (Figure S4a) and then conducted GO and KEGG analyses which revealed that NRGs were mainly associated with immune-related processes which means that they are crucial in the immune regulation of TME (Figures 4(a) and 4(b)). The best number of clusters is three (Figure S4b-S4e). Gene cluster B had the highest survival probability and was related to early stage, early grade, younger age, and better survival status (Figures 4(c)

TABLE 1: Univariate and multivariate Cox regression analyses of the prognosis-related variables.

| Variable | Univariable model | | | | Multivariable model | | | |
|---|---|---|---|---|---|---|---|---|
| | HR | HR.95 L | HR.95H | *P* value | HR | HR.95 L | HR.95H | *P* value |
| Training set | | | | | | | | |
| Age | 1.4679 | 1.0659 | 2.0215 | 0.0187 | 1.4287 | 1.0361 | 1.9702 | 0.0296 |
| Grade | 1.1163 | 0.7573 | 1.6454 | 0.5784 | | | | |
| Stage | 11.9886 | 1.6752 | 85.7985 | 0.0134 | 9.0377 | 1.2600 | 64.8256 | 0.0285 |
| Risk score | 1.9319 | 1.6308 | 2.2885 | 0.0001 | 1.8825 | 1.5865 | 2.2338 | 0.0001 |
| Testing set | | | | | | | | |
| Age | 1.4023 | 1.0347 | 1.9005 | 0.0293 | 1.3802 | 1.0179 | 1.8715 | 0.0380 |
| Grade | 1.2773 | 0.8671 | 1.8814 | 0.2156 | | | | |
| Stage | 2.7702 | 1.2974 | 5.9149 | 0.0085 | 2.4765 | 1.1551 | 5.3098 | 0.0198 |
| Risk score | 1.3176 | 1.1442 | 1.5173 | 0.0001 | 1.2787 | 1.1047 | 1.4800 | 0.0010 |
| All set | | | | | | | | |
| Age | 1.4182 | 1.1398 | 1.7645 | 0.0017 | 1.3772 | 1.1066 | 1.7139 | 0.0041 |
| Grade | 1.2049 | 0.9166 | 1.5839 | 0.1816 | | | | |
| Stage | 3.9317 | 1.9467 | 7.9407 | 0.0001 | 3.3380 | 1.6495 | 6.7551 | 0.0008 |
| Risk score | 1.4904 | 1.3431 | 1.6538 | 0.0001 | 1.4498 | 1.3031 | 1.6130 | 0.0001 |
| GEO set | | | | | | | | |
| Age | 1.5109 | 1.0347 | 2.2063 | 0.0326 | 1.5138 | 1.0358 | 2.2124 | 0.0322 |
| Grade | 1.3183 | 0.8887 | 1.9555 | 0.1696 | | | | |
| Stage | 6.8898 | 2.1835 | 21.7396 | 0.0010 | 5.9038 | 1.8625 | 18.7139 | 0.0026 |
| Risk score | 1.3764 | 1.1772 | 1.6093 | 0.0001 | 1.3164 | 1.1175 | 1.5506 | 0.0010 |

and 4(d)). In addition, patients in gene cluster B and NRGcluster B had similar clinical characteristics. The gene expression in three gene clusters was different (Figure 4(e)). The outcome of ssGSEA showed that the vast immune cells had higher infiltration levels in gene cluster B (Figure 4(f)), which was primarily infiltrated by adaptive immune cells (Figure 4(g)). In addition, gene cluster B tends to have higher TME score and immune checkpoints expression (Figures 4(h)–4(n)). HLA expression levels of gene clusters in the three groups were also the highest expression in gene cluster B (Figure 4(o)). In addition, classical biological progresses were more prominent in gene cluster B (Figure 4(p)). According to these features, we considered that gene cluster B belongs to "hot" tumors.

3.5. Build and Testing the Prognostic Signature. Figure 5(a) displays the spread of patients in three gene clusters and two NRGclusters. 14 OS-related genes were screened as candidate prediction genes (Figure S5a-S5b). Finally 8 achieved genes (GBP2, RARRES3, CD38, LAMP3, GPR34, CLEC5A, RARRES1, and FMO2) were selected. Among them, GBP2, RARRES3, and CD38 were protective genes (Figure S5c). According to the above results, we assessed the NRG score: risk score = $(-0.5265 \times$ GBP2 expression) + $(-0.2565 \times$ RARRES3 expression) + $(-0.2786 \times$ CD38 expression) + $(0.2053 \times$ LAMP3 expression) + $(0.4263 \times$ GPR34 expression) + $(0.2201 \times$ CLEC5A expression) + $(0.2249 \times$ RARRES1 expression) + $(0.1025 \times$ FMO2 expression). We noticed that the score of gene cluster B was lower than that of gene

cluster A (Figure 5(b)). In the grouping according to NRGcluster, there was no diversity of risk score between the two subgroups (Figure 5(c)). Most of the 8NRGs had different expression between the two risk subgroups in training set (Figure 5(d)). Individuals with low scores had higher OS when compared with higher score group (Figure 5(e)). In addition, the AUC of NRG score for 1, 3, and 5 years were 0.623, 0.686, and 0.771, respectively (Figure 5(f)). Then, we use the testing group, all groups from TCGA, and the data from GEO to verify the above results. Figure S6 presents the difference of NRGs expression, survival analysis, and AUC in two risk groups in testing set, all sets, and GEO set, respectively. There were distinctions in the eight gene expressions between the risk groups. The AUC of NRG score for 1, 3, and 5 years is still high, which means that the model had excellent predict ability.

3.6. The Difference of Risk Score between Different Feature Patients. To analyze the relationship between NRG score and clinical features, we compared the risk scores of different individuals. It was found that the NRG score of OV patients in stages I-II was lower than in stages III-IV (Figure S7a). Moreover, the risk score of OV patients with better survival status was also lower than that of OV patients with worse survival status (Figure S7b). The NRG score was proved to be an independent prognostic variable (Table 1). In different age, stage, and grade subgroups, the OS of the high NRG score group tends to be lower (Figure S7c). Furthermore, for BRCA1 wild

(a)



(b)



(c)



(d)



(e)



(f)
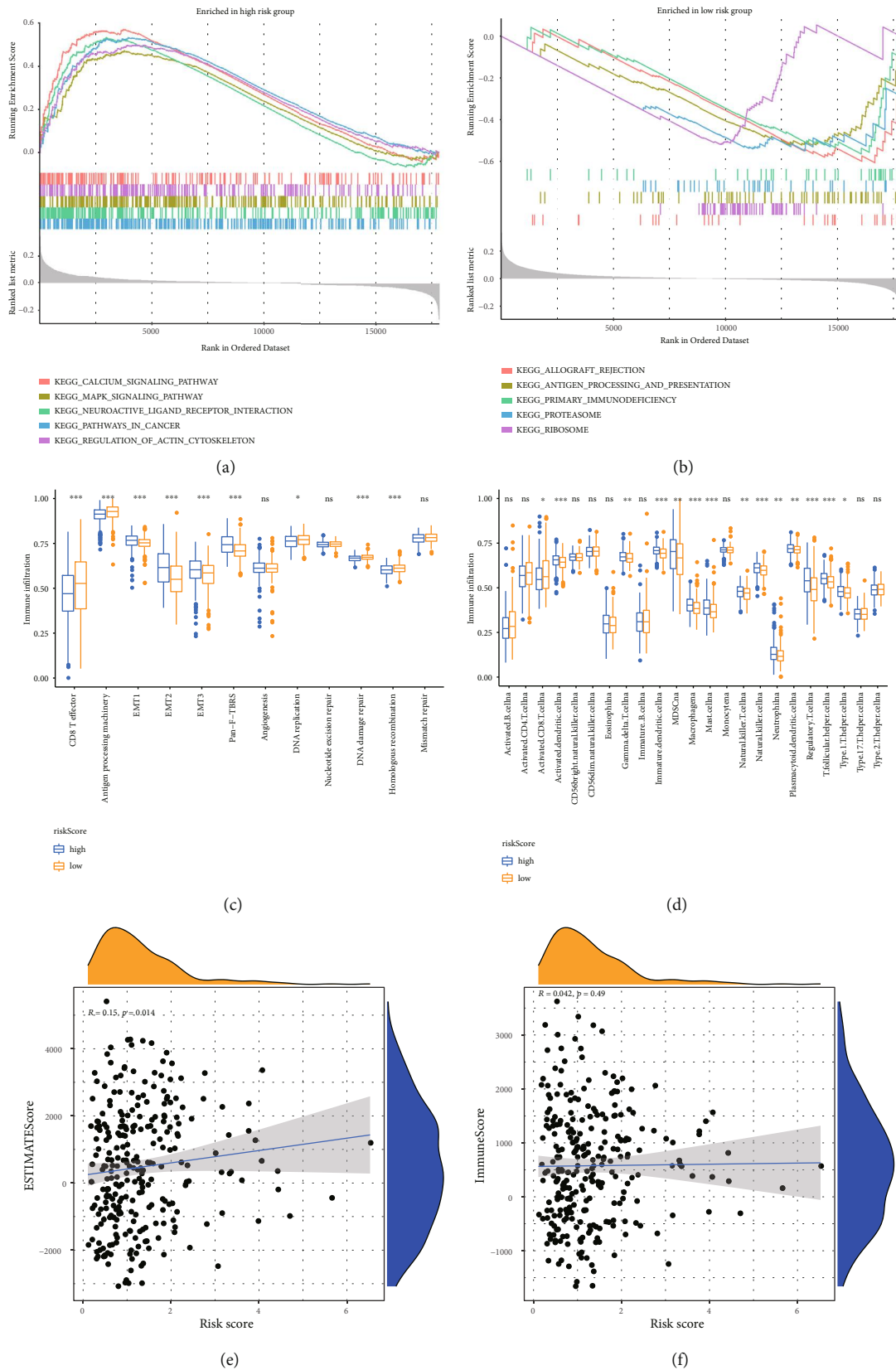
Figure 6: Continued.

(g)



(h)



(i)

Figure 6: Continued.

(j)
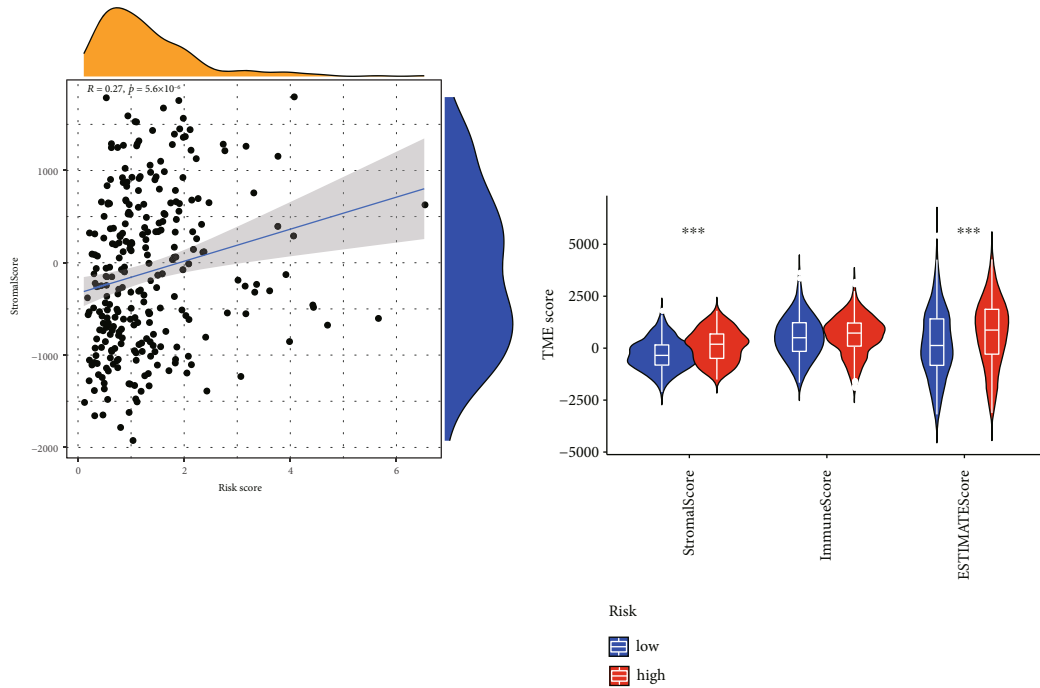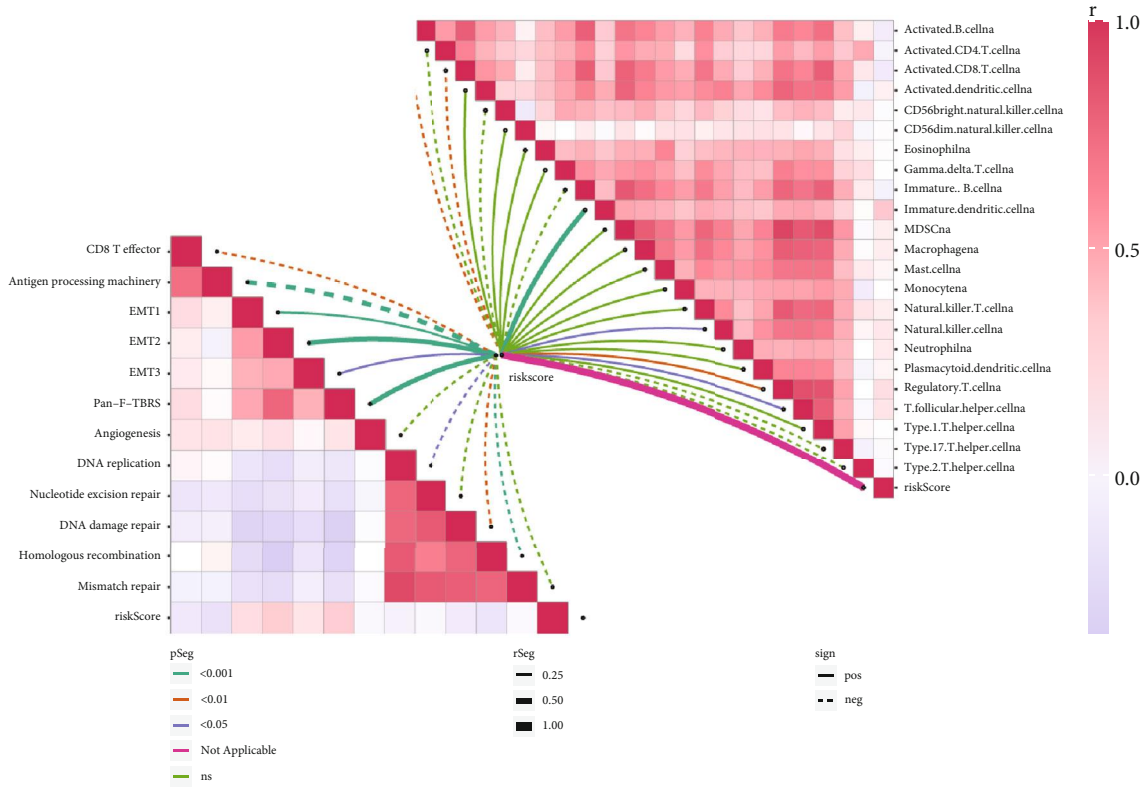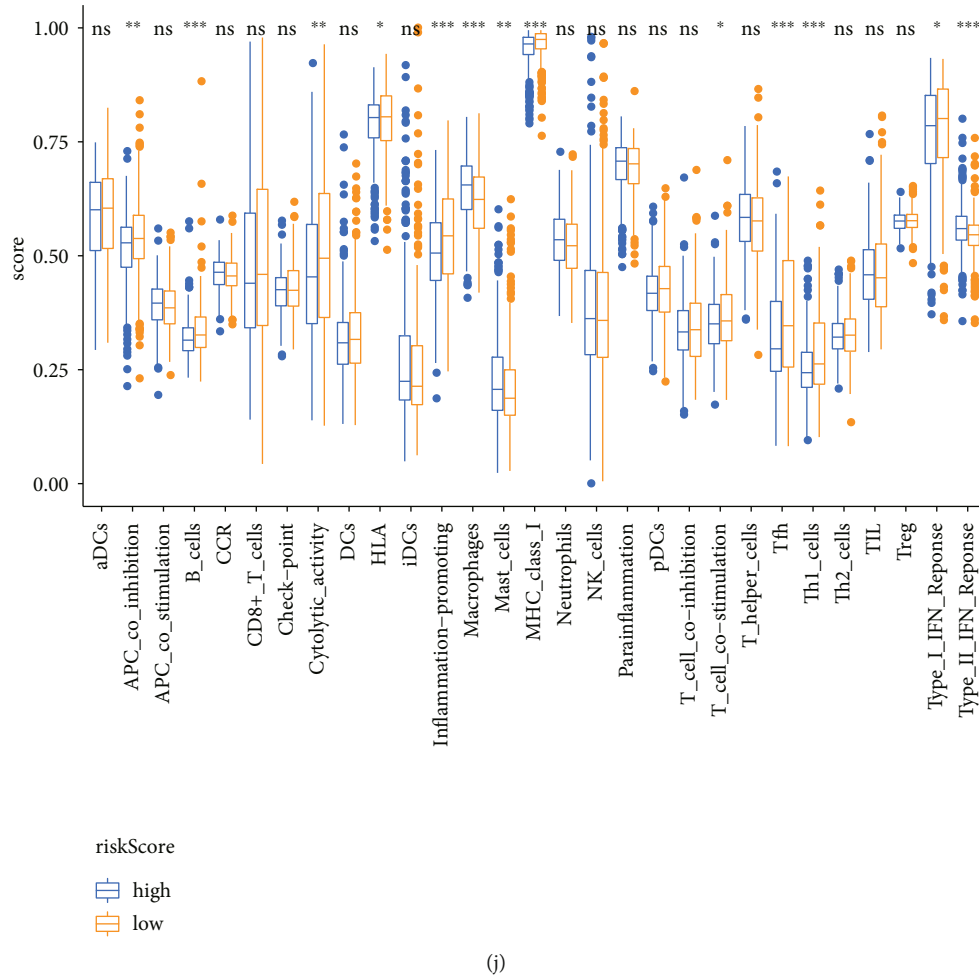
Figure 6: Evaluate TME of different risk groups. Main enriched biological pathways of (a) low NRG score group and (b) high-score group. (c, d) The ssGSEA score and immune infiltration score of risk groups. (e–h) The NRG score had positive association with stromal cells, immune cells, and estimated score. (i) The relevance of risk score and immune cells as well as classical biological pathway score. (c, j) The difference of immune function score betweenthe groups.

patients and chemotherapy treated patients, the survival probability of patients with higher risk score was lower.

3.7. Assessment of TME in Terms of the NRGs. For the purpose of deepen understanding the TME of subgroups, we first conducted GSEA and found that the low NRG score group was mainly related with some immune-related processes (Figure 6(a)), while high-score individuals were associated with cancer-related processes (Figure 6(b)). Further studies indicated that pan-F-TBRS was obviously activated in the high group (Figure 6(c)). People with higher risk score had higher immune score and stromal score (Figures 6(e)–6(g)). Not surprisingly, the combined estimated score of these two scores was also higher in the high-risk group. We then identified the relationships between immune cells and risk score as well as the correlations between risk score and the score of classical biological pathways enrichment (Figures 6(h) and 6(i)). The immune function score of B cells, T helper cells, and Tfh was significantly higher in low-risk groups (Figure 6(j)). Then, we calculated the correlation between risk score and immune cell abundance (Figure 7(a)). The NRG score was positively

associated with macrophages M2, mast cells activated, monocytes, neutrophils, and T cells CD4 memory resting and had negative relationship with macrophages M1, plasma cells, etc. (Figure 7(b)). In terms of oxidative stress, the expression of oxidative stress-related genes was low in low-risk group, especially CYBB (Figure S8). Moreover, we noticed that a great deal of immune cells was related to the genes (Figure 7(c)). Human leukocyte antigen expression was also higher in lower risk cohorts (Figure 7(d)). Figure 7(e) shows that many immune checkpoints were overexpressed in patients in high risk. There was discrepancy of immune checkpoints expression between the groups. CTLA4, CD274, PDCD1, and IDO1 had negatively correlation with risk score, and HAVCR2 has positive relationship with the score (Figure 7(f)). The outcome of IPS score indicates that low-risk score was associated with higher immunogenicity (Figure 7(g)).

3.8. Predict the Curative Effect of Immunotherapy. Figure 8(a) shows the relationships between the NRG score and CSC index. We noticed that the NRG score was negatively associated with CSC index ($r = -0.35$, $P < 0.001$)

(a)



(b)

FIGURE 7: Continued.

(c)



(d)



(e)



(f)



(g)
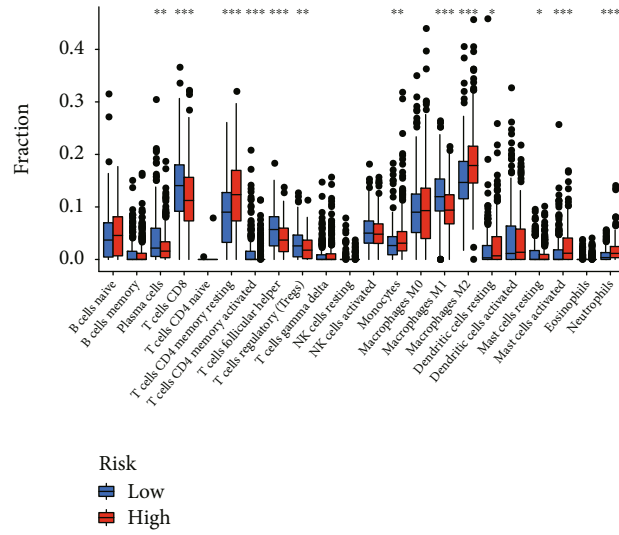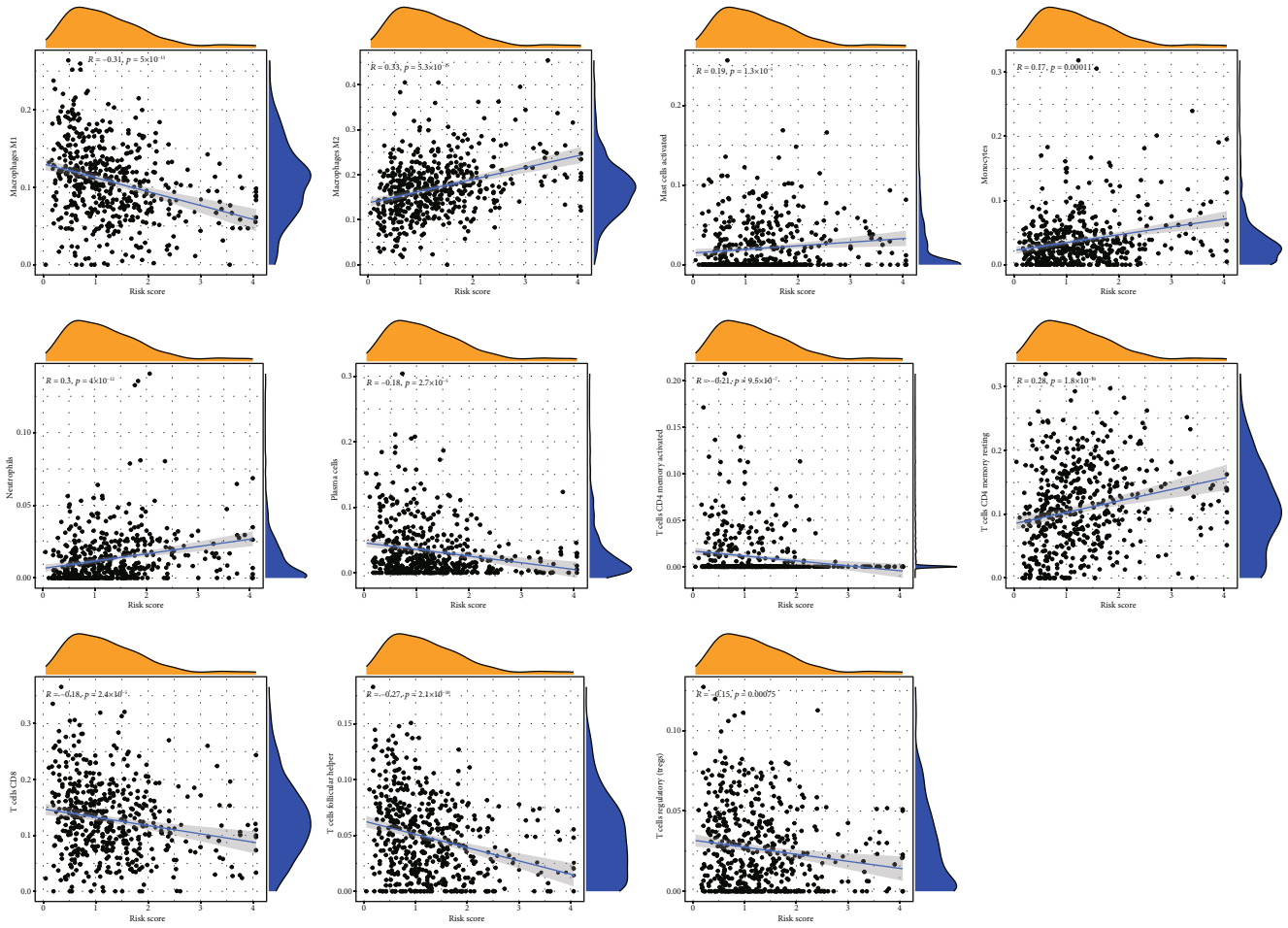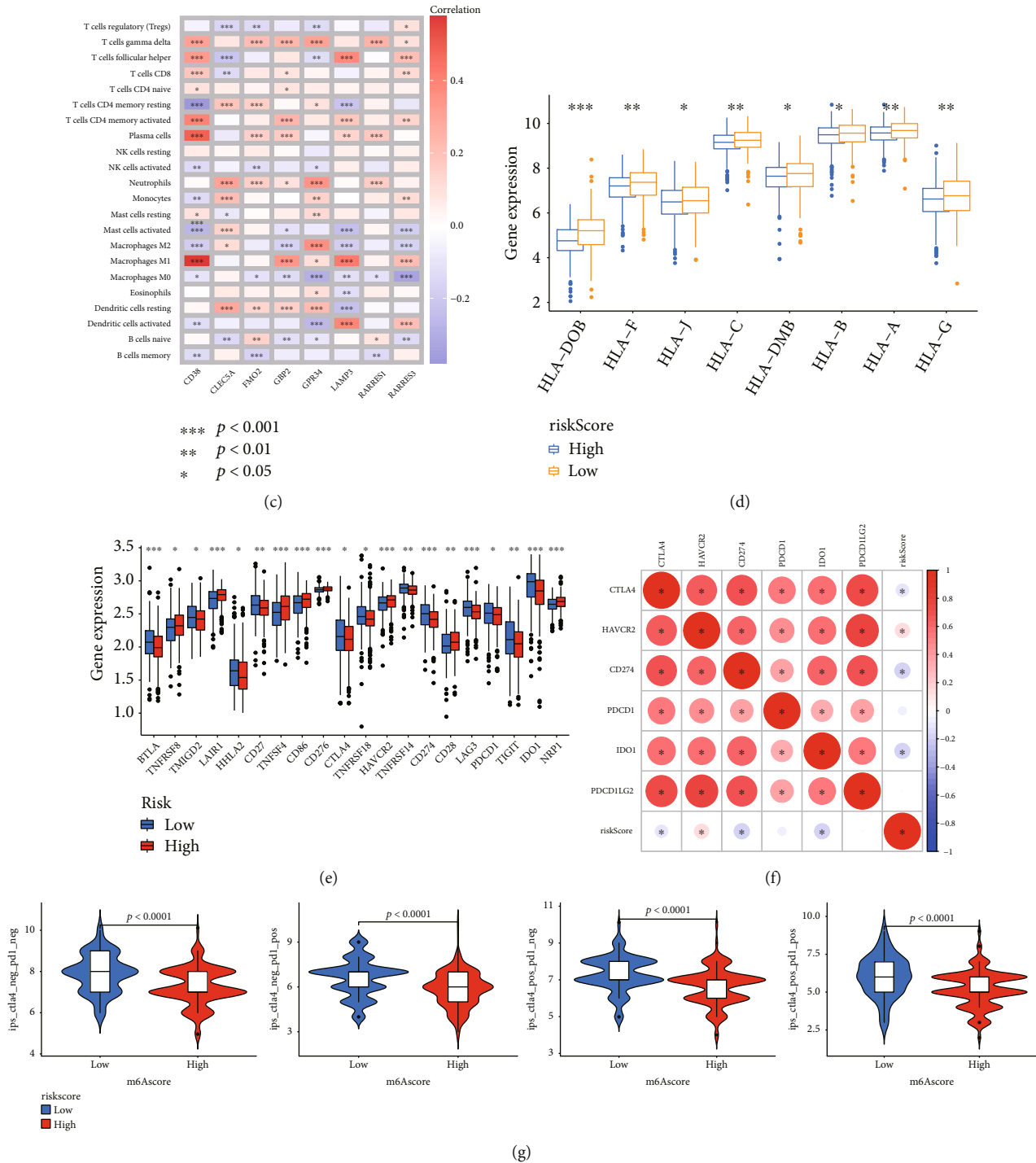
FIGURE 7: Immune infiltration situations of the subtypes. (a) The difference of immune cell abundance between the groups. (b) The correlation between 8 genes and immune cell abundance. (c) The relationship between risk score and immune cell abundance. (d) The difference of HLA expression between the groups. (e) Twenty immune checkpoints with differential expression in the two groups were depicted. (f) The correlation between immune checkpoints and risk score. (g) The differences of IPS cell expression. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

which showed that CRC cells with lower risk score have more obvious stem cells and lower cell differentiation. L-TMB accompanied with high risk means a significant poorer survival probability than other groups (Figures 8(b) and 8(c)). Mutation frequencies of TP53 and TTN were high in both cohorts (Figures 8(d) and 8(e)). Accompanied with high-risk score, OS becomes lower (Figures 8(f) and 8(g)).

Figure 8(h) shows the distribution of GISTIC scores on all chromosomes. Focal amplification and deletion of different chromosome regions were found (Figures 8(i) and 8(j)).

*3.9. Estimation of Drug Sensitivity.* We chose drugs usually applied in the remedy of OV to assess the sensitivity of patients to these drugs. IC50 values of docetaxel in high-

(a)



(b)



(c)



(d)



(e)



(f)

FIGURE 8: Continued.

(g)



(h)



(i)

FIGURE 8: Continued.

(j)

FIGURE 8: Assess the efficacy of immunotherapy. (a) RNAss was negatively correlated with NRG score. (b, c) Survival probability of people with different TMB and risk score. (d, e) The situation of gene mutation in different risk groups. (f, g) Survival probability of people with TP53 mutation and TTN mutation. (h) Copy number score for the groups. (i, j) Cytoband shows amplification (left) and deletion (right).

risk patients were lower, while IC50 of A.443654 and pacli-taxel was lower in people with low-risk score (Figure S8a–S8c). We also calculated the relationships between some common drugs and 8 genes. Taken together, the above results suggested that NRGs are correlated with drug sensitivity (Figure S8d).

*3.10. Construct Nomograms for Clinical Application.* Taking care of the practical utilize of NRG score in predicting, we built a nomogram containing NRG score and clinical factors (Figure 9(a)), predictors including NRG risk score, age, nomogram risk, and stage. Our signature had higher $C$-index (Figure 9(b)). The AUC corresponding to NRG risk was generally higher which indicates great prediction perfor-mance, and it will be better when considering age and stage (Figures 9(c)–9(e)). DCA indicated that the NRG risk score or nomogram risk combined with clinical features had a higher benefit in predicting the OS of OV patients at 1, 3, and 5 years (Figures 9(f)–9(h)). A subsequent calibration diagram proved it again (Figure 9(i)).

*3.11. NRG Model Has Great Prognostic Performance.* To con-trast the prognosis ability of our signature with other signa-tures, we screened four prognosis models from the previous literatures. We used the multivariate Cox regression analysis to assess the estimate score, based on specific genes expres-sion (Figure S10a). Figure S10b indicated that the prognosis of high-risk individuals was worse in all four models. Obviously, our model has the highest $C$-index which was 0.65 (Figure S10c). Therefore, our genetic characteristics performed best in about six years (Figure S10d).

## 4. Discussion

Despite progress in study and remedy of OV, the 5-year sur-vival rate is still low [46], and more than half of the patient

relapse and develop drug resistance [47–50]. Cell death inhi-bition is the ultimate cause of drug resistance in OV [51]. As the main type of cell death, previous studies mostly focused on the drug resistance of apoptosis in OV [52]. Necroptosis is a newly noticed type of regulatory necrosis which has been proven to have great effect on cancer, especially in drug resistance [53]. Therefore, our exploration may improve the poorer outcome of OV.

Patients in NRGcluster A had more advanced clinical characteristics and poorer survival than patients in NRGcluster B. There are also distinctions in the features of TME between the two NRGclusters. The TME score and immune checkpoint expression was higher in NRGcluster B patients. We screened three gene clusters in the light of DEGs. Then, we established an effective prognostic risk score and verified its predict performance. There were great differences in clinical features, TMB, TME, immune check-point, $C$-index, CNV, and drug sensitivity between the risk groups. Finally, the nomogram was established to further enhance the performance and promote use of NRG score.

The prediction model has a close correlation with redox stress and immune environment. The high-risk group has higher levels of redox stress, which may be closely related to their poor prognosis. Despite advances has been achieved in immunotherapy recently, outcomes of OV patients have still been heterogeneous, indicating the effect of TME in the occur-rence and development of OV [54].TME is an ecosystem con-sisting of tumor cells, infiltrating immune cells and stromal cells intertwined with noncellular components. In this study, the necroptosis pattern with immune inhibition was related to higher NRG score, and the necroptosis pattern with immune activation was related to lower NRG score. Macro-phages M1, also known as "classic activated macrophages," has a proinflammatory effect. Their high expression was associated with a better prognosis in patients with OV [55]. OV-associated memory T cells are also associated with

(a)



(b)



(c)

Figure 9: Continued.

(d)

(e)

(f)

(g)

(h)

(i)
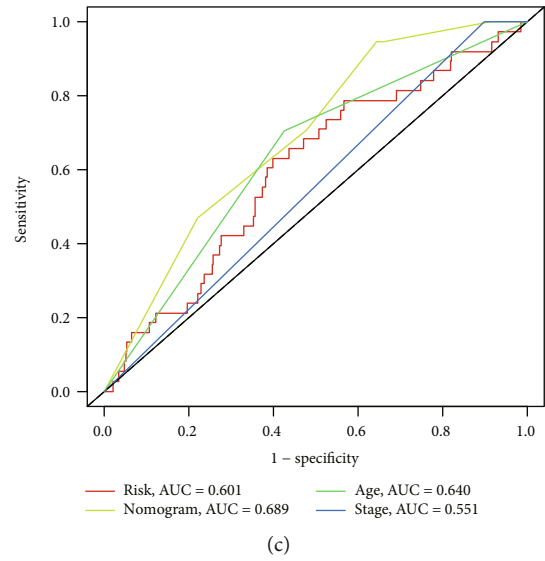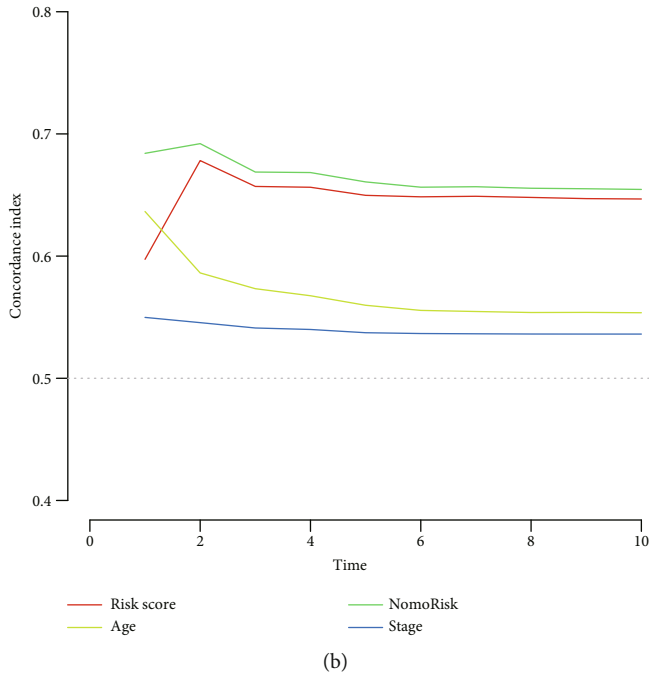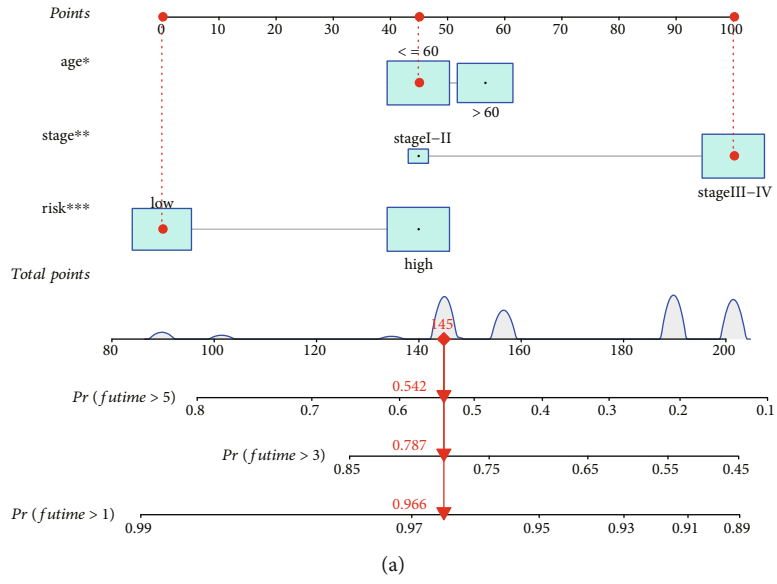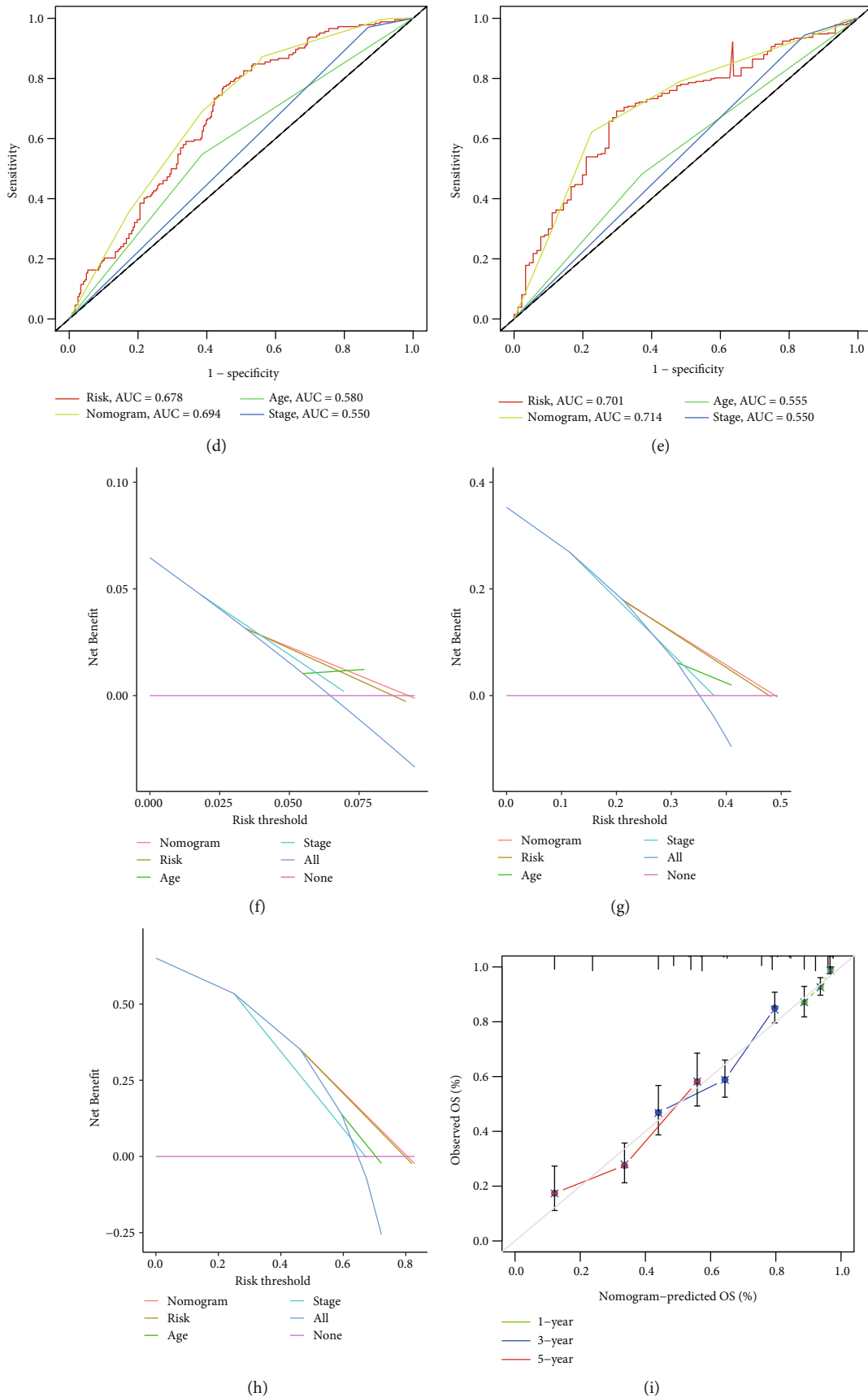
FIGURE 9: Built and verification of nomogram. (a) Nomograms used to predict OS in 1, 3, and 5 years. (b) C-index of prognostic factors including risk score. (c–h) ROC and DCA curves of 1, 3, and 5 years. (i) Nomogram calibration curve of the model.

chemotherapy response and longer survival [56]. CD4 T cells have crucial effect on almost every aspect of immunity and are considered an important component needed in tumor immunotherapy [57]. Plasma cell infiltration is related with high CD4 and CD8 T cell response and great prognosis [58, 59], while NRG score was positively associated with macrophage M2, mast cell activation, monocyte, neutrophil, and T cell CD4 memory rest. The more macrophages M2, the worse the prognosis of patients with advanced OV [60]. Tumors with high mast cell are associated with immunosuppressive OV TME and are potentially insensitive to immunotherapy [61]. Monocytes are recruited around the cancer and differentiated into macrophages, which can be used as biomarkers of OV progression [62]. Neutrophils are key players in OV and have been considered new biomarkers of cancers or as immunotherapy targets to promote tumor progression [14].

TME is closely correlated with the response of patients with various cancers to immunotherapy, and patients with immunodominant TME subtypes benefit the most from immunotherapy [63]. In the past decades, immunotherapy, especially the treatment using ICIs, was developing rapidly. The researches on ICI are booming, and clinical studies have proved their safety and effectiveness. In this study, we observed that CTLA4, CD274, PDCD1, and IDO1 had a negative correlation with risk score, and HAVCR2 has a positive association with risk score. Among them, researchers have a deeper understanding of CTLA4 and PDCD1. Evidence from non-OV shows that patients with hot tumors infiltrated by immunogenic T cells have lasting clinical benefits in PD-1/PD-L1 blocking response compared with individuals with cold tumors [64]. However, their effect in OV is not clear. Whether they can be used as targets of immunotherapy in clinical still needs further research. Immunotherapy conducted on these patients may obtain better curative effect.

In addition, we found that OS decreased significantly when TP53 and TTN mutation particularly combined with high-risk score. Proteins encoded by some major target genes regulated by TP53 are essential for maintaining genomic integrity and cell life cycle [65]. TTN mutations are closely associated with the response to immune checkpoint blockade (ICB) [66, 67]. However, previous articles have not clearly discussed whether TTN or TP53 mutation has an impact on the immunotherapy effect of OV. Human CNV is a repetitive or missing DNA fragment relative to the reference genome, which may lead to genomic imbalance and diseases such as tumor. So, it is correlated with the process of diagnosis and prognosis [68, 69]. CNV has been tested to be related to the prognosis of OV [70]. Low-risk patients have more gene mutations, and CNV load belongs to immune activation subgroup.

However, immunotherapy using ICB alone is less effective in the treatment of OV [71]. Therefore, it is necessary to treat OV patients by ICB combined with chemotherapy, radiotherapy, and other therapeutic methods. Cisplatin and its derivatives are commonly used in OV chemotherapy. It has been determined that cisplatin can induce necroptosis and significantly increase the death of OV cells, to improve the anticancer effect of chemotherapeutics [72, 73]. Taxane cytotoxic drugs such as docetaxel have become one of the most effective drugs for the immunotherapy of gynecological cancer recently. It has been approved for the remedy of OV, breast cancer, and so on [74, 75]. Paclitaxel can induce immunogenic cell death in OV and achieve therapeutic effect [76]. We found IC50 values of docetaxel in high-risk patients were lower, while IC50 values of A.443654 and paclitaxel were lower in low-risk patients. Therefore, different chemotherapy drugs can be used for patients in different risk groups, which may get better therapeutic effect.

This study also has some shortcomings. Firstly, the necroptosis genes included in this study were achieved from previous articles. Some unreported NRGs might be ignored. Secondly, the prognostic model constructed for OV in this study needs to be verified in clinical application. Therefore, we need to screen new genes related to necroptosis in more cohorts and collect enough cases and clinical information of OV in the future, to ensure that the model is effective for clinical application.

## 5. Conclusions

Our comprehensive analysis of NRG reveals its impact on TME, clinical characteristics, and prognosis of OV. The therapy role of NRGs in immunotherapy was also analyzed. The above results emphasize the significance of NRGs and offer new orientation for guiding precision therapy strategy of OV patients.

## Abbreviations

NRGs:    Necroptosis-related genes
OV:      Ovarian cancer
TME:     Tumor microenvironment
LASSO:   Least absolute shrinkage and selection operator
CNV:     Copy number variations
TIL:     Tumor infiltrating lymphocyte
DEG:     Differentially expressed gene
TCGA:    The cancer genome atlas
GEO:     Gene Expression Omnibus
OS:      Overall survival
PCA:     Principal component analysis
GSVA:    Gene set variation analysis
ROC:     Receiver operating characteristic curve
GO:      Gene Ontology
KEGG:    Kyoto Encyclopedia of Genes and Genomes
GSEA:    Gene set enrichment analysis
ssGSEA:  Single-sample gene set enrichment analysis
CSC:     Cancer stem cell
IPS:     Induced pluripotent stem
TMB:     Tumor mutation load
ICI:     Immune checkpoint inhibitor
IC50:    Half maximum inhibitory concentration
NCI-60:  60 types of cancer cell
NCI:     National Cancer Institute's Center for Cancer Research
AUC:     Area under curve
DCA:     Decision curve analysis
RMS:     Restricted mean survival
ICB:     Immune checkpoint blockade
TIIC:    Tumor immune infiltrating cells.

## Data Availability

The data used to support the findings of this study are available from the corresponding author on reasonable request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Authors' Contributions

Jinhui Liu, Jianling Bai, and Hao Yu contributed to the conception of the study; Rui Geng, Zihang Zhong, and Senmiao Ni performed the data analyses and wrote the manuscript; Wen Liu and Zhiqiang He contributed significantly to analysis and manuscript preparation; Shilin Gan and Qinghao Huang helped perform the analysis with constructive discussions. Rui Geng, Zihang Zhong, and Senmiao Ni contributed equally to this work.

## Acknowledgments

## Supplementary Materials

Figure S1: flow chart of this study. Figure S2: OS of patients with different level of NRGs. Figure S3: clustering analysis determine the best number of clusters. (a-b) Uniform clustering cumulative distribution function (CDF) and the change of area under CDF curve, $k = 2 - 9$ ($k$ means the number of clusters). (c) The samples were divided into two clusters when $k = 2$. (d) Tracking plot of the cluster. Figure S4: clustering analysis determine the best number of clusters. (a-b) CDF and the area under CDF curve from k equal two to $k$ equal 9. (c) The samples were divided into two clusters when $k = 3$. (d) Tracking plot of the cluster. Figure S5: identifying genes with predict potential. (a-b) LASSO regression analysis and partial likelihood deviance on the prognostic genes. (c) Forest plot shows the result of multivariate Cox regression analysis. Figure S6: validate prognostic ability of NRG model expression level of NRGs in testing set (a), all set (b), and GEO set (c). Kaplan–Meier survival corves of samples in testing set (d), all set (e), and GEO set (f). ROC curves of 1, 3, and 5 year survival in testing set (g), all set (h), and GEO set (i). Figure S7: attest the prognosis performance of the model by stratified analysis. Compare the risk scores of individuals in different stages (a) and had different fustat (b). (c) Comparing survival time of patients with different clinical features. Figure S8: expression of oxidative stress-related genes in the groups. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Figure S9: evaluate drug sensitivity. The IC50 of A.443654 (a), docetaxel (b), and paclitaxel (c) in samples. (d) Correlativity of common drugs and gene expression level. Figure S10: compare NRG model with some established models (a-b) ROC curves and KM curves of four published gene-signatures. (c-d) C-index and RMS time curve of prognostic models. (Supplementary Materials)

## References

[1] L. Kuroki and S. R. Guntupalli, "Treatment of epithelial ovarian cancer," *BMJ*, vol. 371, article m3773, 2020.

[2] H. Sung, J. Ferlay, R. L. Siegel et al., "Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[3] S. Lheureux, M. Braunstein, and A. M. Oza, "Epithelial ovarian cancer: evolution of management in the era of precision medicine," *CA: a Cancer Journal for Clinicians*, vol. 69, no. 4, pp. 280–304, 2019.

[4] R. C. Bast Jr., U. A. Matulonis, A. K. Sood et al., "Critical questions in ovarian cancer research and treatment: report of an American Association for Cancer Research Special Conference," *Cancer*, vol. 125, no. 12, pp. 1963–1972, 2019.

[5] L. C. Peres, K. L. Cushing-Haugen, M. Köbel et al., "Invasive epithelial ovarian cancer survival by histotype and disease stage," *Journal of the National Cancer Institute*, vol. 111, no. 1, pp. 60–68, 2019.

[6] F. Reid, N. Bhatla, A. M. Oza et al., "The world ovarian cancer coalition every woman study: identifying challenges and opportunities to improve survival and quality of life," *International Journal of Gynecological Cancer*, vol. 31, no. 2, pp. 238–244, 2021.

[7] K. Odunsi, "Immunotherapy in ovarian cancer," *Annals of Oncology*, vol. 28, suppl_8, p. viii1-viii7, 2017.

[8] C. Yang, B. R. Xia, Z. C. Zhang, Y. J. Zhang, G. Lou, and W. L. Jin, "Immunotherapy for ovarian cancer: adjuvant, combination, and neoadjuvant," *Frontiers in Immunology*, vol. 11, article 577869, 2020.

[9] P. Vandenabeele, L. Galluzzi, T. Vanden Berghe, and G. Kroemer, "Molecular mechanisms of necroptosis: an ordered cellular explosion," *Nature Reviews. Molecular Cell Biology*, vol. 11, no. 10, pp. 700–714, 2010.

[10] A. Najafov, H. Chen, and J. Yuan, "Necroptosis and cancer," *Trends Cancer*, vol. 3, no. 4, pp. 294–301, 2017.

[11] Y. Gong, Z. Fan, G. Luo et al., "The role of necroptosis in cancer biology and therapy," *Molecular Cancer*, vol. 18, no. 1, p. 100, 2019.

[12] M. E. Choi, D. R. Price, S. W. Ryter, and A. M. K. Choi, "Necroptosis: a crucial pathogenic mediator of human disease," *Insight*, vol. 4, no. 15, 2019.

[13] X. Qin, D. Ma, Y. X. Tan, H. Y. Wang, and Z. Cai, "The role of necroptosis in cancer: a double-edged sword?," *Biochimica Et Biophysica Acta. Reviews on Cancer*, vol. 1871, no. 2, pp. 259–266, 2019.

[14] D. Baci, A. Bosi, M. Gallazzi et al., "The ovarian cancer tumor immune microenvironment (TIME) as target for therapy: a focus on innate immunity cells as therapeutic effectors," *International Journal of Molecular Sciences*, vol. 21, no. 9, p. 3125, 2020.

[15] L. M. E. Janssen, E. E. Ramsay, C. D. Logsdon, and W. W. Overwijk, "The immune system in cancer metastasis: friend or foe?," *Journal for Immunotherapy of Cancer*, vol. 5, no. 1, p. 79, 2017.

[16] N. M. Anderson and M. C. Simon, "The tumor microenvironment," *Current Biology*, vol. 30, no. 16, pp. R921–R925, 2020.

[17] F. Pagès, B. Mlecnik, F. Marliot et al., "International validation of the consensus immunoscore for the classification of colon

cancer: a prognostic and accuracy study," *Lancet*, vol. 391, no. 10135, pp. 2128–2139, 2018.

[18] F. Petitprez, M. Meylan, A. de Reynies, C. Sautes-Fridman, and W. H. Fridman, "The tumor microenvironment in the response to immune checkpoint blockade therapies," *Frontiers in Immunology*, vol. 11, p. 784, 2020.

[19] T. Yamauchi, F. Fujishima, M. Hashimoto et al., "Necroptosis in esophageal squamous cell carcinoma: an independent prognostic factor and its correlation with tumor-infiltrating lymphocytes," *Cancers*, vol. 13, no. 17, p. 4473, 2021.

[20] H. H. Park, H. R. Kim, S. Y. Park et al., "RIPK3 activation induces TRIM28 derepression in cancer cells and enhances the anti-tumor microenvironment," *Molecular Cancer*, vol. 20, no. 1, p. 107, 2021.

[21] Z. Y. Liu, M. Zheng, Y. M. Li et al., "RIP3 promotes colitis-associated colorectal cancer by controlling tumor cell proliferation and CXCL1-induced immune suppression," *Theranostics*, vol. 9, no. 12, pp. 3659–3673, 2019.

[22] A. W. Zhang, A. McPherson, K. Milne et al., "Interfaces of malignant and immunologic clonal dynamics in ovarian cancer," *Cell*, vol. 173, no. 7, pp. 1755–1769.e22, 2018.

[23] C. Florean, S. Song, M. Dicato, and M. Diederich, "Redox biology of regulated cell death in cancer: a focus on necroptosis and ferroptosis," *Free Radical Biology & Medicine*, vol. 134, pp. 177–189, 2019.

[24] P. Poprac, K. Jomova, M. Simunkova, V. Kollar, C. J. Rhodes, and M. Valko, "Targeting free radicals in oxidative stress-related human diseases," *Trends in Pharmacological Sciences*, vol. 38, no. 7, pp. 592–607, 2017.

[25] B. M. Oxidants, "Oxidants, Antioxidants and thiol redox switches in the control of regulated cell death pathways," *Antioxidants*, vol. 9, no. 4, p. 309, 2020.

[26] T. Ismail, Y. Kim, H. Lee, D. S. Lee, and H. S. Lee, "Interplay between mitochondrial peroxiredoxins and ROS in cancer development and progression," *International Journal of Molecular Sciences*, vol. 20, no. 18, p. 4407, 2019.

[27] Z. Zhao, H. Liu, X. Zhou et al., "Necroptosis-related lncRNAs: predicting prognosis and the distinction between the cold and hot tumors in gastric cancer," *Journal of Oncology*, vol. 2021, Article ID 6718443, 16 pages, 2021.

[28] M. D. Wilkerson and D. N. Hayes, "ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, pp. 1572-1573, 2010.

[29] Y. Aihaiti, Y. Song Cai, X. Tuerhong et al., "Therapeutic effects of Naringin in rheumatoid arthritis: network pharmacology and experimental validation," *Frontiers in Pharmacology*, vol. 12, article 672054, 2021.

[30] R. K. Powers, A. Goodspeed, H. Pielke-Lombardo, A. C. Tan, and J. C. Costello, "GSEA-InContext: identifying novel and common patterns in expression experiments," *Bioinformatics*, vol. 34, no. 13, pp. i555–i564, 2018.

[31] 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, p. 2612, 2013.

[32] J. Wu, L. Li, H. Zhang et al., "A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma," *Oncogene*, vol. 40, no. 26, pp. 4413–4424, 2021.

[33] Y. Jiang, J. Chen, J. Ling et al., "Construction of a glycolysis-related long noncoding RNA signature for predicting survival in endometrial cancer," *Journal of Cancer*, vol. 12, no. 5, pp. 1431–1444, 2021.

[34] B. Chen, M. S. Khodadoust, C. L. Liu, A. M. Newman, and A. A. Alizadeh, "Profiling tumor infiltrating immune cells with CIBERSORT," *Methods in Molecular Biology*, vol. 1711, pp. 243–259, 2018.

[35] J. Zhong, Z. Liu, C. Cai, X. Duan, T. Deng, and G. Zeng, "M(6) a modification patterns and tumor immune landscape in clear cell renal carcinoma," *Journal for Immunotherapy of Cancer*, vol. 9, no. 2, p. e001646, 2021.

[36] F. Li, J. Xu, and S. Liu, "Cancer stem cells and neovascularization," *Cells*, vol. 10, no. 5, p. 1070, 2021.

[37] Z. Wang, Y. Wang, T. Yang et al., "Machine learning revealed stemness features and a novel stemness-based classification with appealing implications in discriminating the prognosis, immunotherapy and temozolomide responses of 906 glioblastoma patients," *Briefings in Bioinformatics*, vol. 22, no. 5, 2021.

[38] Q. Xu, S. Chen, Y. Hu, and W. Huang, "Landscape of immune microenvironment under immune cell infiltration pattern in breast cancer," *Frontiers in Immunology*, vol. 12, article 711433, 2021.

[39] D. M. Merino, L. M. McShane, D. Fabrizio et al., "Establishing guidelines to harmonize tumor mutational burden (TMB): in silico assessment of variation in TMB quantification across diagnostic platforms: phase I of the Friends of Cancer Research TMB Harmonization Project," *Journal for Immunotherapy of Cancer*, vol. 8, no. 1, p. e000147, 2020.

[40] N. Waddell, M. Pajic, A. M. Patch et al., "Whole genomes redefine the mutational landscape of pancreatic cancer," *Nature*, vol. 518, no. 7540, pp. 495–501, 2015.

[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.

[42] N. Hoshino, K. Hida, Y. Sakai et al., "Nomogram for predicting anastomotic leakage after low anterior resection for rectal cancer," *International Journal of Colorectal Disease*, vol. 33, no. 4, pp. 411–418, 2018.

[43] K. Nie, Z. Zheng, Y. Wen et al., "Construction and validation of a TP53-associated immune prognostic model for gastric cancer," *Genomics*, vol. 112, no. 6, pp. 4788–4795, 2020.

[44] J. Wu, H. Zhang, L. Li et al., "A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: a population-based analysis," *Cancer Communications*, vol. 40, no. 7, pp. 301–312, 2020.

[45] L. Li, J. Liang, T. Song et al., "A nomogram model to predict prognosis of patients with genitourinary sarcoma," *Frontiers in Oncology*, vol. 11, article 656325, 2021.

[46] L. A. Torre, B. Trabert, C. E. DeSantis et al., "Ovarian cancer statistics, 2018," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 4, pp. 284–296, 2018.

[47] E. Lengyel, "Ovarian cancer development and metastasis," *The American Journal of Pathology*, vol. 177, no. 3, pp. 1053–1064, 2010.

[48] S. Banerjee and S. B. Kaye, "New strategies in the treatment of ovarian cancer: current clinical perspectives and future potential," *Clinical Cancer Research*, vol. 19, no. 5, pp. 961–968, 2013.

[49] P. T. Kroeger Jr. and R. Drapkin, "Pathogenesis and heterogeneity of ovarian cancer," *Current Opinion in Obstetrics & Gynecology*, vol. 29, no. 1, pp. 26–34, 2017.

[50] U. Testa, E. Petrucci, L. Pasquini, G. Castelli, and E. Pelosi, "Ovarian cancers: genetic abnormalities, tumor heterogeneity and progression, clonal evolution and cancer stem cells," *Medicines*, vol. 5, no. 1, p. 16, 2018.

[51] S. Dilruba and G. V. Kalayda, "Platinum-based drugs: past, present and future," *Cancer Chemotherapy and Pharmacology*, vol. 77, no. 6, pp. 1103–1124, 2016.

[52] M. Binju, M. A. Amaya-Padilla, G. Wan, H. Gunosewoyo, Y. Suryo Rahmanto, and Y. Yu, "Therapeutic inducers of apoptosis in ovarian cancer," *Cancers*, vol. 11, no. 11, p. 1786, 2019.

[53] Y. Li, X. Gong, T. Hu, and Y. Chen, "Two novel prognostic models for ovarian cancer respectively based on ferroptosis and necroptosis," *BMC Cancer*, vol. 22, no. 1, p. 74, 2022.

[54] A. Ghoneum, H. Afify, Z. Salih, M. Kelly, and N. Said, "Role of tumor microenvironment in the pathobiology of ovarian cancer: insights and therapeutic opportunities," *Cancer Medicine*, vol. 7, no. 10, pp. 5047–5056, 2018.

[55] Y. An and Q. Yang, "MiR-21 modulates the polarization of macrophages and increases the effects of M2 macrophages on promoting the chemoresistance of ovarian cancer," *Life Sciences*, vol. 242, article 117162, 2020.

[56] M. Paroli, F. Bellati, M. Videtta et al., "Discovery of chemotherapy-associated ovarian cancer antigens by interrogating memory T cells," *International Journal of Cancer*, vol. 134, no. 8, pp. 1823–1834, 2014.

[57] M. J. Dobrzanski, K. A. Rewers-Felkins, K. A. Samad et al., "Immunotherapy with IL-10- and IFN-γ-producing CD4 effector cells modulate "Natural" and "Inducible" CD4 TReg cell subpopulation levels: observations in four cases of patients with ovarian cancer," *Cancer Immunology, Immunotherapy*, vol. 61, no. 6, pp. 839–854, 2012.

[58] S. Biswas, G. Mandal, K. K. Payne et al., "IgA transcytosis and antigen recognition govern ovarian cancer immunity," *Nature*, vol. 591, no. 7850, pp. 464–470, 2021.

[59] L. Li, Y. Ma, and Y. Xu, "Follicular regulatory T cells infiltrated the ovarian carcinoma and resulted in CD8 T cell dysfunction dependent on IL-10 pathway," *International Immunopharmacology*, vol. 68, pp. 81–87, 2019.

[60] M. H. Barros, F. Hauck, J. H. Dreyer, B. Kempkes, and G. Niedobitek, "Macrophage polarisation: an immunohistochemical approach for identifying M1 and M2 macrophages," *PLoS One*, vol. 8, no. 11, article e80908, 2013.

[61] K. Cao, G. Zhang, X. Zhang et al., "Stromal infiltrating mast cells identify immunoevasive subtype high-grade serous ovarian cancer with poor prognosis and inferior immunotherapeutic response," *Oncoimmunology*, vol. 10, no. 1, p. 1969075, 2021.

[62] M. Prat, A. Le Naour, K. Coulson et al., "Circulating CD14(high) CD16 (low) intermediate blood monocytes as a biomarker of ascites immune status and ovarian cancer progression," *Journal for Immunotherapy of Cancer*, vol. 8, no. 1, p. e000472, 2020.

[63] A. Bagaev, N. Kotlov, K. Nomie et al., "Conserved pan-cancer microenvironment subtypes predict response to immunotherapy," *Cancer Cell*, vol. 39, no. 6, pp. 845–865.e7, 2021.

[64] 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.

[65] S. Pal, M. Garg, and A. K. Pandey, "Deciphering the mounting complexity of the p53 regulatory network in correlation to long non-coding RNAs (lncRNAs) in ovarian cancer," *Cell*, vol. 9, no. 3, p. 527, 2020.

[66] D. Xue, H. Lin, L. Lin, Q. Wei, S. Yang, and X. Chen, "TTN/TP53 mutation might act as the predictor for chemotherapy response in lung adenocarcinoma and lung squamous carcinoma patients," *Translational Cancer Research*, vol. 10, no. 3, pp. 1284–1294, 2021.

[67] Q. Jia, J. Wang, N. He, J. He, and B. Zhu, "Titin mutation associated with responsiveness to checkpoint blockades in solid tumors," *Insight*, vol. 4, no. 10, 2019.

[68] A. Chattopadhyay, Z. H. Teoh, C. Y. Wu et al., "CNVIntegrate: the first multi-ethnic database for identifying copy number variations associated with cancer," *Database: The Journal of Biological Databases and Curation*, vol. 2021, article baab044, 2021.

[69] F. Martignano, U. Munagala, S. Crucitta et al., "Nanopore sequencing from liquid biopsy: analysis of copy number variations from cell-free DNA of lung cancer patients," *Molecular Cancer*, vol. 20, no. 1, p. 32, 2021.

[70] R. P. Graf, R. Eskander, L. Brueggeman, and D. G. Stupack, "Association of copy number variation signature and survival in patients with serous ovarian cancer," *JAMA Network Open*, vol. 4, no. 6, article e2114162, 2021.

[71] S. Grabosch, M. Bulatovic, F. Zeng et al., "Cisplatin-induced immune modulation in ovarian cancer mouse models with distinct inflammation profiles," *Oncogene*, vol. 38, no. 13, pp. 2380–2393, 2019.

[72] L. Liu, J. Fan, G. Ai et al., "Berberine in combination with cisplatin induces necroptosis and apoptosis in ovarian cancer cells," *Biological Research*, vol. 52, no. 1, p. 37, 2019.

[73] X. L. Zheng, J. J. Yang, Y. Y. Wang et al., "RIP1 promotes proliferation through G2/M checkpoint progression and mediates cisplatin-induced apoptosis and necroptosis in human ovarian cancer cells," *Acta Pharmacologica Sinica*, vol. 41, no. 9, pp. 1223–1233, 2020.

[74] S. Halabi, S. Dutta, C. M. Tangen et al., "Overall survival of Black and White men with metastatic castration-resistant prostate cancer treated with docetaxel," *Journal of Clinical Oncology*, vol. 37, no. 5, pp. 403–410, 2019.

[75] Y. Hu, M. Ran, B. Wang, Y. Lin, Y. Cheng, and S. Zheng, "Co-delivery of docetaxel and curcumin via nanomicelles for enhancing anti-ovarian cancer treatment," *International Journal of Nanomedicine*, vol. Volume 15, pp. 9703–9715, 2020.

[76] T. S. Lau, L. K. Y. Chan, G. C. W. Man et al., "Paclitaxel induces immunogenic cell death in ovarian cancer via TLR4/IKK2/SNARE-dependent exocytosis," *Cancer Immunology Research*, vol. 8, no. 8, pp. 1099–1111, 2020.

Hindawi

*Research Article*

# Combined scRNAseq and Bulk RNAseq Analysis to Reveal the Dual Roles of Oxidative Stress-Related Genes in Acute Myeloid Leukemia

**Jing Qi** [iD],[1] **Jingyu Lin** [iD],[2] **Changjiang Wu** [iD],[3] **Hesheng He,**[1] **Junping Yao,**[1] **Youhai Xu,**[1] **Yuqiong Yang,**[1] **Yuanfeng Wei,**[1] **Dongping Huang** [iD],[1] **and Yiming Mao** [iD][2,4]

[1]*Department of Hematology, The First Affiliated Hospital of Wannan Medical College, Wuhu, 241001 Anhui, China*
[2]*Department of Science & Education, Suzhou Kowloon Hospital, Shanghai Jiao Tong University School of Medicine, Suzhou, 215028 Jiangsu, China*
[3]*Department of Intensive Care Unit, Suzhou Kowloon Hospital, Shanghai Jiao Tong University School of Medicine, Suzhou, 215028 Jiangsu, China*
[4]*Department of thoracic surgery, Suzhou Kowloon Hospital, Shanghai Jiao Tong University School of Medicine, Suzhou, 215028 Jiangsu, China*

Correspondence should be addressed to Dongping Huang; hdp_9713@163.com and Yiming Mao; ym_mao0302@suda.edu.cn

Jing Qi, Jingyu Lin, and Changjiang Wu contributed equally to this work.

*Background*. Oxidative stress (OS) can either lead to leukemogenesis or induce tumor cell death by inflammation and immune response accompanying the process of OS through chemotherapy. However, previous studies mainly focus on the level of OS state and the salient factors leading to tumorigenesis and progression of acute myeloid leukemia (AML), and nothing has been done to distinguish the OS-related genes with different functions. *Method*. First, we downloaded single-cell RNA sequencing (scRNAseq) and bulk RNA sequencing (RNAseq) data from public databases and evaluated the oxidative stress functions between leukemia cells and normal cells by the ssGSEA algorithm. Then, we used machine learning methods to screen out OS gene set A related to the occurrence and prognosis of AML and OS gene set B related to treatment in leukemia stem cells (LSCs) like population (HSC-like). Furthermore, we screened out the hub genes in the above two gene sets and used them to identify molecular subclasses and construct a model for predicting therapy response. *Results*. Leukemia cells have different OS functions compared to normal cells and significant OS functional changes before and after chemotherapy. Two different clusters in gene set A were identified, which showed different biological properties and clinical relevance. The sensitive model for predicting therapy response based on gene set B demonstrated predictive accuracy by ROC and internal validation. *Conclusion*. We combined scRNAseq and bulk RNAseq data to construct two different transcriptomic profiles to reveal the different roles of OS-related genes involved in AML oncogenesis and chemotherapy resistance, which might provide important insights into the mechanism of OS-related genes in the pathogenesis and drug resistance of AML.

## 1. Background

Oxidative stress (OS) is a series of adaptive responses caused by the imbalance between the reactive oxygen species (ROS) and the antioxidant system of all aerobic organisms. The integrated antioxidant system of an aerobic organism can block the damage caused by redundant ROS [1]. However, the abnormal redox state of cells can produce toxic effects of peroxides and free radicals, thereby damaging cellular proteins, lipids, and DNA, which can lead to aging, diseases, and tumorigenesis [2]. Some chronic diseases and cancers have been reported as a consequence of oxidative stress

[3–5], including hypertension, Alzheimer's dementia, diabetes, breast cancer, and renal cell carcinoma [6, 7], while ROS shows bidirectional effects in cancer cells. On one hand, the ROS initiates tumorigenesis and supports the proliferation of cancer cells promoting the progression of cancer. On the other hand, high levels of ROS are also toxic to cancer cells causing cell death, which has been applied as a mechanism of chemotherapy [8]. Recent study showed that the combination of retinoic acid, tunicamycin, and arsenic trioxide can cause OS-induced cell death in FLT3-ITD+ AML cell lines [9]. Interestingly, tumor cells can increase their antioxidant capacity to accommodate high ROS-mediated cell proliferation while avoiding apoptosis and senescence triggered by excessive ROS, which might explain the chemotherapy resistance in cancers [10].

AML, the most common type of leukemia in adults worldwide, is characterized by the infiltration of abnormally clonal and undifferentiated leukemia cells in bone marrow, blood, and other tissues [11]. The enormous cytogenetic and molecular contexts make it a highly heterogeneous disease with distinct biological process and prognosis. Although the advance in treatment has been achieved among younger patients with a cure rate of 35-40%, the emergence of drug resistance often leads to treatment failure and the prognosis of elderly patients remains dismal; most patients died within one year of diagnosis [12, 13]. Therefore, the development of new therapies based on the individual genomic landscape risk stratification and understanding the mechanisms underlying drug resistance in leukemia are supposed to improve the remission rate and overall survival of AML patients. Currently, OS has been reported playing a role in the development of several hematologic cancers, including AML, chronic myeloid leukemia (CML), and acute lymphoblastic leukemia (ALL) [14, 15]. CML cells showed a higher oxidative stress with significant lower SOD activity, which is correlated with the altered intracellular calcium homeostasis. The increased oxidative stress and decreased antioxidants have been reported in ALL and AML patients in a recent study [14]. Dong et al. [16] also constructed an effective risk model for predicting the prognosis of AML patients based on the OS-related genes. However, previous studies mainly focus on the level of OS state and the salient factors leading to tumorigenesis and progression of AML, nothing has been done to explore the dual roles and potential mechanism of OS in AML.

In our present work, we selected the OS-related genes, which might contribute to the carcinogenesis of AML through scRNAseq combined with bulk RNA sequencing (RNAseq) data, and then established a molecular phenotype based on the above genes, which could predict prognosis and provide different OS profiles of AML patients. In addition, we established an OS-related gene signature for predicting the chemotherapy response of AML patients. All these could provide more information and genomic evidence for individualized precision therapy.

## 2. Material and Methods

*2.1. scRNAseq and Bulk RNAseq Data Obtaining.* The paired bone marrow scRNAseq data and clinical information of 8

AML patients (AML314, AML371, AML997, AML707B, AML475, AML329, AML556, and AML722B) and 4 healthy donors (BM1, BM2, BM3, and BM4) were obtained from Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) using the accession number GSE11625. We used the UCSC Xena (https://xenabrowser.net/) to download the GDC TCGA Acute Myeloid Leukemia (LAML) transcription expression data ($n = 151$) and phenotype data. The bulk RNAseq data and clinical information of AML patients (GSE106291, $n = 250$) were also downloaded from the GEO database. 1399 oxidative stress-related genes were selected according to a previous study [17].

*2.2. scRNAseq Data Processing.* The scRNAseq data analysis was performed in R version 4.1.3 as follows: (1) Seurat R package [18] was used to convert scRNAseq data as a Seurat object, and "NormalizeData" were used to preprocess and standardize the data; (2) the top 2000 highly variable genes after quality control were selected by "FindVariableFeatures"; (3) principal component analysis (PCA) was performed based on the 2000 genes to analyze the scRNAseq data; (4) uniform manifold approximation and projection (UMAP) was applied to explore the scRNAseq data; (5) a total of 21 different cell types were defined according to the original data and reference [19]; and (6) the "FindMarkers" function was used to find all markers of different cell types with the criterion FDR $< 0.05$ and $|\log 2FC| > 0.5$.

*2.3. ssGSEA Algorithm to Evaluate the Oxidative Stress Functions.* GO BP terms "RESPONSE_TO_OXIDATIVE_STRESS" and "CELL_DEATH_IN_RESPONSE_TO_OXIDATIVE_STRESS" were downloaded from the MsigDB database (https://www.gsea-msigdb.org/gsea/index.jsp) to evaluate oxidative stress-related functions and the R package "GSVA" to score each cell with the single-sample GSEA (ssGSEA) algorithm, setting as abs.ranking = T and parallel .sz = 3. The higher the total score, the higher the gene expression of each sample.

*2.4. Functional Enrichment and Pathway Analysis.* Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Set Enrichment Analysis (GSEA) functional enrichment analyses were performed using the cluster profiler package [20]. Three categories were included in the GO enrichment analysis, i.e., biological process (BP), cellular component (CC), and molecular function (MF). Default parameters are selected as the setting of "enrichgo" function, "enrichkegg" function, and "GSEA" function.

*2.5. Nonnegative Matrix Factorization (NMF) Algorithm to Identify Molecular Subclasses.* First, the differential genes derived from normal cells and tumor cells were intersected with OS-related genes. The significant prognostic genes in LAML dataset were selected by univariate Cox regression. Then, we used unsupervised NMF to cluster the data by using the package "NMF," with the standard "brunet" for 10 iterations. The cluster value $K$ was set at 2 to 10, and the optimal number of clusters was based on the cophenetic correlation coefficient. The relationship among clusters,

(a)



(b)

Figure 1: Continued.

(c)
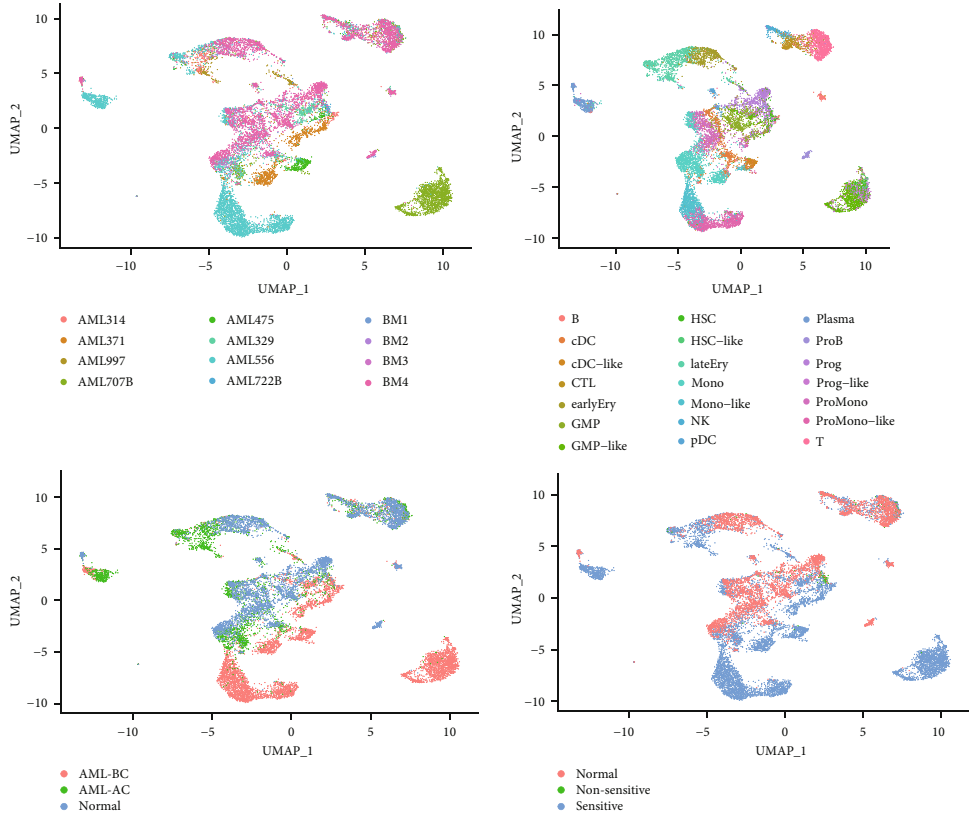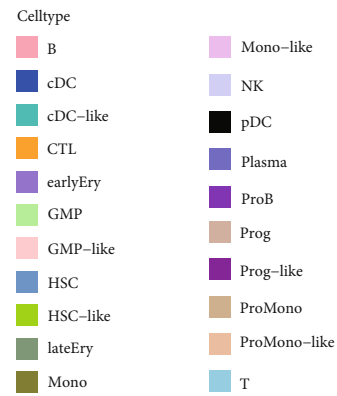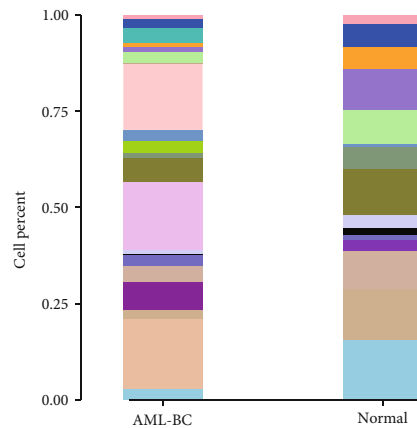


(d)
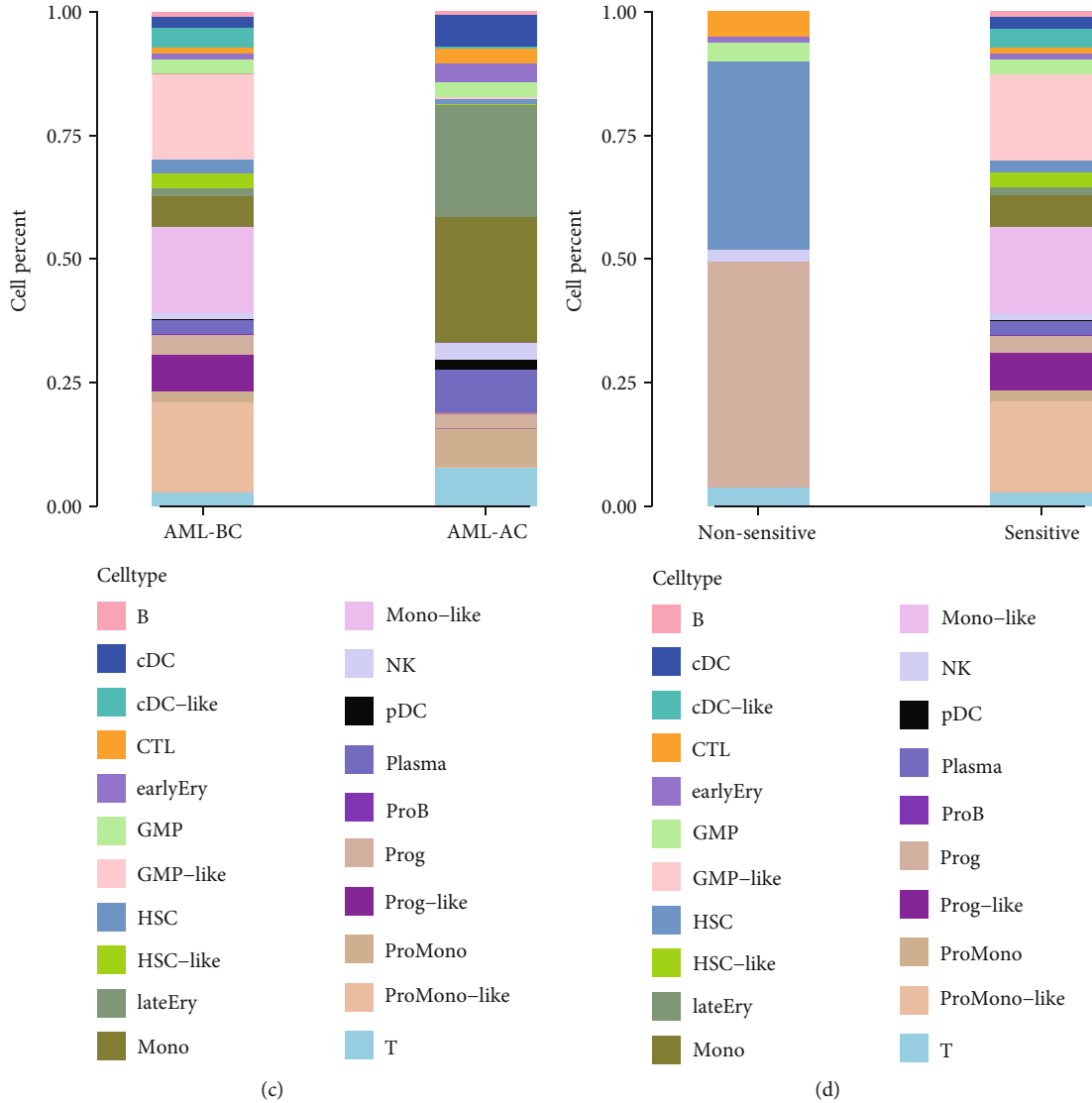
FIGURE 1: Single-cell profiling of AML patients and healthy donors. A total of 13593 cells and 21 cell types from 8 AML patients before and after induction chemotherapy and 4 healthy BM samples of GSE11625 are visualized by UMAP (a). Cell type composition in normal and AML patients before chemotherapy (b). Cell type composition in the AML-BC and AML-AC groups (c). Cell type composition in the sensitive and nonsensitive groups (d).

clinical variables, and hub genes gene expression was shown by heatmap with R package "ComplexHeatmap."

### 2.6. Construction and Validation of OS-Related Gene Risk Model.

To select hub genes based on the above candidate genes, LASSO regression (LR) was first used selecting the minimal lambda value. Bootstrap_multicox regression and risk regression model were built according to the multivariate analysis results, and the formula of the model was as follows:

$$\text{Risk score} = \sum \frac{\text{Coef } i * \text{Gene } i}{\text{SD bootstrap}}. \quad (1)$$

LAML dataset was treated as a training set, and GSE106291 was treated as a validated set. Patients were clas-sified into a high- or low-risk group according to the risk score. Overall survival was defined as the endpoint. The R package "glmnet," "boot," "survival," "survminer," and "ggplot2" were used.

### 2.7. Selection of OS-Related Genes for Chemotherapy-Sensitive Model.

First, the differential genes derived from AML-BC and AML-AC of the HSC-like population were intersected with OS-related genes. Then, a support vector machine (SVM) with "AvgRank" is sorted by the average ranking of 10-folds; target genes are obtained according to the best point of accuracy and error rate value. Random for-est (RF) is used to select top 20 genes according to importancescore > 2.25, and the parameters of LASSO regression (LR) are nlambda = 100, alpha = 1, and nfolds = 10 in the "cv.glmnet" function sets. Then, select coef

(a)



(b)
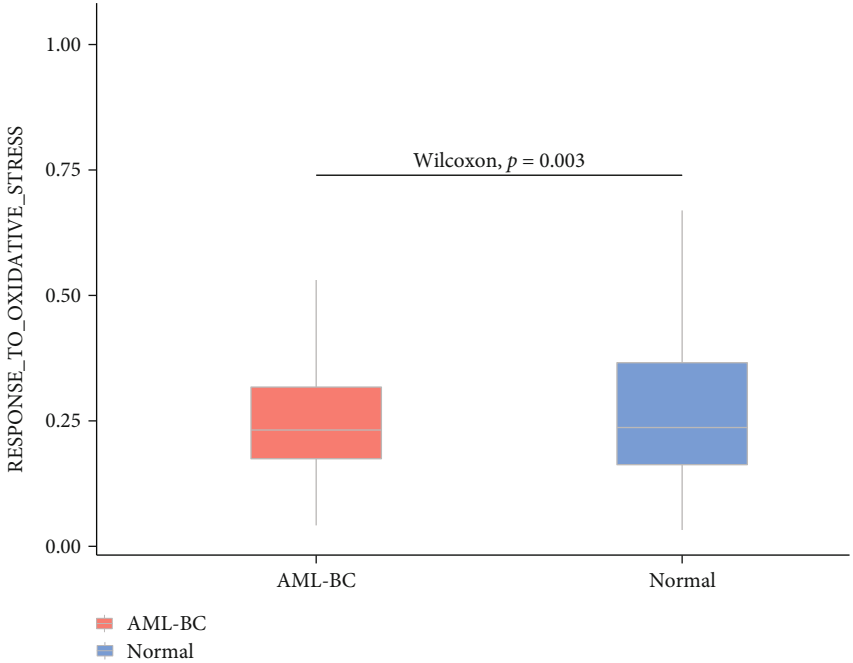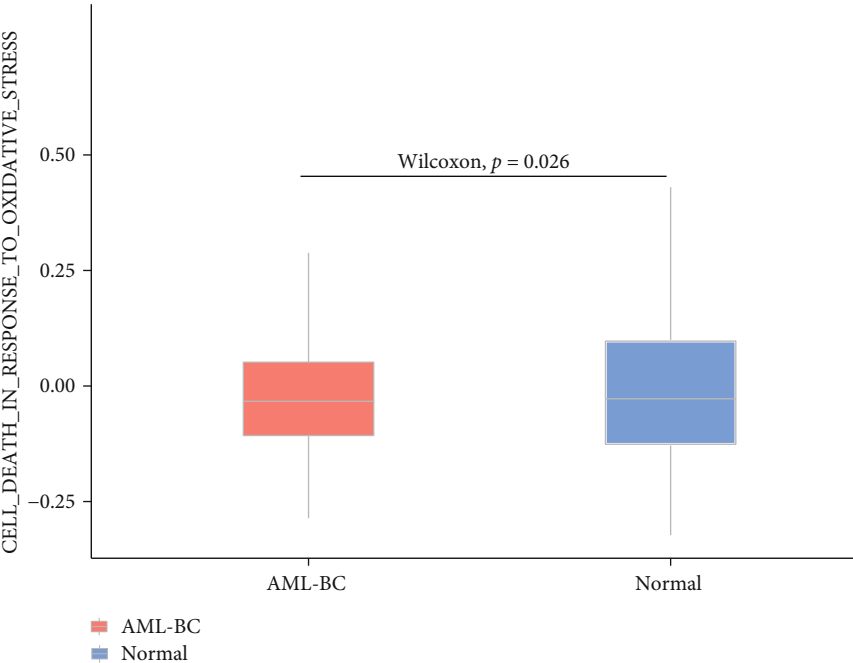
FIGURE 2: Continued.
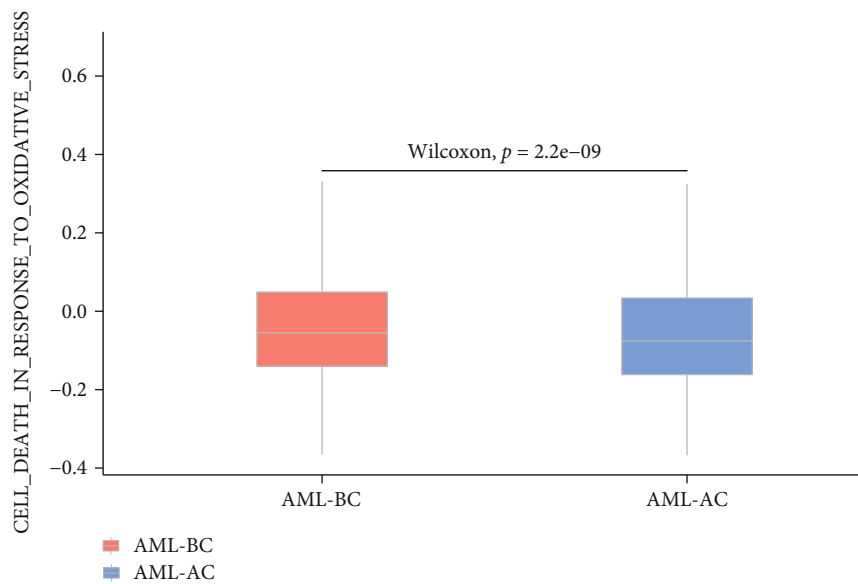
(c)
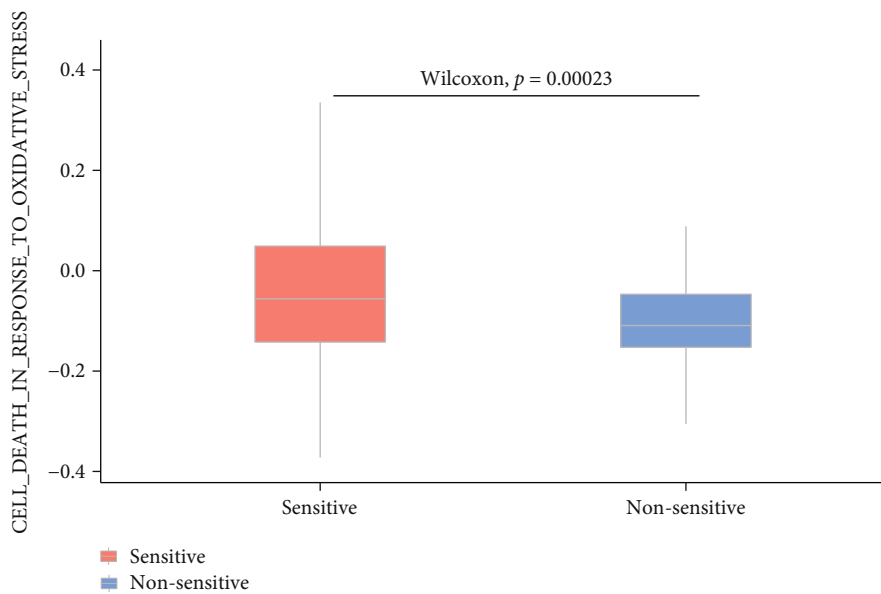


(d)

Figure 2: Continued.

(e)
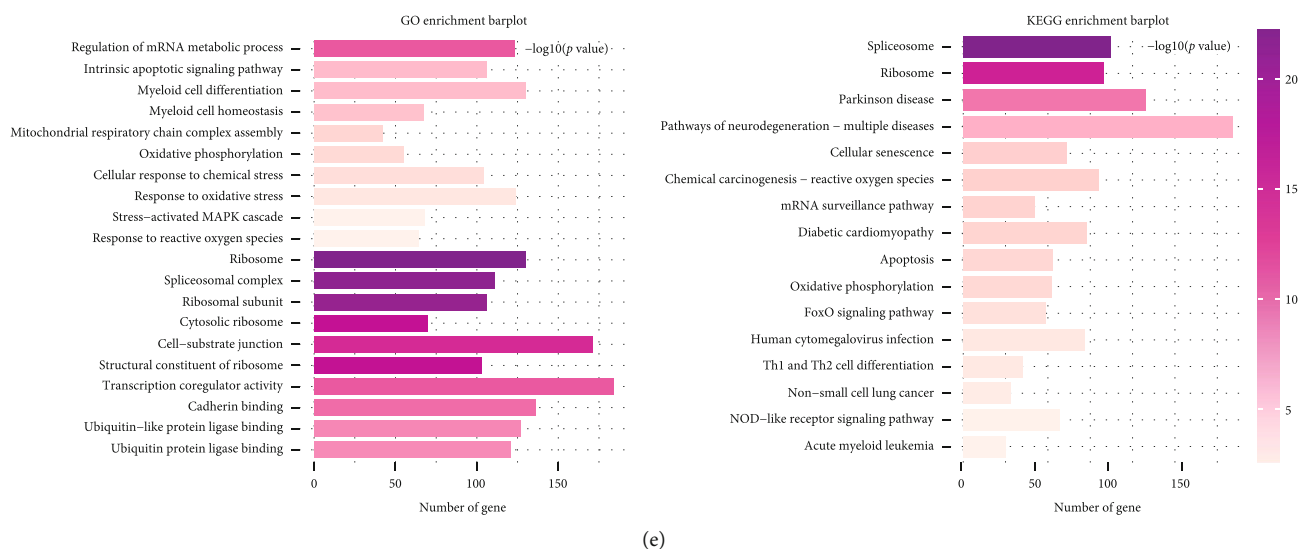
FIGURE 2: Differential oxidative stress functions in AML single cells. Normal cells have a higher response oxidative stress score ($P = 0.003$) and higher cell death in response to oxidative stress score than AML cells ($P = 0.026$) (a, b). Higher cell death in response to oxidative stress score in the AML-BC group than in the AML-AC group ($P = 2.2e - 09$) (c). Higher cell death in response to oxidative stress score in the sensitive group than the nonsensitive group ($P = 0.00023$) (d). GO and KEGG enrichment analyses of DEGs of the HSC-like subgroup pre- and postchemotherapy (e).

according to lambda.min, and finally, filter out the target genes. The intersection of three machine learning genes is used for constructing the therapy response predictive model. After multivariate logistic regression model analysis, a model formula was set.

*2.8. Construction and Validation of a Chemotherapy-Sensitive Model.* GSE106291 was the training set, and 50% of the samples were selected as the internal validation set randomly. Patients were divided into two groups (high-risk vs. low-risk group) according to the sensitive score. The receiver operating characteristic (ROC) was used to evaluate the model's predictability.

*2.9. Statistical Analysis.* All statistical analyses were performed in R version 4.1.3. LASSO regression and Cox regression analyses were conducted by the "glmnet" packages. The Wilcoxon rank sum test was used to compare the two groups. Contingency table variables were analyzed by chi-square test or Fisher's exact tests. Survival analysis was conducted by the KM method and compared by log-rank via "survival" packages. A two-tailed $P$ value $< 0.05$ was indicative of a statistically significant difference.

## 3. Results

*3.1. Single-Cell Profiling of AML Patients and Healthy Donors.* We integrated paired bone marrow scRNAseq data from 8 newly diagnosed AML patients before and after induction chemotherapy and 4 bone marrow samples from healthy donors. A total of 27899 features and13593 cells were included in this study. Among the 8 AML patients, AML 722B achieved morphological remission after two courses of induction chemotherapy and was

included in the nonsensitive group; the left 7 AML patients achieved morphological remission after one standard course of chemotherapy, and they were included in the sensitive group. The cells of AML patients at initial diagnosis were classified as AML before chemotherapy (AML-BC) group, and the cells after induction chemotherapy were treated as AML after chemotherapy (AML-AC) group (Figure 1(a)). A total of 4625 malignant cells of six cell types (HSC-like, progenitor-like, GMP-like, promonocyte-like, monocyte-like, and cDC-like) were identified according to the defined cell types [19]. We can observe highly proliferative leukemia cells in AML patients and suppressed immune cells in the TME, compared with normal bone marrow samples (Figure 1(b)). Significant recovery of hematopoietic and immune cells was seen in these patients when they achieved morphological remission after induction chemotherapy (Figure 1(c)). Although fewer cells are in the nonsensitive group, differences in cell populations between the sensitive group and nonsensitive group can be observed (Figure 1(d)).

*3.2. Differential Oxidative Stress Response in AML and Chemotherapy-Induced Changes in HSC-Like Subgroup.* To assess oxidative stress in leukemia cells, we scored each cell in the AML-BC group and cells from normal bone marrow according to the relevant functional items. First, we evaluated the "GOBP_ RESPONSE_TO_OXIDATIVE_STRESS" and "GOBP_CELL_DEATH_IN_RESPONSE_TO_OXIDA-TIVE_STRESS" of the two types of samples. We found that normal cells have a higher response oxidative stress score ($P = 0.003$) and higher cell death in response to oxidative stress score than leukemia cells ($P = 0.026$) (Figures 2(a) and 2(b)). In addition, to find the subgroup most sensitive to oxidative stress caused by chemotherapy, we assessed

(a)



(b)
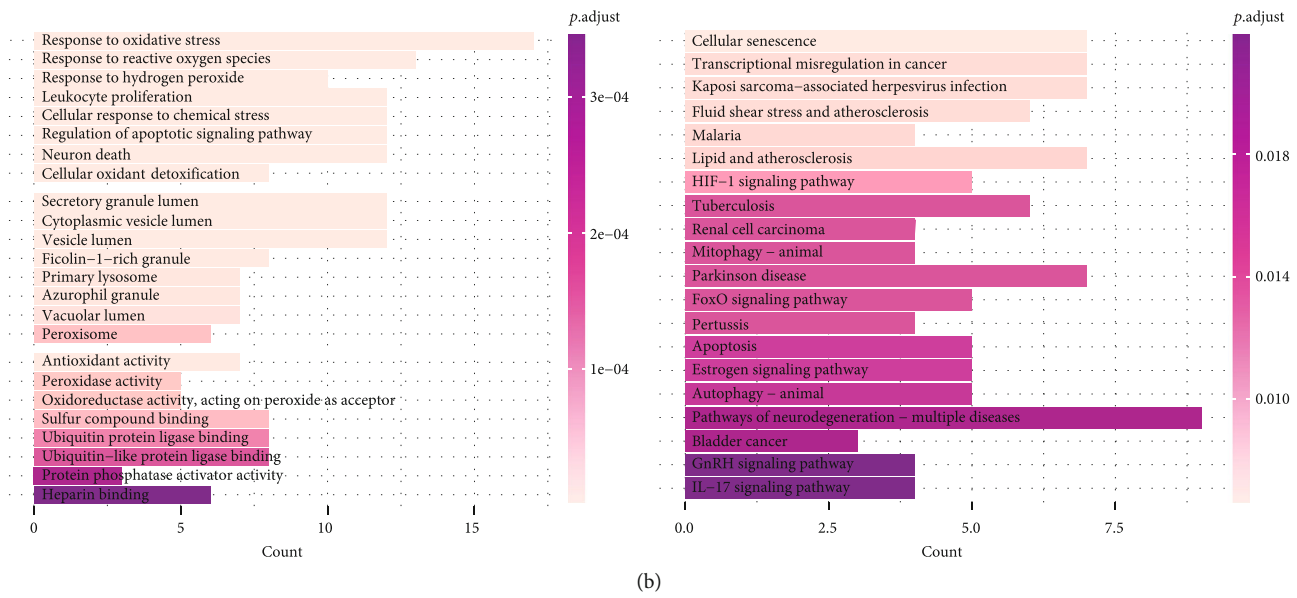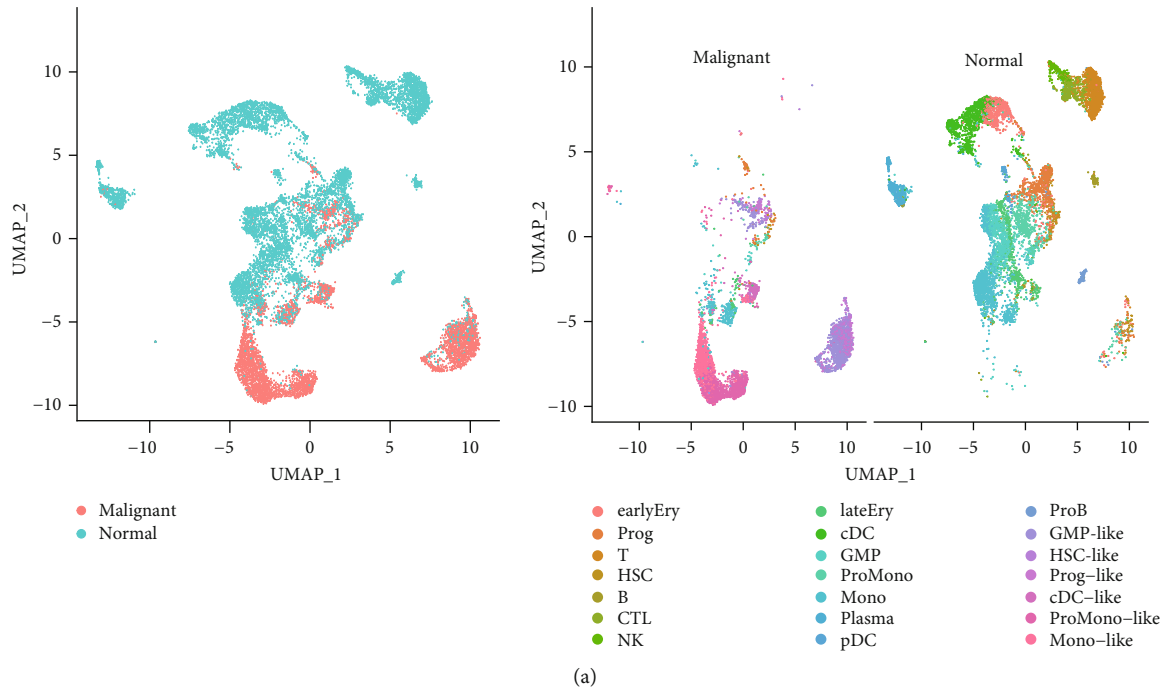
Figure 3: Continued.
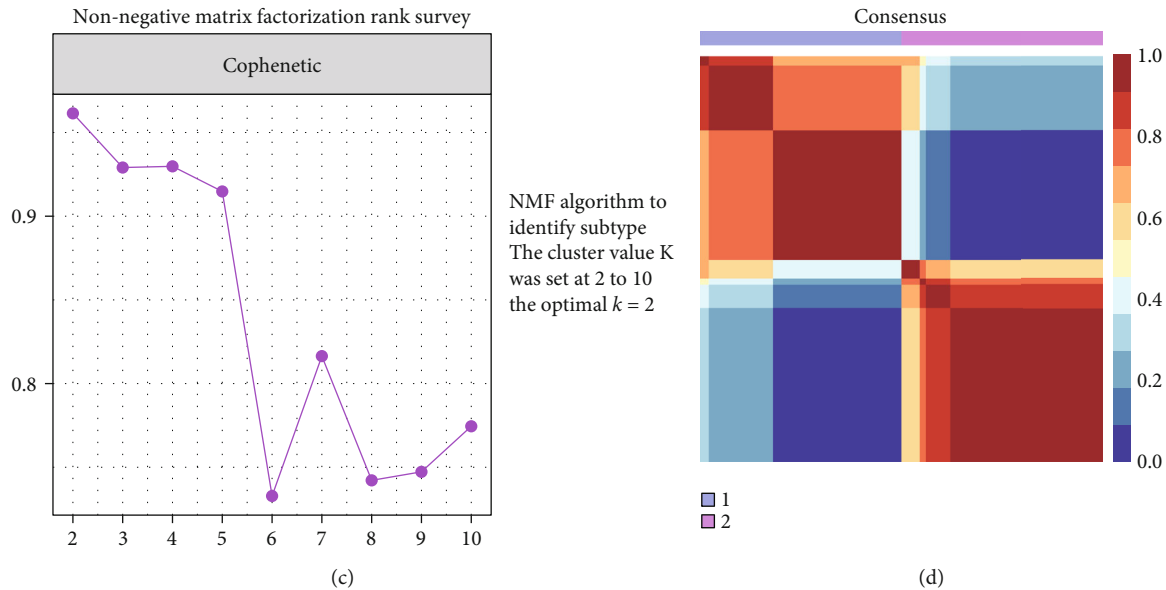
Figure 3: Identification of oncogenesis-related OS gene set A and function annotation. Cells of AML-BC and normal samples are visualized by UMAP (a). GO and KEGG enrichment analyses of oncogenesis-related OS gene set A (b). NMF to classify subclasses of OS gene set A in AML bulk RNAseq dataset LAML (c, d).

the cell death in response to oxidative stress score of different leukemia subgroups before and after treatment. Although a significant difference in cell death in response to oxidative stress score of leukemia cells before and after chemotherapy was observed ($P = 2.2e - 09$) (Figure 2(c)), there was no significant difference in the scores of various leukemia cell subsets pre- and posttreatment (Supplementary Figure 1). Furthermore, the sensitive group also showed higher cell death in response to oxidative stress score than the nonsensitive group ($P = 0.00023$) (Figure 2(d)). Given that leukemia stem cells are the main source of drug resistance and disease progression, it is of great significance to explore the transcriptomic changes before and after treatment. We enriched the DEGs of the HSC-like subgroup pre- and posttreatment and found that these DEGs are mainly involved in the regulation of mRNA metabolic process, intrinsic apoptotic signaling pathway, myeloid cell differentiation, and homeostasis, along with the oxidative stress-related biological processes, and are mainly enriched in the spliceosome, ribosome, Parkinson's disease, pathways of neurodegeneration (multiple diseases), cellular senescence, and chemical carcinogenesis (reactive oxygen species pathways). GSEA also showed an activated OXIDATIVE_PHOSPHORYLATION pathway ($P = 2e - 04$) (Figure 2(e) and Supplementary Figure 2). These results implied that OS has taken part in the development of AML, and the elimination of leukemic stem cell populations is associated with chemotherapy-induced OS response.

3.3. Oncogenesis-Related OS Gene Set A in AML Single Cells. By intersecting the DEGs between initial leukemia cells and normal cells with OS-related genes, we screened out 59 AML oncogenesis-related OS genes, termed OS gene set A

(Figure 3(a)) (Supplementary Table 1). The functional annotation also confirmed that these genes mainly take part in response to oxidative stress, response to reactive oxygen species, and some other redox-related pathways (Figure 3(b)).

3.4. Identification Subclasses of OS Gene Set A in AML Bulk RNAseq. To explore the possible mechanism of the above genes involved in AML, we used the LAML dataset as a training set to screen out 11 prognosis-related OS genes by Cox regression and classify them into 2 clusters (C1 and C2) according to the NMF cophenetic (Figures 3(c) and 3(d)). C2 showed a higher response to oxidative stress score than C1 ($P = 0.03$) (Figure 4(a)). Then, we further explore the relationship between the two clusters and the prognosis of AML, patients in C2 have worse overall survival than C1 ($P < 0.001$) (Figure 4(b)), and C2 also correlated with some clinical risk factors (age, risk category, and status) (Figure 4(c)) (Table 1). All of these suggest that these two clusters have different molecular and biological characteristics. The GSVA analysis of the two clusters group showed the activated INFLAMMATORY_RESPONSE pathway ($P$ adjust $= 2e - 04$), IL6_JAK_STAT3_SIGNALING ($P$ adjust $= 2e - 04$), KRAS_SIGNALING_UP ($P$ adjust $= 9e - 04$), and TNFA_SIGNALING_VIA_NFKB ($P$ adjust $= 9e - 04$) (Figure 4(d)). All these confirmed the oncogenic and prognostic roles of OS-related gene set A in AML.

3.5. Validation of the Prognostic Role of OS Gene Set A. To further validate the relationship between OS gene set A and prognosis, we used LASSO regression to get 6 hub genes (AIF1, ELANE, ENO1, GPX1, MPO, and THBS1) related to prognosis in the LAML dataset to construct a prognostic risk model (Supplementary Figure 3). Then, we got the risk

(a)



(b)

FIGURE 4: Continued.

(c)



(d)
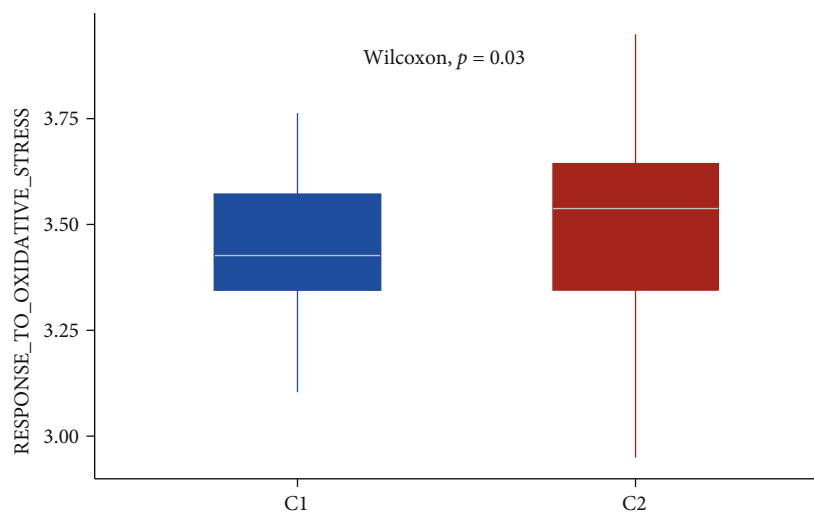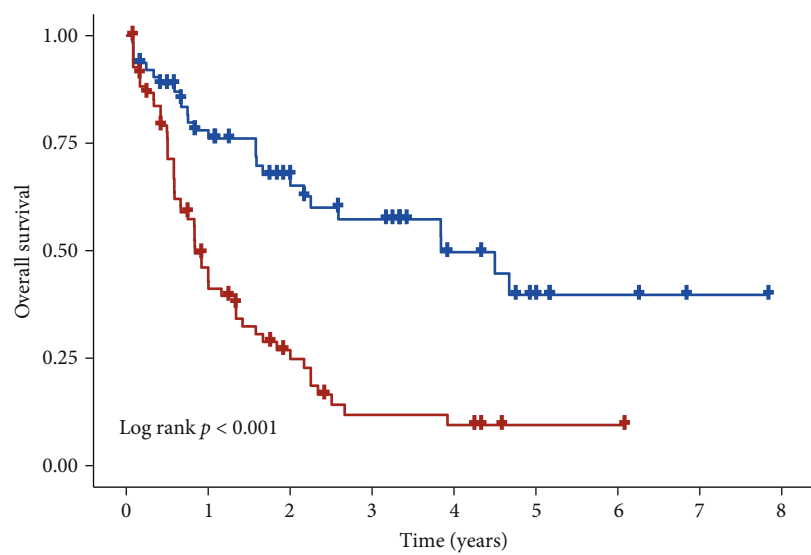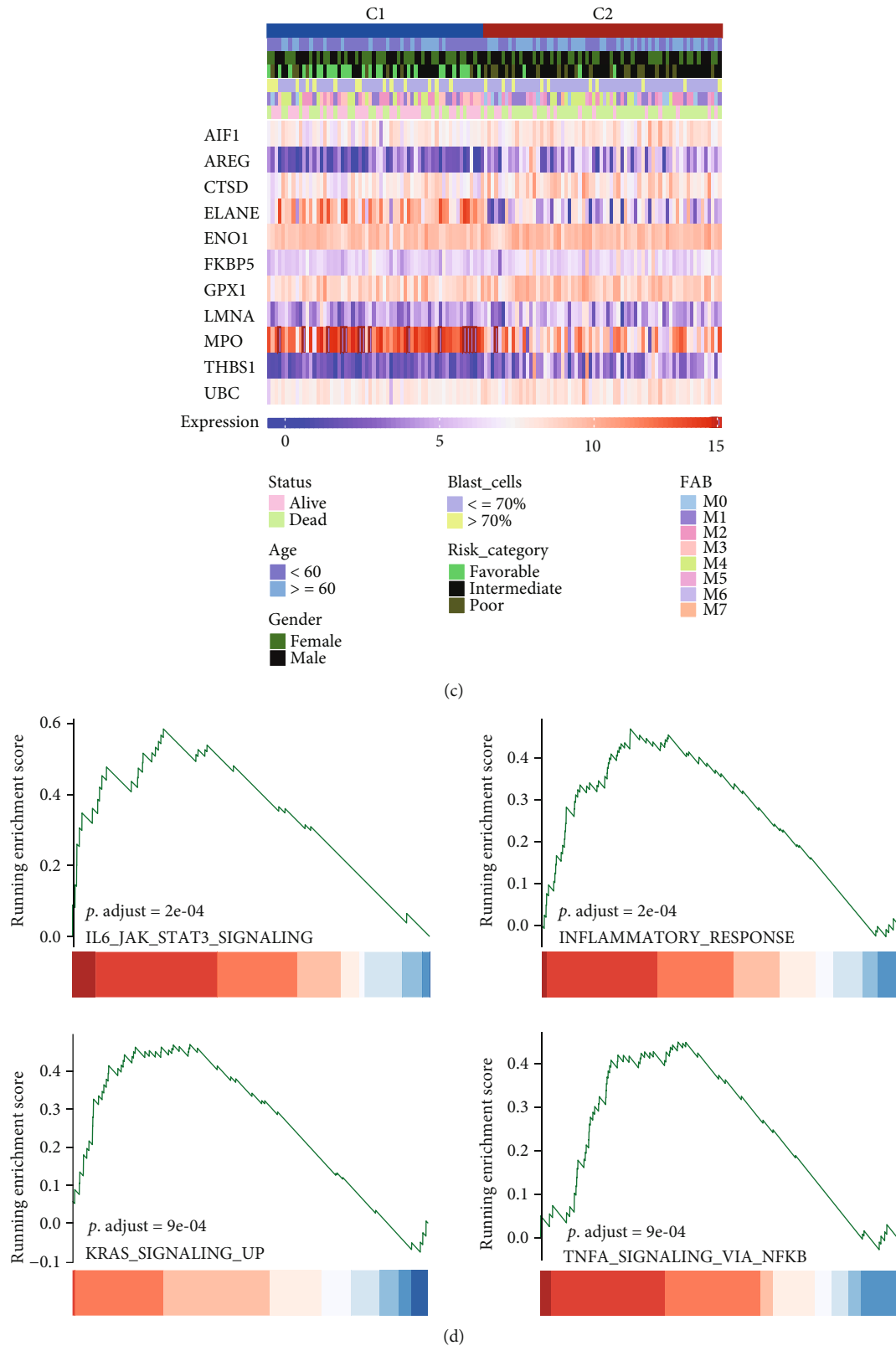
Figure 4: Continued.
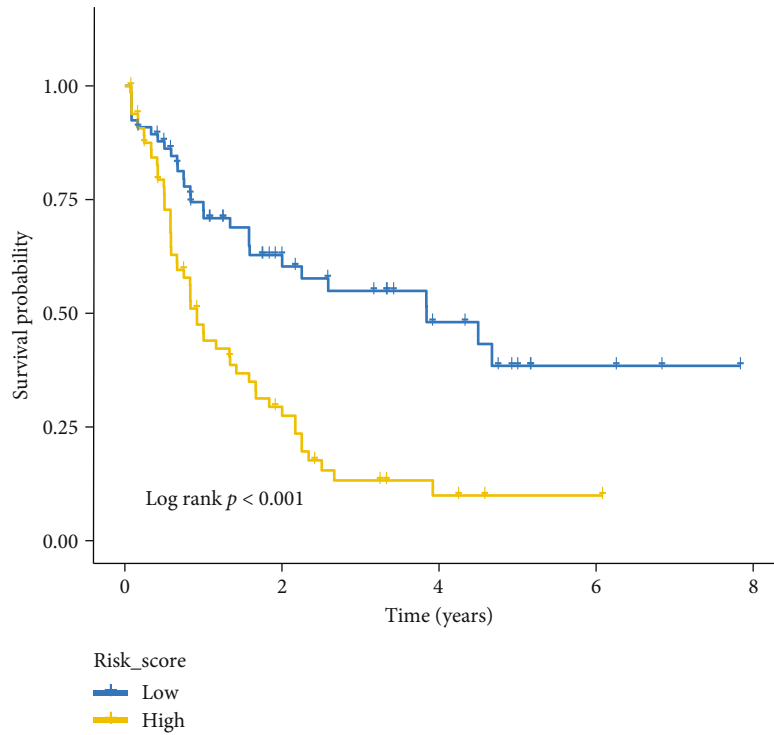
(e)
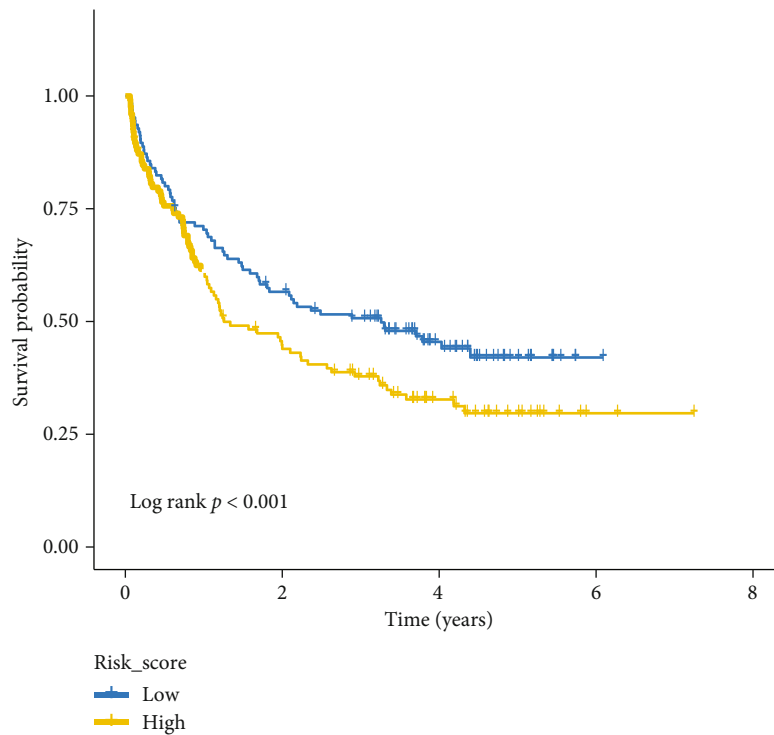


(f)

FIGURE 4: Different biological and clinical features of clusters in oncogenesis-related OS gene set A. Higher response to oxidative stress score by ssGSEA of C2 than C1 ($P = 0.03$) (a). AML patients in C2 have worse overall survival than C1 ($P < 0.001$) (b). The relationship between subclasses, clinical variables, and hub genes gene expression is visualized by heatmap (c). The significant enriched pathways of the two clusters by GSVA (d). AML patients in the high-risk score group have worse in LAML training dataset ($P < 0.001$) (e). AML patients in High-risk score group have worse in GSE106291 validation dataset ($P < 0.001$) (f).

TABLE 1: The different clinical characteristics between the two subtypes of OS gene set A.

| Characteristics | C1 | C2 | P value |
|---|---|---|---|
| *Age (years)* | | | 0.002 |
| <60 | 45 | 31 | |
| ≥60 | 17 | 37 | |
| *Gender* | | | 0.313 |
| Female | 31 | 28 | |
| Male | 31 | 40 | |
| *Risk category* | | | <0.001 |
| Favorable | 27 | 3 | |
| Intermediate/normal | 27 | 46 | |
| Poor | 8 | 19 | |
| *Blast cells* | | | 0.17 |
| ≤70% | 47 | 58 | |
| >70% | 15 | 10 | |
| *FAB* | | | <0.001 |
| M0 | 2 | 10 | |
| M1 | 14 | 17 | |
| M2 | 17 | 14 | |
| M3 | 14 | 0 | |
| M4 | 13 | 14 | |
| M5 | 2 | 10 | |
| M6 | 0 | 2 | |
| M7 | 0 | 1 | |
| *Status* | | | <0.001 |
| Alive | 37 | 15 | |
| Dead | 25 | 53 | |

model formula according to the coefficient of the bootstrap_multicox model.

$$\text{Risk score} = \sum \frac{\text{Coef } i * \text{Gene } i}{\text{SD bootstrap}}. \tag{2}$$

AML patients are divided into a high-risk group and a low-risk group according to the median of each score. We found that patients have a different prognosis between the two risk groups ($P < 0.001$) (Figure 4(e)). At last, this survival difference was also validated in the GSE106291 dataset ($P < 0.001$) (Figure 4(f)).

*3.6. Identification of Chemotherapy-Related OS Gene Set B in AML Single Cells.* To identify genes involved in chemotherapy-induced OS, we intersected the DEGs before and after chemotherapy of leukemia stem cell- (LSC-) like cells (HSC-like) with OS-related genes; 44 chemotherapy-related OS genes were screened out as OS gene set B (Figure 5(a)) (Supplementary Table 2).

*3.7. Construct a Sensitive Model to Predict Therapy Response Based on OS Gene Set B in AML Bulk RNAseq.* GSE106291 dataset was used as a training set. RF, SVM, and LASSO regression models were used to select model genes for pre-

dicting therapy response. We performed the RF algorithm to select a set of 20 candidate genes (Figures 5(b) and 5(c)), the LASSO algorithm to identify a set of 17 candidate genes (Figure 5(d)), and the SVM algorithm to select a set of 25 candidate genes (Figure 5(e)). At last, we choose the intersected 9 genes obtained by the above three machine learning algorithms to construct a model for predicting treatment response (Figure 5(f)). After multivariate logistic regression model analysis, we got a sensitive score for each patient according to the model formula.

*3.8. Validation of the Predictive Role of OS Gene Set B.* Patients in the training set were divided into low- or high-sensitive groups based on a value of 0.5. We used ROC to evaluate the predictive accuracy of the model; the AUC of the training set is 0.819 (Figure 6(a)). In addition, we found that patients in the high-sensitive group have better overall survival than patients in the low-sensitive group ($P < 0.001$) (Figure 6(b)). Besides, some clinical variables, like age, gender, and life status, are correlated with the sensitive score (Figure 6(c)). After validation in the internal dataset, the AUC of the ROC in the validation dataset is 0.784 (Figure 6(d)), and patients in the high-sensitive score group are more sensitive to chemotherapy (Figure 6(e)). All these supported the effective predictive ability of the model.

## 4. Discussion

High levels of ROS can either lead to the death of normal cells through the process of programmed cell death or activate redox-sensitive transcription factors, like forkhead box class O (FoxO) transcription factor, which can regulate cell proliferation, apoptosis, and differentiation and lead to tumor formation and cancer progression [21]. As a clonally malignant disease, the role of OS in the pathogenesis of AML is not fully understood. Previous studies reported that the abnormal mutants in leukemia cells, like FIT3-ITD and Ras (N-Ras and H-Ras), could increase the DNA double-strand breaks and induce the production of superoxide, which lead to the occurrence of leukemia and promote the proliferation of these malignant cells [22, 23]. Conversely, alterations of redox homeostasis in both normal and oncogenic cells can lead to death. Studies have been initiated to explore redox-related mechanisms and proteins to potentially target leukemia cells. Cytarabine and azanucleoside DNA methyltransferase (DNMT) inhibitors are widely used in AML patients, which have been revealed to cause a substantial increase in ROS in both resting and leukemia cells and trigger cell cycle arrest and apoptosis [24, 25]. Therefore, improving the current understanding of the underlying mechanisms of OS generation in leukemogenesis and antioxidant therapy will facilitate the progress in AML risk stratification and therapeutic area.

Benefit from the bulk RNAseq and scRNAseq technology, we can interpret the genomic profile of disease at the single-cell level. The normal cells have more response to oxidative stress score and are more prone to oxidative stress-induced cell death than leukemia cell populations in our data, which was consistent with the high proliferation of

(a)



(b)
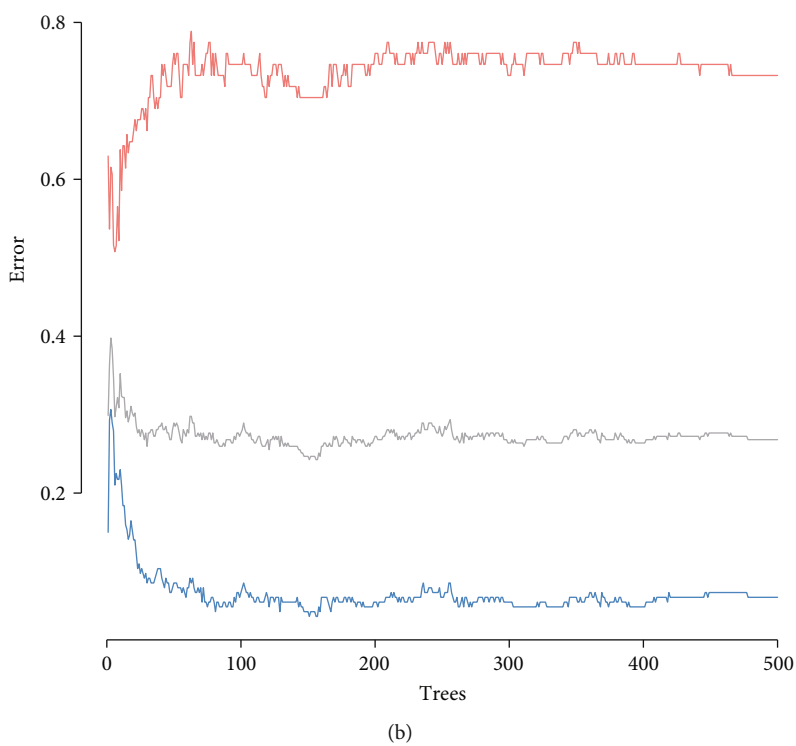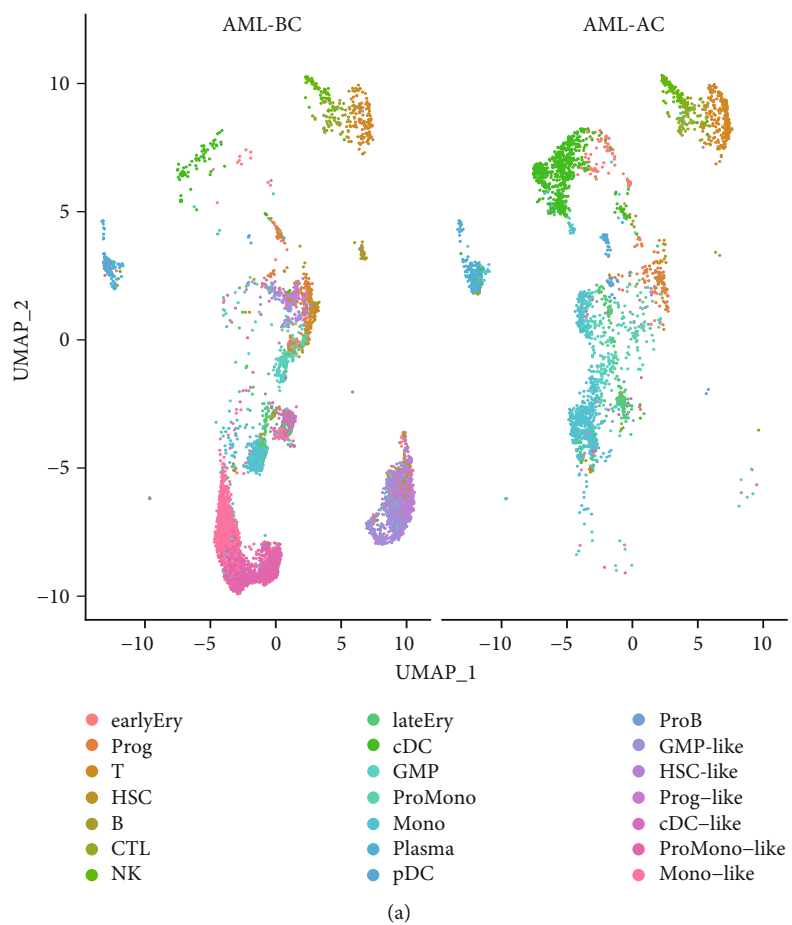
FIGURE 5: Continued.

(c)





(d)
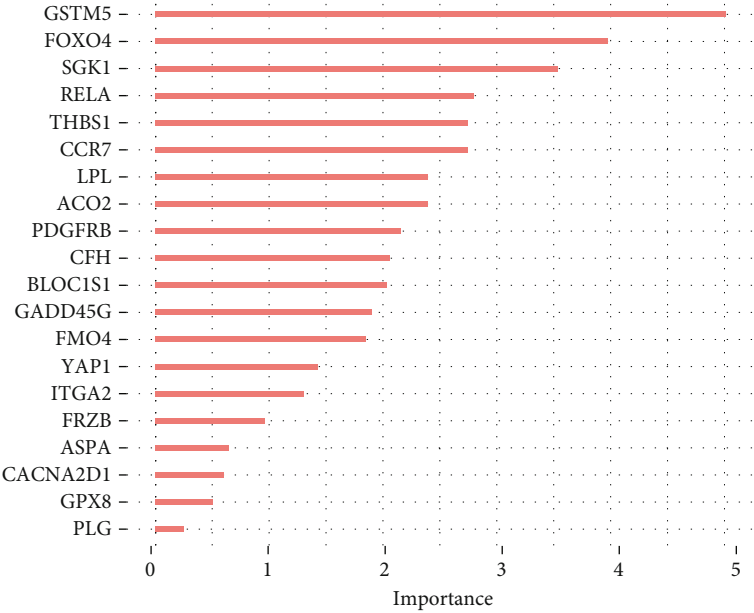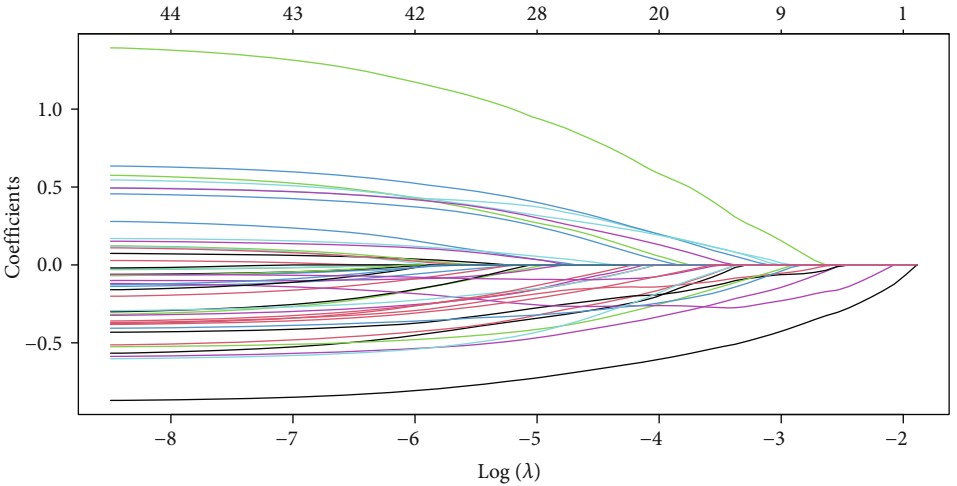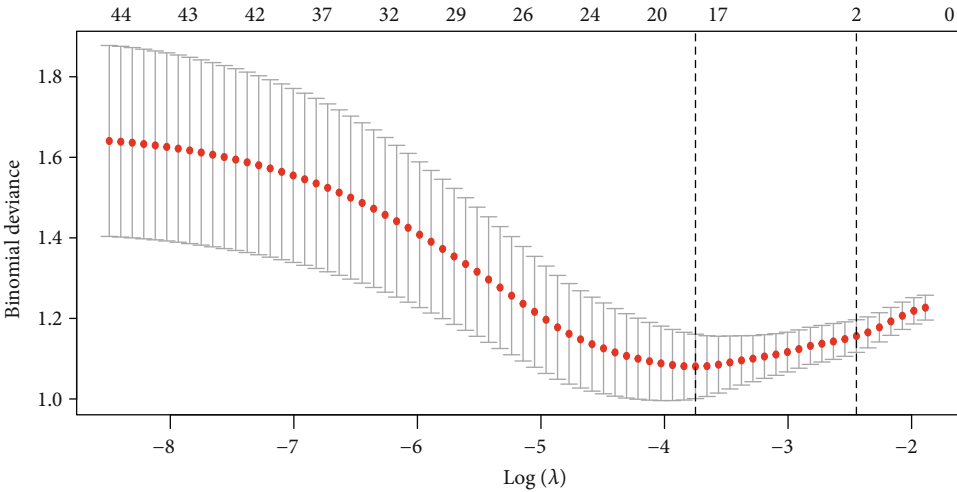
Figure 5: Continued.
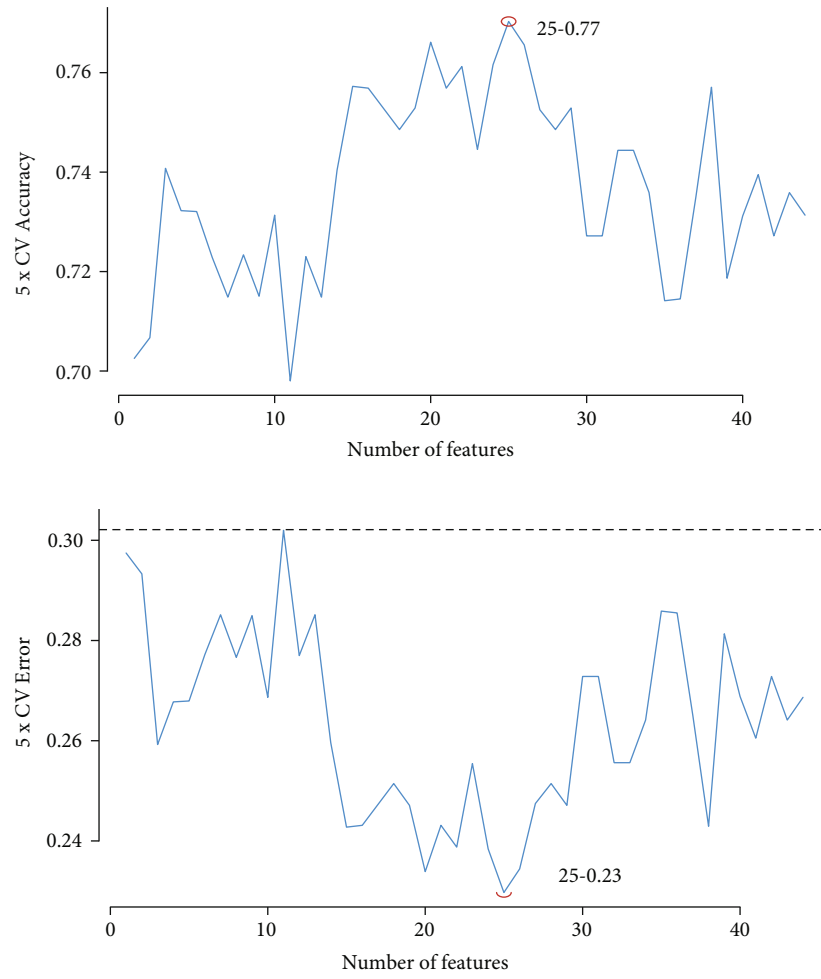
(e)



(f)

FIGURE 5: Selection of chemotherapy-related OS gene set B. Visualization of single cells in 21 subgroups of AML-BC and AML-AC (a). Random forest (RF), LASSO regression, and support vector machine (SVM) selected the chemotherapy-related OS gene set B (b–f).

(a)



(b)



(c)



(d)

Figure 6: Continued.

(e)
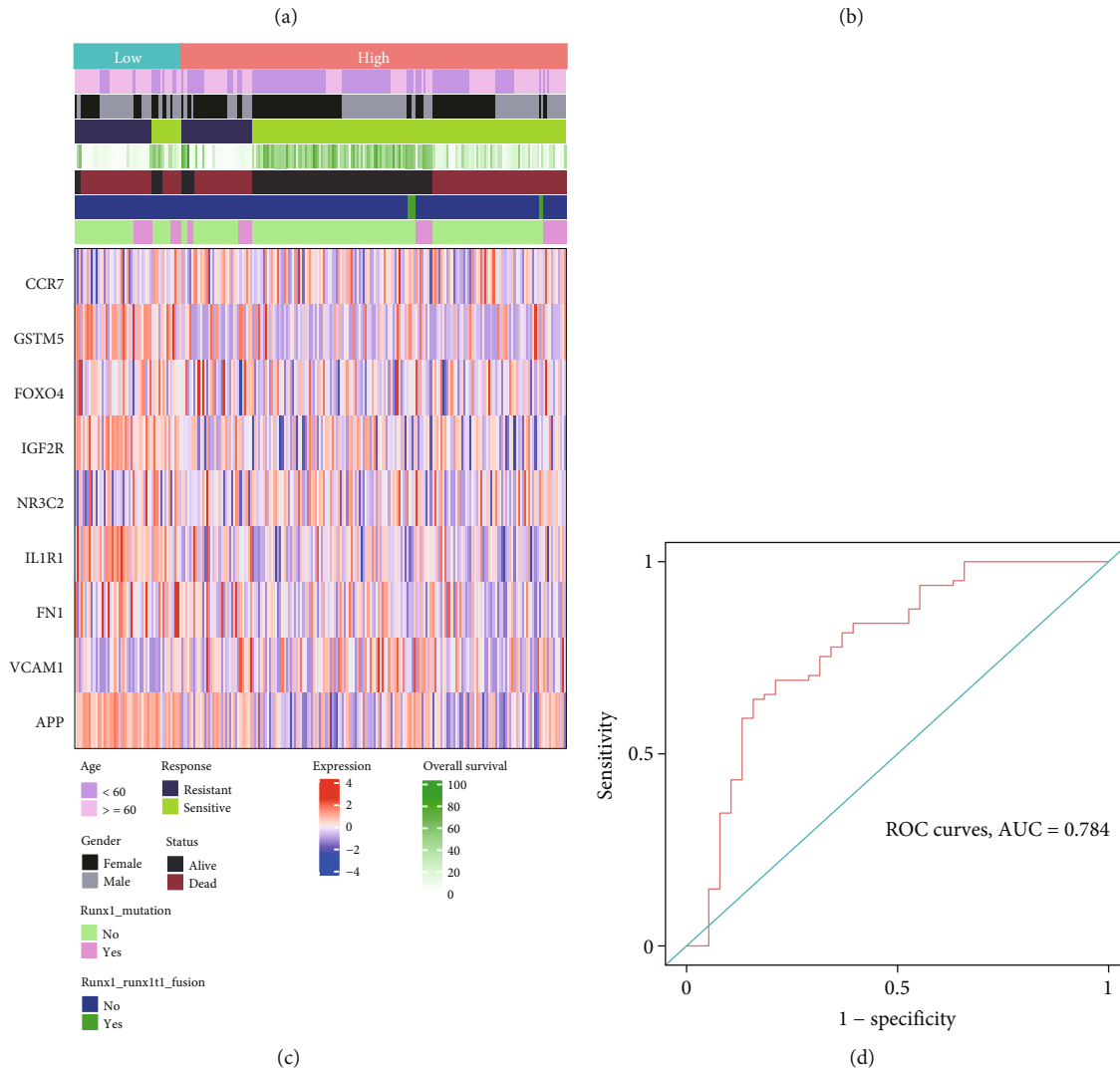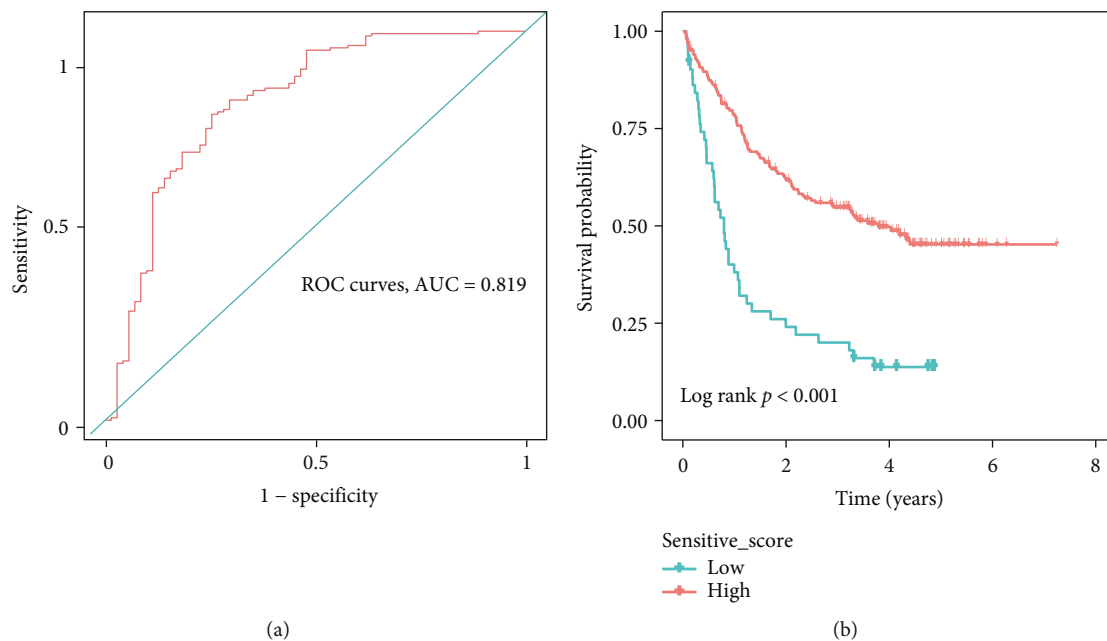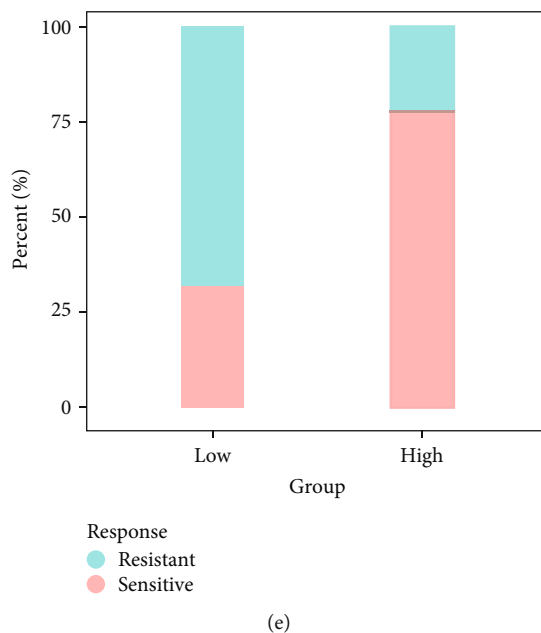
FIGURE 6: The predictive role of OS gene set B. ROC to evaluate the predictive accuracy of the training dataset GSE106291 (AUC = 0.819) (a). Patients in the high-sensitive group have better overall survival than patients in the low-sensitive group ($P < 0.001$) (b). The relationship between high- and low-sensitive groups, clinical variables, and hub genes gene expression are visualized by heatmap (c). ROC to evaluate the predictive accuracy of the internal validation of GSE106291 (AUC = 0.784) (d). Patients in the high-sensitive score group are more sensitive to chemotherapy (e).

tumor cells and the character that tumor cells can produce more peroxidative substances to escape the toxic effects of ROS by-products [26]. Although our results showed a significant difference of cell death in response to oxidative stress score between AML-BC and AML-AC groups, there is no significant difference in different leukemia cell subsets, which might be due to that the tumor microenvironment includes immune cells, stromal cells, and soluble molecules in addition to tumor cells, and the crosstalk network of these components may affect the cellular response to oxidative stress. In addition, a significant difference between the sensitive group and the resistant group can be observed, which also suggested that OS-induced cell death was related to chemotherapy resistance. The maintenance of the quiescent state of hematopoietic stem cells depends on the low production of ROS and high antioxidant defense, while leukemia stem cells are more sensitive to OS compared with normal hematopoietic stem cells [27, 28]. The function annotation results of different leukemia populations also provided evidence that HSC-like populations play an important role in the chemotherapeutic drug-induced OS. Given the bidirectional effect of ROS, we screened the oncogenesis-related OS gene set A and the chemotherapy-related OS gene set B from the scRNAseq data. Both two gene sets have biological processes, like response to oxidative stress, response to reactive oxygen species, and some other redox-related items by GO enrichment analysis. The results of KEGG also suggested that the oncogenesis-related gene set might involve cellular senescence, transcriptional misregulation in cancer, HIF-1 signaling pathway, and FoxO signaling pathway, which have been revealed by

previous studies to be involved in the tumorigenesis of AML [29–32]. In addition, the PI3K-Akt signaling pathway and NF-kappa B signaling pathway were enriched by the OS gene set B. Growing evidence suggested that targeting PI3K-Akt signaling pathway may represent an effective treatment to kill AML LSCs [33, 34], and NF-kappa B inhibitor LC1 also showed an inhibitory effect on primary AML cells in vitro [35].

The two different clusters of prognosis-related OS genes displayed distinct molecular and clinical features. It seems that the higher the oxidative stress score, the worse the prognosis, and the higher oxidative stress score group was also related to some clinical risk factors, like age and risk category [36]. We found that these differential genes were mainly enriched in the INFLAMMATORY_RESPONSE pathway, IL6_JAK_STAT3_SIGNALING, KRAS_SIGNALING_UP, and TNFA_SIGNALING_VIA_NFKB through GSVA analysis, which might be important targets for future antiperoxidative therapy in AML. Notably, myeloperoxidase (MPO), as a common marker to distinguish between the myeloid and lymphoid lineages, was mainly expressed in the low oxidative stress score group according to our result. Previous contradictory study has addressed its clinical significance in pediatric B-acute lymphoblastic leukemia (B-ALL) patients [37], and Kuang et al. [38] identified MPO as a risk gene in AML patients. This contradictory result may be due to the discrepancy between individuals and the biological functions of MPO involved in different pathways. In addition, we constructed a nine-OS-related gene signature that can effectively predict the therapy response from the chemotherapy-related OS gene set B by different machine

learning methods. Among these genes, C-C chemokine receptor type 7 (CCR7) encodes the protein CCR7, which is a member of the G protein-coupled receptor (GPR) family and expressed in various activated B and T lymphocytes, and has lower expression in nonresponders compared to responders involved in therapy resistance of AML which is revealed by a recent study [39]. Interleukin 1 receptor type I (IL1R1), also known as CD121a, is an important mediator involved in many cytokine-induced immune and inflammatory responses. Stratmann et al. [40] reported IL1R1 as a biomarker associated with AML progression and attenuated cellular growth and disease progression in primary AML cells, and AML murine models can be observed when suppressing IL-1 signaling [41]. All these suggest that the chemotherapy-related signature might involve the OS-induced immune response and inflammation to mediate chemotherapy resistance, and further stratified therapy could be adapted according to our signature.

In conclusion, our study is the first one to combine scRNAseq and bulk RNAseq data to construct two different transcriptomic profiles to reveal the different roles of OS-related genes involved in AML oncogenesis and chemotherapy resistance, which not only provide potential biological targets for the treatment of AML but also provide important insights into the mechanism of OS-related genes in the pathogenesis and drug resistance of AML. However, there are some limitations to our study. First, the occurrence of AML and drug resistance are closely related to the tumor microenvironment; our study mainly focused on the leukemia cell populations while not assessing the cell communication between leukemia cells and tumor-infiltrating immune cells mediated by the OS-related genes from the single-cell perspective. Second, more AML cohorts with scRNAseq data, which contain more patients with resistant information, should be included to reveal the role of OS-related genes in AML drug resistance. Third, further experiments should be done to explore the roles of hub OS genes in AML.

## Data Availability

All raw data included in the study are available in TCGA and GEO databases.

## Conflicts of Interest

The authors declare that there is no conflict of interest.

## Authors' Contributions

YM and DH were responsible for the conceptualization. JQ, JL, and CW were responsible for the data collection and formal analysis. JQ, HH, JY, and YX wrote the first report. YY and YW were responsible for the data visualization. All authors read and approved the final manuscript. Jing Qi, Jingyu Lin, and Changjiang Wu contributed equally to this work.

## Supplementary Materials

*Supplementary 1.* Supplementary Figure 1: cell death in response to oxidative stress score of different leukemia cells subsets between AML-BC and AML-AC: HSC-like (A), prog-like (B), GMP-like (C), ProMono-like (D), mono-like (E), and cDC-like (F).

*Supplementary 2.* Supplementary Figure 2: DEGs of AML-BC and AML-AC in HSC-like subset were enriched in OXI-DATIVE_PHOSPHORYLATION pathway of by GSEA analysis.

*Supplementary 3.* Supplementary Figure 3: LASSO regression to get 6 hub genes related to prognosis in the LAML dataset to construct a prognostic risk model (A, B).

*Supplementary 4.* Supplementary Table 1: 59 AML oncogenesis-related OS genes.

*Supplementary 5.* Supplementary Table 2: 44 AML chemotherapy-related OS genes.

## References

[1] E. Birben, U. M. Sahiner, C. Sackesen, S. Erzurum, and O. Kalayci, "Oxidative stress and antioxidant defense," *World Allergy Organization Journal*, vol. 5, no. 1, pp. 9–19, 2012.

[2] H. J. Forman and H. Zhang, "Targeting oxidative stress in disease: promise and limitations of antioxidant therapy," *Nature Reviews Drug Discovery*, vol. 20, no. 9, pp. 689–709, 2021.

[3] D. G. Harrison and M. C. Gongora, "Oxidative stress and hypertension," *Medical Clinics*, vol. 93, no. 3, pp. 621–635, 2009.

[4] A. Gella and N. Durany, "Oxidative stress in Alzheimer disease," *Cell Adhesion & Migration*, vol. 3, no. 1, pp. 88–93, 2009.

[5] L. Piconi, L. Quagliaro, and A. Ceriello, "Oxidative Stress in Diabetes," *Clinical Chemistry and Laboratory Medicine*, vol. 41, no. 9, 2003.

[6] F. Tas, H. Hansel, A. Belce et al., "Oxidative stress in breast cancer," *Medical Oncology*, vol. 22, no. 1, pp. 011–016, 2005.

[7] R. Ganesamoni, S. Bhattacharyya, S. Kumar et al., "Status of oxidative stress in patients with renal cell carcinoma," *The Journal of Urology*, vol. 187, no. 4, pp. 1172–1176, 2012.

[8] J. D. Hayes, A. T. Dinkova-Kostova, and K. D. Tew, "Oxidative stress in cancer," *Cancer Cell*, vol. 38, no. 2, pp. 167–197, 2020.

[9] S. Masciarelli, E. Capuano, T. Ottone et al., "Retinoic acid synergizes with the unfolded protein response and oxidative stress to induce cell death in FLT3-ITD+ AML," *Blood Advances*, vol. 3, no. 24, pp. 4155–4160, 2019.

[10] S. Arfin, N. K. Jha, S. K. Jha et al., "Oxidative stress in cancer cell metabolism," *Antioxidants*, vol. 10, no. 5, p. 642, 2021.

[11] H. Döhner, D. J. Weisdorf, and C. D. Bloomfield, "Acute myeloid leukemia," *The New England Journal of Medicine*, vol. 373, no. 12, pp. 1136–1152, 2015.

[12] H. Bai, M. Zhou, M. Zeng, and L. Han, "PLA2G4A is a potential biomarker predicting shorter overall survival in patients with non-M3/NPM1 wildtype acute myeloid leukemia," *DNA and Cell Biology*, vol. 39, no. 4, pp. 700–708, 2020.

[13] E. H. Estey, "Acute myeloid leukemia: 2021 update on risk-stratification and management," *American Journal of Hematology*, vol. 95, no. 11, pp. 1368–1398, 2020.

[14] M. Rasool, S. Farooq, A. Malik et al., "Assessment of circulating biochemical markers and antioxidative status in acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) patients," *Saudi journal of biological sciences*, vol. 22, no. 1, pp. 106–111, 2015.

[15] R. Ciarcia, D. d'Angelo, C. Pacilio et al., "Dysregulated calcium homeostasis and oxidative stress in chronic myeloid leukemia (CML) cells," *Journal of Cellular Physiology*, vol. 224, no. 2, pp. 443–453, 2010.

[16] C. Dong, N. Zhang, and L. Zhang, "The multi-omic prognostic model of oxidative stress-related genes in acute myeloid leukemia," *Frontiers in Genetics*, vol. 12, article 1781, 2021.

[17] Z. Wu, L. Wang, Z. Wen, and J. Yao, "Integrated analysis identifies oxidative stress genes associated with progression and prognosis in gastric cancer," *Scientific Reports*, vol. 11, no. 1, article 3292, 2021.

[18] E. Z. Macosko, A. Basu, R. Satija et al., "Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets," *Cell*, vol. 161, no. 5, pp. 1202–1214, 2015.

[19] P. van Galen, V. Hovestadt, M. H. Wadsworth II et al., "Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity," *Cell*, vol. 176, no. 6, pp. 1265–1281.e24, 2019.

[20] 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.

[21] M. A. Essers, S. Weijzen, A. M. M. de Vries-Smits et al., "FOXO transcription factor activation by oxidative stress mediated by the small GTPase Ral and JNK," *The EMBO Journal*, vol. 23, no. 24, pp. 4802–4812, 2004.

[22] A. Sallmyr, J. Fan, K. Datta et al., "Internal tandem duplication of FLT3 (FLT3/ITD) induces increased ROS production, DNA damage, and misrepair: implications for poor prognosis in AML," *Blood, The Journal of the American Society of Hematology*, vol. 111, no. 6, pp. 3173–3182, 2008.

[23] P. S. Hole, L. Pearn, A. J. Tonks et al., "Ras-induced reactive oxygen species promote growth factor–independent proliferation in human CD34$^+$ hematopoietic progenitor cells," *Blood, The Journal of the American Society of Hematology*, vol. 115, no. 6, pp. 1238–1246, 2010.

[24] T. E. Fandy, A. Jiemjit, M. Thakar, P. Rhoden, L. Suarez, and S. D. Gore, "Decitabine induces delayed reactive oxygen species (ROS) accumulation in leukemia cells and induces the expression of ROS generating enzymes," *Clinical Cancer Research*, vol. 20, no. 5, pp. 1249–1258, 2014.

[25] M. Hosseini, H. R. Rezvani, N. Aroua et al., "Targeting myeloperoxidase disrupts mitochondrial redox balance and overcomes cytarabine resistance in human acute myeloid leukemia," *Cancer Research*, vol. 79, no. 20, pp. 5191–5203, 2019.

[26] H. Takeuchi, Y. Kondo, K. Fujiwara et al., "Synergistic augmentation of rapamycin-induced autophagy in malignant glioma cells by phosphatidylinositol 3-kinase/protein kinase B inhibitors," *Cancer Research*, vol. 65, no. 8, pp. 3336–3346, 2005.

[27] C. L. Bigarella, R. Liang, and S. Ghaffari, "Stem cells and the impact of ROS signaling," *Development*, vol. 141, no. 22, pp. 4206–4218, 2014.

[28] U. Testa, C. Labbaye, G. Castelli, and E. Pelosi, "Oxidative stress and hypoxia in normal and leukemic stem cells," *Experimental Hematology*, vol. 44, no. 7, pp. 540–560, 2016.

[29] J. R. Dörr, Y. Yu, M. Milanovic et al., "Synthetic lethal metabolic targeting of cellular senescence in cancer therapy," *Nature*, vol. 501, no. 7467, pp. 421–425, 2013.

[30] X. Yin, S. Huang, R. Zhu, F. Fan, C. Sun, and Y. Hu, "Identification of long non-coding RNA competing interactions and biological pathways associated with prognosis in pediatric and adolescent cytogenetically normal acute myeloid leukemia," *Cancer Cell International*, vol. 18, no. 1, pp. 1–11, 2018.

[31] J. Zhang and G.-Q. Chen, "Hypoxia-HIF-1$\alpha$-C/EBP$\alpha$/Runx1 signaling in leukemic cell differentiation," *Pathophysiology*, vol. 16, no. 4, pp. 297–303, 2009.

[32] S. M. Sykes, S. W. Lane, L. Bullinger et al., "AKT/FOXO signaling enforces reversible differentiation blockade in myeloid leukemias," *Cell*, vol. 146, no. 5, pp. 697–708, 2011.

[33] A. M. Martelli, C. Evangelisti, F. Chiarini, C. Grimaldi, L. Manzoli, and J. A. McCubrey, "Targeting the PI3K/AKT/mTOR signaling network in acute myelogenous leukemia," *Expert Opinion on Investigational Drugs*, vol. 18, no. 9, pp. 1333–1349, 2009.

[34] S. Darici, H. Alkhaldi, G. Horne, H. G. Jørgensen, S. Marmiroli, and X. Huang, "Targeting PI3K/Akt/mTOR in AML: rationale and clinical evidence," *Journal of Clinical Medicine*, vol. 9, no. 9, p. 2934, 2020.

[35] C. Jenkins, C. Pepper, K. Mills, and A. Burnett, "NF kappa B as a therapeutic target in AML," *Blood*, vol. 108, no. 11, p. 2587, 2006.

[36] A. Bataller, A. Garrido, F. Guijarro et al., "European LeukemiaNet 2017 risk stratification for acute myeloid leukemia: validation in a risk-adapted protocol," *Blood Advances*, vol. 6, no. 4, pp. 1193–1206, 2022.

[37] M. M. Fawzy, A. Abd el-Hafez, S. el-Ashwah et al., "Isolated myeloperoxidase immunohistochemical expression in bone marrow biopsy depicts clinical outcomes in adults with typical B-acute lymphoblastic leukemia," *Asian Pacific Journal of Cancer Prevention*, vol. 22, no. 7, pp. 2143–2152, 2021.

[38] Y. Kuang, Y. Wang, X. Cao, C. Peng, and H. Gao, "New prognostic factors and scoring system for patients with acute myeloid leukemia," *Oncology Letters*, vol. 22, no. 6, pp. 1–15, 2021.

[39] K. Yamatani, Y. Tabe, A. Maiti et al., "Diverse mechanisms of resistance to decitabine and venetoclax therapy in newly diagnosed and relapsed/refractory AML inferred by transcriptome analysis," *Blood*, vol. 138, Supplement 1, p. 2244, 2021.

[40] S. Stratmann, S. A. Yones, M. Garbulowski et al., "Transcriptomic analysis reveals proinflammatory signatures associated with acute myeloid leukemia progression," *Blood Advances*, vol. 6, no. 1, pp. 152–164, 2022.

[41] A. Carey, D. Edwards, C. A. Eide et al., "Targeted suppression of interleukin-1 signaling attenuates leukemic cellular growth and disease progression in primary human cells and murine models of acute myeloid leukemia," *Blood*, vol. 124, no. 21, p. 702, 2014.

*Research Article*

# Upregulated SSB Is Involved in Hepatocellular Carcinoma Progression and Metastasis through the Epithelial-Mesenchymal Transition, Antiapoptosis, and Altered ROS Level Pathway

**Hao Wu** [1], **Zhixin Zhang** [1], **Xinyu Han**,[2] **Sai Zhang** [3] **Jinrui Zhang** [1] **Pinsheng Han**,[1] **Youcheng Zhang**,[1] **Yi Bai**,[4] **and Yamin Zhang** [4]

[1]*The First Central Clinical School, Tianjin Medical University, Tianjin 300070, China*
[2]*Department of Genetics, School of Basic Medical, Tianjin Medical University, Tianjin 300070, China*
[3]*School of Medicine, Nankai University, Tianjin 300192, China*
[4]*Department of Hepatobiliary Surgery, Tianjin First Central Hospital, School of Medicine, Nankai University, Tianjin 300192, China*

Correspondence should be addressed to Yamin Zhang; 5020200824@nankai.edu.cn

Hepatocellular carcinoma (HCC) is one of the most common malignant tumors with high morbidity and mortality. Therefore, finding new diagnostic and therapeutic targets is vital for HCC patients. Recent studies have shown that dysregulation of RNA-binding proteins is often associated with cancer progression. Several studies have reported that the RNA-binding protein SSB can promote cancer occurrence and progression and is linked to tumor epithelial-mesenchymal transition (EMT), which could be a new diagnostic marker and therapeutic target. However, the expression and function of SSB in HCC remain to be elucidated. Therefore, this study is aimed at clarifying the expression and biological function of SSB in HCC through bioinformatics analysis combined with in vitro experiments. We found that SSB is highly expressed in HCC and is associated with the poor prognosis of HCC patients, and it can serve as an independent unfavorable prognostic factor. Knockdown of SSB can inhibit the growth of HCC cells in vitro, increase the level of apoptosis and the expression of pro-apoptosis-related proteins, and decrease the expression of antiapoptotic proteins. Meanwhile, SSB knockdown reduced HCC cell invasiveness, and the expression of EMT-related proteins changed significantly. We also found that the gene SSB was associated with the level of oxidative stress in liver cancer cells, and the level of intracellular reactive oxygen species (ROS) increased after knockdown of SSB. The results of bioinformatics analysis also showed that high expression of SSB may affect the effect of checkpoint blockade (ICB) therapy. In conclusion, we found that SSB is highly expressed in HCC and that upregulated SSB can promote the proliferation and metastasis of HCC through antiapoptotic, altered intracellular oxidative stress level, and EMT pathways, which can serve as a new diagnostic marker and therapeutic target, and patients with high SSB expression may not have obvious ICB therapy effect.

## 1. Introduction

According to the most recent statistics, more than 8 million new hepatocellular carcinoma (HCC) cases are diagnosed yearly, causing more than 4 million deaths [1]. The incidence of HCC is highest in Asia, especially in China, which accounts for almost half of the global cases [2]. While several

therapies have been developed in recent years, including liver transplantation, hepatectomy, targeted therapy, ablation, and transcatheter arterial chemoembolization, the survival rate in HCC patients is only 18% for five years [3]. This is mainly because the early clinical manifestations of HCC patients are not apparent, and most treatments are limited to early-stage patients. Thus, exploring novel therapeutic

targets based on molecular biomarkers is urgently needed for HCC patients, which is crucial for the survival of HCC patients.

SSB, also known as RNA-binding protein La or La-related protein 3 (LARP3), belongs to the LA-related protein family and is an RNA-binding protein mainly expressed in the nucleus [4, 5]. RNA-binding proteins (RBPs) participate in posttranscriptional regulation, such as splicing, polyadenylation, and stabilization [6]. Abnormal posttranscriptional regulatory processes may lead to tumorigenesis, and the mechanisms underlying this have been elucidated, including genomic changes and posttranscriptional modifications [6, 7]. RBPs also influence oncogenes and tumor suppressor genes' expression and function [8]. Studies have shown that different members of LARP family are involved in the occurrence and progression of cancer [9]. The SSB is highly expressed in various tumors, including chronic myelogenous leukemia, ovarian cancer, and head and neck cancer [10–12]. Aberrant expression of SSB proteins contributes to increased cancer cell proliferation, migration, invasion, and chemoresistance and promotes tumor growth in mice [12–14]. In addition, the SSB protein has RNA chaperone activity that promotes the processing of noncoding precursor RNAs but also stimulates the translation of selective mRNAs that encode cancer-promoting and antiapoptotic genes [15]. Oxidative stress plays an important role in the occurrence and development of cancer, and it also affects the prognosis of patients [16]. Thus, SSB may also affect the level of oxidative stress in tumor cells. However, there are still few reports and in-depth studies on the role of SSB in HCC. Therefore, the specific molecular mechanism of SSB in HCC needs to be explored urgently.

This study is aimed at investigating the prognostic role and cancer-promoting molecular mechanisms of SSB in HCC to identify SSB as a potential therapeutic target for HCC patients. Through bioinformatics analysis combined with in vitro experiments, we confirmed for the first time that SSB was significantly upregulated in HCC tissues and correlated with poor prognosis. It could promote the progression and metastasis of HCC through epithelial-mesenchymal transition (EMT), altered intracellular oxidative stress level, and antiapoptotic pathways and may affect the efficacy of immunotherapy.

## 2. Materials and Methods

*2.1. Data Acquisition.* The transcriptome sequencing and clinicopathological data were retrieved from TCGA-LIHC dataset (https://www.portal.gdc.cancer.gov/), and GTEx data (https://gtexportal.org/) was also downloaded. In addition, the GSE121248 dataset was downloaded from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) for external expression validation. The differential expression of SSB protein in normal liver and HCC tissue was analyzed using immunohistochemistry (IHC) data from the Human Protein Atlas database (HPA, https://www.proteinatlas.org/). In addition, we performed SSB pancancer expression and survival analysis using TIMER 2.0 (http://timer.cistrome.org/) and GEPIA (http://gepia.cancer-pku.cn/) databases.

*2.2. Independent Prognostic and Clinical Characteristic Analysis of Gene SSB.* The patients were divided into high- and low-expression groups according to the median expression level of gene SSB. Survival analysis was performed using Kaplan-Meier and log-rank tests. We also assessed the relationship between SSB and clinical characteristics such as age, sex, stage, grade, and TNM stage. In addition, univariate and multivariate Cox analyses were performed for SSB and clinical characteristics to determine independent prognostic indicators. Analyses were performed using R software's "survival" and "survminer" packages (v 4.0.2). $P < 0.05$ was considered significant.

*2.3. Gene Set Enrichment Analyses (GSEA).* SSB expression was classified into high- and low-expression phenotypes based on the median value of gene SSB expression. Enrichment analysis was performed with default parameter settings. The random number set is 1000 times. The FDR $< 0.05$ are significantly enriched in GSEA.

*2.4. Cell Culture.* The human normal liver cell line (L-02) and hepatocarcinoma cell lines (HepG2 and SMMC-7721) were purchased from ICell Bioscience Inc (Shanghai, China) and the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). The L-02 and SMMC-7721 cell lines were cultured in RPMI-1640 medium (Gibco, Grand Island, USA), and HepG2 cell line was cultured in DMEM medium (Gibco, Grand Island, USA) and all media containing 10% fetal bovine serum (FBS, Gibco, Grand Island, USA) and 1% penicillin/streptomycin (Beyotime, Shanghai, China). All cells were preserved in a humidified incubator (37°C, 5% $CO_2$).

*2.5. Transfection of Hepatocarcinoma Cell Lines.* Tianjin Medical University's Basic Experiment Center extracted and synthesized the pLKO.1-SSB knockdown and pLKO.1-SSB (scramble) plasmid. Three different SSB short hairpin RNAs (shRNAs) were designed and transfected into hepatoma cells using lentiviral packaging plasmids. The specific information is as follows: sh-scramble (5′-TGTGAGGAA CTTGAGATCT-3′), sh-SSB1 (5′-CCTGCATCCAAACA ACAGAAA-3′) and sh-SSB2 (5′-GCTGAAATGAAATCTC TAGAA-3′). The shSSB base sequence is based on previously studied and verified sequence information [17].

*2.6. RNA Extraction and Quantitative Real-Time Polymerase Reaction (qRT-PCR) Assay.* A total RNA extraction kit (Solarbio, Beijing, China) was used to extract total RNA from cells following the manufacturer's instructions. Afterward, the RNA was quantified using the One-Step SYBR Prime Script RT-PCR kit (Takara, Japan). The expression of SSB was normalized to GAPDH and was analyzed using the $2^{-\triangle\triangle CT}$ method [18]. The sequence information of the SSB primers is in Supplementary Table 1.

*2.7. Cell Viability Assay.* Cell viability was measured by cell counting kit-8 (CCK-8, Boster, Wuhan, China) assay. Following transfection, each group's cells were seeded into 96-well plates ($1 \times 10^4$ cells/well) and cultured in the

incubator for 24, 48, and 72 h. After that, $10 \mu$L of CCK-8 was added to each well and placed in the incubator (37°C, 5% $CO_2$) for 1 h. Then, the absorbance at 450 nm was measured using a microplate reader (EnSpire, USA).

2.8. *Colony Formation Assay.* Cells in each group were digested with 0.25% trypsin, seeded into 6-well plates, and cultured in a medium containing 10% FBS for two weeks. When visible colonies were formed, cells were fixed in 20% methanol for 10–15 min, then stained with 0.1% crystal violet for 10 min at room temperature, and finally, the colonies in each well were counted [19].

2.9. *Wound Healing Assay.* Cells from each group were seeded into 6-well plates for 24 h before replacing the medium with a serum-free medium. The cells were scratched with a yellow pipette tip and cultured for 48 h. The cells were photographed at 0, 6, 24, and 48 h. Image-J (V7.0) software was used to quantify the distance between two edges of the wound surface at each time point.

2.10. *Apoptosis Detection by Flow Cytometry.* Cell apoptosis was detected using the Annexin-V-FITC/PI Apoptosis Detection Kit (Solarbio, Beijing, China). After digesting the cells of each group, $500 \mu$L of binding buffer, $5 \mu$L of FITC-Annexin V, and $5 \mu$L of PI working solution were added. Then, the cells were incubated at room temperature in the dark for 15 min. Finally, the level of apoptosis was detected by flow cytometry (BD Accuri C6 Plus, USA) [20]. Data were processed using FlowJo (V10.0) software.

2.11. *Cell Cycle Detection by Flow Cytometry.* Cell cycle distribution was assessed using the Cell Cycle Detection Kit (KeyGEN, Nanjing, China). Cells in each group were first fixed with 70% ethanol at 4°C overnight, then collected by centrifugation, mixed with $500 \mu$L of RNase A/PI working solution, and incubated at room temperature for 30 min. Finally, flow cytometry was used for detection. Data were processed using ModFit LT (V3.1) software.

2.12. *Measurement of ROS Levels by Flow Cytometry.* The ROS levels were measured using the ROS assay kit (Solarbio, Beijing, China). Briefly, cells were resuspended in a serum-free medium containing 10 mM DCFH-DA. Subsequently, flow cytometry was performed to detect the DCF fluorescence intensity. Finally, the mean fluorescence intensity (MFI) was calculated.

2.13. *Transwell Assay.* Cell invasion assays were performed using a transwell cell culture chamber coated with Matrigel (Corning Costar, Inc.) according to the manufacturer's instructions. First, the cells of each group were inoculated in a serum-free medium in the upper chamber, and the medium containing 20% FBS was filled in the lower chamber. After that, cells were cultured in an incubator (37°C, 5% $CO_2$) for 24 h. Transwell migration experiments are similar to the above steps, except that Matrigel is not added to the migration experiments. Data were processed using ImageJ (V7.0) software.

2.14. *Western Blotting.* Total protein was extracted from cells in each group using RIPA lysate (Solarbio, Beijing, China). The protein concentration was measured by the BCA (KeyGEN, Nanjing, China) method. Then, the proteins were separated by SDS-PAGE (Boster, Wuhan, China). After the target proteins were transferred to the nitrocellulose filter (NC, Boster, Wuhan, China) membrane, it was blocked with 5% skimmed milk at room temperature for 1 h. Next, these membranes were incubated with primary antibodies such as SSB (1:1000, Boster, Wuhan, China), E-cadherin (1:1000, Cell Signaling Technology, USA), N-cadherin (1:1000, Cell Signaling Technology, USA), ZO-1 (1:3000, Proteintech, Wuhan, China), Vimentin (1:1000, Immuno-Way, USA), GAPDH (1:20000, Affinity Biosciences, China), Caspase3 (1:1000, Santa Cruz, USA), BCL2 (1:1000, Abcam, USA), BCLXL(1:1000, Abcam, USA), and BAX (1:2000, Abcam, USA) overnight at 4°C. Subsequently, these membranes were incubated with the corresponding secondary antibody for 1 h at room temperature. The membranes were exposed using an imaging system (Bio-Rad, Hercules, USA) and quantified using ImageJ (V7.0) software.

2.15. *Estimation of Correlations between SSB Expression Level and Immune Cells.* Spearman's correlation analysis was used to describe the correlation between the SSB and various immune cells. The correlation analyses between SSB expression and immune cells were performed using the R package "ggstatsplot" (v 4.0.2). $P < 0.05$ was considered significant.

2.16. *Prediction of Immune Checkpoint Blockade (ICB) Therapy Response.* According to the median SSB expression, HCC samples were divided into two groups, and Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was used to predict the potential response of ICB in the two groups [21].

2.17. *Evaluation of the Relationship between SSB Expression and Immune Checkpoint-Related Genes.* SIGLEC15, TIGIT, CD274, HAVCR2, PDCD1, CTLA4, LAG3, and PDCD1LG2 are genes related to immune checkpoints. The expression values of these eight genes were extracted to clarify the relationship between immune checkpoint-related genes and SSB expression. The above results were statistically analyzed by R software (v 4.0.2). $P < 0.05$ was considered significant.

2.18. *Statistical Analysis.* The bioinformatics data analyses of this study were performed using R software (v 4.0.2). The Wilcoxon rank-sum test determined the comparison between the two groups. Analysis of in vitro experimental data was conducted using GraphPad Prism 8.0 statistical software. A one-way analysis of variance (ANOVA) compared the differences among three or more groups, and the Student $t$-test evaluates the significance between the two groups. $P < 0.05$ was considered to be statistically significant.

# 3. Results

3.1. *SSB Is Upregulated in HCC and Associated with Poor Prognosis of HCC Patients.* A total of 374 tumor samples

(a)



(b)



(c)



(d)



(e)
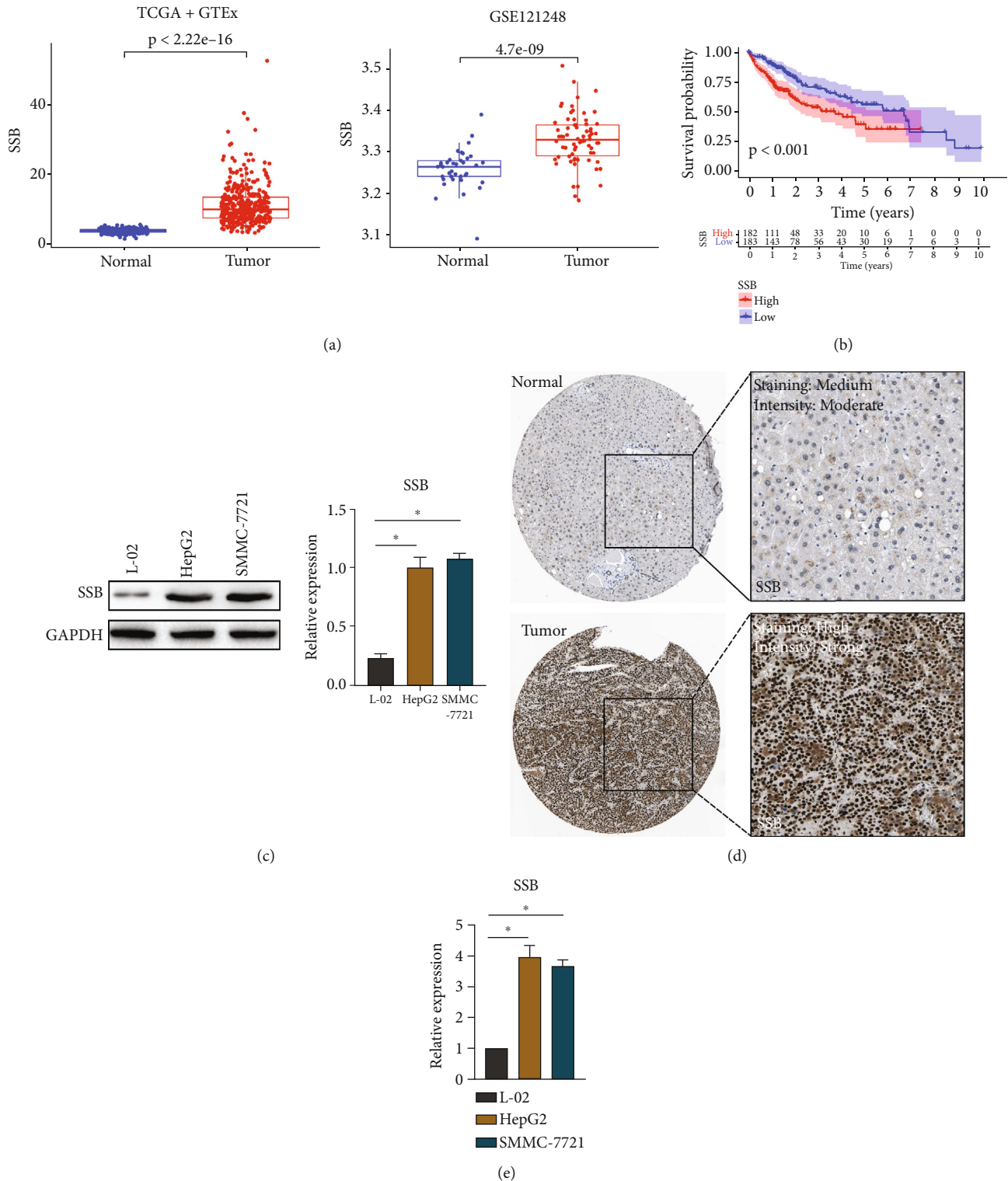
FIGURE 1: SSB is highly expressed in HCC tissues or cells. (a) SSB is highly expressed in HCC tissues compared to normal tissues based on TCGA and GEO databases. (b) High expression of SSB is associated with poor prognosis in HCC patients. (c) SSB is highly expressed in HepG2 and SMMC-7721 HCC cells compared with normal hepatocyte L-02. (d) SSB is highly expressed in HCC tissues compared to normal liver tissues based on the HPA IHC database. (e) qRT-PCR results showed that SSB expression was significantly upregulated in HepG2 and SMMC-7721 cells compared with L-02 cells. Data are shown as mean ± SD. $^{*}P < 0.05$.

(TCGA) and 276 normal samples (TCGA + GTEx) were obtained. In addition, 70 HCC and 37 normal samples were obtained from the GEO database (GSE121248). As displayed in Figure 1(a), in TCGA and GEO databases, SSB was highly expressed in HCC tissues compared with normal samples ($P < 0.05$). Survival analysis results suggested that patients

with high SSB expression had a poor prognosis (Kaplan-Meier and log-rank test, $P < 0.05$, Figure 1(b)). In vitro experiments, western blot, and qRT-PCR showed that SSB was highly expressed in HCC cell lines HepG2 and SMMC-7721 compared with normal liver cell line L-02 ($P < 0.05$, Figures 1(c) and 1(e)). The IHC data of SSB in HCC tissues and normal tissues were downloaded from the HPA database. The results are depicted in Figure 1(d), SSB was highly expressed in HCC tissues compared with normal liver tissues.

### 3.2. Evaluation of the Correlation between SSB Gene Expression and Clinicopathological Characteristics. 
The clinical information of patients in TCGA-LIHC cohort is shown in Table 1. Increased SSB expression correlated with age ($P = 0.021$, Figure 2(a)), histological grade (G3/G4 vs. G1/G2, $P = 0.0004$, Figure 2(c)), T stage (T3/T4 vs. T1/T2, $P = 0.0033$, Figure 2(d)), N stage ($P = 0.028$, Figure 2(e)), and AJCC stage (III/IV vs. I/II, $P = 0.00032$, Figure 2(g)). However, there was no correlation between gender, M stage, and SSB expression ($P > 0.05$, Figures 2(b) and 2(f)).

### 3.3. SSB Is an Independent Prognostic Risk Factor and Is Involved in Multiple Cancer-Related Signaling Pathways. 
The potential KEGG pathway between high and low SSB expression was analyzed using GSEA method. The results are described in Figure 2(h). The high expression of SSB enriches a variety of cancer-related signaling pathways, such as "colorectal cancer," "glioma," "pancreatic cancer," "small cell and non-small-cell lung cancer," and other cancer signaling pathways. In addition, the high expression of SSB also affects the cell cycle and enriches signaling pathways related to cancer invasion and metastasis, such as "adherens junction" and "gap junction." On the other hand, the low expression of SSB mainly enriches the metabolism-related signaling pathways.

Independent prognostic analyses of SSB and clinicopathological characteristics were assessed using univariate and multivariate Cox regression methods. Univariate Cox analysis showed that AJCC stage, T stage, M stage, and SSB expression were relevant to overall survival (all $P < 0.05$, Figure 2(i)), and the multivariate Cox analysis showed that only SSB expression was relevant to overall survival, which also indicated that SSB was an independent prognostic factor ($P < 0.05$, Figure 2(j)).

### 3.4. Pancancer Analysis of SSB. 
The pancancer expression analysis results are shown in Figure 3(a). SSB is differentially expressed in other 14 cancer types ($P < 0.05$). In 9 of the 14 types of cancer, SSB was highly expressed in tumor tissues compared to normal tissues, including urothelial bladder carcinoma (BLCA), cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), rectum adenocarcinoma (READ), and stomach adenocarcinoma (STAD). In the remaining five cancer types, SSB was lowly expressed in tumor tissues compared with normal tissues or metastasis tissues, including kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC),

TABLE 1: The clinical characteristics in TCGA-LIHC cohort.

| Parameter | Type | Patients |
|---|---|---|
| Status | Alive | 241 (64.96%) |
| | Dead | 130 (35.04%) |
| Age | Mean (SD) | 59.4 (13.5) |
| | Median [Min, Max] | 61 [16, 90] |
| Gender | Female | 121 (32.61%) |
| | Male | 250 (67.39%) |
| Histologic grade | G1 | 55 (14.82%) |
| | G2 | 178 (47.98%) |
| | G3 | 120 (32.35%) |
| | G4 | 13 (3.50%) |
| | Unknown | 5 (1.35%) |
| Pathologic M | M0 | 269 (72.51%) |
| | M1 | 3 (0.81%) |
| | Unknown | 99 (26.68%) |
| Pathologic N | N0 | 253 (68.19%) |
| | N1 | 4 (1.08%) |
| | Unknown | 114 (30.73%) |
| Pathologic T | T1 | 184 (49.60%) |
| | T2 | 92 (24.80%) |
| | T3 | 79 (21.29%) |
| | T4 | 13 (3.50%) |
| | Unknown | 3 (0.81%) |
| AJCC stage | Stage I | 174 (46.90%) |
| | Stage II | 85 (22.91%) |
| | Stage III | 84 (22.58%) |
| | Stage IV | 4 (1.08%) |
| | Unknown | 24 (6.53%) |

kidney renal papillary cell carcinoma (KIRP), skin cutaneous melanoma (SKCM), and thyroid carcinoma (THCA).

According to the median value of SSB gene expression, patients were divided into low- and high-expression groups, and survival analysis was performed using Kaplan-Meier and log-rank tests. The results of the SSB pancancer survival analysis are illustrated in Figure 3(b). SSB has survival significance in 6 cancer types [$P < 0.05$]. The high expression of SSB in adrenocortical carcinoma (ACC), HNSC, KIRP, brain lower grade glioma (LGG), and LUAD has a poor prognosis, while high SSB expression in KIRC has a good prognosis.

### 3.5. SSB Knockdown Inhibited the Proliferation and Migration Ability of HepG2 and SMMC-7721 Cell Lines and Increased Intracellular ROS Levels. 
Three types of shRNA (sh-scramble, sh-SSB1, and sh-SSB2) were designed to transfect HepG2 and SMMC-7721 cell lines. The purpose of sh-SSB1 and sh-SSB2 is to knock down the gene SSB. After determining the knockdown efficiency, we performed the CCK-8 experiment. The results are shown in Figures 4(a) and 4(b). In HepG2 and SMMC-

Figure 2: Continued.

| | pvalue | Hazard ratio |
|---|---|---|
| Age | 0.583 | 1.005 (0.987–1.023) |
| Stage | < 0.001 | 1.843 (1.440–2.357) |
| Grade | 0.0961 | 0.992 (0.730–1.349) |
| T | < 0.001 | 1.788 (1.422–2.249) |
| M | 0.020 | 3.969 (1.244–12.661) |
| N | 0.306 | 2.088 (0.510–8.548) |
| SSB | < 0.001 | 2.103 (1.378–3.209) |

Univariate Cox analysis

(i)

| | pvalue | Hazard ratio |
|---|---|---|
| Age | 0.242 | 1.011 (0.992–1.030) |
| Stage | 0.968 | 1.020 (0.384–2.714) |
| Grade | 0.885 | 1.025 (0.732–1.435) |
| T | 0.262 | 1.660 (0.685–4.024) |
| M | 0.866 | 1.118 (0.305–4.095) |
| N | 0.508 | 1.838 (0.304–11.123) |
| SSB | 0.009 | 1.841 (1.164–2.911) |

Multivariate Cox analysis

(j)

FIGURE 2: SSB is a prognostic risk factor involved in multiple cancer-related signaling pathways (a–g). The correlation of SSB expression with clinicopathological variables. (h) GSEA suggests that SSB is involved in multiple cancer-related pathways. (i, j) Univariate and multivariate Cox analyses confirmed SSB as an independent prognostic risk factor.

7721 cell lines, SSB expression was significantly reduced in both sh-SSB1 and sh-SSB2 groups ($P < 0.05$) compared to CON and sh-scramble groups, demonstrating that both sh-SSB1 and sh-SSB2 have good knockdown efficiency. The subsequent CCK8 experiments showed that in HepG2 and SMMC-7721 cell lines, knockdown of SSB could significantly inhibit the proliferation of HCC cell lines compared with CON and sh-scramble groups ($P < 0.05$, Figures 4(c) and 4(d)). However, there was no significant difference in cell viability between CON and sh-scramble groups ($P > 0.05$).

We only selected sh-SSB1 as sh-SSB to knock down SSB for subsequent in vitro experiments. The results of clone formation experiments showed that in HepG2 and SMMC-7721 cell lines, the number of cell clones formed after knockdown of SSB was significantly lower than that in CON and sh-scramble group cells ($P < 0.05$, Figure 4(e)). The results of cell scratch experiments are shown in Figure 4(f). In HepG2 and SMMC-7721 cell lines, there was no significant difference in cell migration rates between groups at 6 hours ($P > 0.05$), while at 24 and 48 hours, the cell migration rate of the sh-SSB group was significantly lower than that of the CON and sh-scramble groups ($P < 0.05$). The 24-hour transwell migration assay confirmed that the cell migration ability of the knockdown SSB group decreased ($P < 0.05$, Supplementary Figure 1). In addition, we used flow cytometry to detect the cell cycle of the cells in each group. The results showed that in the HepG2 and SMMC-7721 cell lines, the proportion of cells in the proliferative G2 phase in the sh-SSB group decreased compared with that in the CON and sh-scramble groups ($P < 0.05$, Figure 4(g)).

We used flow cytometry to detect ROS levels, and the results showed that the intracellular ROS levels were increased in hepatoma cells in the SSB knockdown group compared with the CON and sh-scramble groups ($P < 0.05$, Figure 4(h)).

### 3.6. Knockdown of SSB Promotes Apoptosis in HepG2 and SMMC-7721 Cell Lines. After collecting the cells of each group, the apoptosis level was detected by flow cytometry.

The results are shown in Figure 5(a). In the HepG2 and SMMC-7721 cell lines, the apoptosis rate was increased in the sh-SSB group compared with the CON and sh-scramble groups (all $P < 0.05$). However, there was no significant difference in apoptosis rate between the CON group and the sh-scramble group ($P > 0.05$). In addition, we extracted the total protein of each group of cells for western blot experiments. The results showed that in HepG2 and SMMC-7721 cell lines, knockdown of SSB significantly increased the expression of cleaved-Caspase3 and BAX proapoptotic protein and decreased the antiapoptotic protein BCL2 and BCLXL expression compared with the CON group and sh-scramble group ($P < 0.05$, Figures 5(b) and 5(c)).

### 3.7. SSB Is Involved in the Invasion and Metastasis of HCC through the EMT Pathway. Transwell invasion assay was used to detect the effect of knockdown of SSB on the invasion function of HepG2 and SMMC-7721 HCC cell lines. The results are depicted in Figure 6(a). In HepG2 and SMMC-7721 cells, the invasive ability of the cells after knockdown of SSB decreased compared with CON and sh-scramble groups ($P < 0.05$). In addition, we performed western blot analysis to clarify the expression of EMT-related proteins. The results are displayed in Figures 6(b) and 6(c). In HepG2 and SMMC-7721 cells, after the knockdown of SSB, the expression of N-cadherin, MMP-2, MMP-9, Vimentin, and Snail proteins was decreased compared with CON and sh-scramble groups, and the expression of E-cadherin and ZO-1 proteins was increased ($P < 0.05$).

### 3.8. High Expression of SSB May Affect the Effect of Immunotherapy. The expression of SSB was negatively correlated with macrophages and NK cells but significantly positively correlated with CD4+T cells ($P < 0.05$, Figure 7(a)). All eight immune checkpoint-related genes showed high expression levels in patients with high SSB expression ($P < 0.05$, Figure 7(b)). In TCGA-LIHC cohort, the high-expression SSB group had significantly higher TIDE scores than the low-expression group ($P < 0.05$, Figure 7(c)). As

(a)



(b)
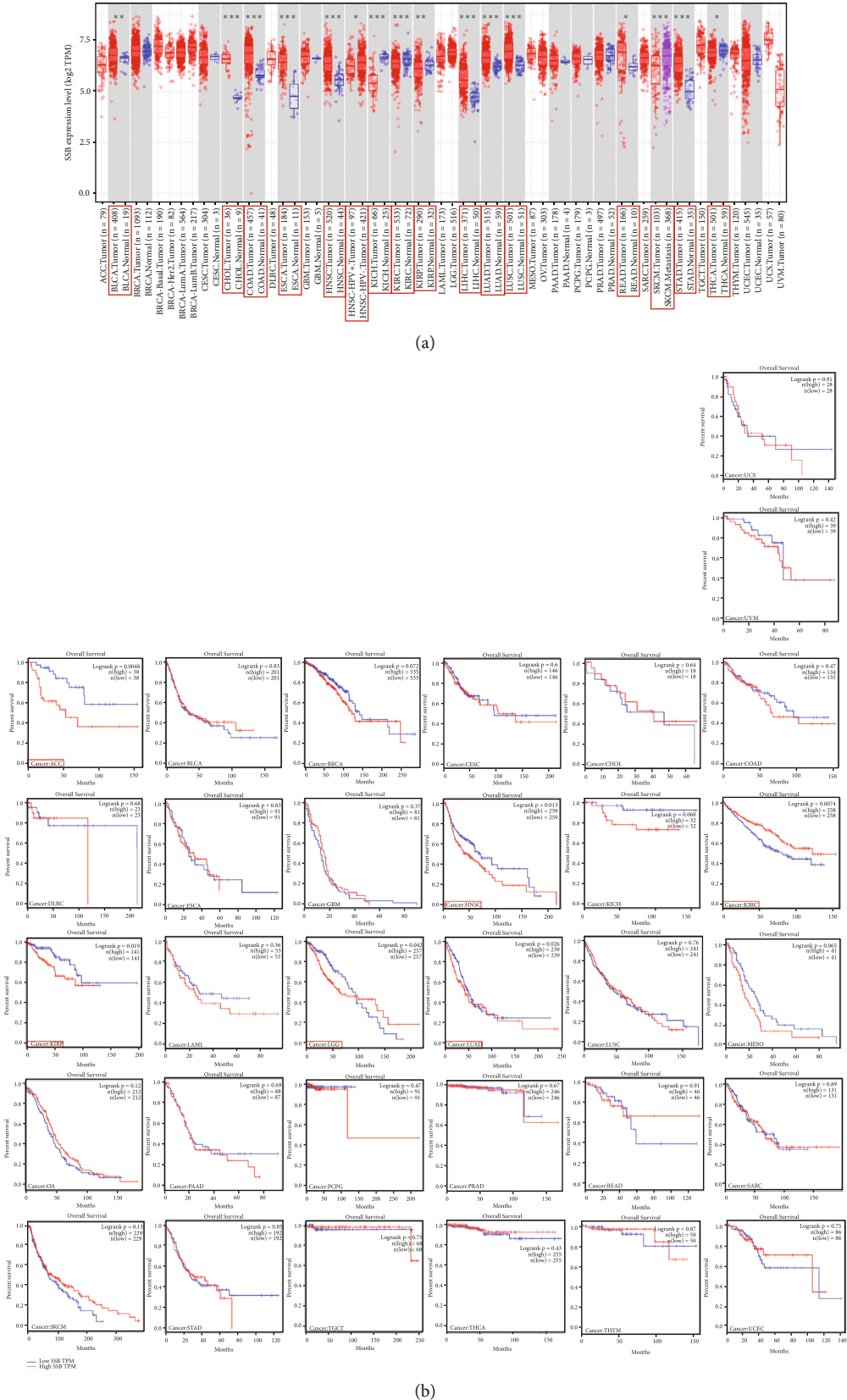
Figure 3: Expression and Kaplan-Meier survival analysis of SSB in pancancer. (a) SSB pancancer expression analysis. (b) Kaplan-Meier survival analysis of SSB in pancancer.

(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 4: Continued.

(g)



(h)
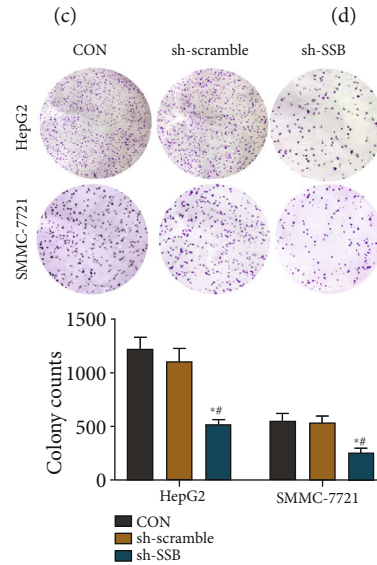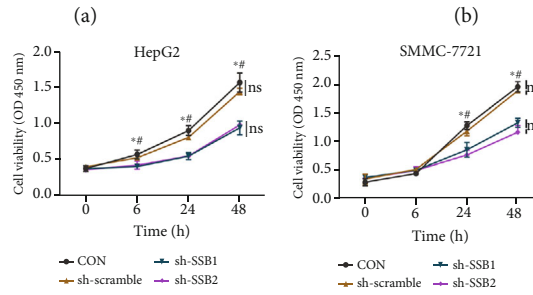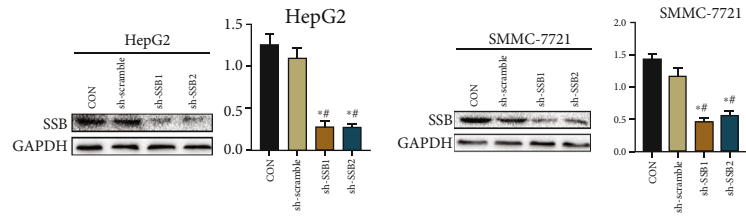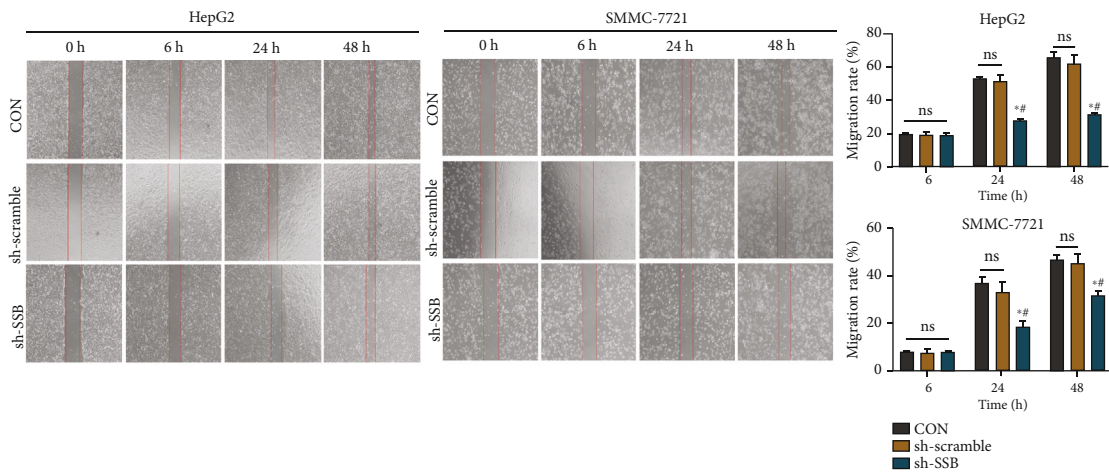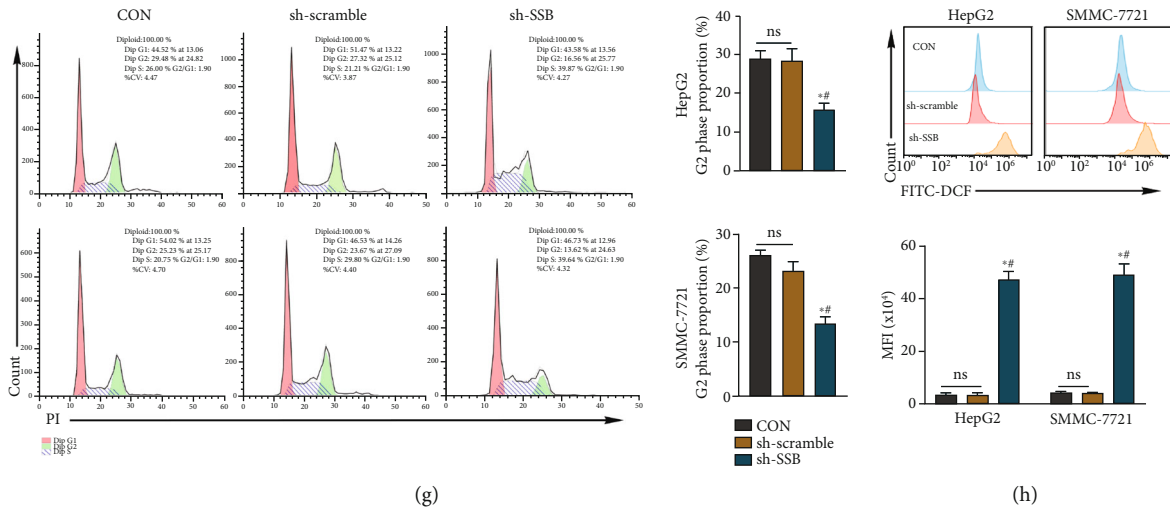
Figure 4: SSB promoted HepG2 and SMMC-7721 cell proliferation in vitro. (a, b) The knockdown efficiency of SSB in hepatoma cell lines HepG2 and SMMC-7721. (c, d) The cell viability of hepatoma cell lines SMMC-7721 and HepG2 decreased after SSB knockdown. (e) Cell clone formation assay showed that the proliferation of hepatoma cells decreased after the knockdown of SSB. (f) Cell scratch assay showed that hepatoma cell migration ability decreased after the knockdown of SSB. (g) The flow cytometry cell cycle results showed that the proportion of cells in the proliferative phase decreased after the knockdown of SSB. (h) The levels of ROS in hepatocellular carcinoma cell lines increased after SSB knockdown. Data are shown as mean ± SD. * refers to a statistically significant difference in the sh-SSB group compared to the CON group, $P < 0.05$. # refers to the statistically significant difference in the sh-SSB group compared to the sh-scramble group, $P < 0.05$. NS: no statistical difference.

judged by the TIDE score, patients with high expression of SSB may be insensitive to ICB treatment. In other words, patients with low expression of SSB may be more sensitive to ICB treatment.

## 4. Discussion

HCC remains a disease with a poor prognosis and high mortality due to its high late diagnosis rate, metastasis rate, and rapid malignant progression [22]. However, early diagnosis and effective treatment measures can significantly improve the survival rate of HCC patients. Therefore, it is vital to find new diagnostic markers and therapeutic targets for the early progression of HCC [23, 24]. In recent years, due to the rapid development of omics technology, we have gained a deeper understanding of the pathogenesis, diagnosis, and treatment of HCC [25].

In this study, we first analyzed the transcriptome data of HCC samples in TCGA database, concluding that SSB is highly expressed in HCC tissues and is an independent prognostic risk factor for HCC patients. Afterward, we verified the expression of SSB using GEO and HPA IHC databases. Furthermore, we used western blot and qRT-PCR assays to confirm the high expression of SSB in HCC. SSB expression and clinicopathological characteristics of HCC patients were analyzed by univariate and multivariate Cox analysis, suggesting that SSB was an independent prognostic risk factor. In addition, GSEA suggested that the high expression of SSB was associated with different tumor-related signaling pathways. The pancancer research project was initiated by TCGA in 2012, and this research relies on multiomics high-throughput database information mining to find the similarities and differences among tumors and

provide guidance for the subsequent diagnosis, prognosis, and other treatment plans of tumors [26]. We subsequently performed pancancer expression and survival analysis of SSB, and the results showed that SSB was upregulated in multiple cancers and affected patient prognosis, which was also consistent with previous studies [10–12]. Although the incidence of different cancers and the depth of research are inconsistent, through pancancer analysis, the mechanism, and effective drugs can be compared in-depth, and the same target can be used for different cancers. In this study, we mainly studied the prognostic significance and related mechanisms of SSB in patients with liver cancer and proved that the high expression of SSB is not conducive to the prognosis of patients. From the perspective of HCC treatment, specific antibodies can be designed to try to block the function of target molecules such as SSB, to achieve the possibility of treatment and prevention of migration and recurrence. Now, the development of tumor therapeutic drugs has crossed the boundaries of tumor tissue types, and the development of specific drugs for the target molecules of pancancer analysis can treat a variety of cancers of different tissue origins.

To further explore the biological function of SSB in HCC cells, we knocked down SSB in two HCC cell lines (HepG2 and SMMC-7721). The results showed that after the knockdown of SSB, the proliferation ability of HepG2 and SMMC-7721 HCC cell lines decreased, and the cell cycle results also showed that the proportion of cells in the proliferative phase decreased. Apoptosis is a form of programmed cell death that plays a crucial role in homeostasis, infection, injury, and clearance of senescent cells [27]. After the knockdown of SSB, the level of apoptosis increased, the expression levels of anti-apoptotic-related proteins decreased, and the expression levels of pro-apoptotic-related proteins increased in HCC cells. This
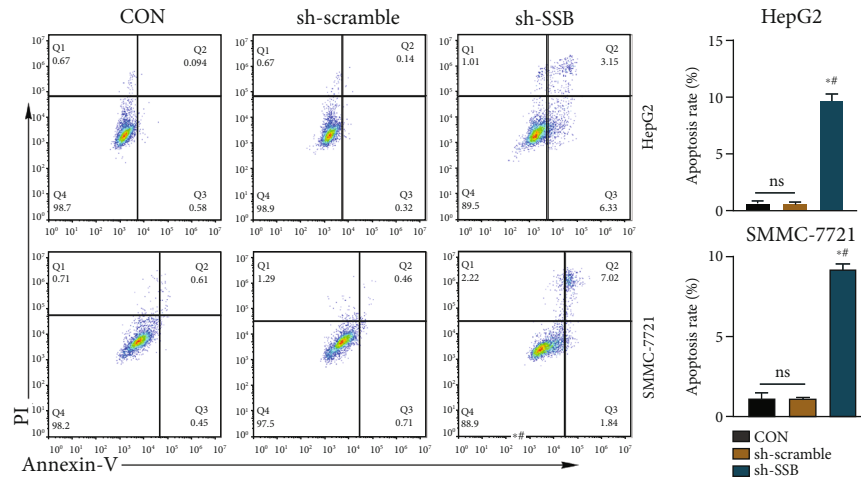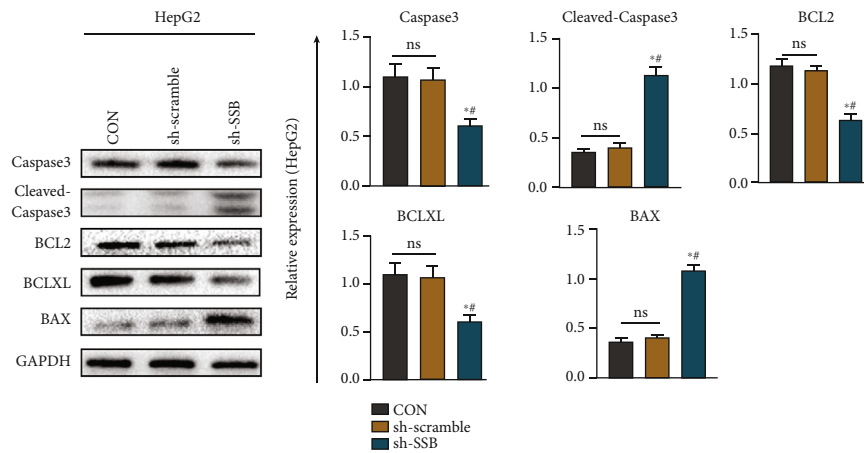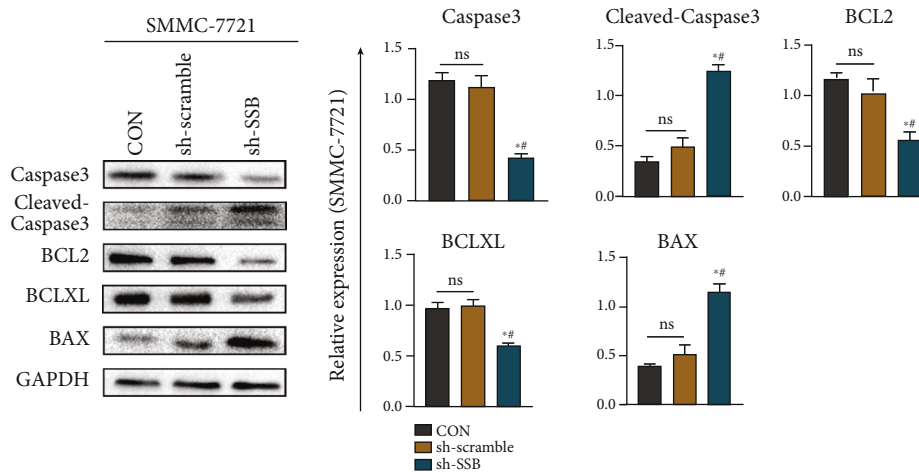
FIGURE 5: Knockdown of SSB can promote the apoptosis of HCC cells. (a) After the knockdown of SSB, the level of apoptosis in hepatoma cell lines HepG2 and SMMC-7721 increased. (b) Expression of apoptosis-related proteins in hepatoma cell line HepG2. (c) Expression of apoptosis-related proteins in hepatoma cell line SMMC-7721. Data are shown as mean ± SD. * refers to a statistically significant difference in the sh-SSB group compared to the CON group, $P < 0.05$. # refers to the statistically significant difference in the sh-SSB group compared to the sh-scramble group, $P < 0.05$. NS: no statistical difference.

(a)



(b)



(c)
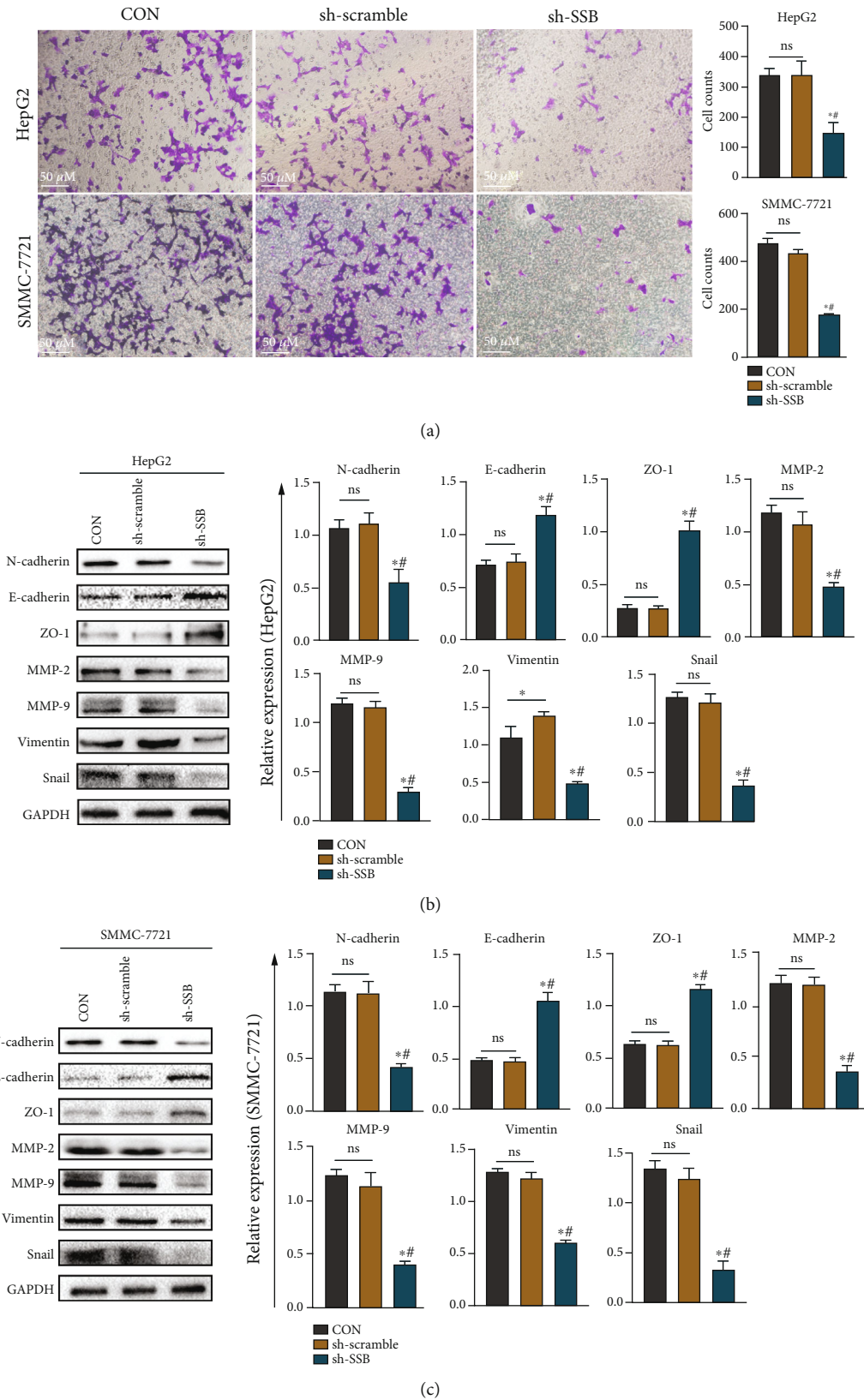
FIGURE 6: SSB can promote the invasion of hepatoma cells through the EMT pathway. (a) Cell invasion assay of HepG2 and SMMC-7721 cells. (b) EMT-related protein expression in the hepatoma cell line HepG2. (c) Expression of EMT-related proteins in hepatoma cell line SMMC-7721. Data are shown as mean ± SD. * refers to a statistically significant difference in the sh-SSB group compared to the CON group, $P < 0.05$. # refers to the statistically significant difference in the sh-SSB group compared to the sh-scramble group, $P < 0.05$. NS: no statistical difference.

(a)



(b)



(c)
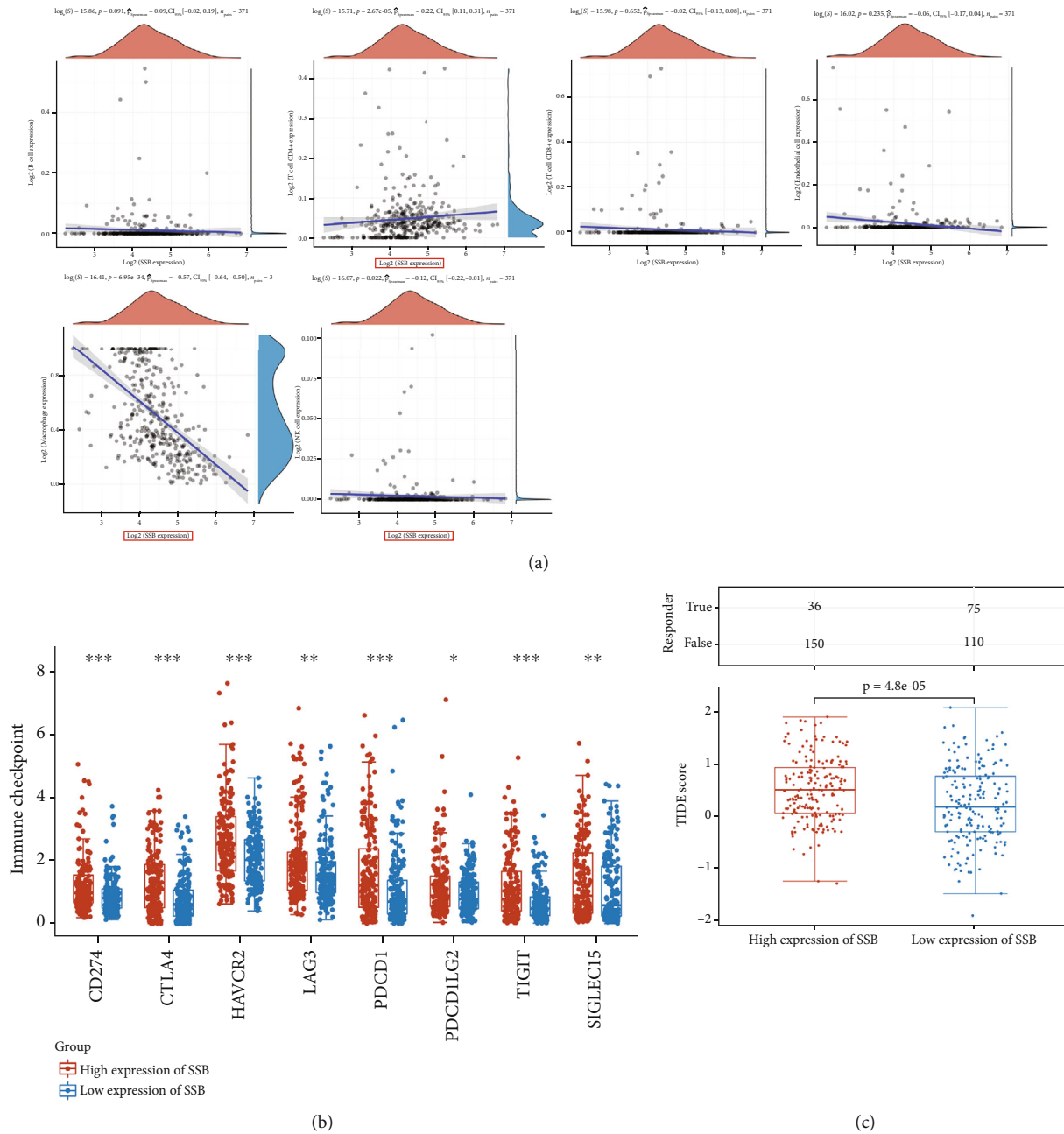
FIGURE 7: High expression of SSB may affect the effect of ICB therapy. (a) Spearman correlation analysis between SSB and immune cells. (b) Expression analysis of genes related to immune checkpoints and SSB. (c) Statistical table of immune response of samples in different groups in the prediction results and the TIDE scores in different groups in the prediction results. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

also indicates that the gene SSB may promote the proliferation of HCC cells through the antiapoptotic pathway.

Oxidative stress is associated with many physiological and pathological processes [28, 29]. Disruption of cellular oxidative stress homeostasis has been shown to be associated with the development of HCC [30, 31]. Related studies have shown that ROS play multiple roles in cancer [16, 32, 33]. On the one hand, ROS is crucial for cancer cell survival and tumor growth, and on the other hand, excess ROS can lead to cancer cell death [34]. Importantly, tumor cells utilize cellular antioxidant systems to counteract the prodeath effects of ROS. There is increasing evidence that proteins with antioxidant activity are involved in tumorigenesis and metastasis [35]. By conducting a series of experiments, we verified that modulation of SSB affects ROS levels and may thereby lead to the proliferation and metastasis of HCC.

Epithelial-mesenchymal transition (EMT) is the biological process of transforming epithelial cells into cells with a mesenchymal phenotype through a specific program. It plays a vital role in embryonic development, chronic inflammation, tissue remodeling, cancer metastasis, and various fibrotic diseases [36–38]. In addition, studies have shown that the malignant progression of many cancer types is likely to depend entirely on the activation of EMT in tumor cells

[39–41]. After EMT activation, E-cadherin and ZO-1 were inhibited, resulting in loss of epithelial cell morphology. The cells change to a spindle-shaped mesenchymal morphology and express markers associated with the mesenchymal cell state, such as N-cadherin, Vimentin, and fibronectin [42, 43]. Previous studies have shown that the gene SSB is associated with EMT, and the motility and invasion ability of squamous cell carcinoma cells are decreased after knockdown of SSB, and the expression of matrix metalloproteinase 2 (MMP-2) protein is significantly decreased [12, 17].

Tumor invasion and metastasis are a complex, continuous process involving multiple molecules, especially matrix metalloproteinases (MMPs) [44]. Degradation of basement membrane and extracellular matrix (ECM) by MMPs promotes tumor cell invasion and proliferation [45]. MMP-2 and MMP-9, among all MMP members, have been linked to tumor metastasis [46, 47]. Therefore, the invasive ability of tumor cells should be proportional to the expression of MMP-2 and MMP-9. Snail is an essential inducer of EMT and can strongly inhibit the expression of E-cadherin [48]. In addition, studies have revealed that Snail expression levels are elevated in metastatic lesions of ovarian cancer [49]. In the present study, we confirmed by transwell invasion assay that the invasive ability of HepG2 and SMMC-7721 cell lines was decreased after SSB knockdown. Afterward, a western blot was performed to detect the expression of EMT marker proteins. The results showed that after the knockdown of SSB, the expression of epithelial marker protein N-cadherin decreased, E-cadherin and other mesenchymal proteins increased, and MMP-2 and MMP-9 expression also decreased. This is also consistent with the results of previous studies. All the above results indicate that SSB can participate in the invasion and progression of HCC through the EMT pathway.

ICB therapy has revolutionized the treatment of cancer in humans, and it can significantly improve patient outcomes by reversing the immunosuppressive microenvironment by reducing the likelihood of tumor immune escape [50, 51]. TIDE uses a panel of gene expression signatures to assess two distinct tumor immune escape mechanisms, including tumor-infiltrating cytotoxic T lymphocyte (CTL) dysfunction and CTL rejection by immunosuppressive factors. The higher the TIDE prediction score, the higher the possibility of immune evasion, indicating that patients are less likely to benefit from ICB treatment [21]. To further evaluate the potential immune mechanism of SSB in LIHC, we analyzed the level of SSB-related immune infiltration. The results showed that the level of SSB expression was positively correlated with the infiltration level of CD4+ T cells in LIHC. In addition, the expression of SSB was negatively correlated with macrophages and NK cells. We investigated the association of SSB with immune checkpoints, including SIGLEC15, TIGIT, CTLA4, CD274, HAVCR2, LAG3, PDCD1, and PDCD1LG2, which are associated with ICB responses. PD-1, CTLA4, LAG3, and HAVCR2 are T cell depletion markers [52], and the T cell depletion is a major factor contributing to immune dysfunction in cancer patients. In this study, all these marker genes were positively correlated with SSB expression in LIHC. Since high expression of immune checkpoints is associated with T cell exhaus-

tion and poor prognosis, this also partially explains the cancer-promoting role of SSB. The TIDE score of the high-expressing SSB group was higher, which also suggested that it is not suitable for ICB treatment in the HCC patients with high SSB expression.

In conclusion, we confirmed that SSB is highly expressed in HCC tissues, which is a prooncogene and may be involved in the proliferation, invasion, and metastasis of HCC through antiapoptosis, changing the level of cellular oxidative stress, and EMT pathway. We have demonstrated for the first time that the gene SSB can affect the level of cellular oxidative stress and it can be used as a new target for the diagnosis and treatment of HCC patients. However, HCC patients with high SSB expression may be insensitive to ICB therapy.

## Data Availability

The databases were downloaded from TCGA database (https://www.portal.gdc.cancer.gov), GEO database (https://www.ncbi.nlm.nih.gov/geo/), TIMER2.0 (http://timer.cistrome.org/), GEPIA (http://gepia.cancer-pku.cn/index.html), and HPA (https://www.proteinatlas.org/). Further inquiries can be directed to the corresponding author.

## Conflicts of Interest

The authors have declared that they had no conflict of interest.

## Authors' Contributions

Hao Wu contributed to the experimental operation and data analysis. Hao Wu drafted the manuscript. Zhixin Zhang and Xinyu Han extracted the plasmid and embedded lentivirus. Sai Zhang, Jinrui Zhang, Pinsheng Han, Youcheng Zhang, Yi Bai, and Yamin Zhang participated in the revision of the manuscript. All authors read and approved the manuscript. Hao Wu and Zhixin Zhang have contributed equally to this work and share first authorship.

## Acknowledgments

## Supplementary Materials

Supplementary Table 1: the sequence information of the SSB primers. Supplementary Figure 1: hepatoma cell transwell migration assay. (*Supplementary Materials*)

## References

[1] 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.

[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. Villanueva, "Hepatocellular Carcinoma. Reply," *The New England Journal of Medicine*, vol. 381, no. 1, article e2, 2019.

[4] F. Brenet, N. D. Socci, N. Sonenberg, and E. C. Holland, "Akt phosphorylation of La regulates specific mRNA translation in glial progenitors," *Oncogene*, vol. 28, no. 1, pp. 128–139, 2009.

[5] V. Fok, K. Friend, and J. A. Steitz, "Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes nucleocytoplasmic shuttling," *The Journal of Cell Biology*, vol. 173, no. 3, pp. 319–325, 2006.

[6] B. Pereira, M. Billaud, and R. Almeida, "RNA-binding proteins in cancer: old players and new actors," *Trends Cancer*, vol. 3, no. 7, pp. 506–528, 2017.

[7] K. Iino, Y. Mitobe, K. Ikeda et al., "RNA-binding protein NONO promotes breast cancer proliferation by post- transcriptional regulation of SKP2 and E2F8," *Cancer Science*, vol. 111, no. 1, pp. 148–159, 2020.

[8] D. A. Lujan, J. L. Ochoa, and R. S. Hartley, "Cold-inducible RNA binding protein in cancer and inflammation," *RNA*, vol. 9, no. 2, 2018.

[9] C. Stavraka and S. Blagden, "The La-related proteins, a family with connections to cancer," *Biomolecules*, vol. 5, no. 4, pp. 2701–2722, 2015.

[10] R. Trotta, T. Vignudelli, O. Candini et al., "BCR/ABL activates mdm2 mRNA translation via the La antigen," *Cancer Cell*, vol. 3, no. 2, pp. 145–160, 2003.

[11] X. Huang and J. Tang, "Human La protein: an RNA-binding protein involved in ovarian cancer development and multidrug resistance," *Oncotargets and Therapy*, vol. 13, pp. 10721–10727, 2020.

[12] G. Sommer, C. Rossa, A. C. Chi, B. W. Neville, and T. Heise, "Implication of RNA-binding protein La in proliferation, migration and invasion of lymph node-metastasized hypopharyngeal SCC cells," *PLoS One*, vol. 6, no. 10, article e25402, 2011.

[13] G. Sommer, J. Dittmann, J. Kuehnert et al., "The RNA-binding protein La contributes to cell proliferation and CCND1 expression," *Oncogene*, vol. 30, no. 4, pp. 434–444, 2011.

[14] M. Petz, N. C. Them, H. Huber, and W. Mikulits, "PDGF enhances IRES-mediated translation of Laminin B1 by cytoplasmic accumulation of La during epithelial to mesenchymal transition," *Nucleic Acids Research*, vol. 40, no. 19, pp. 9738–9749, 2012.

[15] G. Sommer and T. Heise, "Role of the RNA-binding protein La in cancer pathobiology," *RNA Biology*, vol. 18, no. 2, pp. 218–236, 2021.

[16] S. Prasad, S. C. Gupta, and A. K. Tyagi, "Reactive oxygen species (ROS) and cancer: role of antioxidative nutraceuticals," *Cancer Letters*, vol. 387, pp. 95–105, 2017.

[17] T. Heise and G. Sommer, "RNA-binding protein La mediates TGFβ-induced epithelial to mesenchymal transition and cancer stem cell properties," *Cancers*, vol. 13, no. 2, p. 343, 2021.

[18] X. Ouyang, L. Lv, Y. Zhao et al., "ASF1B serves as a potential therapeutic target by influencing cell cycle and proliferation in hepatocellular carcinoma," *Frontiers in Oncology*, vol. 11, p. 801506, 2022.

[19] X. Zhu, P. Shao, Y. Tang, M. Shu, W. W. Hu, and Y. Zhang, "hsa_circRNA_100533 regulates GNAS by sponging hsa_miR_933 to prevent oral squamous cell carcinoma," *Journal of Cellular Biochemistry*, vol. 120, no. 11, pp. 19159–19171, 2019.

[20] Y. Bai, H. Wu, J. Zhang et al., "IL-22 protects against biliary ischemia-reperfusion injury after liver transplantation via activating STAT3 and reducing apoptosis and oxidative stress levels in vitro and in vivo," *Oxidative Medicine and Cellular Longevity*, vol. 2022, Article ID 9635075, 24 pages, 2022.

[21] P. Jiang, S. Gu, D. Pan et al., "Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response," *Nature Medicine*, vol. 24, no. 10, pp. 1550–1558, 2018.

[22] E. Vo Quang, Y. Shimakawa, and P. Nahon, "Epidemiological projections of viral-induced hepatocellular carcinoma in the perspective of WHO global hepatitis elimination," *Liver International*, vol. 41, no. 5, pp. 915–927, 2021.

[23] J. M. Llovet, R. Montal, D. Sia, and R. S. Finn, "Molecular therapies and precision medicine for hepatocellular carcinoma," *Nature Reviews. Clinical Oncology*, vol. 15, no. 10, pp. 599–616, 2018.

[24] Y. Pan, H. Chen, and J. Yu, "Biomarkers in hepatocellular carcinoma: current status and future perspectives," *Biomedicine*, vol. 8, no. 12, p. 576, 2020.

[25] R. Dhanasekaran, J. C. Nault, L. R. Roberts, and J. Zucman-Rossi, "Genomic medicine and implications for hepatocellular carcinoma prevention and therapy," *Gastroenterology*, vol. 156, no. 2, pp. 492–509, 2019.

[26] Consortium I T P-C A O W G, "Pan-cancer analysis of whole genomes," *Nature*, vol. 578, no. 7793, pp. 82–93, 2020.

[27] H. Flores-Romero, L. Hohorst, M. John et al., "BCL-2-family protein tBID can act as a BAX-like effector of apoptosis," *The EMBO Journal*, vol. 41, no. 2, article e108690, 2022.

[28] J. G. Farias, E. A. Herrera, C. Carrasco-Pozo et al., "Pharmacological models and approaches for pathophysiological conditions associated with hypoxia and oxidative stress," *Pharmacology & Therapeutics*, vol. 158, pp. 1–23, 2016.

[29] S. Datta, M. Cano, K. Ebrahimi, L. Wang, and J. T. Handa, "The impact of oxidative stress and inflammation on RPE degeneration in non- neovascular AMD," *Progress in Retinal and Eye Research*, vol. 60, pp. 201–218, 2017.

[30] B. Foglia and M. Parola, "Of FACT complex and oxidative stress response: a KEAP1/NRF2-dependent novel mechanism sustaining hepatocellular carcinoma progression," *Gut*, vol. 69, no. 2, pp. 195-196, 2020.

[31] D. Lee, I. M. Xu, D. K. Chiu et al., "Induction of oxidative stress through inhibition of thioredoxin reductase 1 is an effective therapeutic approach for hepatocellular carcinoma," *Hepatology*, vol. 69, no. 4, pp. 1768–1786, 2019.

[32] J. N. Moloney and T. G. Cotter, "ROS signalling in the biology of cancer," *Seminars in Cell & Developmental Biology*, vol. 80, pp. 50–64, 2018.

[33] Z. Liao, D. Chua, and N. S. Tan, "Reactive oxygen species: a volatile driver of field cancerization and metastasis," *Molecular Cancer*, vol. 18, no. 1, p. 65, 2019.

[34] C. Wiel, K. le Gal, M. X. Ibrahim et al., "BACH1 stabilization by antioxidants stimulates lung cancer metastasis," *Cell*, vol. 178, no. 2, pp. 330–345.e22, 2019.

[35] L. Lignitto, S. E. LeBoeuf, H. Homer et al., "Nrf2 activation promotes lung cancer metastasis by inhibiting the degradation of Bach1," *Cell*, vol. 178, no. 2, pp. 316–329.e18, 2019.

[36] M. A. Nieto, "Epithelial-mesenchymal transitions in development and disease: old views and new perspectives," *The International Journal of Developmental Biology*, vol. 53, no. 8-9-10, pp. 1541–1547, 2009.

[37] M. A. Nieto, R. Y. Huang, R. A. Jackson, and J. P. Thiery, "EMT: 2016," *Cell*, vol. 166, no. 1, pp. 21–45, 2016.

[38] J. P. Thiery, H. Acloque, R. Y. Huang, and M. A. Nieto, "Epithelial-mesenchymal transitions in development and disease," *Cell*, vol. 139, no. 5, pp. 871–890, 2009.

[39] X. Ye, W. L. Tam, T. Shibue et al., "Distinct EMT programs control normal mammary stem cells and tumour- initiating cells," *Nature*, vol. 525, no. 7568, pp. 256–260, 2015.

[40] A. D. Rhim, E. T. Mirek, N. M. Aiello et al., "EMT and dissemination precede pancreatic tumor formation," *Cell*, vol. 148, no. 1-2, pp. 349–361, 2012.

[41] A. M. Krebs, J. Mitschke, M. Lasierra Losada et al., "The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer," *Nature Cell Biology*, vol. 19, no. 5, pp. 518–529, 2017.

[42] A. Dongre and R. A. Weinberg, "New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer," *Nature Reviews. Molecular Cell Biology*, vol. 20, no. 2, pp. 69–84, 2019.

[43] R. Kalluri and R. A. Weinberg, "The basics of epithelial-mesenchymal transition," *The Journal of Clinical Investigation*, vol. 119, no. 6, pp. 1420–1428, 2009.

[44] L. Yadav, N. Puri, V. Rastogi, P. Satpute, R. Ahmad, and G. Kaur, "Matrix metalloproteinases and cancer - roles in threat and therapy," *Asian Pacific Journal of Cancer Prevention*, vol. 15, no. 3, pp. 1085–1091, 2014.

[45] K. C. Nannuru, M. Futakuchi, M. L. Varney, T. M. Vincent, E. G. Marcusson, and R. K. Singh, "Matrix metalloproteinase (MMP)-13 regulates mammary tumor-induced osteolysis by activating MMP9 and transforming growth factor-$\beta$ signaling at the tumor-bone interface," *Cancer Research*, vol. 70, no. 9, pp. 3494–3504, 2010.

[46] M. Egeblad and Z. Werb, "New functions for the matrix metalloproteinases in cancer progression," *Nature Reviews. Cancer*, vol. 2, no. 3, pp. 161–174, 2002.

[47] R. L. Scherer, J. O. Mcintyre, and L. M. Matrisian, "Imaging matrix metalloproteinases in cancer," *Cancer Metastasis Reviews*, vol. 27, no. 4, pp. 679–690, 2008.

[48] A. Barrallo-Gimeno and M. A. Nieto, "The Snail genes as inducers of cell movement and survival: implications in development and cancer," *Development*, vol. 132, no. 14, pp. 3151–3161, 2005.

[49] H. Jin, Y. Yu, T. Zhang et al., "Snail is critical for tumor growth and metastasis of ovarian carcinoma," *International Journal of Cancer*, vol. 126, no. 9, pp. 2102–2111, 2010.

[50] E. B. Garon, N. A. Rizvi, R. Hui et al., "Pembrolizumab for the treatment of non-small-cell lung cancer," *The New England Journal of Medicine*, vol. 372, no. 21, pp. 2018–2028, 2015.

[51] 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.

[52] D. Li, T. Yu, J. Han et al., "Prognostic value and immunological role of KIFC1 in hepatocellular carcinoma," *Frontiers in Molecular Biosciences*, vol. 8, p. 799651, 2022.

*Research Article*

# TRIM6 Reduces Ferroptosis and Chemosensitivity by Targeting SLC1A5 in Lung Cancer

**Ying Zhang,[1] Ping Dong,[2] Nian Liu,[3] Jun-Yuan Yang,[4] Hui-Min Wang,[5] and Qing Geng [iD][2]**

[1]Department of Vascular Surgery, Renmin Hospital of Wuhan University, Wuhan, 430060 Hubei, China
[2]Department of Thoracic Surgery, Renmin Hospital of Wuhan University, Wuhan, 430060 Hubei, China
[3]Department of Neonatology, Renmin Hospital of Wuhan University, Wuhan, 430060 Hubei, China
[4]Department of Gynecologic Oncology, Zhongnan Hospital of Wuhan University, Wuhan, 430071 Hubei, China
[5]Department of Fever Clinic, Renmin Hospital of Wuhan University, Wuhan, 430060 Hubei, China

Correspondence should be addressed to Qing Geng; gengqingwhu@whu.edu.cn

*Objective*. Ferroptosis, a newly identified form of cell death, plays critical roles in the development and chemoresistance of lung cancer. Tripartite motif 6 (TRIM6) acts as an E3-ubiquitin ligase and can promote the progression of human colorectal cancer. The present study is aimed at investigating its role and potential mechanisms in lung cancer. *Methods*. Lentiviral vectors were used to overexpress or knock down TRIM6 in human lung cancer cells. Cell survival, colony formation, lipid peroxidation, intracellular iron levels, and other ferroptotic markers were examined. The role of TRIM6 on ferroptosis and chemosensitivity was further tested in mouse tumor xenograft models. *Results*. TRIM6 was highly expressed in human lung cancer tissues and cells, and its expression in the lung cancer cells was further increased by ferroptotic stimulation. TRIM6 overexpression inhibited, while TRIM6 silence promoted erastin- and RSL3-induced glutaminolysis and ferroptosis in the lung cancer cells. Mechanistically, TRIM6 directly interacted with solute carrier family 1 member 5 to promote its ubiquitination and degradation, thereby inhibiting glutamine import, glutaminolysis, lipid peroxidation, and ferroptotic cell death. Moreover, we observed that TRIM6 overexpression reduced the chemotherapeutic effects of cisplatin and paclitaxel. In contrast, TRIM6 silence sensitized human lung cancer cells to cisplatin and paclitaxel in vivo and in vitro. *Conclusion*. Our findings for the first time define TRIM6 as a negative regulator of ferroptosis in the lung cancer cells, and TRIM6 overexpression enhances the resistance of human lung cancer cells to chemotherapeutic drugs. Overall, targeting TRIM6 may help to establish novel strategies to treat lung cancer.

## 1. Introduction

Lung cancer is the leading cause of cancer mortality worldwide, and most patients are diagnosed at the advanced stages, with very poor prognosis [1–5]. Cell death plays an important role in regulating tumor growth, progression, and chemotherapeutic response. Ferroptosis is a newly discovered nonapoptotic death mode that involves the accumulation of lipid reactive oxygen species (ROS) and subsequent depletion of plasma membrane polyunsaturated fatty acids [6–10]. Glutathione (GSH) and the associated glutathione peroxidase 4 (GPX4) are intracellular antioxidant defenses to scavenge the toxic lipid ROS [11]. In contrast, iron donates

electrons to oxygen to accelerate lipid ROS formation and ferroptosis [12]. Accordingly, lipophilic or membrane impermeable iron chelators notably prevent lethal lipid peroxidation and ferroptosis [11, 13]. L-Glutamine (Gln) is a major nitrogen source for the synthesis of amino acids, nucleotides, and lipids and also provides carbon source for the tricarboxylic acid (TCA) cycle and cellular energetics, which is required for the growth of cancer cells [14]. Yet, recent findings have found that glutaminolysis promotes productions of oxidizable lipids via the TCA cycle and eventually facilitates ferroptosis [15, 16]. Gln is imported inside the cells by solute carrier family 1 member 5 (SLC1A5) and SLC38A1, converted into glutamate (Glu) by glutaminases (GLS), and then

metabolized into alpha-ketoglutarate ($\alpha$-KG) by either glutamate dehydrogenase- (GLUD1-) mediated glutamate deamination or glutamic-oxaloacetic transaminase 1 (GOT1-) mediated transamination [15]. And supplementing $\alpha$-KG can fuel both energetic and anabolic pathways, mimicking Gln-mediated ferroptotic induction. Therefore, targeting ferroptosis may develop novel therapeutic approaches to treat lung cancer.

Ubiquitination acts as a pivotal posttranslational modification for various proteins. During ubiquitination, polyubiquitin (Ub) chains are attached to the targeted proteins by E1 Ub-activating enzymes, E2 Ub-conjugating enzymes, and E3 Ub-ligases, which then mediate the proteasomal degradation of these proteins [17–20]. Tripartite motif (TRIM) proteins are a family of E3 Ub-ligases and implicated in carcinogenesis and chemoresistance of diverse cancers [21–23]. TRIM6, a member of TRIM proteins, plays critical roles in regulating interferon signaling and antiviral responses [24, 25]. Results from Zeng et al. demonstrated that TRIM6 aggravated cardiomyocyte apoptosis and myocardial ischemia/reperfusion injury [26]. TRIM6 could also interact with protooncogenic Myc to maintain the pluripotency of mouse embryonic stem cells [27]. Moreover, Zheng et al. recently observed that TRIM6 was upregulated in human colorectal cancer (CRC) samples and that TRIM6 overexpression promoted proliferation and chemoresistance of CRC cells [28]. These findings identify TRIM6 as a promising therapeutic target of lung cancer.

## 2. Materials and Methods

### 2.1. Antibodies and Chemicals.
Anti-TRIM6 (#11953-1-AP) and anti-glyceraldehyde-3-phosphate dehydrogenase (GAPDH, #10494-1-AP) were purchased from Proteintech (Chicago, IL, USA). Anti-GPX4 (#ab125066), anti-SLC7A11 (xCT, #ab37185), anti-SLC3A2 (CD98, #ab108300), anti-glutathione synthetase (GSS, #ab124811), anti-transferrin (Tf, #ab109503), anti-Tf receptor (TfR, #ab84036), anti-nuclear factor E2-related factor 2 (NRF2, #ab137550), anti-SLC1A5 (#ab237704), and anti-SLC38A1 (#ab60145) were purchased from Abcam (Cambridge, UK), while anti-ferroportin (FPN, #NBP1-21502) was obtained from Novus Biologicals (Littleton, Colorado, USA). Erastin (#S7242), RSL3 (#S8155), ferrostain-1 (Fer-1, #S7243), and liproxstatin-1 (Lip-1, #S7699) were obtained from Selleck Chemicals (Houston, TX, USA). $2',7'$-dichlorofluorescin diacetate (DCFH-DA, #D6883), superoxide anion assay kit (#CS1000), GSH assay kits (#CS0206), $\alpha$-KG (#349631), L-$\gamma$-glutamyl transpeptidase substrate (SLC1A5 inhibitor; GPNA, #G1135), compound 968 (GLS inhibitor; 968, #352010), bis-2-(5-phenylacetamido-1,3,4-thiadiazol-2-yl) ethyl sulfide (GLS1 inhibitor; BPTES, #SML0601), amino oxyacetate (pan-transaminase inhibitor; AOA, #C13408), cycloheximide (protein synthesis inhibitor; CHX, #01810), MG132 (proteasome inhibitor, #M7449), cisplatin (DDP, #P4394), and paclitaxel (PTX, #1491332) were purchased from Sigma-Aldrich (St. Louis, MO, USA). BODIPY™ 581/591 C11 (BODIPY, #D3861) and tetramethylrhodamine ethyl ester (TMRE, #T669) were obtained from Invitrogen (Carlsbad, CA, USA). Malondialdehyde (MDA) assay kits

(#ab118970) were purchased from Abcam (Cambridge, UK), while CellTiter 96® AQ$_{ueous}$ One Solution Cell Proliferation Assay kit (MTS assay, #G3582) was obtained from Promega (Madison, WI, USA). Lentivirus carrying the short hairpin RNA sequences against human TRIM6 (TRIM6-KD #1 and TRIM6-KD #2), human SLC1A5 (SLC1A5-KD), or the control sequence (CTRL-KD) were generated by Gene Pharma Corporation (Shanghai, China). For TRIM6 overexpression, human TRIM6 cDNA (TRIM6-OE), human SLC1A5 cDNA (SLC1A5-OE), or a negative control (CTRL-OE) sequence was cloned into the lentiviral vectors by Gene Pharma Corporation (Shanghai, China).

### 2.2. Cell Culture.
Human lung cancer cell lines A549, H358, H460, H1299, PC9, and SPC-A-1 and normal human lung epithelial cell BEAS-2B were purchased from American Type Culture Collection and cultured in DMEM medium supplemented with 10% fetal bovine serum (FBS) and 1% antibiotics at 37°C under the humidified atmosphere [29–31]. The cells were preinfected for 12 h with lentiviral vectors carrying two different interfering sequences against TRIM6 at a multiplicity of infection (MOI) of 50 to silence endogenous TRIM6 or with TRIM6-OE virus (MOI = 20) to overexpress TRIM6. And then, the cells were maintained in fresh medium containing 10% FBS for an additional 24 h before further treatment. To induce ferroptosis, the cells were incubated with erastin (5 $\mu$mol/L) or RSL3 (2 $\mu$mol/L) for 24 h after TRIM6 genetic manipulation except special annotation [32]. For ferroptosis suppression, Fer-1 (1 $\mu$mol/L) or Lip-1 (0.2 $\mu$mol/L) was added at 8 h before erastin or RSL3 treatment [33]. In addition, the cells were treated with GPNA (5 mmol/L), 968 (20 $\mu$mol/L), BPTES (10 $\mu$mol/L), or AOA (0.5 mmol/L) at 8 h before erastin or RSL3 stimulation to inhibit Gln uptake or metabolism in the presence or absence of $\alpha$-KG (4 mmol/L) [15, 16]. For SLC1A5 overexpression or silence, the cells were preinfected with SLC1A5-OE (MOI = 20) or SLC1A5-KD (MOI = 50) for 12 h before TRIM6 genetic manipulation. In a separated study, the cells were infected with TRIM6-OE (MOI = 20) or CTRL-OE for 12 h and then cultured in normal medium for an addition 24 h, followed by a stimulation with CHX (20 mmol/L) for indicating times [28]. To clarify the role of TRIM6 on chemosensitivity in human lung cancer cells, the cells were treated with DDP (20 $\mu$mol/L) or PTX (0.3 $\mu$mol/L) for 12 h after TRIM6 genetic manipulation [34].

### 2.3. Cell Survival Assay.
Cell survival was determined using the CellTiter 96® AQ$_{ueous}$ One Solution Cell Proliferation Assay kit (MTS assay) [32]. Briefly, the cells (approximately 200 cells in 96-well plates) were incubated with CellTiter 96® AQ$_{ueous}$ One Solution Reagent (20 $\mu$L per 100 $\mu$L medium) at 37°C for 2 h under the humidified atmosphere, and then, the absorbance was recorded at 490 nm using a 96-well plate reader.

### 2.4. Colony Formation Assay.
For colony formation assay, the cells were seeded into the 6-well plates and incubated for 14 days with the colonies stained by 0.1% crystal violet.

Next, the colonies were carefully rinsed with tap water and dried at room temperature, and the colonies with a diameter more than 0.05 mm were counted by ImageJ software in a blinded manner [32, 35, 36].

2.5. Measurements of Intracellular ROS and Lipid Peroxidation. Intracellular ROS production was measured using the nonfluorescent DCFH-DA reagent that could be converted to the fluorescent DCF by free radicals [37–39]. In brief, the cells were homogenized in the assay buffer and then incubated with DCFH-DA (10 $\mu$mol/L) at 37°C for 30 min. The fluorescent intensity was examined using a spectrofluorometer with an excitation/emission wavelength at 488/525 nm. To detect lipid ROS level, the cells were incubated with BODIPY (10 $\mu$mol/L) at 37°C for 30 min and the fluorescent intensity was recorded by the simultaneous acquisition of green signals (484/510 nm) and red signals (581/610 nm) using the BD FACSAria cytometer [32]. Intracellular MDA content was assessed using the commercial kit following the manufacturer's instructions, and the absorbance was measured at 532 nm [32, 40].

2.6. Evaluations of GSH Level and GPX4 Activity. Intracellular GSH level was evaluated with a commercial kit according to the manufacturer's protocols and assayed colorimetrically at 412 nm. Relative GPX4 activity was determined using the HPTLC method according to previous studies [11, 41]. In brief, the cells were lysed in the reaction buffer, and the supernatants were collected to incubate with 7$\alpha$ cholesterol hydroperoxide (100 $\mu$mol/L) at 37°C. Next, the peroxides were extracted for HPTLC analysis, and analytes were scanned and quantified using ImageJ software.

2.7. Detections of Superoxide Anion Generation and Mitochondrial Membrane Potential (MMP). Superoxide anion generation was assessed with a superoxide anion assay kit via referring to the standard protocols. Briefly, the cells were incubated with luminol solution (5 $\mu$L) and enhancer solution (5 $\mu$L) at 37°C for 15 min, and then, the luminescence intensity was immediately measured. MMP was measured by incubating the cells with TMRE (200 nmol/L) at 37°C for 20 min, and the fluorescence intensity of TMRE was determined at 582 nm [42].

2.8. Iron Assay. Labile iron pool (LIP) was measured by the calcein-acetoxymethyl ester method [43]. In brief, intracellular LIP was loaded with calcein (2 $\mu$mol/L) at 37°C for 30 min, and then, the calcein was removed from iron by deferoxamine (100 $\mu$mol/L). The changes of fluorescence intensity with or without deferoxamine incubation at an excitation/emission wavelength of 485/535 nm were quantified as the amount of LIP. Ferrous iron ($Fe^{2+}$) levels were quantified at 593 nm using a commercial kit.

2.9. Protein Extraction, Immunoblots (IB), and Immunoprecipitation (IP). Cells were lysed in the RIPA lysis buffer containing protease/phosphatase inhibitor cocktail at 4°C, and total protein concentrations were determined by the bicinchoninic acid kit [44–46]. Then, equal amounts of proteins were separated by sodium dodecyl sulfate/poly-acrylamide gels (SDS-PAGE) and electrotransferred to the polyvinylidene difluoride membranes, followed by an incubation with 5% nonfat dried milk to block nonspecific binding. Next, the membranes were incubated with primary antibodies at 4°C overnight and stained by the secondary antibodies at room temperature for an additional 1 h. After that, protein bands were visualized with an ECL reagent and analyzed using the ImageJ software. For IP assay, cells were lysed in IP lysis buffer, and then, the lysates were incubated with indicating primary antibodies or IgG at 4°C overnight with gentle shaking, followed by the incubation with Protein A/G-agarose beads at room temperature for an additional 2 h. The immunoprecipitated proteins were subsequently washed for 5 times using the lysis buffer and boiled before SDS-PAGE electrophoresis.

2.10. RNA Purification and Quantitative Real-Time PCR. Total RNA was extracted using TRIzol reagent and then converted to cDNA using oligo (dT) primers. Quantitative real-time PCR was performed using QuantiNova SYBR Green PCR Kit (Qiagen; Hamburg, Germany) and normalized to GAPDH gene expression [47, 48].

2.11. Gln Uptake Assay. Gln uptake assay was performed using the [$^3$H]-L-Gln according to a previous study [15]. In brief, the cells were incubated with [$^3$H]-L-Gln (200 nmol/L) in Gln-free medium at 37°C for 15 min, which were then harvested for Gln measurements using a liquid scintillation counter.

2.12. Ubiquitination Assay In Vivo and In Vitro. For the in vivo ubiquitination assay, HEK293T cells were transfected with indicating plasmids for 48 h, and then, the cells were harvested in lysis buffer. Next, the samples were incubated with HA beads at 4°C for 2 h and then subjected to IP assay. For the in vitro ubiquitination assay, purified HA-SLC1A5 proteins were incubated with E1, E2 enzymes and human recombinant Ub with or without Flag-TRIM6 proteins in ubiquitination reaction buffer (Boston Biochem) at 30°C for 90 min, and then, the samples were prepared for IP assay [49, 50].

2.13. Mouse Xenograft Tumor Model. All animal experiments were approved by the Animal Ethics Committee of Renmin Hospital of Wuhan University and also complied with the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines. $5 \times 10^6$ TRIM6-manipulated H460 or PC9 cells were subcutaneously inoculated into the right flank of athymic BALB/c nude mice (4-5 weeks old), and the tumor parameters were calculated 4 weeks after cell inoculation [15]. To validate the role of TRIM6 on chemosensitivity, tumor-bearing mice received intraperitoneal injections of DDP (5 mg/kg) or PTX (15 mg/kg) for 3 times every other day at the last week before study termination [51].

2.14. Human Tissue Samples. Lung adenocarcinoma (ADC), squamous cell cancer (SCC), and corresponding adjacent normal tissues (ANT) were obtained from the patients without neoadjuvant or adjuvant therapies after written informed consent signed. ANT was obtained from the same patients and was at least 3 cm away from the tumor tissue. This study was approved by the Institutional

Review Board of Renmin Hospital of Wuhan University and conformed to the principles outlined in the *Declaration of Helsinki*.

*2.15. Statistical Analysis.* All data are reported as the mean $\pm$ SD, and $P < 0.05$ was considered statistically significant. Differences between two groups were compared using Student's two-tailed $t$-test, while one-way ANOVA followed by the Tukey post hoc test was applied for comparison of multiple groups. All statistical analyses were performed using SPSS 19.0 software in a blinded manner.

# 3. Results

*3.1. TRIM6 Expression in the Lung Cancer Samples Is Increased upon Ferroptotic Stimulation.* We first compared TRIM6 expression in human lung cancer tissues and corresponding ANT. As shown in Figures 1(a) and 1(b), human lung ADC and SCC tissues exhibited higher TRIM6 expression. Besides, *TRIM6* mRNA levels were also increased in serials of the lung cancer cell lines (A549, H358, H460, H1299, PC9, and SPC-A-1) in comparison with the normal human lung epithelial cell BEAS-2B (Figure 1(c)). Besides, we found that *TRIM6* mRNA expressions in H460 and PC9 cells were higher than those in other cancer cell lines; therefore, we selected these two cell lines in our further study (Figure 1(c)). Consistent with the mRNA levels, increased TRIM6 protein expressions were also detected in H460 and PC9 cells compared with BEAS-2B cell (Figure 1(d)). Next, we explored whether TRIM6 expression in the lung cancer cells was altered upon ferroptotic stimulation. As shown in Figure 1(e), *TRIM6* mRNA levels in H460 and PC9 cells were increased in the initial phase after erastin or RSL3 treatment, but fell and even decreased at the later stages. Therefore, all cells were incubated with erastin or RSL3 for 24 h except special annotation in our further experiments. At this time, both of the two cell lines had increased TRIM6 expression and also received sufficient intensities of ferroptotic stimulation (Figure 1(e)). Results from IB further confirmed that TRIM6 expression in the lung cancer cells was increased upon ferroptotic stimulation (Figures 1(f) and 1(g)). Collectively, these data demonstrate a potential involvement of TRIM6 in ferroptosis of the lung cancer cells.

*3.2. TRIM6 Overexpression Inhibits Erastin- and RSL3-Induced Ferroptosis in the Lung Cancer Cells.* We then overexpressed TRIM6 in H460 and PC9 cells using lentiviral vectors, and the efficiency was confirmed in Figure 2(a). Interestingly, TRIM6 overexpression significantly enhanced the survival and colony formation of the lung cancer cells upon ferroptotic stimulation (Figures 2(b) and 2(c)). Lipid peroxidation is an important feature of ferroptosis [13]. As expected, erastin and RSL3 treatment provoked significant increases of cellular and lipid ROS production, which were inhibited in TRIM6-overexpressed cells (Figures 2(d) and 2(e)). The levels of intracellular superoxide anion and MDA generation were also decreased by TRIM6 overexpression (Figures 2(f) and 2(g)). Consistent with previous studies, the lung cancer cells with erastin and RSL3 treatment

exhibited higher MMP levels that were inhibited by TRIM6 overexpression (Figure 2(h)) [52]. GSH and GPX4 are essential for reducing lipid hydroperoxides to lipid alcohols, thereby preventing lipid peroxidation and ferroptotic cell death [11]. As shown in Figures 2(i) and 2(j), the cells with erastin or RSL3 stimulation exhibited lower levels of GSH and GPX4 activities, which were preserved by TRIM6 overexpression. However, TRIM6 overexpression did not affect GPX4 protein abundances upon ferroptotic stimulation (Figure 2(k)). Iron, especially LIP and $Fe^{2+}$, is essential for the execution of ferroptosis, and we thus evaluated the effect of TRIM6 on intracellular LIP and $Fe^{2+}$ levels [13, 43]. We observed that TRIM6 overexpression slightly but significantly reduced iron accumulation following the treatment with erastin or RSL3 (Figures 2(l) and 2(m)). These findings suggest that TRIM6 overexpression inhibits erastin- and RSL3-induced ferroptosis in the lung cancer cells.

*3.3. TRIM6 Silence Promotes Erastin- and RSL3-Induced Ferroptosis in the Lung Cancer Cells.* Next, we used two lentiviral vectors to knock down endogenous TRIM6 expression, and the efficiency was confirmed in Figure 3(a). As expected, TRIM6 silence further decreased the survival and colony formation of the lung cancer cells upon erastin and RSL3 treatment (Figures 3(b) and 3(c)). Lipid ROS level and MDA generation were also augmented in TRIM6-deficient cells (Figures 3(d) and 3(e)). GSH depletions in H460 and PC9 cells by erastin or RSL3 incubation were more obvious after TRIM6 knockdown (Figure 3(f)). Intracellular LIP and $Fe^{2+}$ levels were increased in the lung cancer cells by ferroptotic stimulation, which were further enhanced in those with TRIM6 silence (Figures 3(g) and 3(h)). However, TRIM6 knockdown-associated cell death could be remarkably suppressed by ferroptosis inhibitors, Fer-1 and Lip-1, indicating an involvement of ferroptosis (Figure 3(h)). These data imply that TRIM6 silence promotes erastin- and RSL3-induced ferroptosis in the lung cancer cells.

*3.4. TRIM6 Modulates Ferroptosis via Affecting SLC1A5-Mediated Glutaminolysis.* We then examined the possible molecular basis underlying TRIM6-mediated ferroptotic actions. Unexpectedly, TRIM6 silence did not affect the molecules essential for Glu uptake, GSH synthesis, and iron transport (Figure S1A). In view of the unchangeable GPX4 proteins and slight alterations of the iron accumulation, we speculated that GSH/GPX4-mediated antioxidant defenses and iron overload might not be the primary mechanisms for TRIM6-mediated ferroptotic actions. NRF2 is a major redox-dependent transcription factor and negatively regulates ferroptosis [53]. However, TRIM6 knockdown also unaffected NRF2 expression and its transcription activity, as confirmed by expressions of the downstream heme oxygenase 1 (*HMOX1*), NAD(P)H quinone dehydrogenase 1 (*NQO1*), and glutamate-cysteine ligase modifier subunit (*GCLM*) (Figures S1A and S1B). Gln provides nutrition for the growth of cancer cells; however, recent studies have reported that glutaminolysis is linked to ferroptosis of the cancer cells via inducing the accumulation of lipid ROS
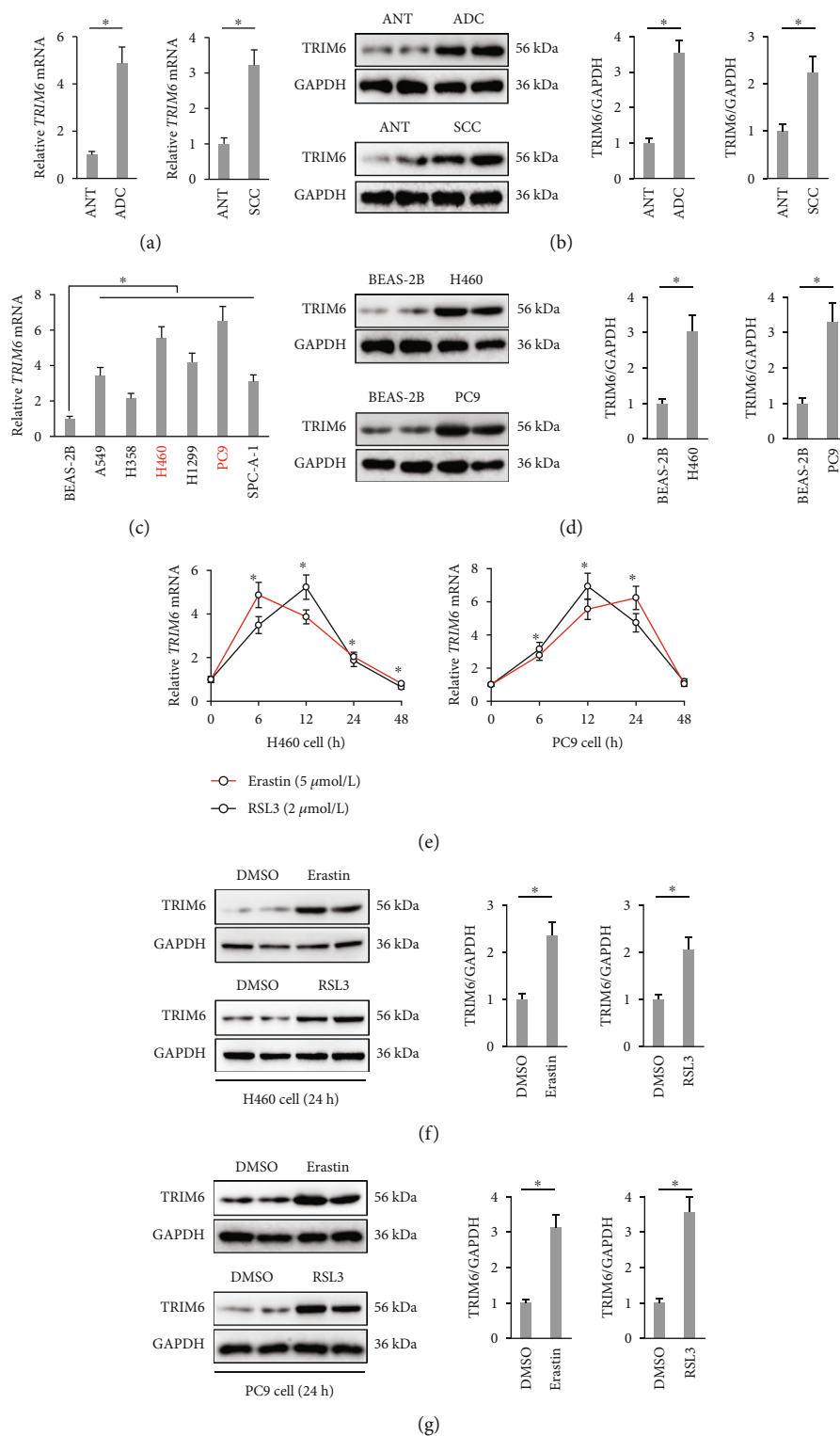
FIGURE 1: TRIM6 expression in lung cancer samples is increased upon ferroptotic stimulation. (a) Relative *TRIM6* mRNA levels in ADC, SCC, and corresponding ANT ($n = 10$). (b) Protein levels of TRIM6 in ADC, SCC, and corresponding ANT ($n = 6$). (c) Relative *TRIM6* mRNA levels in human lung cancer cell lines and normal epithelial cell ($n = 6$). (d) Protein levels of TRIM6 in H460, PC9, and BEAS-2B cells ($n = 6$). (e) Relative *TRIM6* mRNA levels in erastin- or RSL3-treated human lung cancer cells ($n = 6$). (f, g) Protein levels of TRIM6 in erastin- or RSL3-treated human lung cancer cells ($n = 6$). All data are reported as the mean $\pm$ SD, $^*P < 0.05$ versus corresponding groups.
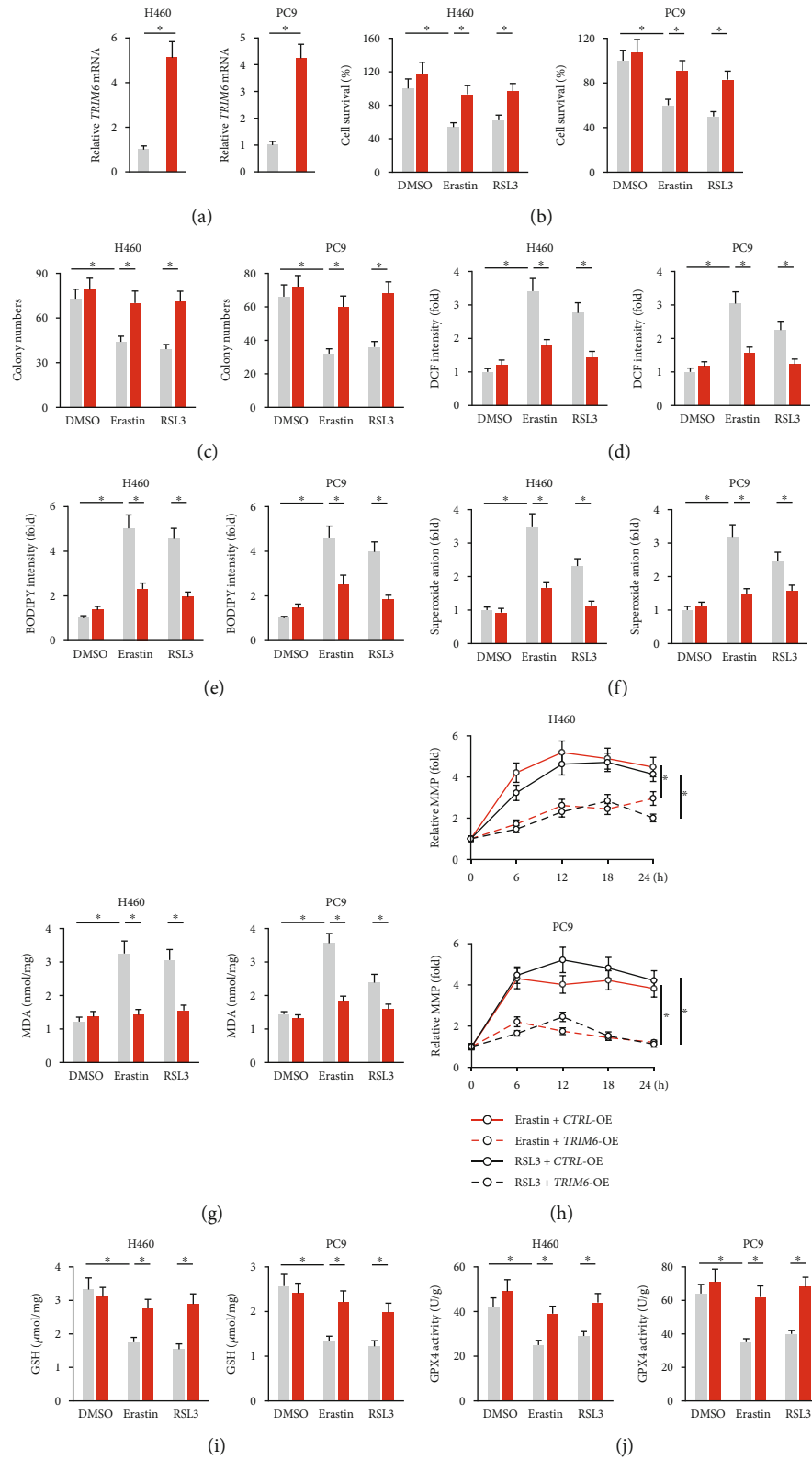
FIGURE 2: Continued.
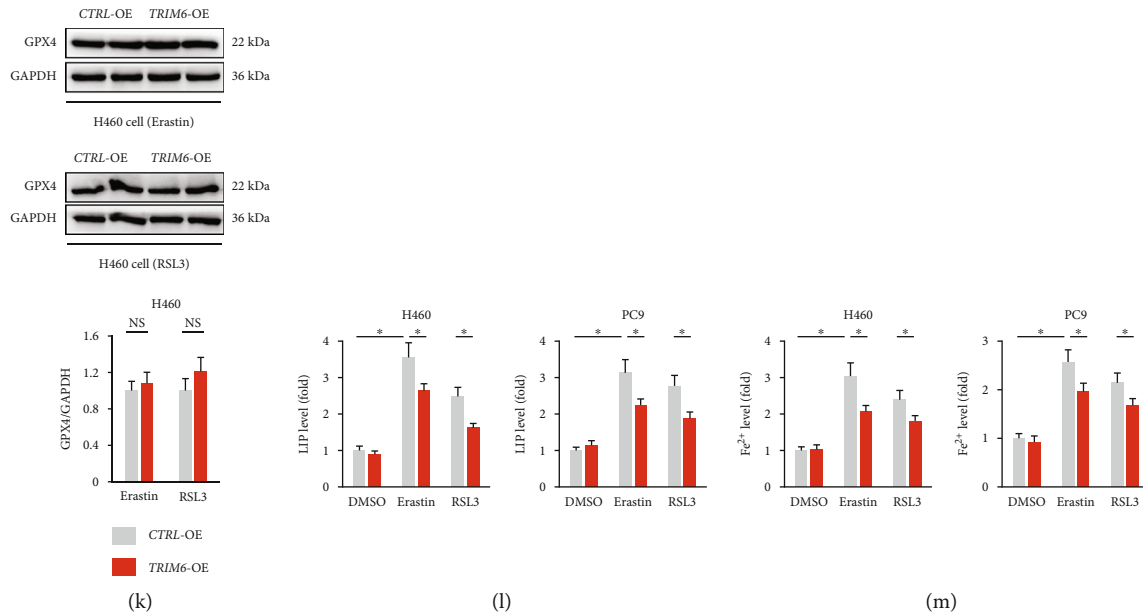
FIGURE 2: TRIM6 overexpression inhibits erastin- and RSL3-induced ferroptosis in the lung cancer cells. (a) Relative *TRIM6* mRNA levels in H460 and PC9 cells with or without *TRIM6*-OE infection ($n = 6$). (b) Cell survival status in the presence or absence of erastin/RSL3 stimulation after TRIM6 overexpression ($n = 6$). (c) Colony formation in the cells with or without TRIM6 overexpression upon ferroptotic stimulation ($n = 6$). (d, e) Intracellular ROS and lipid peroxidation levels ($n = 6$). (f, g) The levels of intracellular superoxide anion and MDA formation ($n = 6$). (h) Relative MMP levels in indicating times after erastin/RSL3 stimulation ($n = 5$). (i, j) The levels of intracellular GSH and GPX4 activities ($n = 6$). (k) Protein levels of GPX4 in erastin- or RSL3-treated lung cancer cells after TRIM6 overexpression ($n = 6$). (l, m) Relative LIP and $Fe^{2+}$ levels ($n = 6$). All data are reported as the mean ± SD, $^{*}P < 0.05$ versus corresponding groups.

accumulation [14–16]. To investigate whether TRIM6 regulated ferroptosis via affecting glutaminolysis, TRIM6-deficient cells were treated with different pharmacological inhibitors of Gln metabolism upon erastin stimulation (Figure 4(a)). As shown in Figures 4(b) and 4(c), cell death and MDA formation in TRIM6-deficient H460 cells were markedly suppressed by inhibitors of Gln metabolism, except BPTES, a specific GLS1 inhibitor. In contrast, supplementation of α-KG, the final product of glutaminolysis, reinduced ferroptosis of erastin-treated lung cancer cells in the presence of GPNA, 968, and AOA (Figures 4(b) and 4(c)). SLC1A5 and SLC38A1 are two critical Gln importers and play critical roles in regulating ferroptosis and lung cancer [15]. We found that TRIM6 silence increased, while TRIM6 overexpression decreased SLC1A5 protein levels in erastin-treated H460 cells, with no impact on SLC38A1 expressions (Figures 4(d) and 4(e)). Accordingly, Gln uptake was enhanced in the lung cancer cells with TRIM6 silence, but inhibited in those with TRIM6 overexpression (Figure 4(f)). To further confirm the involvement of SLC1A5, H460 cells were preinfected with *SLC1A5*-OE lentivirus, and the efficiency was confirmed in Figure 4(g). As shown in Figure 4(h), Gln uptake in erastin-treated H460 cells was decreased by TRIM6 overexpression, yet restored by SLC1A5 overexpression, which was inhibited by GPNA incubation. The decreases of MDA and lipid ROS generation in TRIM6-overexpressed cells with erastin stimulation were increased after the overexpression of SLC1A5, which were then suppressed by GPNA treatment (Figure 4(i)). Accordingly, TRIM6 overexpression-mediated restorations

of cell survival and colony formation were prevented in SLC1A5-overexpressed H460 cells, but not in those treated with GPNA (Figure 4(j)). In contrast, the increased Gln uptake in TRIM6-deficient cells was significantly inhibited by *SLC1A5*-KD infection (Figures S2A and S2B). Correspondingly, TRIM6 silence-elicited ferroptosis was attenuated in SLC1A5-deficient H460 cells (Figures S2C and S2D). These data indicate that TRIM6 modulates ferroptosis via affecting SLC1A5-mediated glutaminolysis.

### 3.5. TRIM6 Directly Interacts with SLC1A5 to Promote Its Degradation.
We also investigated how TRIM6 regulated SLC1A5 in H460 cancer cells. As shown in Figure 5(a), TRIM6 overexpression made no alteration on *SLC1A5* mRNA level. This finding suggested that SLC1A5 protein might be destabilized in TRIM6-overexpressed cells, and we thus assessed the half-life of SLC1A5 by treating TRIM6-manipulated H460 cells with CHX. As shown in Figure 5(b), TRIM6 overexpression significantly shortened the half-life of SLC1A5 protein. TRIM6 functions as an E3 Ub-ligase, while it is unclear whether TRIM6 affects SLC1A5 protein stability via regulating its ubiquitination. Intriguingly, we found that TRIM6 overexpression enhanced SLC1A5 ubiquitination in erastin-treated H460 cells (Figure 5(c)). The catalytic ability of TRIM6 on SLC1A5 ubiquitination was also confirmed in vivo and in vitro (Figure 5(d)). To determine the Ub-dependent proteasomal degradation of SLC1A5, TRIM6-overexpressed H460 cells were incubated with MG132 upon erastin treatment. The data implied that MG132 treatment blocked the

(a)

(b)

(c)

(d)

(e)

(f)

Figure 3: Continued.

(g)   (h)



(i)
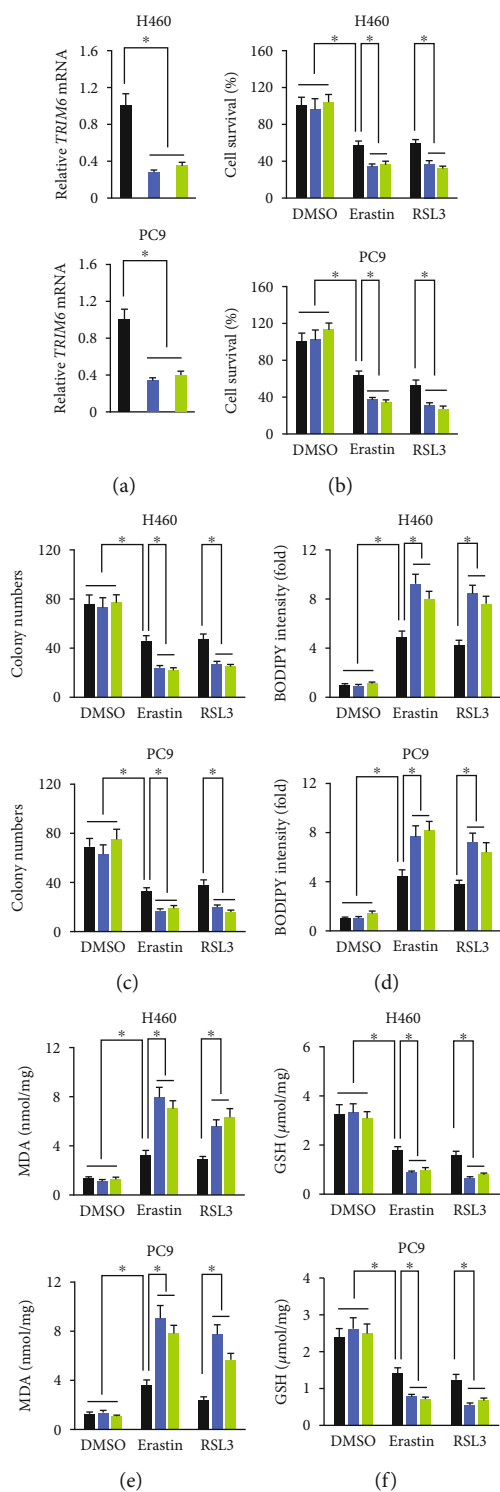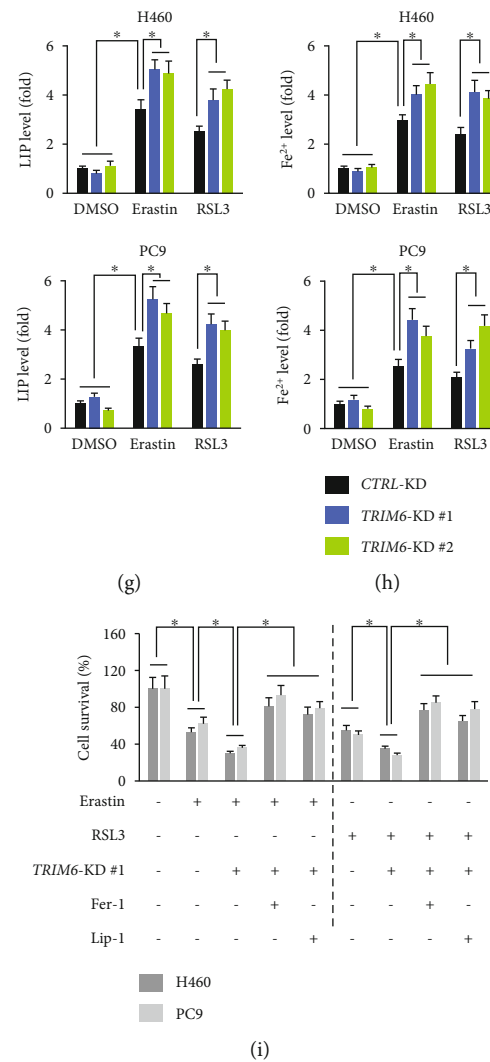
FIGURE 3: TRIM6 silence promotes erastin- and RSL3-induced ferroptosis in the lung cancer cells. (a) Relative *TRIM6* mRNA levels in H460 and PC9 cells with or without *TRIM6*-KD infection ($n = 6$). (b) Cell survival status in the presence or absence of erastin/RSL3 stimulation after TRIM6 knockdown ($n = 6$). (c) Colony formation in the cells with or without TRIM6 silence upon ferroptotic stimulation ($n = 6$). (d, e) Intracellular lipid ROS levels and MDA formation ($n = 6$). (f) Intracellular GSH levels ($n = 6$). (g, h) Relative LIP and $Fe^{2+}$ levels ($n = 6$). (i) Cell survival status ($n = 6$). All data are reported as the mean ± SD, $^*P < 0.05$ versus corresponding groups.

reduction of SLC1A5 proteins caused by TRIM6 overexpression (Figure 5(e)). Accordingly, the decreased Gln uptake was also prevented by MG132 incubation (Figure 5(f)). We next explored whether this ubiquinated process depended on the direct interaction between TRIM6 and SLC1A5. The endogenous physical interaction was confirmed by IP assay using H460 lysates (Figures 5(g) and S3A). To further validate this reciprocal binding, lysates prepared from HEK293T cells transiently transfected with Flag-tagged TRIM6 and HA-tagged SLC1A5 were subjected to IP assay. Immunoprecipitation with anti-Flag or anti-HA antibodies brought down both Flag-TRIM6 and HA-SLC1A5, indicating that the two tagged proteins were associated with each other in HEK293T cells (Figures 5(h) and S3B). Taken together, we conclude that TRIM6 directly interacts with SLC1A5 to promote its degradation.

3.6. TRIM6 Regulates the Chemosensitivity of the Lung Cancer Cells In Vivo and In Vitro. Given its effective role in regulating ferroptosis, we finally determined whether TRIM6 manipulation affected the chemosensitivity of the lung cancer cells in vivo and in vitro. As shown in Figures 6(a) and 6(b), TRIM6 overexpression significantly reduced DDP- and PTX-mediated toxic effects to H460 cells in vitro, as evidenced by the increased cell survival and colony formation. Conversely, TRIM6 silence potentiated the chemotherapeutic effects of DDP and PTX in H460 cells (Figures 6(c) and 6(d)). We also examined the role of TRIM6 on DDP- and PTX-mediated tumor-killing actions in mouse xenograft tumor models. In line with the in vitro findings, we observed that TRIM6 overexpression promoted, while TRIM6 knockdown further inhibited tumor growth upon DDP or PTX chemotherapy (Figures 6(e)–6(h)). To

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)
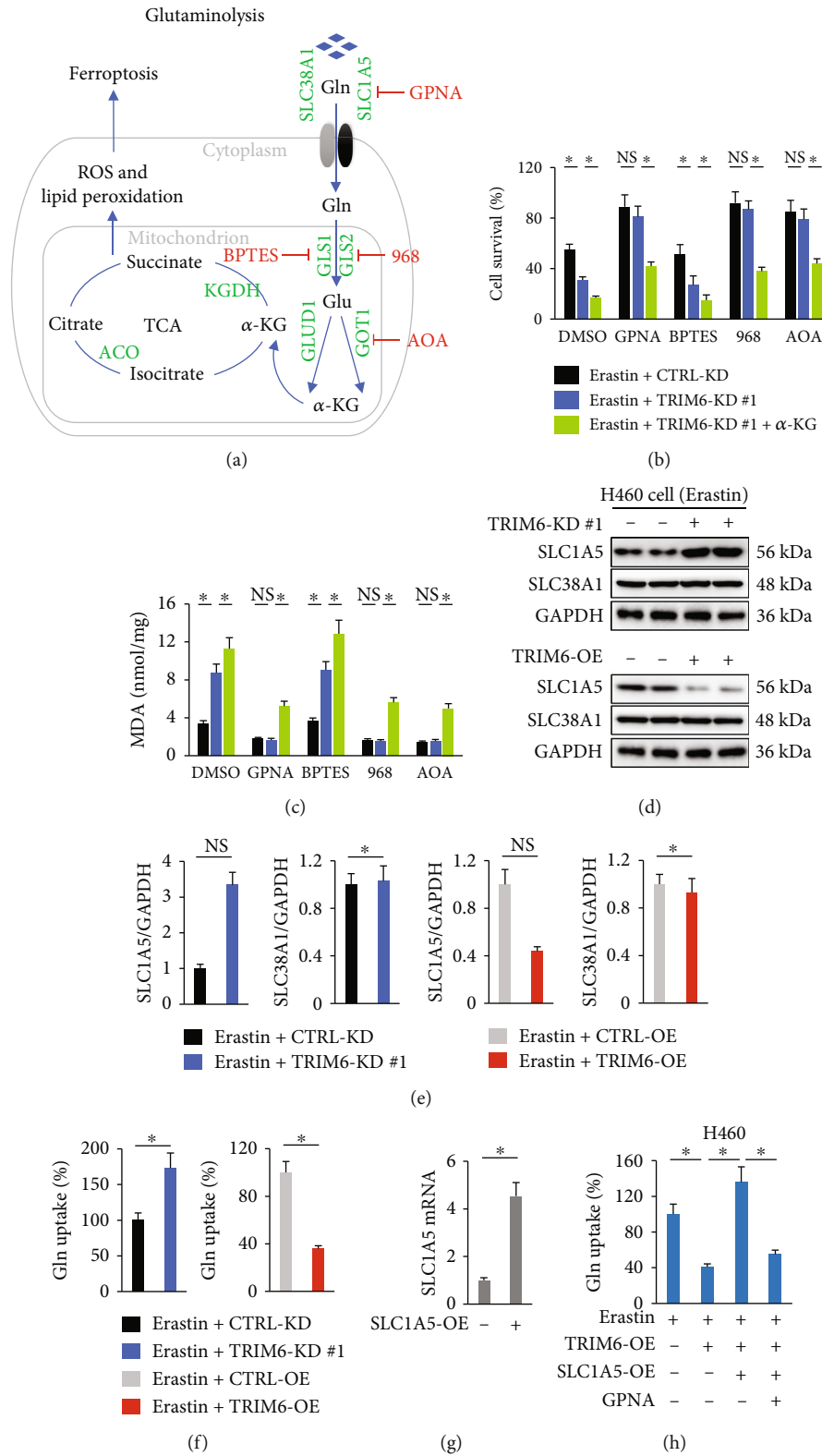
Figure 4: Continued.
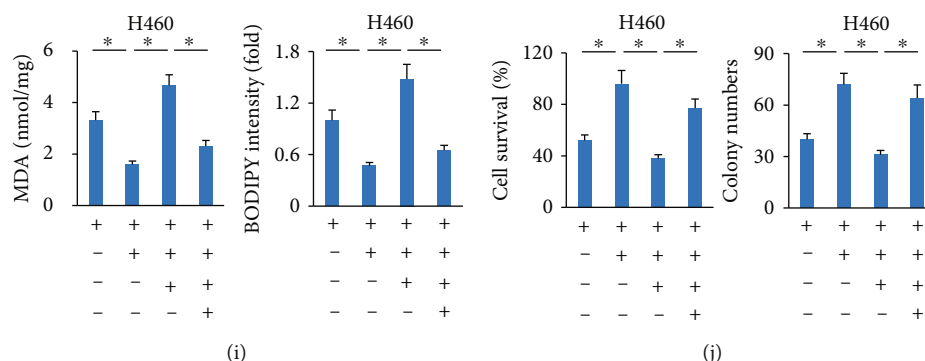
FIGURE 4: TRIM6 modulates ferroptosis via affecting SLC1A5-mediated glutaminolysis. (a) Schematic overview of the glutaminolysis pathway and TCA cycle in ferroptosis. Gln is imported inside the cells by SLC1A5/SLC38A1 and then converted to Glu by GLS in mitochondria. The GOT1 and GLUD1 ultimately converts Glu to $\alpha$-KG, which contributes to ROS accumulation via the TCA cycle. The small molecule inhibitors are indicated in red: L-Gln transporter inhibitor, GPNA; GLS1 inhibitor, BPTES; GLS inhibitor, 968; pan-transaminase inhibitor, AOA. (b, c) Cell survival and MDA formation in erastin-treated H460 cells ($n = 6$). (d, e) Protein levels of SLC1A5 and SLC38A1 in erastin-treated H460 cells with TRIM6 knockdown or overexpression ($n = 6$). (f) Relative Gln uptake in erastin-treated H460 cells ($n = 8$). (g) Relative *SLC1A5* mRNA levels in H460 cells with or without *SLC1A5*-OE infection ($n = 6$). (h) Relative Gln uptake in erastin-treated H460 cells ($n = 8$). (i) Intracellular lipid ROS levels and MDA formation in erastin-treated H460 cells ($n = 6$). (j) Cell survival status and colony formation in erastin-treated H460 cells ($n = 6$). All data are reported as the mean $\pm$ SD, * $P < 0.05$ versus corresponding groups. NS indicates no significance.

enhance the translational value of our findings, we also analyzed the predictive role of TRIM6 and SLC1A5 on patient survival in LUAD database. As shown in Figures S4A and S4B, both TRIM6 and SLC1A5 expressions negatively correlated with patient survival in LUAD database, indicating a clinical role of TRIM6 and SLC1A4 of lung cancer. These observations define TRIM6 as a promising therapeutic target for the treatment of lung cancer.

## 4. Discussion

The present study shows the role of TRIM6 on ferroptosis and chemosensitivity of lung cancer, and our major findings are presented as below. Firstly, TRIM6 is highly expressed in human lung cancer tissues and cells, and its expression in the lung cancer cells is further increased by ferroptotic stimulation. Secondly, TRIM6 overexpression inhibits, while TRIM6 silence promotes erastin- and RSL3-induced glutaminolysis and ferroptosis in the lung cancer cells. Thirdly, TRIM6 directly interacts with SLC1A5 to promote its ubiquitination and degradation, thereby inhibiting Gln import, glutaminolysis, lipid peroxidation, and ferroptotic cell death. Finally, TRIM6 overexpression reduces the chemotherapeutic effects of DDP and PTX. In contrast, TRIM6 silence sensitized human lung cancer cells to DDP and PTX in vivo and in vitro. Overall, our research for the first time defines TRIM6 as a negative regulator of ferroptosis in the lung cancer cells, and TRIM6 overexpression enhances the resistance of human lung cancer cells to chemotherapeutic drugs. Overall, TRIM6 is a promising therapeutic target for the treatment of lung cancer.

Ferroptosis, a newly identified form of cell death, plays critical roles in the development and chemoresistance of lung cancer. Wang et al. found that inhibiting ferroptosis facilitated the proliferation of human lung cancer cells,

thereby promoting tumor progression [32], while inducing ferroptosis by erianin suppressed the growth and migration of the lung cancer cells [33]. Iron-related accumulation of lethal lipid ROS is the predominant feature during ferroptosis; however, we found that TRIM6 genetic manipulation did not affect Glu uptake, GSH synthesis, and iron transport. Gln is the most abundant amino acid in human tissues and plasma and provides nitrogen source for the biosynthesis of amino acids, nucleotides, and lipids. Besides, Gln is also an important carbon source and replenishes the intermediates for TCA cycle via glutaminolysis [14]. However, recent studies have reported that fueling of the TCA cycle by glutaminolysis accelerates lipid peroxidation and ferroptosis and that inhibiting glutaminolysis prevents ferroptotic cell death [15, 16]. SLC1A5, a membranous importer, is required for the uptake of neutral amino acids (e.g., Gln) and contributes to metabolic reprogramming of cancer cells [54]. Luo et al. proved that SLC1A5 suppression decreased Gln uptake, lipid peroxidation, and ferroptosis, thereby facilitating the survival of melanoma cells and tumor progression [15]. Consistently, we also found that TRIM6 directly interacted with SLC1A5 to promote its protein degradation and then inhibit erastin- or RSL3-mediated ferroptotic cell death. In contrast, TRIM6 silence elevated SLC1A5 expression and ferroptosis of the lung cancer cells.

TRIM family proteins function as kinds of E3 Ub-ligases and are implicated in the pathogenesis of lung cancer. Results from Chen et al. implied that TRIM28 reduced the proliferation of the lung cancer cell lines and that TRIM28 depletion led to increased cell proliferation [55]. Liu et al. found that TRIM29 knockdown suppressed the proliferation and invasion of human lung squamous cancer cells and also enhanced the chemosensitivity of DDP [56]. And knockdown of TRIM65 suppressed survival of DDP-resistant lung cancer cell lines and tumor growth [57]. TRIM6 belongs to
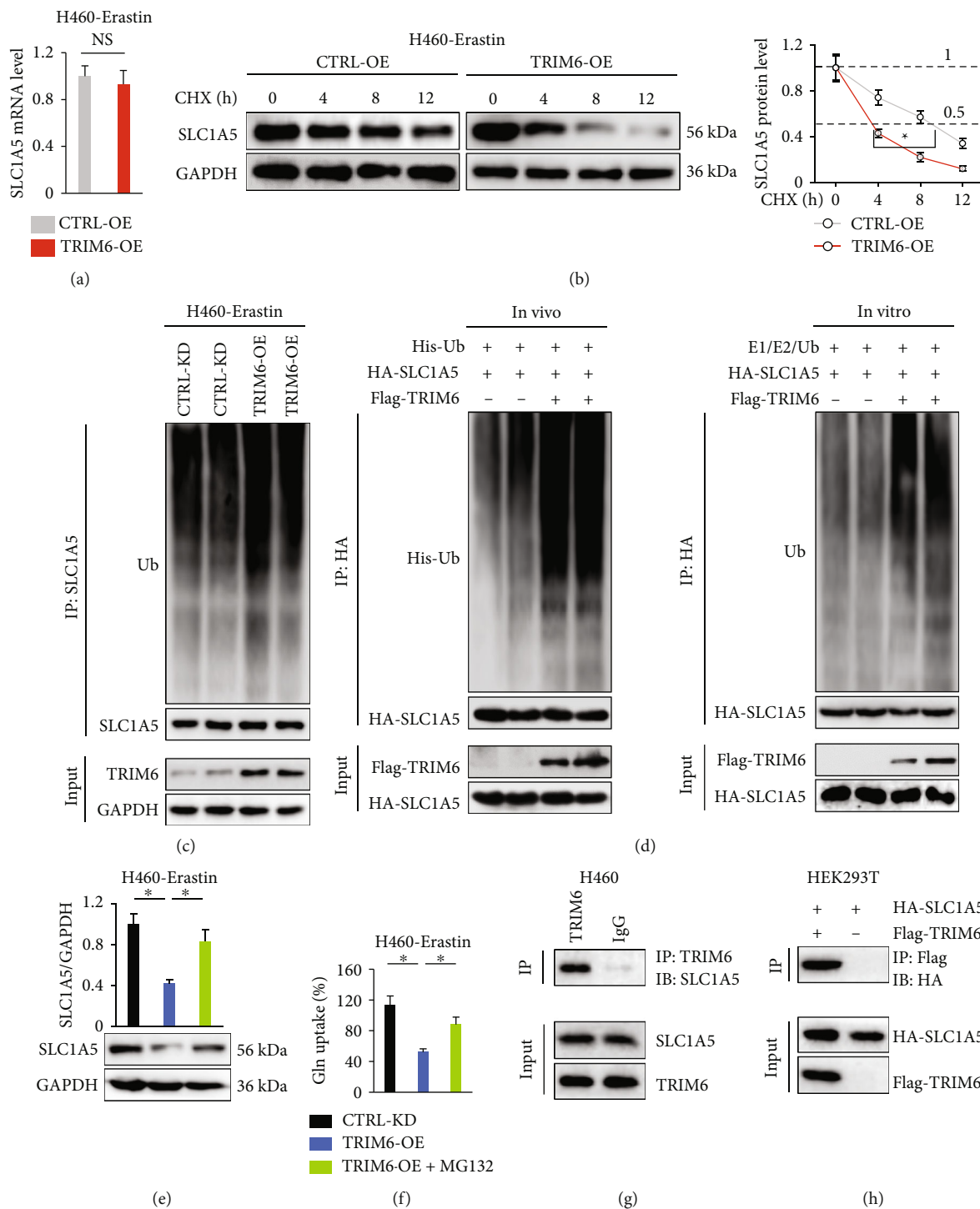
Figure 5: TRIM6 directly interacts with SLC1A5 to promote its degradation. (a) Relative *SLC1A5* mRNA levels in erastin-treated H460 cells with or without *TRIM6*-OE infection ($n = 6$). (b) Protein levels of SLC1A5 in erastin-treated H460 cells after CHX incubation ($n = 6$). (c) Ubiquinated levels of SLC1A5 in erastin-treated cells with or without TRIM6 overexpression ($n = 6$). (d) Ubiquitination assay in vivo and in vitro ($n = 4$). (e) Protein levels of SLC1A5 in erastin-treated H460 cells after MG132 incubation ($n = 6$). (f) Relative Gln uptake in erastin-treated H460 cells after MG132 incubation ($n = 6$). (g, h) IP assay for examining the interaction between TRIM6 and SLC1A5 ($n = 6$). All data are reported as the mean ± SD, $^{*}P < 0.05$ versus corresponding groups. NS indicates no significance.

the TRIM family and is well known for its role in the antiviral responses [24, 25]. Yet, recent studies revealed some additional actions of TRIM6, including the regulation on tumor progression [26–28]. Herein, we found that human lung cancer tissues and cells exhibited higher TRIM6 expression compared with the ANT or normal lung epithelial cell and that its expression in the lung cancer cells was further increased by ferroptotic stimulation. Consistently, Zheng
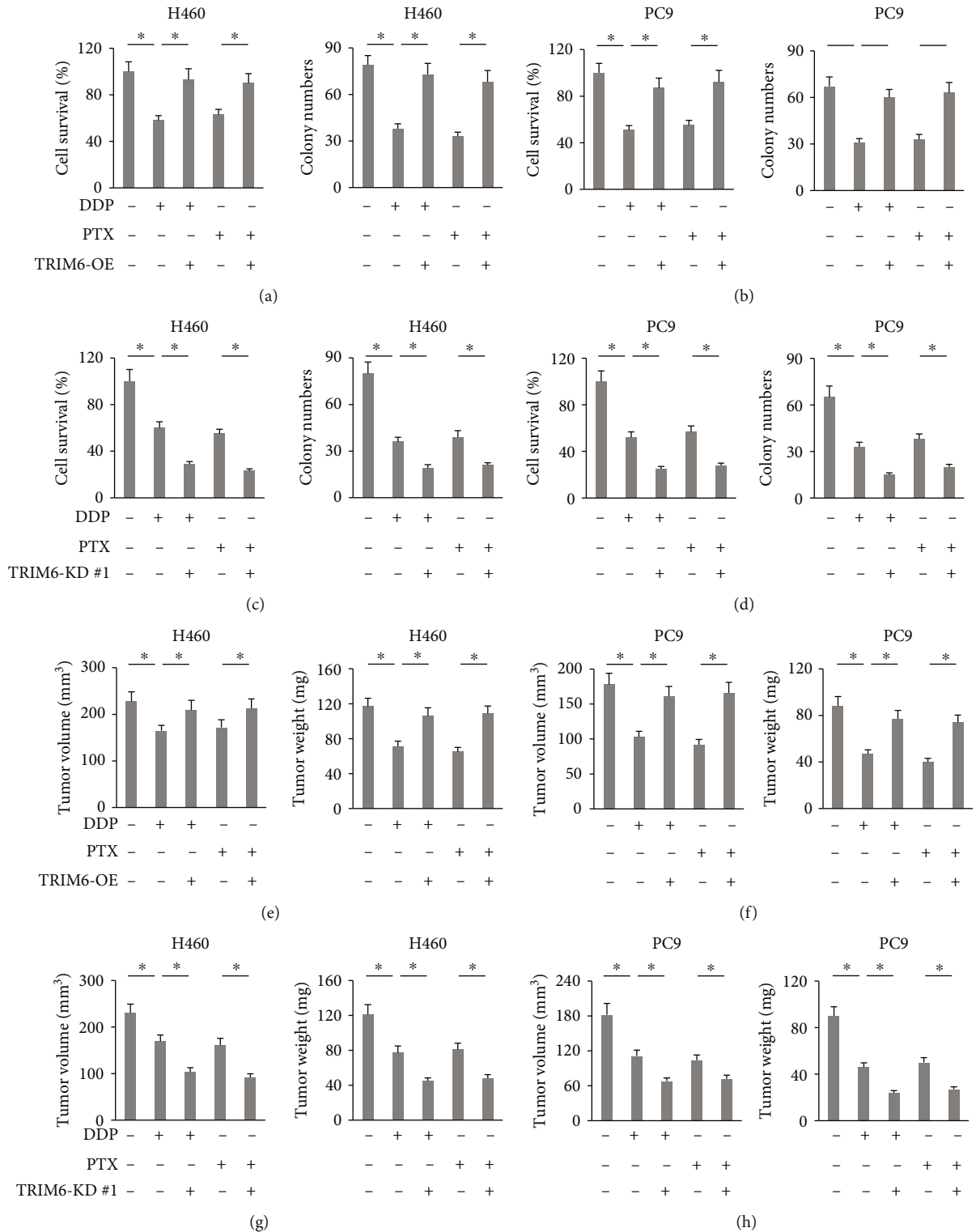
FIGURE 6: TRIM6 regulates the chemosensitivity of the lung cancer cells in vivo and in vitro. (a, b) Cell survival status and colony formation in DDP/PTX-treated lung cancer cells with or without TRIM6 overexpression ($n = 6$). (c, d) Cell survival status and colony formation in DDP/PTX-treated lung cancer cells with or without TRIM6 silence ($n = 6$). (e, f) Tumor volumes and weights in DDP/PTX-treated xenografts models inoculated with TRIM6-overexpressed lung cancer cells ($n = 6$). (g, h) Tumor volumes and weights in DDP/PTX-treated xenograft models inoculated with TRIM6-silenced lung cancer cells ($n = 6$). All data are reported as the mean ± SD, $^*P < 0.05$ versus corresponding groups.

et al. previously also detected an upregulated TRIM6 expression in human CRC samples [28]. Liu et al. determined that TRIM6 was highly expressed in angiotensin II- (Ang II-) stimulated fibrotic kidneys and positively correlated with the severity of renal fibrosis. Mechanistically, Ang II-induced ROS generation activated nuclear factor-$\kappa$B pathway, which subsequently elevated TRIM6 expression through binding to its promoter directly [58]. As we know, ROS overproduction is a key feature of myocardial ischemia/reperfusion (I/R) injury. Results from Zeng et al. revealed that cardiac TRIM6 mRNA and protein levels were significantly upregulated following I/R injury [26]. Based on these findings, we speculated that TRIM6 upregulated in ferroptotic condition might be associated with the increased oxidative stress. Meanwhile, we found that increased TRIM6 expression in the lung cancer cells upon ferroptotic stimulation could provide cytoprotective effects against chemotherapeutic reagents. Further detections revealed that TRIM6 reduced glutaminolysis via targeting SLC1A5-mediated Gln uptake. However, relatively little is known about how TRIM6 modulates SLC1A5 currently. TRIM6 acts as an E3 Ub-ligase and is essential for protein stability via regulating the ubiquitinated processes [24, 28]. In line with these studies, we proved that TRIM6 directly bound to SLC1A5 and promoted its ubiquinated modification at the posttranscriptional levels, thereby shortening the half-life of SLC1A5 protein and reducing ferroptotic cell death. Moreover, TRIM6 knockdown potentiated the lung cancer cells to DDP and PTX treatment in vivo and in vitro.

In summary, our findings determine a novel regulatory role of TRIM6 on ferroptosis and tumor progression of lung cancer. Genetic or pharmacological inhibition of TRIM6 may provide promising strategies for the treatment of lung cancer.

## Data Availability

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 there are no conflicts of interests.

## Authors' Contributions

Ying Zhang, Ping Dong, and Qing Geng conceived the hypothesis and designed the study. Ying Zhang, Nian Liu, and Jun-Yuan Yang carried out the experiments and acquired the data. Ying Zhang and Hui-Min Wang conducted the data analysis. Ying Zhang, Ping Dong, and Qing Geng drafted the manuscript. Ying Zhang and Qing Geng revised the manuscript. Ying Zhang and Ping Dong contributed equally to this work.

## Acknowledgments

## Supplementary Materials

Figure S1: TRIM6 knockdown does not affect the molecules essential for Glu uptake, GSH synthesis, and iron transport. Figure S2: SLC1A5 knockdown abrogates the protective effects against ferroptosis in TRIM6-deficient H460 cells upon erastin stimulation. Figure S3: TRIM6 directly interacts with SLC1A5. Figure S4: expressions of TRIM6 and SLC1A5 negatively correlate with patient survival in the LUAD database. (Supplementary Materials)

## References

[1] 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.

[2] X. Jie, W. P. Fong, R. Zhou et al., "USP9X-mediated KDM4C deubiquitination promotes lung cancer radioresistance by epigenetically inducing TGF-$\beta$2 transcription," Cell Death and Differentiation, vol. 28, no. 7, pp. 2095–2111, 2021.

[3] W. Du, J. Zhu, Y. Zeng et al., "KPNB1-mediated nuclear translocation of PD-L1 promotes non-small cell lung cancer cell proliferation via the Gas6/MerTK signaling pathway," Cell Death and Differentiation, vol. 28, no. 4, pp. 1284–1300, 2021.

[4] C. Chen, L. Gong, X. Liu et al., "Identification of peroxiredoxin 6 as a direct target of withangulatin A by quantitative chemical proteomics in non-small cell lung cancer," Redox Biology, vol. 46, p. 102130, 2021.

[5] M. Wang, S. Chen, Y. Wei, and X. Wei, "DNA-PK inhibition by M3814 enhances chemosensitivity in non-small cell lung cancer," Acta Pharmaceutica Sinica B, vol. 11, no. 12, pp. 3935–3949, 2021.

[6] H. He, L. Wang, Y. Qiao, B. Yang, D. Yin, and M. He, "Epigallocatechin-3-gallate pretreatment alleviates doxorubicin-induced ferroptosis and cardiotoxicity by upregulating AMPK$\alpha$2 and activating adaptive autophagy," Redox Biology, vol. 48, p. 102185, 2021.

[7] W. Cui, D. Liu, W. Gu, and B. Chu, "Peroxisome-driven ether-linked phospholipids biosynthesis is essential for ferroptosis," Cell Death and Differentiation, vol. 28, no. 8, pp. 2536–2551, 2021.

[8] W. D. Bao, P. Pang, X. T. Zhou et al., "Loss of ferroportin induces memory impairment by promoting ferroptosis in Alzheimer's disease," Cell Death and Differentiation, vol. 28, no. 5, pp. 1548–1562, 2021.

[9] X. Zhang, J. X. Zhu, Z. G. Ma et al., "Rosmarinic acid alleviates cardiomyocyte apoptosis via cardiac fibroblast in doxorubicin-induced cardiotoxicity," International Journal of Biological Sciences, vol. 15, no. 3, pp. 556–567, 2019.

[10] Y. Liu, Z. Song, Y. Liu et al., "Identification of ferroptosis as a novel mechanism for antitumor activity of natural product derivative a2 in gastric cancer," Acta Pharmaceutica Sinica B, vol. 11, no. 6, pp. 1513–1525, 2021.

[11] W. S. Yang, R. SriRamaratnam, M. E. Welsch et al., "Regulation of ferroptotic cancer cell death by GPX4," Cell, vol. 156, no. 1-2, pp. 317–331, 2014.

[12] C. W. Brown, J. J. Amante, P. Chhoy et al., "Prominin2 drives ferroptosis resistance by stimulating iron export," Developmental Cell, vol. 51, no. 5, pp. 575–586.e4, 2019.

[13] S. J. Dixon, K. M. Lemberg, M. R. Lamprecht et al., "Ferroptosis: an iron-dependent form of nonapoptotic cell death," *Cell*, vol. 149, no. 5, pp. 1060–1072, 2012.

[14] R. J. DeBerardinis, J. J. Lum, G. Hatzivassiliou, and C. B. Thompson, "The biology of cancer: metabolic reprogramming fuels cell growth and proliferation," *Cell Metabolism*, vol. 7, no. 1, pp. 11–20, 2008.

[15] M. Luo, L. Wu, K. Zhang et al., "miR-137 regulates ferroptosis by targeting glutamine transporter SLC1A5 in melanoma," *Cell Death and Differentiation*, vol. 25, no. 8, pp. 1457–1472, 2018.

[16] M. Gao, P. Monian, N. Quadri, R. Ramasamy, and X. Jiang, "Glutaminolysis and transferrin regulate ferroptosis," *Molecular Cell*, vol. 59, no. 2, pp. 298–308, 2015.

[17] B. Zhao, Y. C. Tsai, B. Jin et al., "Protein engineering in the ubiquitin system: tools for discovery and beyond," *Pharmacological Reviews*, vol. 72, no. 2, pp. 380–413, 2020.

[18] J. Zhao, B. Cai, Z. Shao et al., "TRIM26 positively regulates the inflammatory immune response through K11-linked ubiquitination of TAB1," *Cell Death and Differentiation*, vol. 28, no. 11, pp. 3077–3091, 2021.

[19] S. Zeng, Z. Zhao, S. Zheng et al., "The E3 ubiquitin ligase TRIM31 is involved in cerebral ischemic injury by promoting degradation of TIGAR," *Redox Biology*, vol. 45, p. 102058, 2021.

[20] Y. Chen, X. Shao, J. Cao et al., "Phosphorylation regulates cullin-based ubiquitination in tumorigenesis," *Acta Pharmaceutica Sinica B*, vol. 11, no. 2, pp. 309–321, 2021.

[21] S. Hatakeyama, "TRIM family proteins: roles in autophagy, immunity, and carcinogenesis," *Trends in Biochemical Sciences*, vol. 42, no. 4, pp. 297–311, 2017.

[22] J. Ji, K. Ding, T. Luo et al., "TRIM22 activates NF-$\kappa$B signaling in glioblastoma by accelerating the degradation of I$\kappa$B$\alpha$," *Cell Death and Differentiation*, vol. 28, no. 1, pp. 367–381, 2021.

[23] M. Di Rienzo, A. Romagnoli, M. Antonioli, M. Piacentini, and G. M. Fimia, "TRIM proteins in autophagy: selective sensors in cell damage and innate immune responses," *Cell Death and Differentiation*, vol. 27, no. 3, pp. 887–902, 2020.

[24] R. Rajsbaum, G. A. Versteeg, S. Schmid et al., "Unanchored K48-linked polyubiquitin synthesized by the E3-ubiquitin ligase TRIM6 stimulates the interferon-IKK$\varepsilon$ kinase - mediated antiviral response," *Immunity*, vol. 40, no. 6, pp. 880–895, 2014.

[25] P. Bharaj, C. Atkins, P. Luthra et al., "The host E3-ubiquitin ligase TRIM6 ubiquitinates the Ebola virus VP35 protein and promotes virus replication," *Journal of Virology*, vol. 91, no. 18, 2017.

[26] G. Zeng, C. Lian, P. Yang, M. Zheng, H. Ren, and H. Wang, "E3-ubiquitin ligase TRIM6 aggravates myocardial ischemia/reperfusion injury via promoting STAT1-dependent cardiomyocyte apoptosis," *Aging (Albany NY)*, vol. 11, no. 11, pp. 3536–3550, 2019.

[27] T. Sato, F. Okumura, T. Ariga, and S. Hatakeyama, "TRIM6 interacts with Myc and maintains the pluripotency of mouse embryonic stem cells," *Journal of Cell Science*, vol. 125, no. 6, pp. 1544–1555, 2012.

[28] S. Zheng, C. Zhou, Y. Wang, H. Li, Y. Sun, and Z. Shen, "TRIM6 promotes colorectal cancer cells proliferation and response to thiostrepton by TIS21/FoxM1," *Journal of Experimental & Clinical Cancer Research*, vol. 39, no. 1, p. 23, 2020.

[29] M. La Montagna, L. Shi, P. Magee, S. Sahoo, M. Fassan, and M. Garofalo, "AMPK$\alpha$ loss promotes KRAS-mediated lung tumorigenesis," *Cell Death and Differentiation*, vol. 28, no. 9, pp. 2673–2689, 2021.

[30] M. Luo, Y. Xia, F. Wang et al., "PD0325901, an ERK inhibitor, enhances the efficacy of PD-1 inhibitor in non- small cell lung carcinoma," *Acta Pharmaceutica Sinica B*, vol. 11, no. 10, pp. 3120–3133, 2021.

[31] W. Liu, Y. Zhou, W. Duan et al., "Glutathione peroxidase 4-dependent glutathione high-consumption drives acquired platinum chemoresistance in lung cancer-derived brain metastasis," *Clinical and Translational Medicine*, vol. 11, no. 9, p. e517, 2021.

[32] M. Wang, C. Mao, L. Ouyang et al., "Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA," *Cell Death and Differentiation*, vol. 26, no. 11, pp. 2329–2343, 2019.

[33] P. Chen, Q. Wu, J. Feng et al., "Erianin, a novel dibenzyl compound in dendrobium extract, inhibits lung cancer cell growth and migration via calcium/calmodulin-dependent ferroptosis," *Signal Transduction and Targeted Therapy*, vol. 5, no. 1, p. 51, 2020.

[34] G. H. Fan, T. Y. Zhu, and J. Huang, "FNDC5 promotes paclitaxel sensitivity of non-small cell lung cancers via inhibiting MDR1," *Cellular Signalling*, vol. 72, p. 109665, 2020.

[35] J. Zhou, L. Wang, Q. Sun et al., "Hsa_circ_0001666 suppresses the progression of colorectal cancer through the miR-576-5p/PCDH10 axis," *Clinical and Translational Medicine*, vol. 11, no. 11, p. e565, 2021.

[36] X. Zheng, W. Li, H. Xu et al., "Sinomenine ester derivative inhibits glioblastoma by inducing mitochondria- dependent apoptosis and autophagy by PI3K/AKT/mTOR and AMPK/mTOR pathway," *Acta Pharmaceutica Sinica B*, vol. 11, no. 11, pp. 3465–3480, 2021.

[37] C. Hu, X. Zhang, N. Zhang et al., "Osteocrin attenuates inflammation, oxidative stress, apoptosis, and cardiac dysfunction in doxorubicin-induced cardiotoxicity," *Clinical and Translational Medicine*, vol. 10, no. 3, p. e124, 2020.

[38] X. Zhang, C. Hu, C. Y. Kong et al., "FNDC5 alleviates oxidative stress and cardiomyocyte apoptosis in doxorubicin-induced cardiotoxicity via activating AKT," *Cell Death and Differentiation*, vol. 27, no. 2, pp. 540–555, 2020.

[39] Y. X. Ji, P. Zhang, X. J. Zhang et al., "The ubiquitin E3 ligase TRAF6 exacerbates pathological cardiac hypertrophy via TAK1-dependent signalling," *Nature Communications*, vol. 7, no. 1, p. 11267, 2016.

[40] C. Hu, X. Zhang, W. Wei et al., "Matrine attenuates oxidative stress and cardiomyocyte apoptosis in doxorubicin-induced cardiotoxicity via maintaining AMPK $\alpha$ /UCP2 pathway," *Acta Pharmaceutica Sinica B*, vol. 9, no. 4, pp. 690–701, 2019.

[41] A. Roveri, M. Maiorino, and F. Ursini, "Enzymatic and immunological measurements of soluble and membrane-bound phospholipid-hydroperoxide glutathione peroxidase," *Methods in Enzymology*, vol. 233, pp. 202–212, 1994.

[42] J. Lu, J. Li, Y. Hu et al., "Chrysophanol protects against doxorubicin-induced cardiotoxicity by suppressing cellular PARylation," *Acta Pharmaceutica Sinica B*, vol. 9, no. 4, pp. 782–793, 2019.

[43] J. Lee, J. H. You, D. Shin, and J. L. Roh, "Inhibition of glutaredoxin 5 predisposes cisplatin-resistant head and neck cancer

cells to ferroptosis," *Theranostics*, vol. 10, no. 17, pp. 7775–7786, 2020.

[44] X. Zhang, C. Hu, N. Zhang et al., "Matrine attenuates pathological cardiac fibrosis via RPS5/p38 in mice," *Acta Pharmacologica Sinica*, vol. 42, no. 4, pp. 573–584, 2021.

[45] Y. Li, Y. Cao, J. Xiao et al., "Inhibitor of apoptosis-stimulating protein of p53 inhibits ferroptosis and alleviates intestinal ischemia/reperfusion-induced acute lung injury," *Cell Death and Differentiation*, vol. 27, no. 9, pp. 2635–2650, 2020.

[46] C. Hu, X. Zhang, M. Hu et al., "Fibronectin type III domain-containing 5 improves aging-related cardiac dysfunction in mice," *Aging Cell*, vol. 21, p. e13556, 2022.

[47] C. Hu, X. Zhang, P. Song et al., "Meteorin-like protein attenuates doxorubicin-induced cardiotoxicity via activating cAMP/PKA/SIRT1 pathway," *Redox Biology*, vol. 37, p. 101747, 2020.

[48] X. Zhang, C. Hu, X. P. Yuan et al., "Osteocrin, a novel myokine, prevents diabetic cardiomyopathy via restoring proteasomal activity," *Cell Death & Disease*, vol. 12, no. 7, p. 624, 2021.

[49] P. Mei, F. Xie, J. Pan et al., "E3 ligase TRIM25 ubiquitinates RIP3 to inhibit TNF induced cell necrosis," *Cell Death and Differentiation*, vol. 28, no. 10, pp. 2888–2899, 2021.

[50] X. Li, J. Yuan, C. Song et al., "Deubiquitinase USP39 and E3 ligase TRIM26 balance the level of ZEB1 ubiquitination and thereby determine the progression of hepatocellular carcinoma," *Cell Death and Differentiation*, vol. 28, no. 8, pp. 2315–2332, 2021.

[51] B. Qiao, P. Sugianto, E. Fung et al., "Hepcidin-induced endocytosis of ferroportin is dependent on ferroportin ubiquitination," *Cell Metabolism*, vol. 15, no. 6, pp. 918–924, 2012.

[52] D. Shin, J. Lee, J. H. You, D. Kim, and J. L. Roh, "Dihydrolipoamide dehydrogenase regulates cystine deprivation-induced ferroptosis in head and neck cancer," *Redox Biology*, vol. 30, p. 101418, 2020.

[53] X. Sun, Z. Ou, R. Chen et al., "Activation of the p62-Keap1-NRF2 pathway protects against ferroptosis in hepatocellular carcinoma cells," *Hepatology*, vol. 63, no. 1, pp. 173–184, 2016.

[54] H. C. Yoo, S. J. Park, M. Nam et al., "A variant of SLC1A5 is a mitochondrial glutamine transporter for metabolic reprogramming in cancer cells," *Cell Metabolism*, vol. 31, no. 2, pp. 267–283.e12, 2020.

[55] L. Chen, D. T. Chen, C. Kurtyka et al., "Tripartite motif containing 28 (Trim28) can regulate cell proliferation by bridging HDAC1/E2F interactions," *The Journal of Biological Chemistry*, vol. 287, no. 48, pp. 40106–40118, 2012.

[56] C. Liu, X. Huang, S. Hou, B. Hu, and H. Li, "Silencing of tripartite motif (TRIM) 29 inhibits proliferation and invasion and increases chemosensitivity to cisplatin in human lung squamous cancer NCI-H520 cells," *Thorac Cancer*, vol. 6, no. 1, pp. 31–37, 2015.

[57] X. Pan, Y. Chen, Y. Shen, and J. Tantai, "Knockdown of TRIM65 inhibits autophagy and cisplatin resistance in A549/DDP cells by regulating miR-138-5p/ATG7," *Cell Death & Disease*, vol. 10, no. 6, p. 429, 2019.

[58] W. Liu, Y. Yi, C. Zhang et al., "The expression of TRIM6 activates the mTORC1 pathway by regulating the ubiquitination of TSC1-TSC2 to promote renal fibrosis," *Frontiers in Cell and Development Biology*, vol. 8, p. 616747, 2021.

*Research Article*

# A Prognostic Risk Model of a Novel Oxidative Stress-Related Signature Predicts Clinical Prognosis and Demonstrates Immune Relevancy in Lung Adenocarcinoma

Xing Huang,[1] Zhichao Lu [iD],[2] Min He [iD],[2] Yipeng Feng,[3,4,5] Shaorong Yu,[6] Bo Shen [iD],[6] Jianwei Lu [iD],[6] Pingping Wu,[6] Banzhou Pan,[6] Hanlin Ding,[3,4,5] Chen Chen [iD],[6,7] and Yidan Sun [iD][8]

[1]Department of Pathology,
Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Nanjing Medical University Affiliated Cancer Hospital, China
[2]Research Center of Clinical Medicine, Affiliated Hospital of Nantong University, Nantong 226001, China
[3]Department of Thoracic Surgery, Nanjing Medical University Affiliated Cancer Hospital & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, 21009 Nanjing, China
[4]Jiangsu Key Laboratory of Molecular and Translational Cancer Research, Cancer Institute of Jiangsu Province, Nanjing, China
[5]The Fourth Clinical College of Nanjing Medical University, Nanjing, China
[6]Department of Oncology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing 210000, China
[7]The Comprehensive Cancer Centre of Nanjing Drum Tower Hospital,
The Affiliated Hospital of Nanjing University Medical School, China
[8]Department of Oncology, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin, China

Correspondence should be addressed to Chen Chen; chenchen881021@outlook.com and Yidan Sun; huiranzhongyi@126.com

Lung adenocarcinoma (LUAD) is among the most prevalent malignant lung cancers with a poor prognosis due to high invasiveness and lethality despite multiple treatments. Since the lung is an important organ associated with oxidative stress, and it has been confirmed that oxidative stress represents a potential cancer-specific depletion, it is of important significance to investigate and evaluate the clinical value of oxidative stress mechanisms regulating tumor cell apoptosis. Furthermore, there are few studies on the impact of the microenvironment on reaction to immune-checkpoint inhibitors (ICIs) in patients with LUAD. Based on the TCGA-LUAD dataset, which is stratified into a training set as well as a validation set in a ratio of 2 : 1, this investigation constructs and validates a prognostic predictive power of a gene signature model of oxidative stress-related prognostic signatures. To ascertain the differences between the high-risk score group and the low-risk score group in tumor-infiltrating lymphocytes and patients' response to ICI therapy. This oxidative stress-related prognostic gene signature is composed of MAP3K19 and NTSR1 and is an independent prognosis-related factor in the LUAD group. The outcome of patients having a low risk score is better, and the difference was statistically significant, and individuals with a low risk score had a larger number of infiltrating immune cell distribution in the tumor microenvironment, which was closely related to clinical outcome. Our study suggests that the synergistic effect of oxidative stress-related prognostic gene markers-MAP3K19 and NTSR1 has clinical significance in the prognosis identification and immunotherapy of LUAD patients. Thus, the results may help to better intersect the oxidative stress-related mechanisms in clinical value in LUAD but requires prospective validation.

# 1. Introduction

Lung cancer is the major important cause of tumor-related mortalities worldwide [1, 2]. LUAD is the most significant subtype. Although many clinical studies have confirmed that multitargeted drugs and immunotherapy can prolong the overall survival (OS) and improve the objective response rate (ORR) of LUAD patients [3–5], the rapid progression of the disease due to multi-drug resistance is currently very difficult. Few targeted treatment options [6–9]. With the in-depth research in many aspects [10, 11], it is beneficial to discover small molecule inhibitors. Furthermore, significant literature confirmed the hypothesis that the tumor microenvironment (TME) enhances tumor growth through paracrine signaling [12]. So, more investigation should be taken to improve the outcome among LUAD patients.

Redox homeostasis is crucial in not only the survival of normal cells but also cancerous cells. Oxidative stress (OS) is predominantly triggered by an imbalance between cellular antioxidant mechanisms and metabolically generated oxidative free radical species [13, 14]. This imbalance eventually causes the excessive buildup of reactive oxygen species (ROS) within body cells, leading to irreversible or reversible injury to the body [15]. Nevertheless, numerous tumors have increased levels of ROS and exhibited signs of chronic oxidative stress as a result of oncogenic injury, hypoxia, metabolic malfunctions, and proteotoxic stress [13]. Increased ROS is hypothesized to enhance the progression of tumors at sublethal levels by inducing the mutations and changing cell signaling [16]. Nonetheless, to block excessive oxidative injury, tumors generally upregulate the antioxidant pathways [17]. Consequently, numerous cancerous cells are hypersensitive to perturbation of ROS levels. Excessive oxidative stress is known to aggravate the cytotoxic impacts of chemotherapy, and efforts are being made to enhance ROS generation in these environments [18, 19].

In this study, we attempted to obtain oxidative stress-related expression profiling data for LUAD from The Cancer Genome Atlas (TCGA) database and aggregated clinical information and transcriptomes from 445 patients in TCGA-LUAD with complete clinical informatics data and express the data, dividing it into a training set (2/3 of the total, $n = 296$) as well as a test set (1/3 of the total, $n = 149$). We then performed univariate Cox proportional hazards regression in both the training and the validation sets, respectively, to identify genes with prognostic values utilizing the expression data of 147 oxidative stress-related genes. Cox $p$ values $< 0.05$ indicated a coexpression network of 35 oxidative stress DEGs associated with LUAD overall survival, with clinical information and transcriptomic expression data. Based on the LASSO algorithm, the prognosis-related gene signatures composed of 3 and 8 oxidative stress-related genes were screened in the training set and the test set, respectively. Therefore, we obtained a LUAD prognostic risk model composed of MAP3K19 and NTSR1 by intersection difference analysis. We further validated the subgroup prognostic risk model by Kaplan-Meier (KM) analysis as well as receiver-operating characteristic curve (ROC) analysis. Additionally, the link between the risk

model and the TME was ascertained by both the ESTIMATE R software package and the CIBERSORT tool. Lastly, differentially infiltrating immune cells were discovered in the two risk groups, and MAP3K19 and NTSR1 were found to be statistically significant with immune-related genes. In conclusion, our study suggests that oxidative stress-related risk models may provide a viable prognostic tool and important function in the modulation of immune cell distribution in the LUAD tumor microenvironment. Our hypothesis is that this oxidative stress-related prognostic model gene signature has the capacity to anticipate the targeting and prognosis of LUAD. The aim of our investigation was to create and validate an oxidative stress-related LUAD prognostic model and to explore its predictive effect on poor prognosis in LUAD. Assess and validate prognostic power and its independent prognostic value. Our goal was to guide the clinical application of this oxidative stress-related prognostic model gene signature in LUAD.

# 2. Materials and Methods

Figure 1 presents the flow chart for the bioinformatics analysis. R language software (Version 4.0.3) [20] accomplished all the statistical analyses, and $p < 0.05$ denoted a statistically significant difference without a designated setting.

## 2.1. Data Acquiring and Cleaning.
Transcriptomic data (reads), as well as relevant clinical data encompassing survival status, age, sex, grade, and stage of LUAD patients, were retrieved from TCGA (https://portal.gdc.cancer.gov/) database. In total, 551 samples were incorporated in this investigation (54 normal lung samples and 497 LUAD samples). The initial expression data were normalized with the aid of the trim mean of $M$ values (TMM) algorithm in the "limma" package [21], and genes with an average expression not exceeding 1 were excluded. The "limma" package was additionally utilized for MTG differential expression analysis. The criteria for identifying differentially expressed genes in our study were $|logFC| > 1$ and adj.$p < 0.05$. The read counts were transformed to TPM values, and a $log2(x + 1)$ conversion was conducted for further analyses since the TPM values were identical to the microarray values.

We use the $k$-fold cross-validation method to divide the sample set into $k$ mutually exclusive subsets of similar size, each of which keeps the data distribution as consistent as possible. Then, each time the union of $k − 1$ subsets is used as the training set, and the remaining subset is used as the test set, so that $k$ sets of training/testing sets can be obtained. Perform $k$ training and testing times, and return the average of the $k$ test results.

## 2.2. Identification of Differentially Expressed Oxidative Stress-Related Genes with Prognostic Value in LUAD Tissue and Normal Lung Tissue.
As per the search term "oxidative stress," from the OMIM database, NCBI gene function module, and the GeneCard database, a total of 9469 human genes which are linked to oxidative stress were collected. Based on this, we intersected these oxidative stress-related genes with
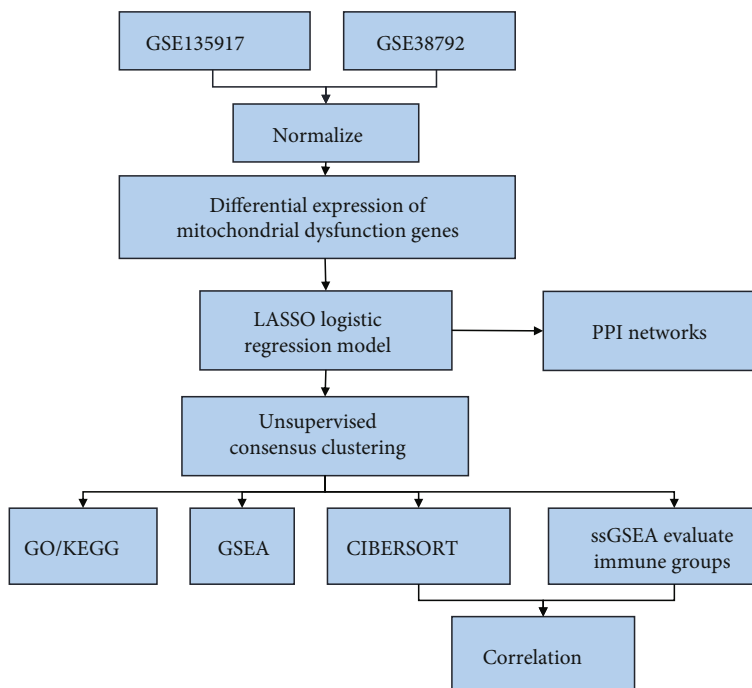
FIGURE 1: Flowchart for constructing and validating a prognostic model for TCGA-LUAD overall survival.

the differentially expressed genes in the TCGA-LUAD database to set and screen 147 oxidative stress-related genes with significant expression differences for subsequent analysis. We used a string-generated protein interaction (PPI) network database (version 11.0) of 147 oxidative stress-related genes differentially expressed in the TCGA-LUAD dataset to construct a molecular interaction network for analyzing closely interacting differential genes. Then, the PPI was exported, and the Cytoscape [22] software was used for further analysis, the network properties of each node were calculated, and the MCODE [23] and Cytohubba [24] were used to mine the hub nodes based on the degree of the nodes. The high level of linkage may have an extremely important function in the modulation of the whole biological process, which deserves further study.

*2.3. Creation and Validation of a Prognostic Oxidative Stress-Related Signature for LUAD.* We randomly divided 445 patients with complete clinical informatics data in TCGA-LUAD into a training set (2/3 of the total, $n = 296$) and a test set (1/3 of the total, $n = 149$). We then performed univariate Cox proportional hazards regression in the training set and validation set, respectively, to discover genes with prognostic value, utilizing the expression data of 147 DE-oxidative stress-related genes. A Cox $p$ value $< 0.05$ connoted a substantial correlation with overall survival (OS). The degree of correlation and prognostic value were selected using the "Venn" R package for cross-analysis and genes affecting prognostic value. To avert overfitting, all genes with $p$ values $< 0.05$ were subjected to a least absolute shrinkage and selection operator (LASSO) analysis using the package glmnet. LASSO regression is usually a regularization method for

high-dimensional predictor selection. The hazard system used LASSO Cox proportional hazards to build a score-identifying gene signature to anticipate OS models for LUAD patients. Prediction scores are weighted sums using developed genes, with coefficient regularization by LASSO. After being filtered by the LASSO model, the selected genes are constructed by the multivariate Cox proportional hazards model to construct an immune-related risk model: risk score = level of gene $a^*$ coefficient $a$ + level of gene $b^*$ coefficient $b$ + level of gene $c^*$ coefficient $c$ + $\cdots$ +level of gene $n^*$ coefficient $n$. The risk score in the model represents the prognosis of LUAD patients; the smaller the risk score, the better the prognosis. Patients were categorized into two risk groups utilizing the median risk score that served as a cutoff value. The "pheatmap" R package used scatter diagrams to show how the risk scores and survival durations of all patients were distributed. The "stats" R package used principal component analysis (PCA) to measure the gene expression of an established signature. Kaplan-Meier (K-M) survival analysis and time-dependent ROC analysis based on OS were carried out utilizing the "survival" package, the "survminer" package [25], and the "timeROC" package [26] in R to measure the prognostic accuracy of the gene signature in the derivation set and validate it in the validation set. Kaplan-Meier survival curves were used to derive the predictive power, whereas a log-rank $p$ value $< 0.05$ denoted a statistical significance (utilizing package survival and survminer). To ascertain predictive power of the immune signature's, time-dependent ROC curves (package survivalROC) were employed. Subsequently, logistic regression was used for correlation analysis of dichotomous clinical subgroup variables.

TABLE 1: Clinical characteristics of patients in TCGA-LUAD internal training set and validation set.

| Characteristics | Train ($N = 296$) | Test ($N = 149$) | Total ($N = 445$) | $p$ value | FDR |
|---|---|---|---|---|---|
| Age | | | | | |
| Mean ± SD | 64.86 ± 10.00 | 65.05 ± 10.21 | 64.93 ± 10.06 | | |
| Median [min-max] | 66.00 [33.00, 88.00] | 66.00 [41.00, 86.00] | 66.00 [33.00, 88.00] | | |
| Gender | | | | 1 | 1 |
| Female | 162 (36.40%) | 81 (18.20%) | 243 (54.61%) | | |
| Male | 134 (30.11%) | 68 (15.28%) | 202 (45.39%) | | |
| Stage | | | | 0.43 | 1 |
| Stage I | 157 (35.28%) | 87 (19.55%) | 244 (54.83%) | | |
| Stage II | 68 (15.28%) | 36 (8.09%) | 104 (23.37%) | | |
| Stage III | 54 (12.13%) | 21 (4.72%) | 75 (16.85%) | | |
| Stage IV | 17 (3.82%) | 5 (1.12%) | 22 (4.94%) | | |
| T | | | | 0.88 | 1 |
| T1 | 108 (24.27%) | 49 (11.01%) | 157 (35.28%) | | |
| T2 | 152 (34.16%) | 80 (17.98%) | 232 (52.13%) | | |
| T3 | 24 (5.39%) | 14 (3.15%) | 38 (8.54%) | | |
| T4 | 12 (2.70%) | 6 (1.35%) | 18 (4.04%) | | |
| M | | | | 0.52 | 1 |
| M0 | 198 (44.49%) | 100 (22.47%) | 298 (66.97%) | | |
| M1 | 17 (3.82%) | 5 (1.12%) | 22 (4.94%) | | |
| MX | 81 (18.20%) | 44 (9.89%) | 125 (28.09%) | | |
| N | | | | 0.19 | 0.93 |
| N0 | 186 (41.80%) | 105 (23.60%) | 291 (65.39%) | | |
| N1 | 54 (12.13%) | 27 (6.07%) | 81 (18.20%) | | |
| N2 | 49 (11.01%) | 17 (3.82%) | 66 (14.83%) | | |
| N3 | 2 (0.45%) | 0 (0.0e+0%) | 2 (0.45%) | | |
| NX | 5 (1.12%) | 0 (0.0e+0%) | 5 (1.12%) | | |

*2.4. Pathway and Function Enrichment Analysis.* To investigate the biological value of these differentially expressed IRGs, we utilized DAVID Bioinformatics Resources 6.8 to conduct a pathway and function enrichment analysis. The visualization procedure was carried out via the package "ggplot2" [25].

*2.5. Gene Set Enrichment Analysis (GSEA) of Oxidative Stress Prognostic Model in LUAD.* GSEA is a computational approach utilized to ascertain if a predefined set of genes exhibits statistical differences between two biological states, which is typically employed to obtain the expression in a dataset sample. Changes in the pathway as well as activity of biological processes were analyzed [23]. To investigate the variations in biological processes between the two groups of samples, utilizing the gene expression profile dataset, we retrieved the reference gene set "c2.cp.kegg.v7.4.entrez.gmt" from the MSigDB database, using the R package "clusterProfiler" The GSEA method is included in "enrichment analysis and visualization of datasets. $p$ value < 0.05 was considered statistically significant.

Gene set variation analysis (GSVA) [27] is a nonparametric unsupervised analysis method, which is mainly used to evaluate the microarray by transforming the expression matrix of genes between various samples into the expression matrix of gene sets between samples. Transcriptome gene set enrichment findings assess if various metabolic pathways are enriched across various samples. To investigate the biological process variation occurring in the two groups of samples, we utilized the R package GSVA to carry out the GSVA procedure based on the gene expression profile dataset and downloaded the reference "c2.cp.kegg.v7.4. entrez" gene set from the MSigDB database to compute the enrichment score of every sample in each pathway in the dataset, combined with the R package limma to screen significantly different pathways, the GSVA enrichment results were visualized based on the heatmap utilizing the R package pheatmap, and $p$ value < 0.05 denoted a statistically significant significance.

*2.6. Establishment and Verification of Nomogram.* Nomogram has been extensively utilized for anticipating cancer-related prognosis. This approach allows individualized approximates of the likelihood of recurrence, mortality, or drug adherence. Using the prognostic model, this investigation developed the nomogram in the TCGA-LUAD cohort by incorporating the above-stated clinical parameters to anticipate the OS probably over 1, 3, and 5 years.

| Features | p-value | Hazard ratio (95%CI) |
|---|---|---|
| HAO1 | 1.1e-289 | 21.69 (18.37-25.60) |
| HTR1B | 1.2e-273 | 0.01 (0.01-0.02) |
| APOB | 1.2e-268 | 13.46 (11.64-15.57) |
| SLC6A5 | 4.6e-211 | 11.41 (9.78-13.30) |
| SCARNA22 | 1.3e-209 | 2.06 (1.97-2.16) |
| MUC2 | 5.3e-194 | 7.34 (6.43-8.37) |
| PRODH2 | 9.4e-166 | 2.0e-13 (2.4e-14-1.6e-12) |
| KCNA4 | 2.7e-164 | 4.2e-4 (2.4e-4-7.4e-4) |
| FOXL2 | 1.4e-121 | 138.42 (91.67-209.01) |
| NPSR1 | 1.6e-94 | 92.01 (59.87-141.41) |
| MRGPRX3 | 4.4e-69 | 56298651.63 (7687524.49-412296335.40) |
| TNR | 3.8e-34 | 0.07 (0.04-0.10) |
| FGF5 | 1.2e-33 | 0.37 (0.32-0.44) |
| MAP3K19 | 8.1e-29 | 0.11 (0.07-0.16) |
| NR1H4 | 4.0e-28 | 3.07 (2.51-3.75) |
| GABRR3 | 1.9e-24 | 6.5e-64 (4.8e-76-8.9e-52) |
| DNAH8 | 3.5e-24 | 216326752036.15 (1397552079.89-33485166184404.80) |
| CA6 | 3.5e-23 | 22.50 (12.16-41.62) |
| F11 | 3.6e-22 | 2.33 (1.96-2.76) |
| CHRM1 | 1.1e-20 | 3.70 (2.81-4.88) |
| DDX25 | 7.8e-19 | 8.66 (5.37-13.96) |
| NOS2P2 | 2.0e-17 | 1.8e-14 (1.2e-17-2.6e-11) |
| TTC29 | 2.0e-16 | 3.45 (2.57-4.63) |
| HYDIN | 7.9e-15 | 51.34 (19.01-138.64) |
| GABRA5 | 1.8e-14 | 18.52 (8.78-39.05) |
| GIP | 1.0e-13 | 0.46 (0.38-0.57) |
| C10orf90 | 3.9e-13 | 0.04 (0.02-0.10) |
| TNFRSF13B | 5.7e-12 | 0.41 (0.32-0.53) |
| ANGPTL3 | 3.7e-11 | 5.47 (3.31-9.06) |
| CPN1 | 3.8e-11 | 0.06 (0.03-0.14) |
| KCNK9 | 4.5e-9 | 0.43 (0.32-0.57) |
| NEUROD1 | 3.5e-8 | 0.20 (0.11-0.35) |
| UGT3A1 | 4.5e-8 | 0.32 (0.21-0.48) |
| TRIML2 | 4.8e-8 | 0.27 (0.17-0.43) |
| PTPRQ | 7.1e-5 | 0.34 (0.20-0.57) |
| UPB1 | 8.7e-5 | 0.10 (0.03-0.32) |
| CLCA1 | 9.9e-4 | 4.53 (1.84-11.13) |
| SLC7A3 | 3.1e-3 | 0.28 (0.12-0.65) |

(a)

Figure 2: Continued.

FIGURE 2: Univariate Cox analysis based on TCGA-LUAD internal training set and validation set and screening of differentially oxidative stress genes associated with LUAD prognosis. (a) Forest plot displays the findings of univariate Cox analysis of differential oxidative stress genes associated with LUAD prognosis in the training set; (b) forest plot demonstrates the findings of univariate Cox analysis of differential oxidative stress genes associated with LUAD prognosis in the training set; (c) Venn shows 35 prognosis-related differential oxidative stress genes at the intersection of the two sets.

(a)

(b)

(c)

(d)

(e)

(f)
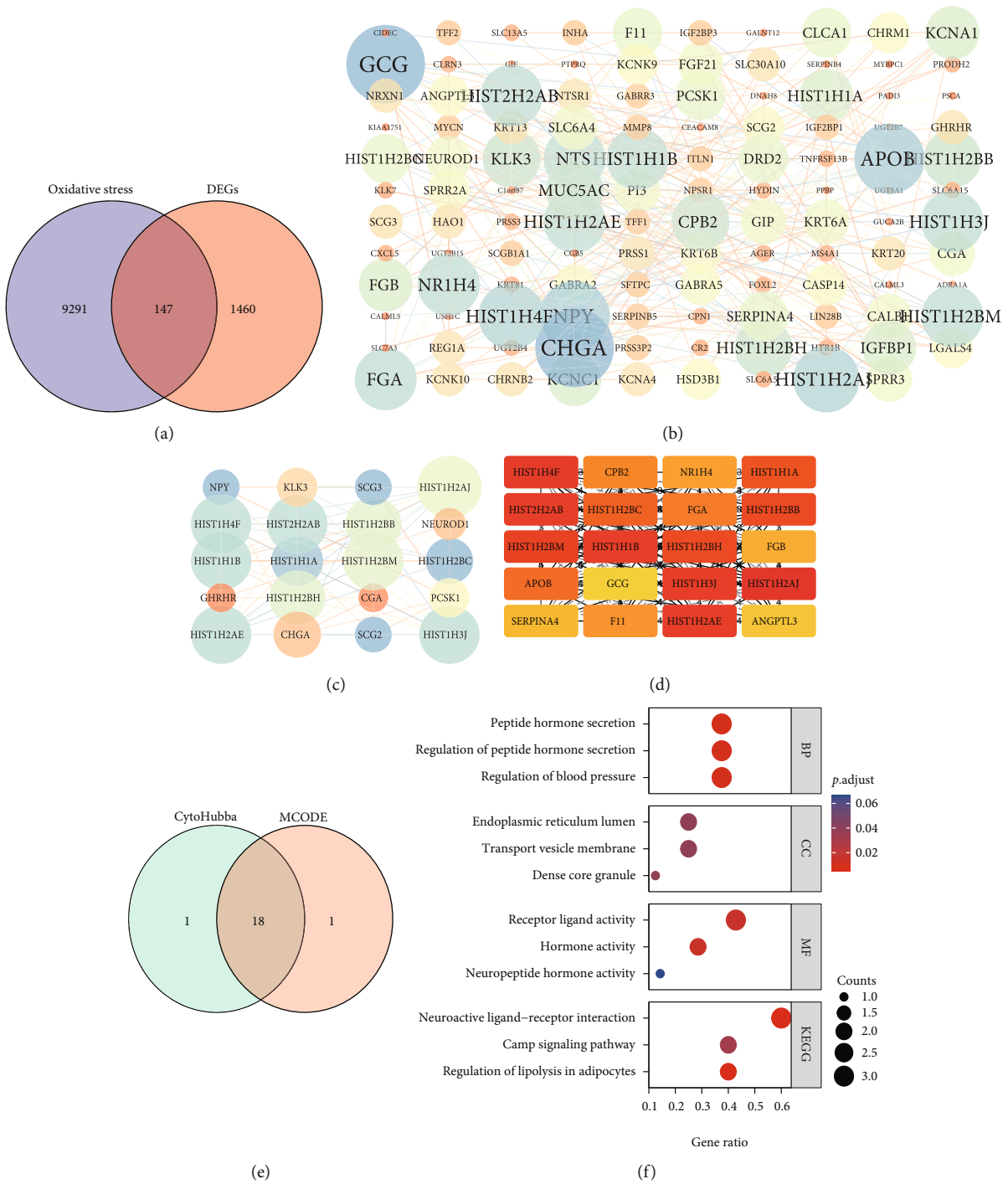
Figure 3: Continued.
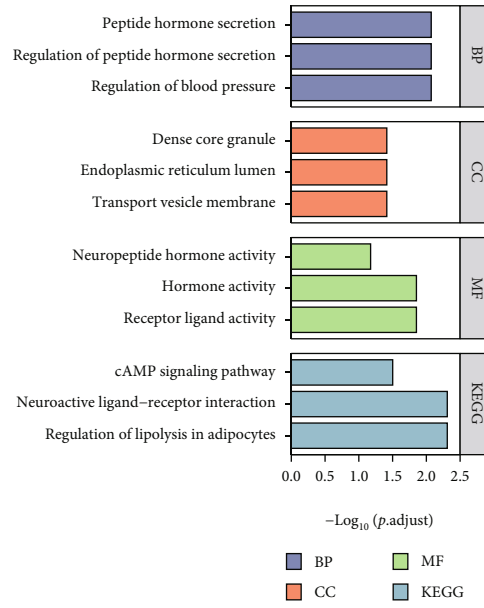
(g)

FIGURE 3: PPI network construction of differentially expressed genes related to oxidative stress and enrichment analysis of hub genes. (a) 1607 differentially expressed genes ($|LogFC| > 2$, adj. $p$ value $< 0.05$) between normal samples and LUAD samples in the TCGA-LUAD dataset, and 9469 oxidative stress genes were intersected, and 147 differentially expressed genes were obtained oxidative stress genes; (b) the Network Analyzer tool of Cytoscape (v3.7.2) visualizes the PPIs of 147 differentially expressed oxidative stress genes with the largest confidence interaction score of 0.4. As the degree of interaction increases, the color gradually changes from yellow to blue, and the font changes from small to large; (c) the MCODE plug-in screened and visualized the closely related genes of the PPI network module; (d) the CytoHubba plug-in was used to screen the top 20 closely related genes; (e) the intersection of the two methods is shown by Venn diagram to obtain 18 closely related differentially oxidative stress-related genes; (f, g) based on 18 Hub Gene Ontology (GO) enrichment analysis of genes as well as KEGG pathway enrichment bubble plots and histograms.

Furthermore, calibration curves evaluated the fitness between actual survival statuses with visualized survival status of the developed nomogram via bootstrap methods (1,000 replicates). The values of prognosis evaluation between risk signature, stage, and the nomogram were compared via ROC curves at 1 year, 3 years, and 5 years, correspondingly.

2.7. Computation of Immune Infiltration Score Was Premised on the Gene Groups Previously Determined. We downloaded the unified as well as the standardized pan-cancer dataset from the UCSC (https://xenabrowser.net/) database: TCGA TARGET GTEx (PANCAN, $N = 19131$, $G = 60499$). Additionally, we extracted gene expression data in every sample and further screened the sample sources: Primary Blood-Derived Cancer-Peripheral Blood (TCGA-LAML), Primary Tumor, Metastatic TCGA-SKCM, Primary Blood-Derived Cancer-Bone Marrow, Primary Solid Tumor, and Recurrent Blood-Derived Cancer-Bone Marrow samples, further log2 $(x + 0.001)$ transformation was carried out on every expression value; in addition, we also extracted the gene expression profile of every tumor, respectively, and mapped the expression profile. On the GeneSymbol, the R software package ESTIMATE (version 1.0.13, https://bioinformatics .mdanderson.org/public-software/estimate/, doi:10.1038/ ncomms3612) [28] was further used to calculate the gene

expression in each tumor, stromal, immune, and ESTI-MATE scores for each patient.

2.8. GO, KEGG, and Immune Infiltration Enrichment Analyses for Risk-Related DEGs. In accordance with the risk grouping, normalized gene expression matrixes of the derivation set as well as the validation set generated above were employed with the "limma" R package to detect risk-related DEGs with the cut-off criteria of $|logFC| \geq 1$ and adj.$p < 0.05$, correspondingly. Risk-related DEGs were analyzed with both GO and KEGG utilizing the "clusterProfiler" R package. Next, single-sample GSEA (ssGSEA) for immune infiltration was adopted with the "GSVA" R package to ascertain the infiltration score of immune cells and the immune-related roles.

2.9. Calculation of Immune Infiltration Score. CIBERSORT played a crucial function in calculating an absolute immune infiltrate score for the primary tumor samples. The default CIBERSORT parameters were instrumental in generating the curated CIBERSORT signature matrix. It roughly yielded the expected relative abundances. We performed a pan-cancer analysis of the CIBERSORT score, including the correlation of immune score in pan-cancer, analyzed the correlation of MAP3K19 and NTSR1 with the distribution of various immune cells for LUAD patients, and visualized them with heatmaps and scatter plots. The tumor purity,

TABLE 2: List of GO and KEGG enrichment analysis results of differential oxidative stress hub genes with close interaction.

| Ontology | ID | Description | $p$ value | $p$.adjust | $Q$ value |
|---|---|---|---|---|---|
| BP | GO:0008217 | Regulation of blood pressure | 4.92e-05 | 0.008 | 0.004 |
| | GO:0090276 | Regulation of peptide hormone secretion | 7.32e-05 | 0.008 | 0.004 |
| | GO:0030072 | Peptide hormone secretion | 1.26e-04 | 0.008 | 0.004 |
| | GO:0007187 | G protein-coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger | 1.36e-04 | 0.008 | 0.004 |
| | GO:0046883 | Regulation of hormone secretion | 1.52e-04 | 0.008 | 0.004 |
| CC | GO:0030658 | Transport vesicle membrane | 0.003 | 0.039 | 0.024 |
| | GO:0005788 | Endoplasmic reticulum lumen | 0.006 | 0.039 | 0.024 |
| | GO:0031045 | Dense core granule | 0.008 | 0.039 | 0.024 |
| | GO:0042629 | Mast cell granule | 0.009 | 0.039 | 0.024 |
| | GO:0030133 | Transport vesicle | 0.010 | 0.039 | 0.024 |
| MF | GO:0048018 | Receptor ligand activity | 6.48e-04 | 0.014 | 0.008 |
| | GO:0005179 | Hormone activity | 9.68e-04 | 0.014 | 0.008 |
| | GO:0005184 | Neuropeptide hormone activity | 0.011 | 0.067 | 0.036 |
| | GO:0071855 | Neuropeptide receptor binding | 0.013 | 0.067 | 0.036 |
| | GO:0042056 | Chemoattractant activity | 0.015 | 0.067 | 0.036 |
| KEGG | hsa04923 | Regulation of lipolysis in adipocytes | 4.83e-04 | 0.005 | 0.001 |
| | hsa04080 | Neuroactive ligand-receptor interaction | 7.00e-04 | 0.005 | 0.001 |
| | hsa04024 | cAMP signaling pathway | 0.007 | 0.032 | 0.009 |
| | hsa04950 | Maturity onset diabetes of the young | 0.016 | 0.056 | 0.017 |
| | hsa04913 | Ovarian steroidogenesis | 0.031 | 0.064 | 0.019 |

as well as stromal scores, were determined utilizing the Estimation of Stromal and Immune Scores ESTIMATE based on RNA-seq data and global proteomic data.

*2.10. Quantitative Reverse Transcription-Polymerase Chain Reaction (qRT-PCR).* Total RNA extracted from lung adenocarcinoma tissue and para-cancerous tissue with Trizol Reagent (Invitrogen, USA) were reverse transcribed with HiScript III 1st Strand cDNA Synthesis Kit (Vzayme, China). Next, HiScript II One Step RT-PCR Kit (Vzayme, China) and qRT-PCR analysis were used for detecting cDNA expression levels, and GAPDH was used as internal reference. Primers were shown as follows: GAPDH, forward (F): 5′-AATGGGCAGCCGTTAGGAAA-3′, reverse(R): 5′-GCCCAATACGACCAAATCAGAG-3′; NTR1, forward (F): 5′-TCATCGCCTTTGTGGTCTGCT-3′, reverse (R): 5′-TGGTTGCTGGACACGCTGTCG-3′; MAP3K19, forward (F): 5′-AGGAGTTCGACCAAGATGGTG-3′, reverse (R): 5′-GGTCGAAAACTCTTCTGTCCTG-3′.

*2.11. Immunohistochemistry (IHC) Staining.* After deparaffinization and dehydrating the tissue sections, they were subjected to epitope retrieval, treated with $H_2O_2$, and blocked against nonspecific bindings. The tissues were then incubated overnight with anti-NTSR1 antibodies (1 : 100, Abcam, ab217134) and anti-MAP3K19 antibodies (1 : 100, Invitrogen, PA5-29285) at 4°C. Subsequently, the tissue sections were incubated with secondary antibodies (1 : 1000, Proteintech, SA00001-2) for two hours at ambient tempera-

ture. The signal was detected with an enhanced DAB staining kit (Proteintech, China).

*2.12. Statistical Analysis.* The Spearman correlation test was performed to investigate the link between two variables that were nonlinearly linked. The Student's $t$-test, on the other hand, was utilized in comparing the normally distributed data, whereas the chi-square test was carried out to contrast pairwise and categorical features in various subgroup. Univariate as well as multivariate Cox regression analyses assessed the influences of the immune signature and numerous clinic-pathological parameters on the survival of patients. The package pheatmap was vital in plotting heatmaps. A two-sided $p < 0.05$ was statistically significant. Kruskal-Wallis test was employed for one independent variable with two or more levels and an ordinal dependent variable. K-M analysis measured the proportion of individuals living for a particular period, whereas the log-rank test evaluated the significance of differences. In these investigations, statistical analysis was performed by R software (4.0.0). A two-tailed $p$ value of <0.05 was statistically significant unless otherwise stated.

# 3. Result

*3.1. The Characteristics of Patients.* RNA-sequencing profiles of a total of 445 LUAD sample as well as clinic-pathological data were obtained from the UCSC Xena TCGA-LUAD dataset. Then, randomization was conducted to divide them
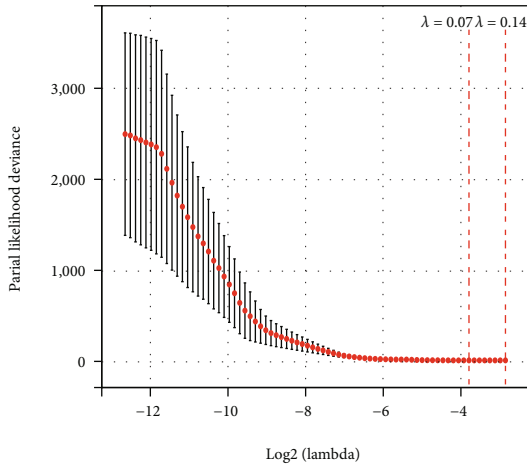
(a)
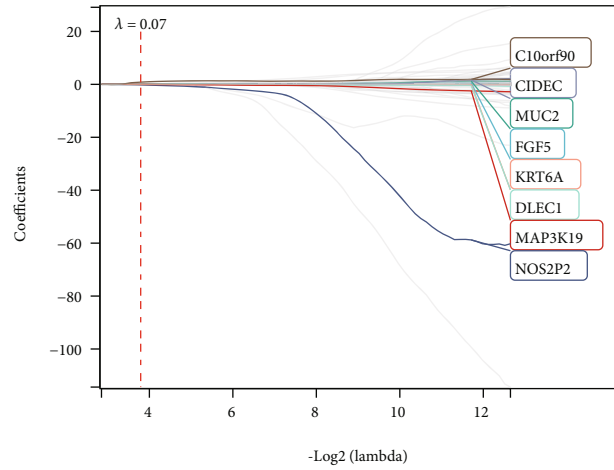


(b)



(c)



(d)

FIGURE 4: Continued.

(e)
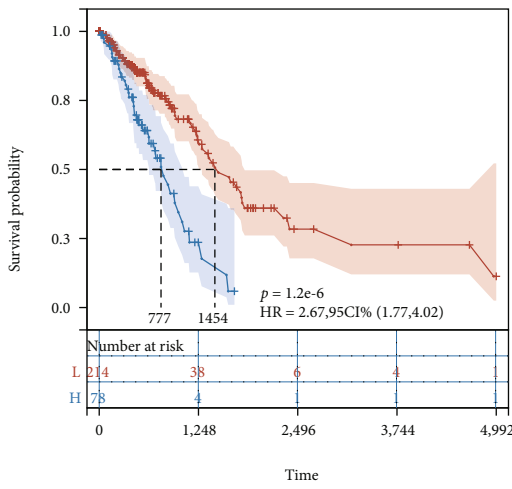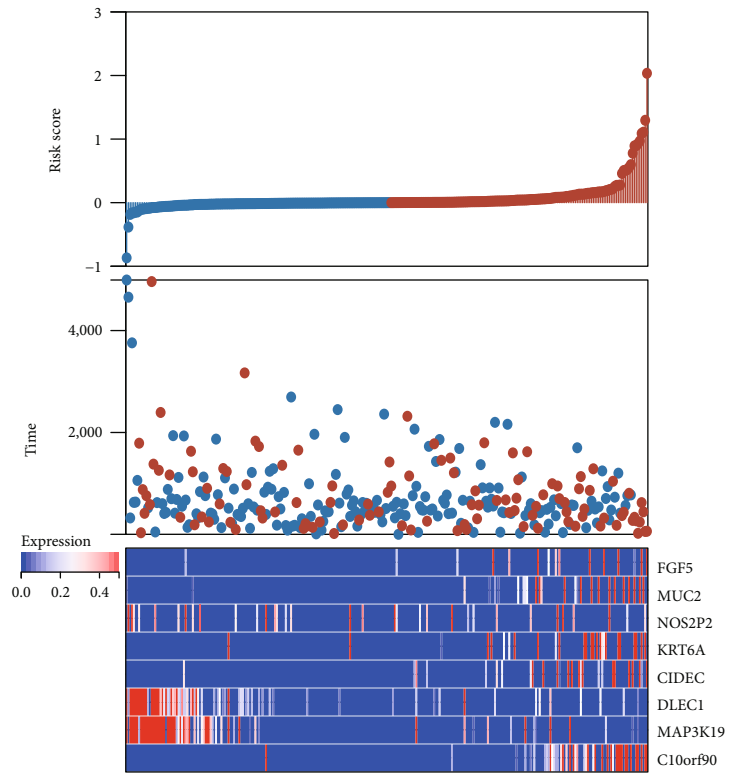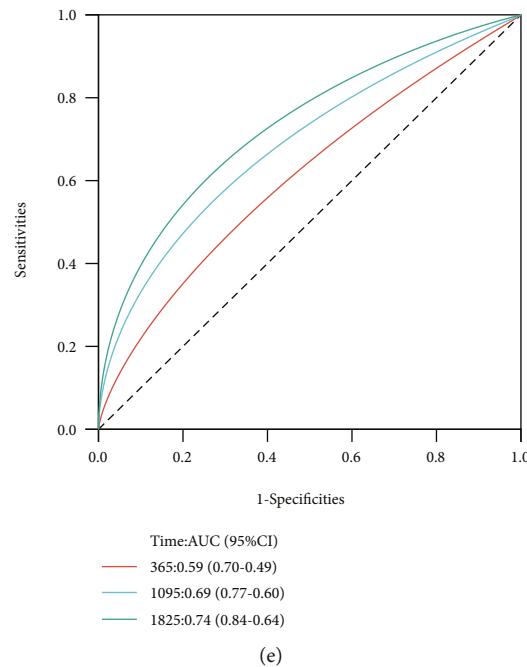
Figure 4: LASSO analysis of 147 differentially oxidative stress genes revealed the distribution and prognostic analysis of 8 gene markers. (a) Cross-validation of tuning parameter selection in the LASSO model; (b) LASSO coefficient spectrum of 8 prognostic differentially oxidative stress-related genes; (c) OS-based K-M survival of patients in the two risk groups in the TCGA-LUAD training set curves; (d) risk analysis graph of 8 prognostic differential oxidative stress gene signatures; the upper panel shows the distribution as well as the median of risk scores in the TCGA-LUAD training set; the middle panel shows the distribution of patients in various risk groups; the lower panel heatmap for the differential expression of 8 genes in the two risk groups; (e) the prognostic performance of risk scores in the TCGA-LUAD training set at 1, 3, and 5 years is validated utilizing AUC of time-dependent ROC curves.

into two groups, namely, the training set (331 patients) and the test set (166 patients) (Table 1).

### 3.2. Prognosis-Related Differential Oxidative Stress Gene Signatures Are Independent Prognostic Factors for LUAD Patients.

To determine the prognostic value of differentially oxidative stress genes, we did a univariate Cox regression analysis. There were 100 and 56 xenooxidative stress genes significantly linked to the overall survival (OS) in the training and validation sets, respectively (Figures 2(a) and 2(b)). A total of 35 differentially oxidative stress genes associated with prognosis were obtained by intersecting them (Figure 2(c)).

### 3.3. Construction of PPI Network of DEGs Related to Oxidative Stress.

We first performed a differential analysis between 59 normal samples and 535 LUAD samples in the TCGA-LUAD dataset using the limma algorithm and obtained a total of 1607 differentially expressed genes ($|logFC| > 2$, adj. $p$ value $< 0.05$). Based on this differential result, 9469 previously identified oxidative stress genes were intersected, and 147 differentially expressed oxidative stress genes were obtained (Figure 3(a)). The STRING database constructed a protein-protein interaction network (PPI) to reflect the intermolecular interactions, and the largest confidence interaction score was established at 0.4, which was analyzed and visualized by the Network Analyzer tool of

Cytoscape (v3.7.2) (Figure 3(b)). PPI network modules were screened for closely linked genes using the MCODE plugin and visualized (Figure 3(c)). At the same time, the Cyto-Hubba plugin was used to screen the Top20 closely linked genes (Figure 3(d)). The intersection of the two approaches was shown by the Venn diagram, obtaining 18 closely related differentially oxidative stress-related genes (Figure 3(e)), including HIST2H2AB, HIST1H2BC, SCG3, HIST1H2AJ, CHGA, HIST1H3J, HIST1H1B, HIST1H2BM, HIST1H1A, SCG2, HIST1H4F, HIST1H2AE, GHRHR, KLK3, NEUROD1, NPY, HIST1H2BB, HIST1H2BH, and CGA.

### 3.4. Functional Enrichment Analysis.

We analyzed these 18 closely related differentially oxidative stress-related genes. Gene Ontology (GO) enrichment analysis employing the above genes suggests that these hub genes exist in the membrane of transport vesicles, endoplasmic reticulum lumen, mast cell granules, and transport vesicles, which can affect blood pressure, hormone secretion, and G protein. Coupling cyclic, coupling receptor signaling pathway, nucleotide second messenger, and other functions play a regulatory role and have certain applications for receptor-ligand activity, neuropeptide hormone activity, neuropeptide receptor binding, and chemotactic activity value. The enrichment of BP set was mostly concentrated in oxidative phosphorylation, mitochondrial translation elongation, mitochondrial

(a)

(b)

(c)

(d)

Figure 5: Continued.

(e)

(f)
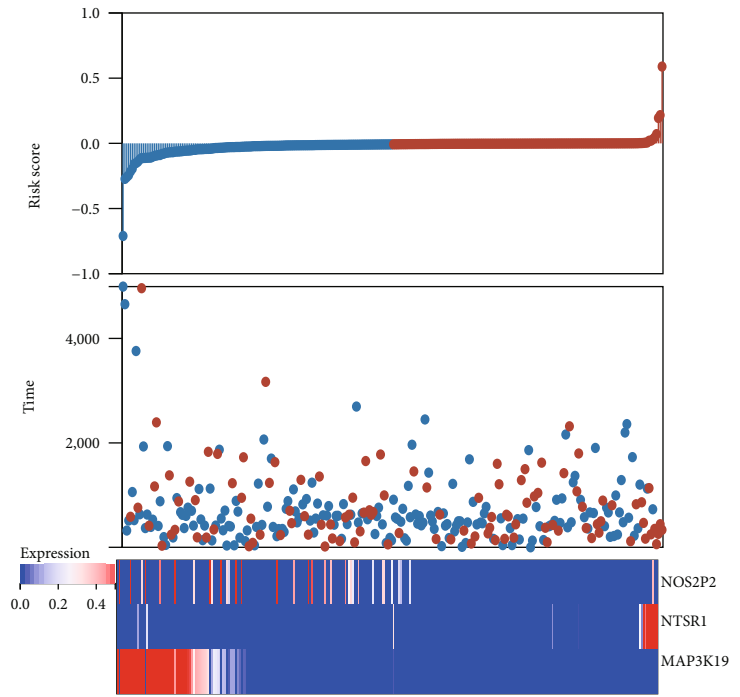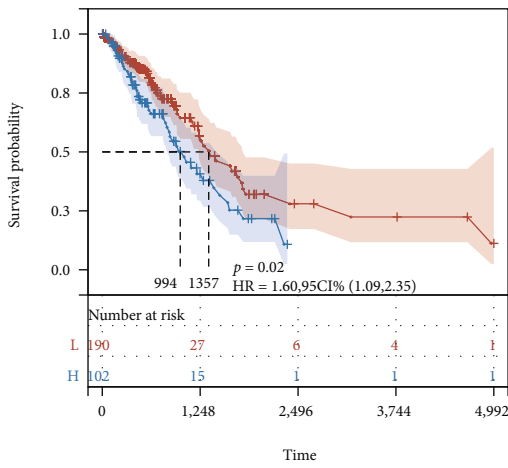
FIGURE 5: LASSO analysis of 35 differentially oxidative stress genes in the TCGA-LUAD training set showing the distribution and prognostic analysis of 3 gene signatures. (a) Cross-validation of tuning parameter selection in the LASSO model; (b) LASSO coefficient spectrum of 3 prognostic differentially oxidative stress-related genes; (c) OS-based K-M survival of patients in the two risk groups in the TCGA-LUAD training set curves; (d) 3 prognostic differential oxidative stress gene signature risk analysis graphs; the upper graph shows the distribution as well as the median of risk scores in the TCGA-LUAD training set; the middle graph shows the distribution of individuals in various risk groups, and the lower graph shows heatmap of differential expression of 3 genes in the two risk groups; (e) the prognostic performance of risk scores in the TCGA-LUAD training set over 1, 3, and 5 years is validated utilizing AUC of time-dependent ROC curves; (f) based on the training set, multivariate Cox proportional hazards regression forest plot for a 3-gene prognostic model.

translation termination, translation termination, and purine ribonucleoside triphosphate metabolism. CC enrichment was mostly concentrated in the inner mitochondrial membrane, mitochondrial protein complexes, mitochondrial matrix, organelle ribosomes, and mitochondrial ribosomes. The enrichment of MF was mostly concentrated in the structural components of ribosomes, proton transmembrane, transporter activity, electron transfer activity, cytochrome-c oxidase activity, and heme-copper terminal oxidase activity. This suggests that the Hub gene may be related to the transmembrane transport of cells in terms of molecular structure. Combined with all the above enrichment results, we speculate that key genes related to oxidative stress may play a role in the interaction process and related mechanisms between cells. The KEGG pathway enrichment results indicated that differential oxidative stress hub genes were enriched in modulation of lipolysis in adipocytes, neuroactive ligand-receptor interaction, cAMP signaling pathway, maturity onset diabetes of the young, ovarian steroidogenesis, and other pathways. This result suggests that the differentially expressed oxidative stress-related genes can affect the occurrence and progression of LUAD via the above potential biological functions and molecular pathways (Figures 3(f) and 3(g) and Table 2).

3.5. Establishment and Validation of Prognostic Gene Signature Associated with Oxidative Stress in TCGA-LUAD. To avert overfitting, we additionally performed a LASSO-

Cox analysis. In order to avoid the influence of confounding factors, we first performed LASSO-Cox analysis on 147 differentially oxidative stress genes in the training set and established 8 gene signatures, including C10orf90, CIDEC, MUC2, FGF5, KRT6A, DLEC1, MAP3K19, and NOS2P2. Patients were categorized into high- and low-risk groups as per the median risk score with each group containing 167 participants. To tune parameter selection through the LASSO model, we utilized cross-validation (Figure 4(a)). The coefficient profiles of LASSO for the eight prognostic differentially oxidative stress-related genes were shown (Figure 4(b)). At the same time, we show eight prognostic differential oxidative stress gene signature risk analysis graphs (Figure 4(d)). The upper graph shows the distribution as well as the median of risk scores in the training set of TCGA-LUAD, and the middle graph shows the distribution of individuals in various risk groups. The figure below is a heatmap of the differential expression of 8 genes in the two risk groups. Further assessment of the gene signatures' prognostic value as well as the predictive performance utilizing both K-M survival and time-dependent ROC analyses was done, both of which yielded remarkable results. To account for survival outcomes, we observed a statistically significantly higher number of dead participants in the high-risk group in contrast with the low-risk group ($p = 1.2e - 6$, HR $= 2.67\,(1.77, 4.02)$) (Figure 4(c)). The AUC reached 0.59 (0.70-0.49) at 1 year, 0.69 (0.77-0.60) at 3 years, and 0.74 (0.84-0.64) at 5 years (Figure 4(e)).

(a)



(b)



(c)



(d)

FIGURE 6: Continued.

(e)
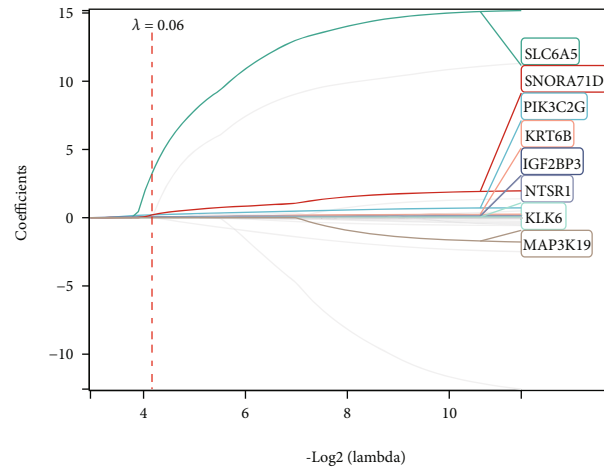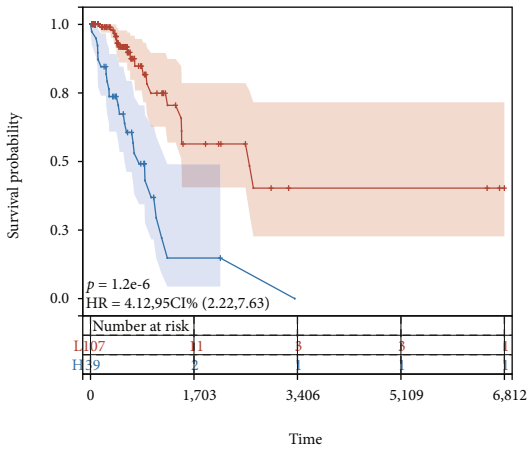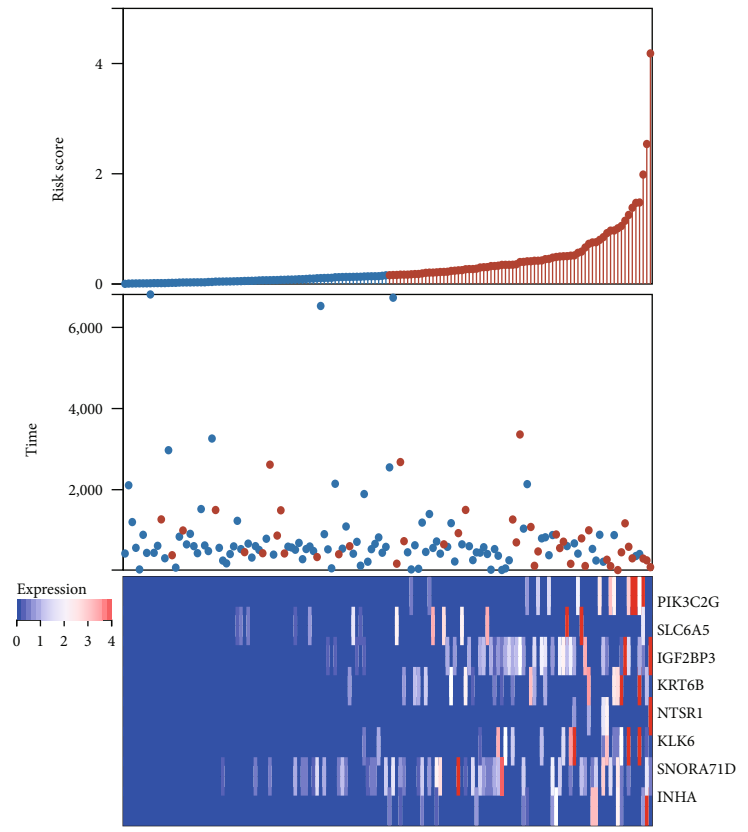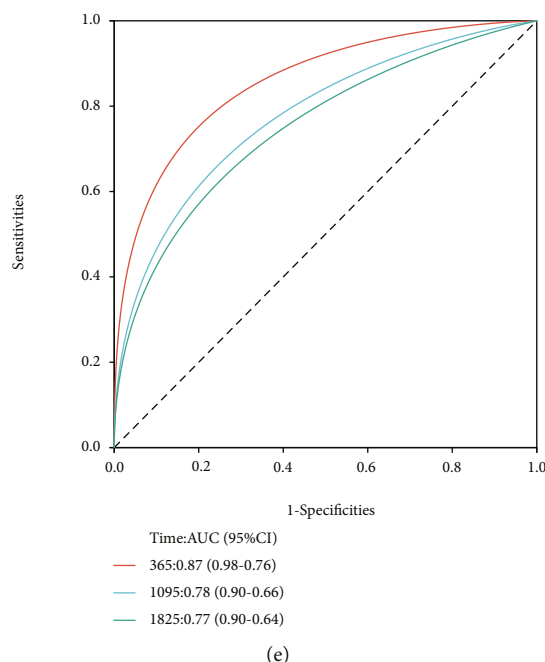
FIGURE 6: LASSO analysis of 35 differentially oxidative stress genes in the TCGA-LUAD validation set showed the distribution and prognostic analysis of 8 gene markers. (a) Cross-validation of tuning parameter selection in the LASSO model; (b) LASSO coefficient spectrum of 8 prognostic differentially oxidative stress-related genes; (c) OS-based KM survival of patients in the two groups in the TCGA-LUAD training set Curves; (d) 8 prognostic differential oxidative stress gene signature risk analysis graphs; the upper graph shows the distribu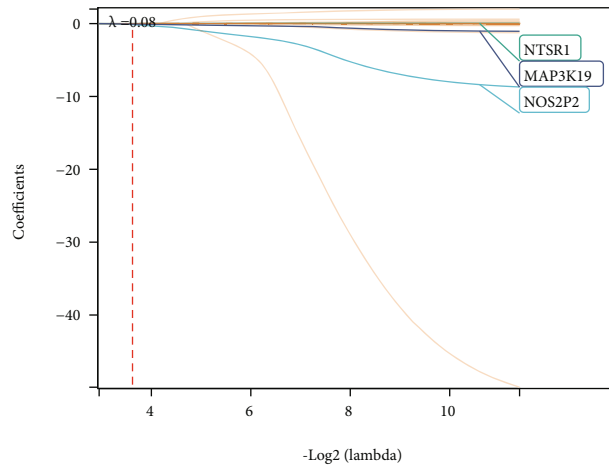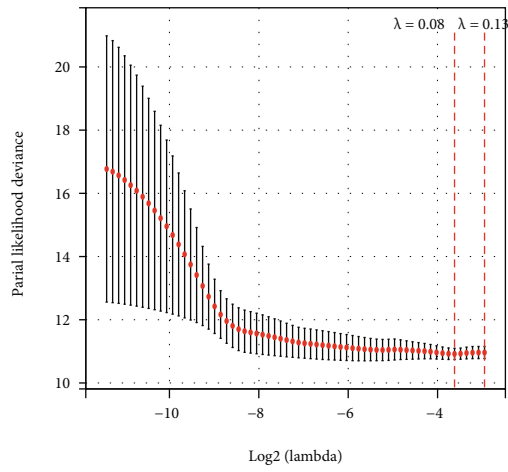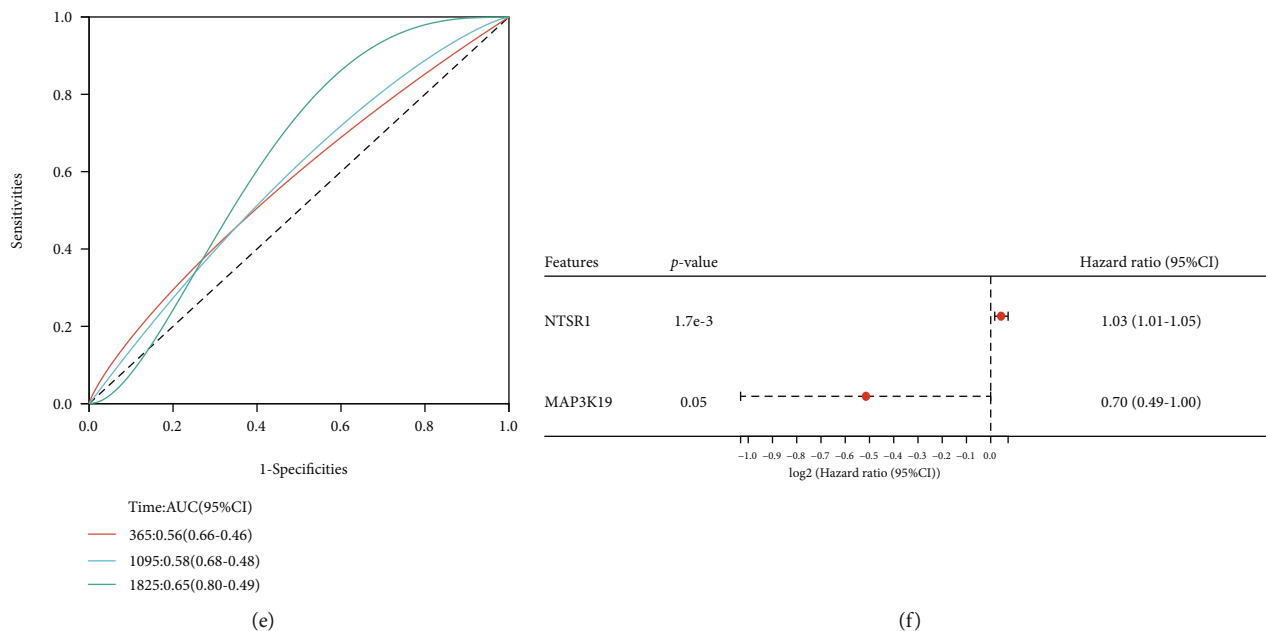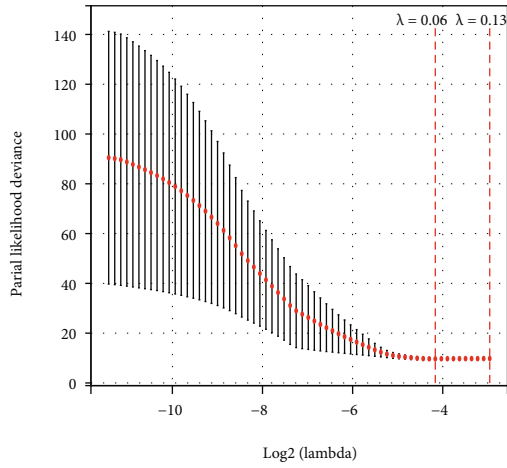tion as well as the median of risk scores in the TCGA-LUAD training set; the middle graph shows the distribution of individuals in various risk groups, and the lower graph shows heatmap of differential expression of 8 genes in both risk groups; (e) AUC of time-dependent ROC curves validates the prognostic performance of risk scores in the TCGA-LUAD training set at 1, 3, and 5 years.

Furthermore, LASSO-Cox analysis was conducted on the above 35 prognosis-related differential oxidative stress genes in the training set and validation set, correspondingly. A 3-gene prognostic model was predicted among the genes in the training set, including NTSR1, MAP3K19, and NOS2P2 as predictors of patient prognosis (Figures 5(a) and 5(b)). Not only K-M survival but also time-dependent ROC analysis findings were significant. The high-risk group had a greater mortality risk in contrast with the low-risk group ($p = 0.02$, HR $= 1.60\,(1.09, 2.35)$) (Figure 5(c)). We show 3 prognostic differential oxidative stress gene signature risk analysis graphs, the upper graph shows the distribution and median of risk scores in the TCGA-LUAD training set, and the heatmap reflects the expression differences of model genes (Figure 5(d)). The AUC reached 0.56 (0.66-0.46) at 1 year, 0.58 (0.68-0.48) at 3 years, and 0.65 (0.80-0.49) at 5 years (Figure 5(e)). In the multivariate Cox regression analysis, NTSR1, MAP3K19, and NOS2P2 were incorporated. Multivariate Cox proportional hazards regression based on the training set suggested that only NTSR1 and MAP3K19 were jointly used as predictors of poor prognosis at high risk of LUAD, with significant statistical significance (Figure 5(f)).

An 8-gene prognostic model was predicted among the genes in the validation set, including SLC6A5, SNORA71D, PIK3C2G, KRT6B, IGF2BP3, NTSR1, KLK6, and MAP3K19 as predictors of patient prognosis (Figures 6(a) and 6(b)). Both K-M survival and time-dependent ROC analysis find-ings were significant. The high-risk group had a greater mortality risk in comparison with the low-risk group ($p = 1.2e - 6$, HR $= 4.12\,(2.22, 7.63)$) (Figure 6(c)). We show risk analysis plots for eight prognostic differential oxidative stress gene signatures, the upper panel shows the distribution and median of risk scores in the TCGA-LUAD training set, the middle panel shows the distribution of individuals in each risk group, and the lower panel shows heatmap of differential expression of 8 genes in the two risk groups (Figure 6(d)). The AUC reached 0.87 (0.98-0.76) at 1 year, 0.78 (0.90-0.66) at 3 years, and 0.77 (0.90-0.64) at 5 years (Figure 6(e)).

In the prediction model constructed based on LASSO-Cox analysis of 147 differentially oxidative stress-related genes and 35 prognostic differentially oxidative stress-related genes in the validation set, both MAP3K19 and NOS2P2 were found to be prognostic predictors, although NOS2P2 in the further multivariate Cox regression analysis results were suggested to have no statistically significant effect on prognosis. More interestingly, LASSO-Cox analysis based on 35 prognostic differentially oxidative stress-related genes found that MAP3K19 and NTSR1 could serve as prognostic predictors in both training and validation sets. Based on the above analysis, we believe that MAP3K19 and NTSR1 may have potential application value in predicting the prognosis of LUAD, and more abundant analysis is urgently needed to evaluate its application value and important clinical significance in LUAD.

(a)



Time:AUC(95%CI)
— 365:0.75(0.84-0.66)
— 1095:0.73(0.81-0.65)
— 1825:0.76(0.86-0.66)

(b)

Figure 7: Continued.

(c)
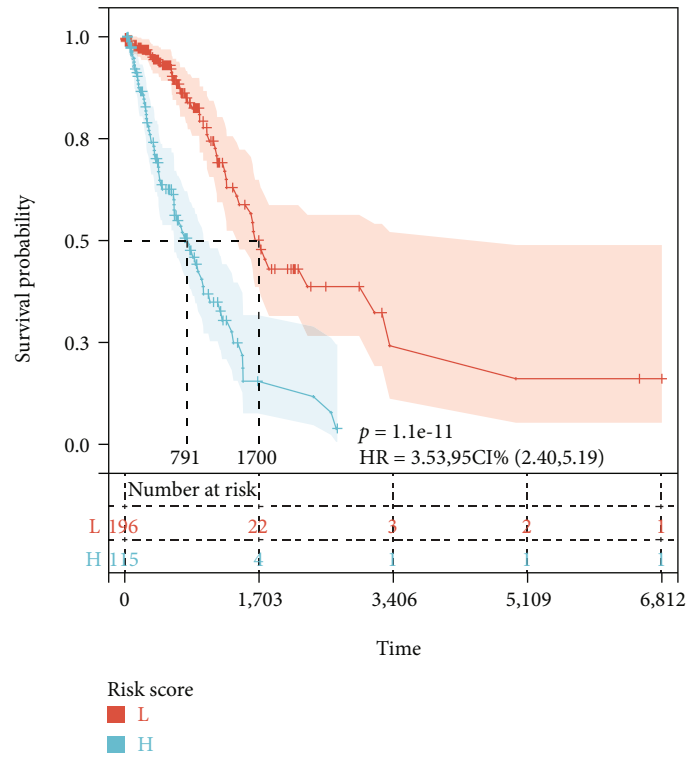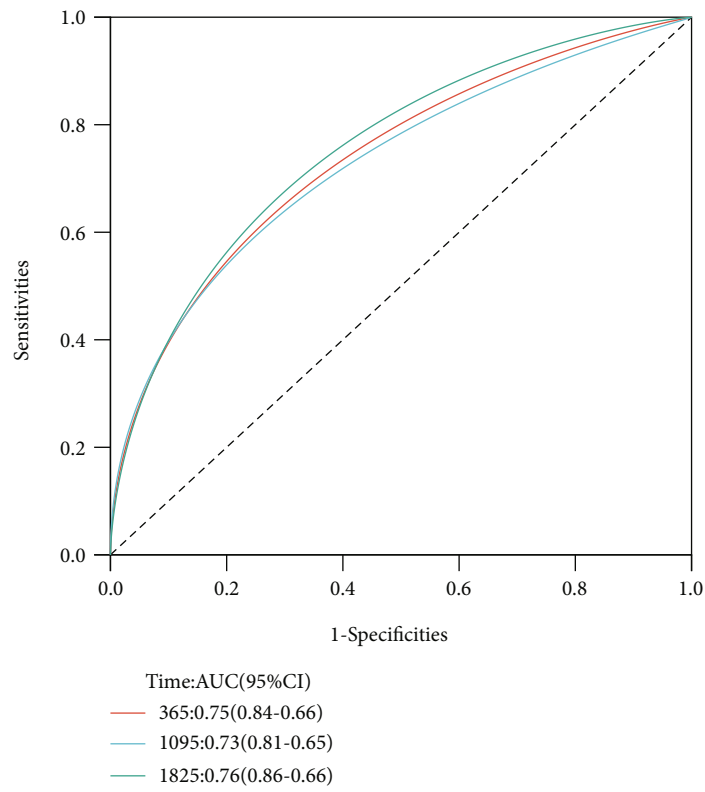


(d)

FIGURE 7: Continued.

(e)



(f)
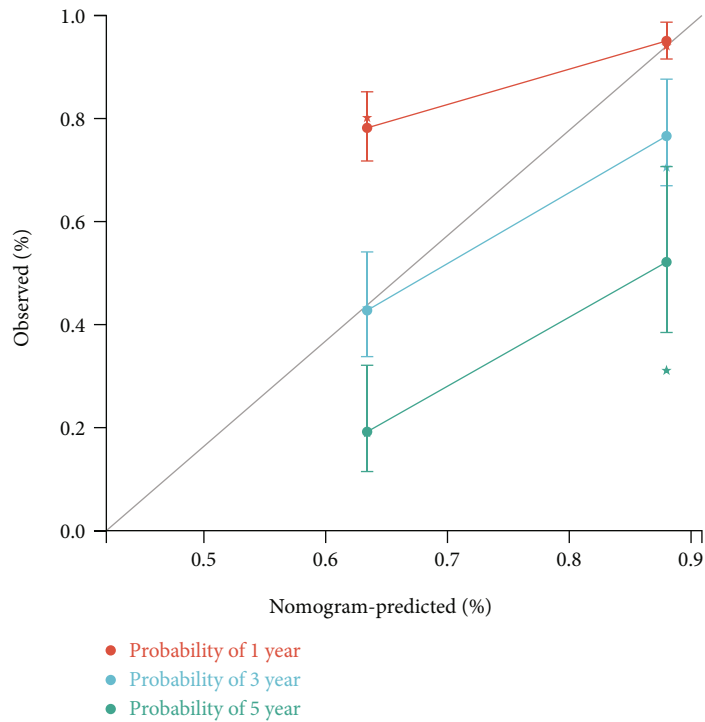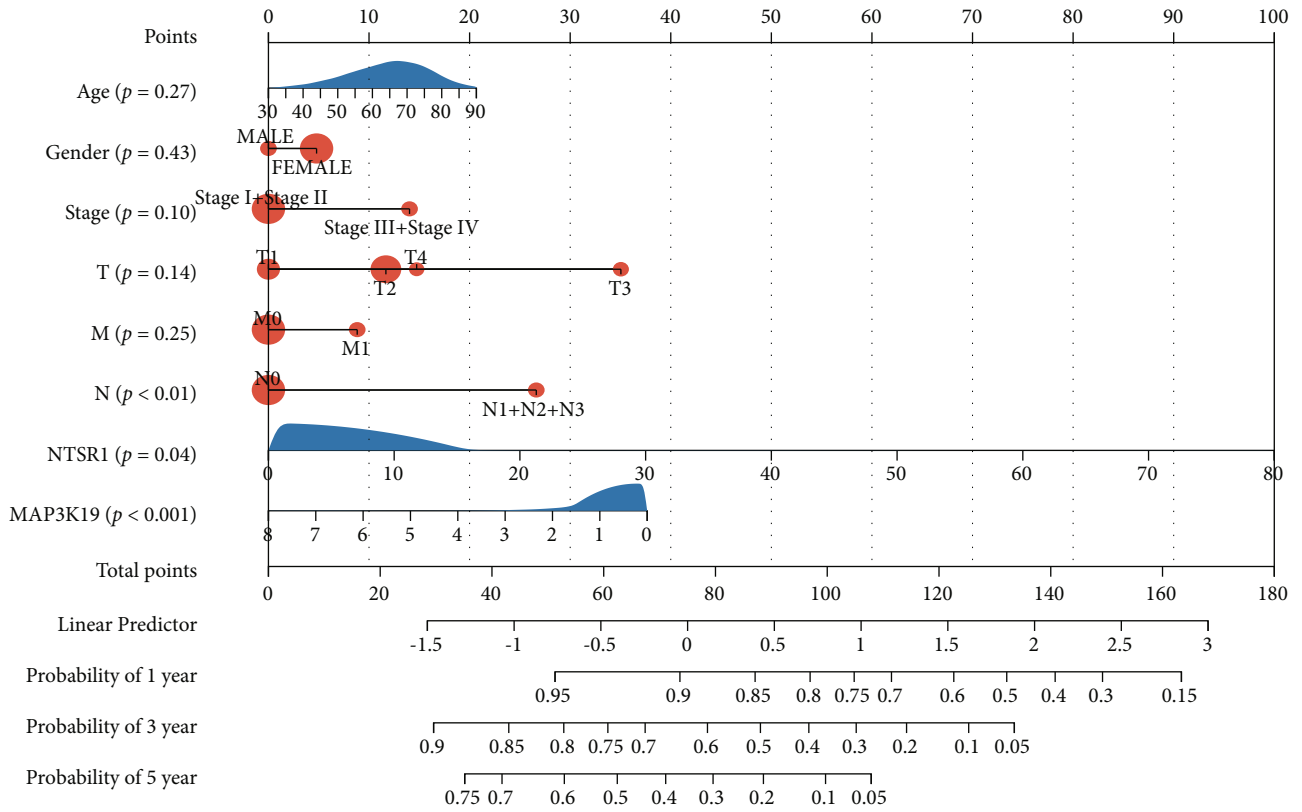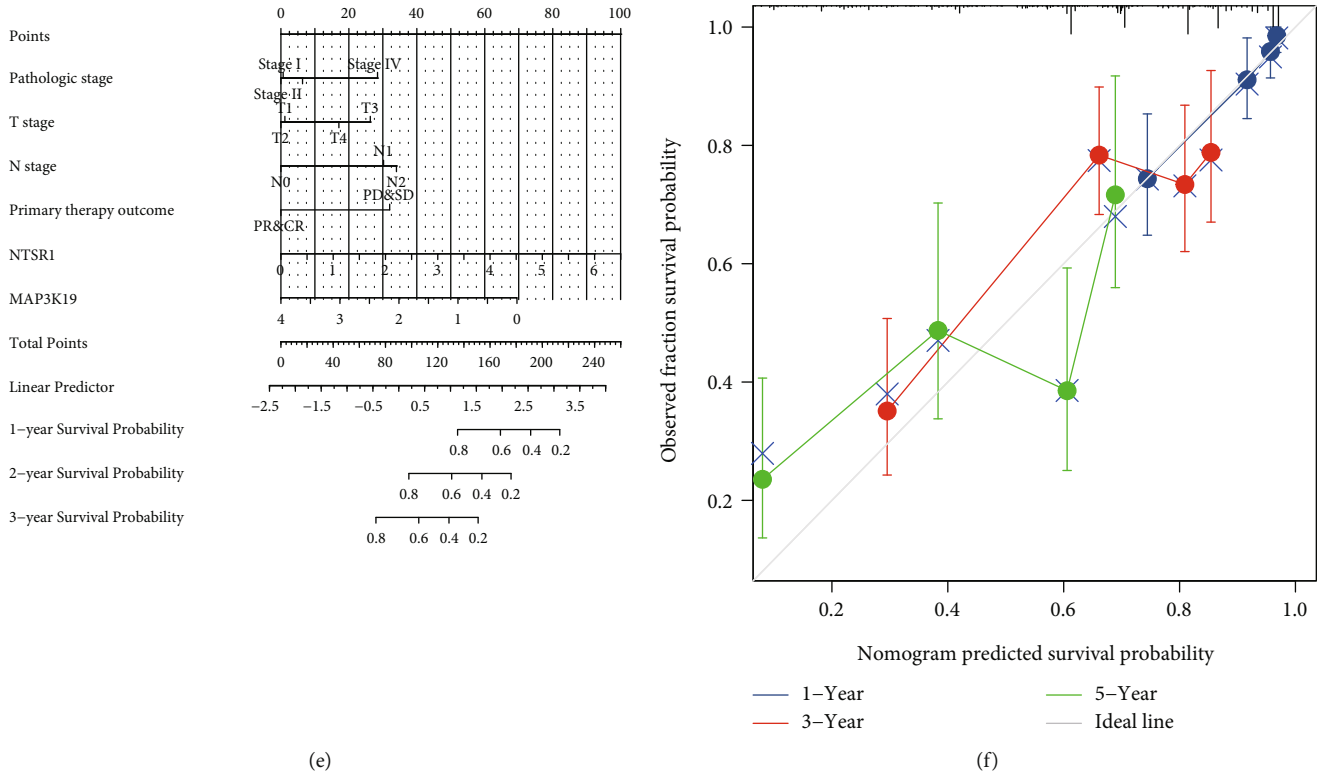
FIGURE 7: Univariate and multivariate Cox regression analyses determined the prognostic value and predictive performance of the 2-gene signature prognostic model, MAP3K19 and NTSR1. (a) K-M survival analysis revealed that the high-risk group had a higher mortality risk than the low-risk group; (b) 1, 3, and 5-year time-dependent ROC curve analysis based on OS in LUAD patients; (c) nomogram effectively integrated and shows the ability of age, sex, tumor pathological status grade, and TNM stage, as well as MAP3K19 and NTSR1 (high- and low-expression groups based on the median value) in the training set-related prognostic variables to predict LUAD overall survival; (d) calibration plot for internal validation of nomograms for overall survival prognostic status for LUAD at 1, 3, and 5 years of training set; (e) nomograms effectively integrate and demonstrate the ability of prognostic variables in TCGA-LUAD to predict overall survival in LUAD; (f) calibration plot for internal validation of nomograms for overall survival prognostic status for LUAD at 1, 3, and 5 years of the TCGA-LUAD dataset.

## 3.6. MAP3K19 and NTSR1 Prognostic Independence of Clinical Characteristics.

To further ascertain the prognostic value and predictive performance of the MAP3K19 and NTSR1 gene signatures, we first conducted a Cox regression analysis based on 445 patient samples with complete clinical information in TCGA-LUAD. We included clinical factors including age at diagnosis, gender, tumor pathological grade, lymph node status, distant metastasis, TNM stage, and gene signatures MAP3K19 and NTSR1 expression (high- and low-expression groups divided by median value). Grouped by the median RiskScore, K-M survival analysis revealed that the high-risk group had a higher mortality risk in contrast with the low-risk group ($p = 1.1e - 11$, HR = 3.53 (2.40, 5.19)) (Figure 7(a)). The AUC of the time-dependent ROC analysis reached 0.75 (0.84-0.66) at 1 year, 0.73 (0.81-0.65) at 3 years, and 0.76 (0.86-0.66) at 5 years (Figure 7(b)). Subsequently, we constructed nomograms to predict the prognostic status of LUAD at 1, 3, and 5 years utilizing the validation set and TCGA-LUAD overall patient samples (Figures 7(e) and 7(f)), respectively, and showed the predicted and actual nomograms with calibration plots. The nomogram effectively integrated the above prognostic variables and improved the ability to predict overall survival in LUAD. Based on the results of the validation set (Figures 7(c) and 7(d)), it was suggested that among the above prognostic factors, N ($p < 0.01$), NTSR1 ($p = 0.04$), and MAP3K19 ($p < 0.001$) were statistically significant. The above prognostic factors were validated in the TCGA-LUAD overall survival with consistent results by nomogram and calibration plot (Figures 7(e) and 7(f)).

The above results suggest that two gene marker prognostic factors-MAP3K19 and NTSR1 have important clinical significance for the prognosis of LUAD. Thus, we performed independent analyses for MAP3K19 and NTSR1 in TCGA-LUAD. Unpaired differential analysis between normal samples and LUAD samples indicated that NTSR1 was significantly overexpressed in LUAD (Figure 8(a)), while MAP3K19 was significantly underexpressed in LUAD (Figure 8(d)). The results of differential expression analysis of paired samples were consistent (Figures 8(b) and 8(e)). Subsequent ROC curve analysis suggested that NTSR1 (AUC: 0.601 (0.532−0.671), Figure 8(c)) and MAP3K19 (AUC: 0.710 (0.633−0.787), Figure 8(f)) had a good diagnostic performance for the differential diagnosis of LUAD samples.
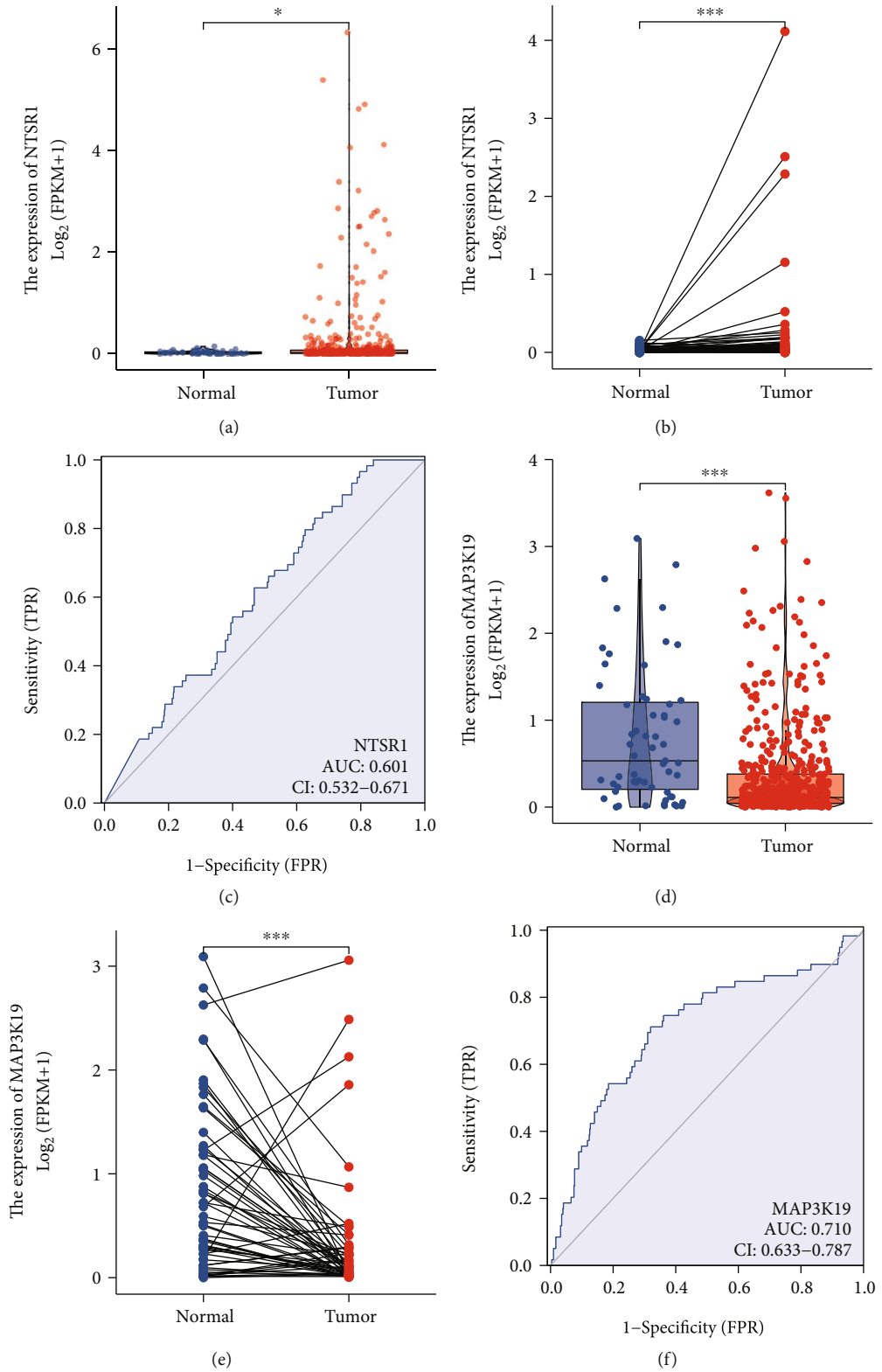
FIGURE 8: Differential expression analysis and diagnostic efficacy verification of oxidative stress-related 2-gene marker prognostic factors MAP3K19 and NTSR1 in TCGA-LUAD. (a) Unpaired differential analysis of NTSR1 in LUAD between normal and LUAD samples indicated significantly high expression; (b) paired differential analysis of NTSR1 in LUAD and normal samples suggested significantly high expression; (c) ROC curve validated MAP3K19. It has good diagnostic performance for the differential diagnosis of LUAD samples; (d) between normal and LUAD samples, the unpaired differential analysis of MAP3K19 in LUAD suggests significantly lower expression in LUAD and significantly lower expression in LUAD; (e) MAP3K19 in LUAD paired difference analysis with normal samples indicated significantly low expression; (f) ROC curve verified that MAP3K19 has good diagnostic performance for the differential diagnosis of LUAD.

(a)



(b)



(c)



(d)



(e)



(f)

FIGURE 9: Continued.

| Characteristics | Total (N) | Odds ratio (CI95%) | | P value |
|---|---|---|---|---|
| T stage (T2 & T3 & T4 vs. T1) | 532 | 0.637 (0.441–0.916) | | 0.015 |
| N stage (N1 & N2 & N3 vs. N0) | 519 | 0.688 (0.475–0.993) | | 0.047 |
| M stage (M1 vs. M0) | 386 | 0.844 (0.365–1.905) | | 0.685 |
| Pathologic stage (Stage III & Stage IV vs. Stage I & Stage II) | 527 | 0.755 (0.493–1.150) | | 0.191 |
| Age (> 65 vs. < = 65) | 516 | 1.344 (0.951–1.901) | | 0.094 |
| Residual tumor (R1 & R2 vs. R0) | 372 | 1.238 (0.463–3.368) | | 0.668 |

(g)                                                                                 (h)
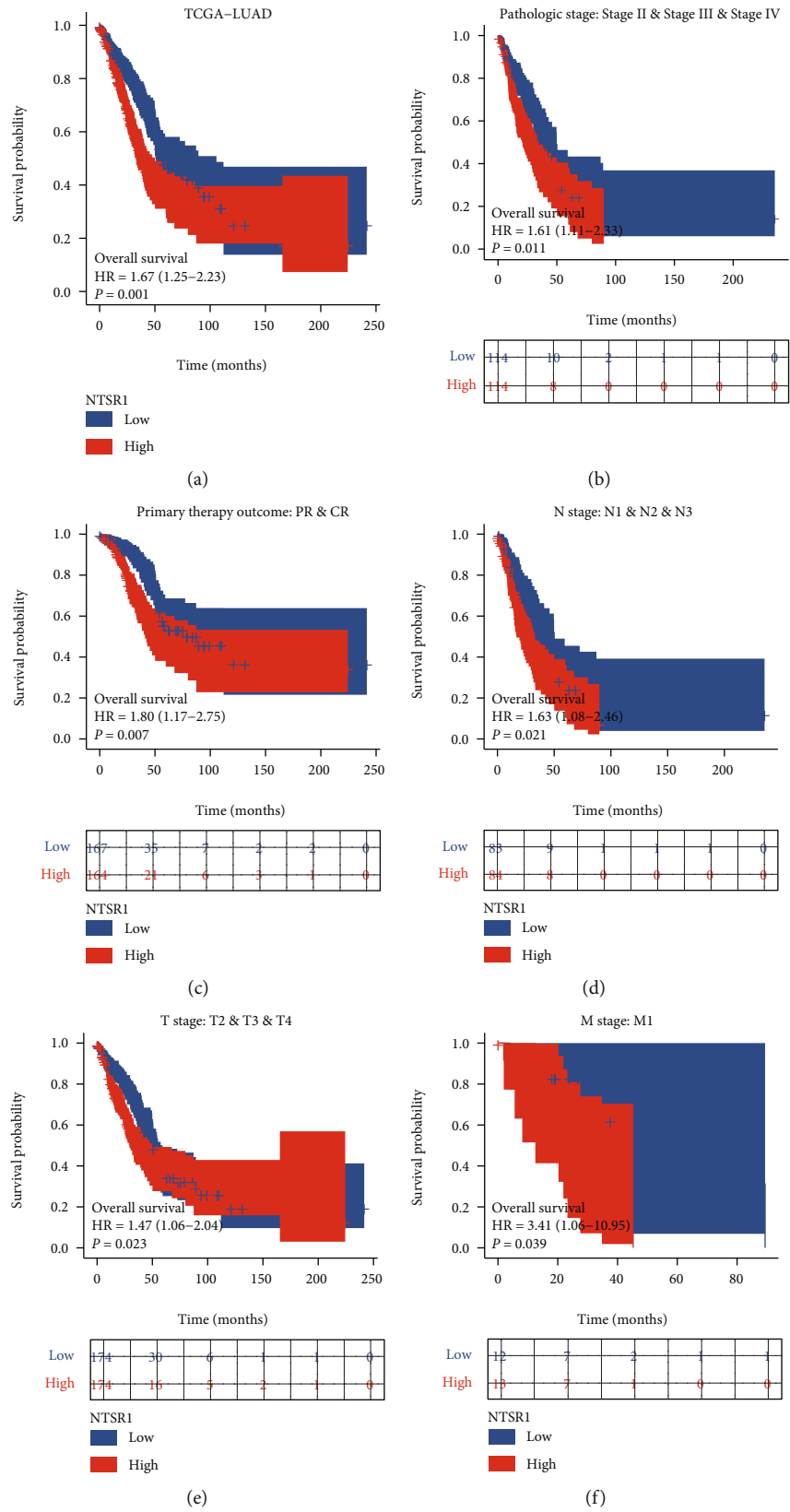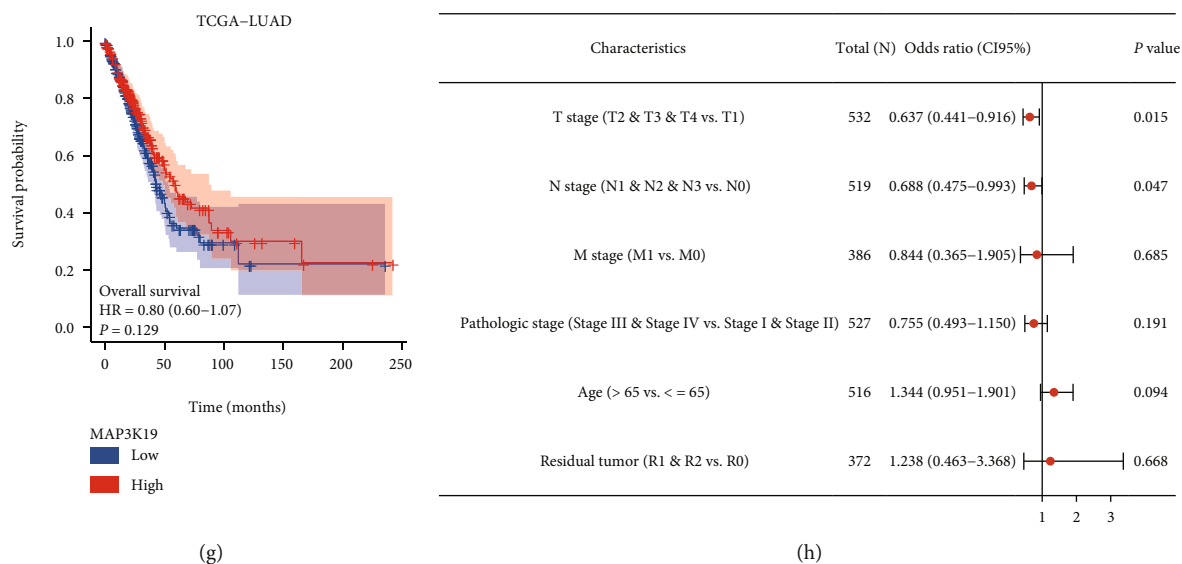
FIGURE 9: Subgroup survival analysis of clinical variables for oxidative stress-related 2 gene signatures MAP3K19 and NTSR1 in TCGA-LUAD. (a) K-M survival curves suggest that elevated NTSR1 expression level is substantially linked to poorer OS outcomes in LUAD; (b–f) K-M survival curves suggest that high NTSR1 expression is linked to poor pathologic outcomes, respectively; stage, primary therapy outcome and TNM stage were significantly correlated; (g) K-M survival curve suggested that low expression of MAP3K19 was remarkably linked to poorer OS outcome in LUAD; (h) logistic regression forest plot of binary variables of clinical subgroups of MAP3K19 showed the correlation sex.

We further analyzed the clinical variable subgroup survival analysis of oxidative stress-related 2 gene signatures MAP3K19 and NTSR1 in TCGA-LUAD, suggesting that high NTSR1 expression and low expression of MAP3K19 were remarkably linked to poorer OS outcomes in LUAD (Figure 9(a)); subgroup survival analysis of clinical variables of primary therapy outcome, pathologic stage, and TNM stage exhibited that high expression of NTSR1 was significantly linked to poor prognosis, respectively (Figures 9(b)–9(f)). In addition, our logistic regression forest plot of dichotomous variables for clinical subgroups of MAP3K19 demonstrated correlations (not statistically significant in NTSR1 analysis) (Figure 9(h)). Meanwhile, in vivo RT-PCR and IHC results against NTSR1 (Figures 10(a)–10(c)) and MAP3K19 (Figures 10(d)–10(f)) were also consistent with our previous analysis.

*3.7. Differential Analysis and Enrichment Analysis Based on Oxidative Stress-Related 2 Gene Signatures.* All of the above analyses confirmed the superior predictive performance of MAP3K19 and NTSR1 for poor prognosis in LUAD, and the following analysis focused on exploring how the oxidative stress-related 2 gene signature might lead to poor prognosis in LUAD.

In order to additionally explore the molecular mechanism engaged in the identification of high-risk populations with poor prognosis of LUAD in the oxidative stress prognostic model composed of MAP3K19 and NTSR1, we separately analyzed the significant differences between two risk groups in the prognostic model constructed by the LASSO algorithm in the training set and the validation set. We performed differential analysis on the two risk groups defined

by the median risk of the training set (Figure 11(a)) and validation set (Figure 11(b)) prediction model, respectively, showing the expression of significantly different genes as a heatmap. The 92 prognostic risk genes of oxidative stress coexisting in both sets were obtained by the Venn diagram (Figure 11(c)). These genes were then subjected to enrichment analysis. GSVA analysis shows enriched entries with $|LogFC| > 0.5$ and adj.$p < 0.05$ as a heatmap. We then analyzed the reference gene set retrieved from the MSigDB database in "c2.cp.v7.2.symbols.gmt," "h.all.v7.2.symbols.gmt" gene set, and immune cell infiltration-related gene set. Enrichment scores for each gene functional pathway were obtained utilizing the GSVA package in R. The heatmap visualizes the differential gene expression between the two risk groups in the training set as well as the validation set. It was found that the gene set enrichment results of "h.all.v7.2.symbols.gmt" (Figure 11(d)) and "c2.cp.v7.2.symbols.gmt" (Figure 11(e)) based on the MSigDB database were concentrated in UNFOLDED PROTEIN RESPONSE, GLYCOLYSIS, MTORC1 SIGNALING, MYC TARGETS V1, MYC TARGETS V2, and other functions and ways. Immune-related "immune.gmt" (Figure 11(f)) gene set enrichment results focused on Eosinophil, Natural.killer.cell, Immature..B.cell, Activated.B.cell, Mast.cell, and Type.1.-T.helper .cell.

Next, the GO (Table 3) and KEGG (Table 4) enrichment analysis results are shown by bubble plots (Figures 12(a) and 12(b)) and bar graphs (Figures 12(c) and 12(d)). BP was significantly enriched in oxidative phosphorylation, mitochondrial translational elongation, mitochondrial translational termination, translational termination, purine ribonucleoside triphosphate metabolic process, and other

(a)



(b)



(c)



(d)



(e)



(f)
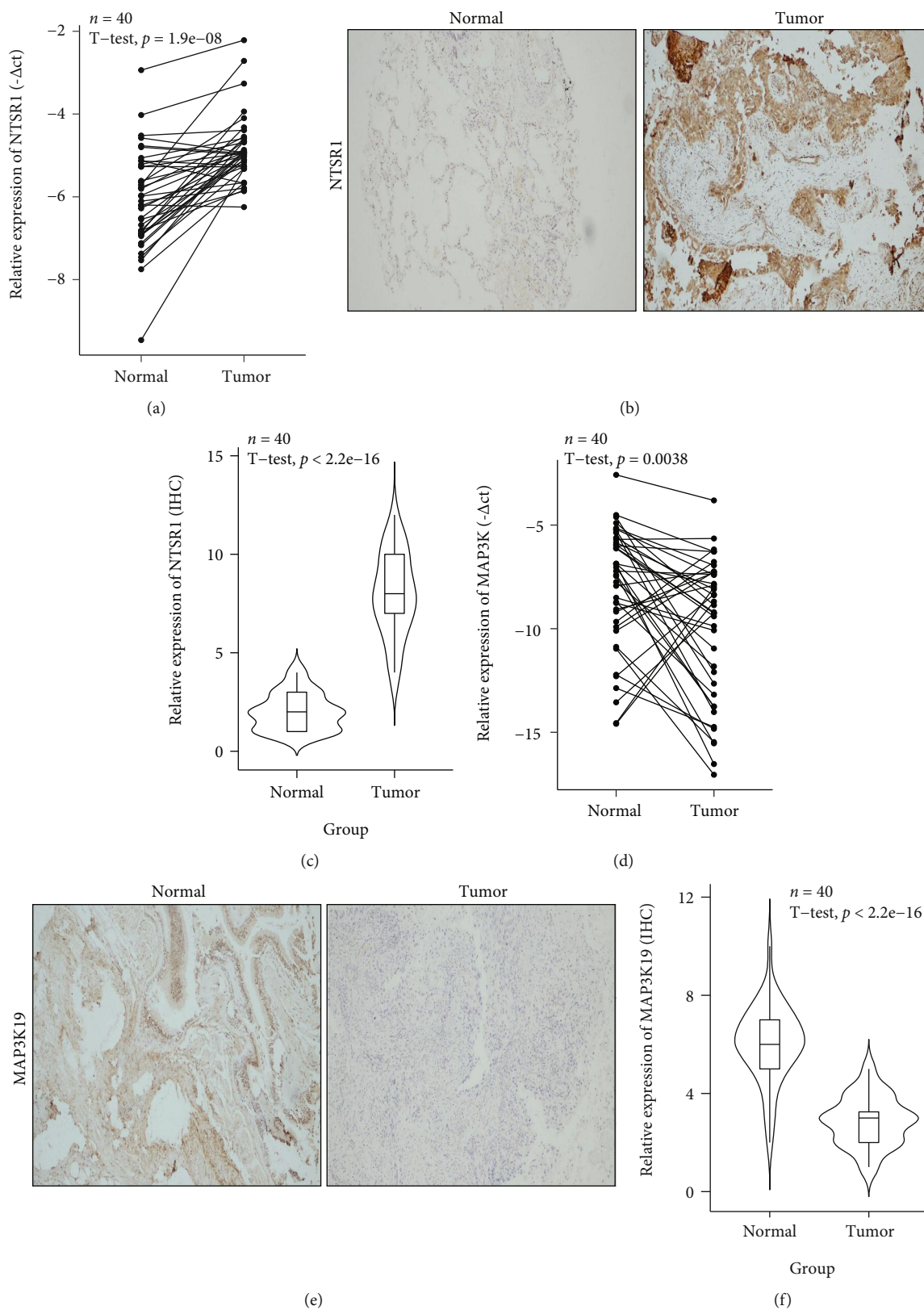
FIGURE 10: qRT-PCR and IHC analysis of model genes. (a–c) NTSR1 mRNA and protein expression were significantly increased in LUAD samples; (d–f) MAP3K19 expression in LUAD samples was significantly lower than that in paratumor normal tissue.
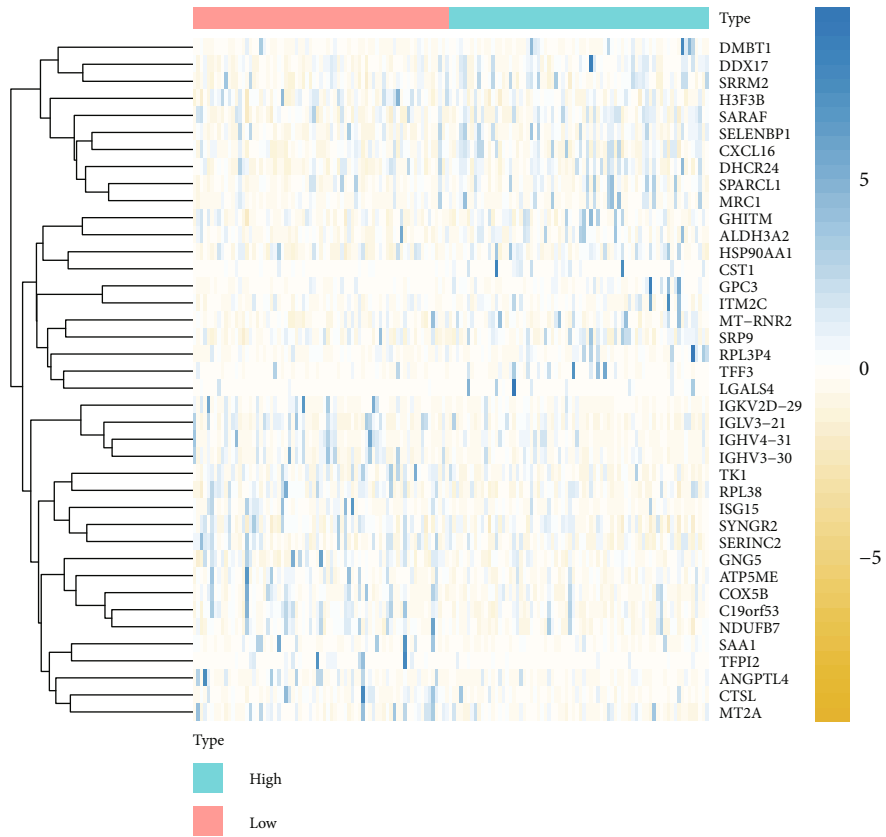
(a)
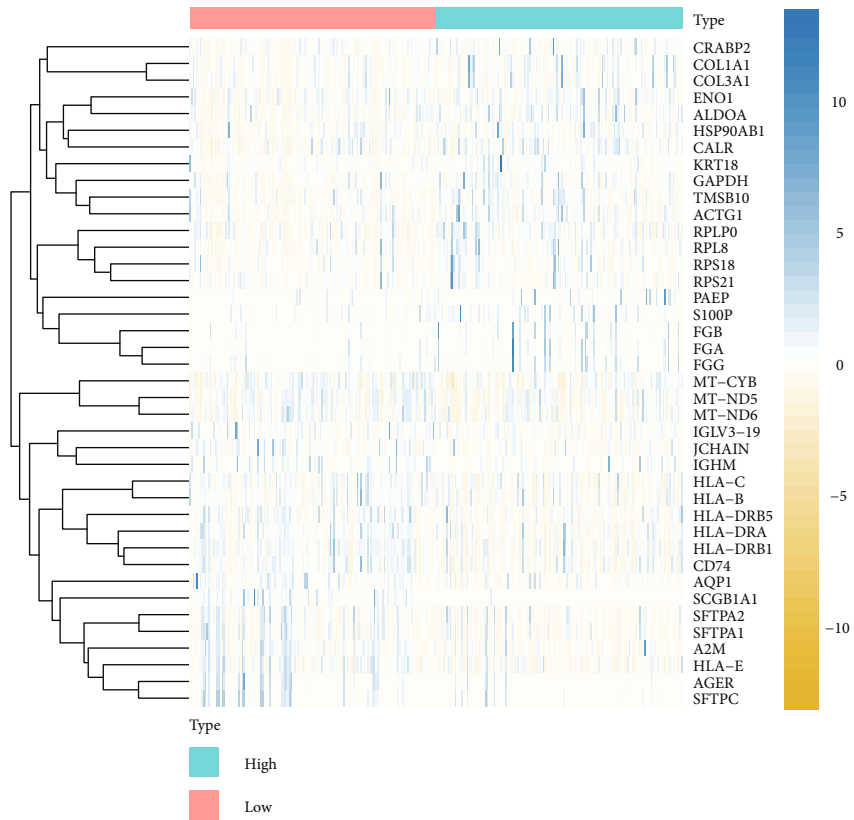


(b)

Figure 11: Continued.
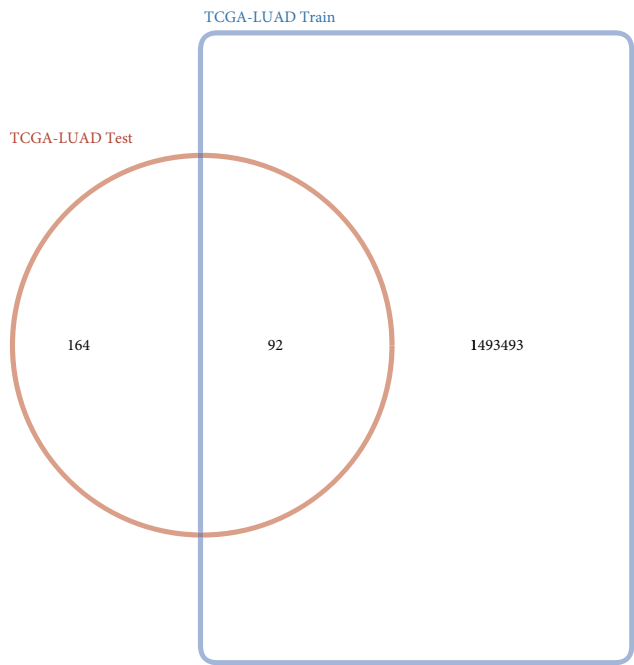
(c)
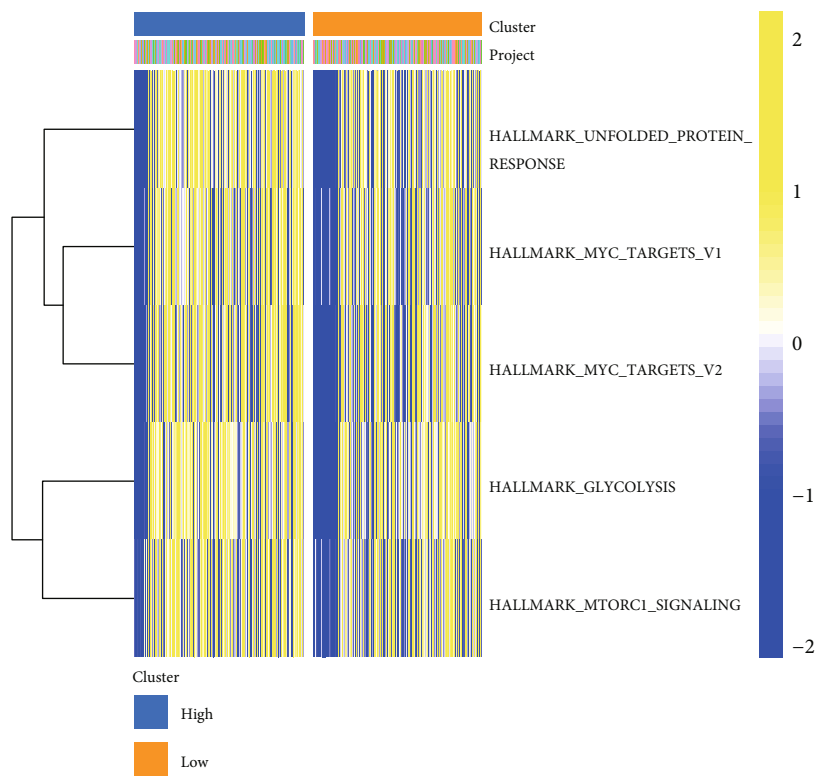


(d)

Figure 11: Continued.
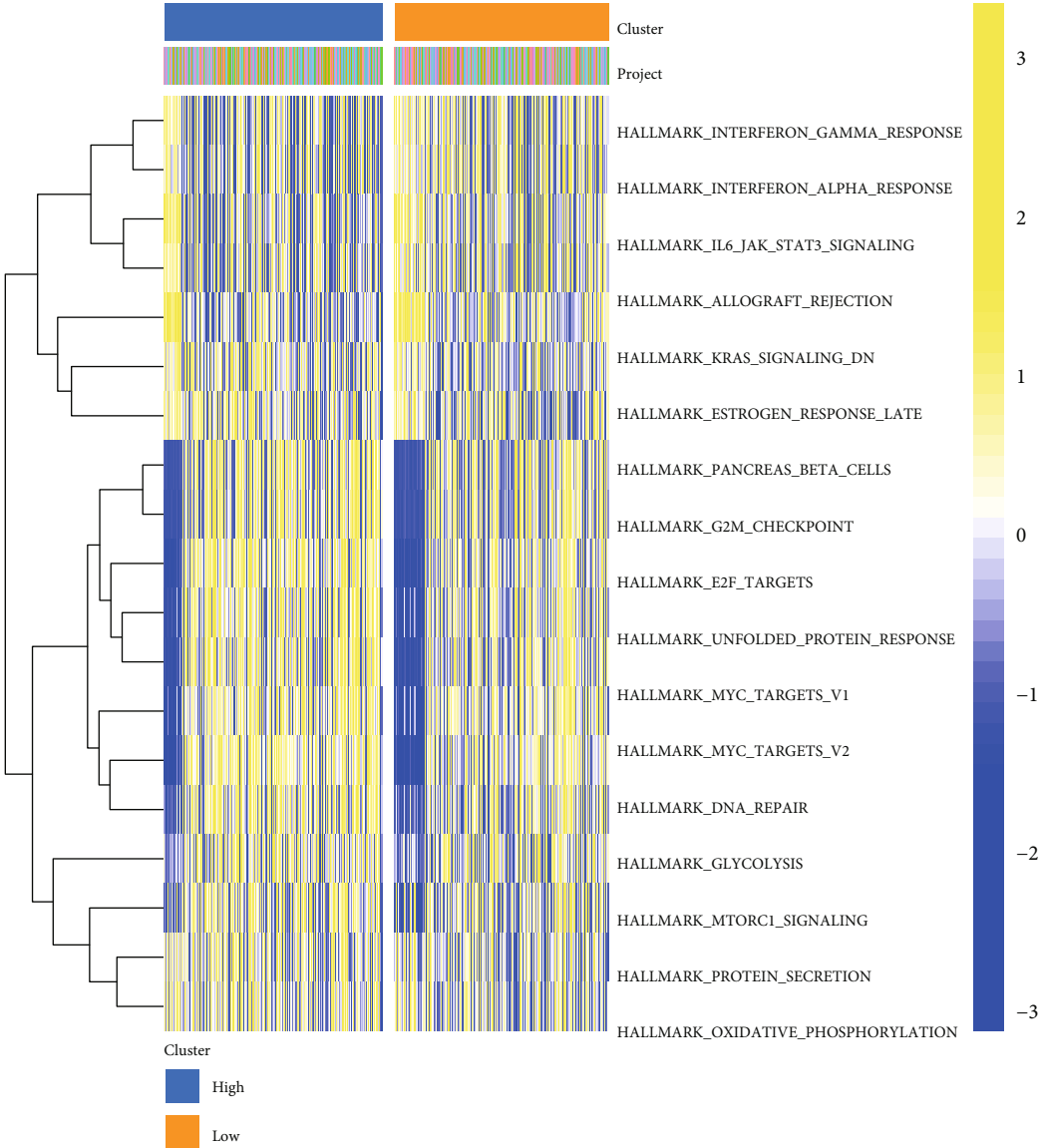
(e)

FIGURE 11: Continued.

(f)
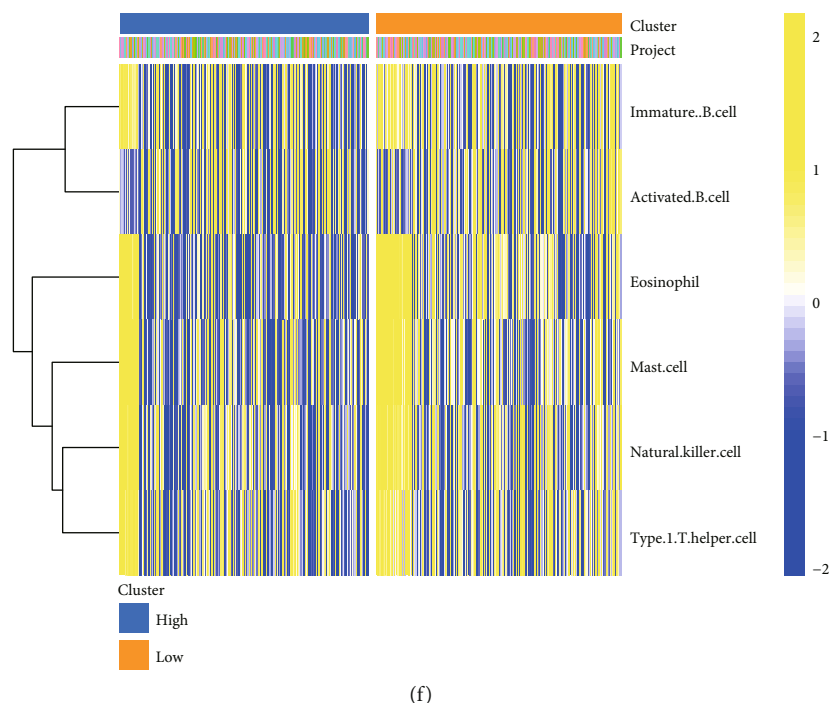
FIGURE 11: Differential analysis based on oxidative stress-related 2 gene signatures and GSVA enrichment analysis. (a, b) We obtained heatmaps of differential gene expression between the training set (a) and validation set (b) between high- and low-risk groups, correspondingly; (c) Venn diagram obtained the coexisting 92 prognostic risk genes of oxidative stress; (d–f) "h.all.v7.2.symbols.gmt" (e), "c2.cp.v7.2.symbols.gmt" with |LogFC| > 0.5 and adj.$p$ < 0.05 as a heatmap (e), enriched entries for immune-related (f) gene set GSVA analysis.

TABLE 3: GO enrichment analysis of coexpressed genes based on the LUAD oxidative stress-related prognostic risk model.

| Ontology | ID | Description | BgRatio | $p$ value | $p$.adjust | $Q$ value |
|---|---|---|---|---|---|---|
| BP | GO:0006119 | Oxidative phosphorylation | 145/18670 | 8.76$e$-10 | 1.28$e$-06 | 1.17$e$-06 |
| | GO:0070125 | Mitochondrial translational elongation | 88/18670 | 5.21$e$-09 | 2.78$e$-06 | 2.53$e$-06 |
| | GO:0070126 | Mitochondrial translational termination | 89/18670 | 5.70$e$-09 | 2.78$e$-06 | 2.53$e$-06 |
| | GO:0006415 | Translational termination | 104/18670 | 1.97$e$-08 | 7.19$e$-06 | 6.56$e$-06 |
| | GO:0009205 | Purine ribonucleoside triphosphate metabolic process | 335/18670 | 2.83$e$-08 | 7.42$e$-06 | 6.76$e$-06 |
| CC | GO:0005743 | Mitochondrial inner membrane | 473/19717 | 3.79$e$-18 | 7.50$e$-16 | 6.10$e$-16 |
| | GO:0098798 | Mitochondrial protein complex | 262/19717 | 1.10$e$-10 | 1.09$e$-08 | 8.85$e$-09 |
| | GO:0005759 | Mitochondrial matrix | 469/19717 | 1.45$e$-08 | 9.60$e$-07 | 7.81$e$-07 |
| | GO:0000313 | Organellar ribosome | 87/19717 | 1.03$e$-07 | 4.09$e$-06 | 3.33$e$-06 |
| | GO:0005761 | Mitochondrial ribosome | 87/19717 | 1.03$e$-07 | 4.09$e$-06 | 3.33$e$-06 |
| MF | GO:0003735 | Structural constituent of ribosome | 202/17697 | 3.60$e$-06 | 9.07$e$-04 | 7.84$e$-04 |
| | GO:0015078 | Proton transmembrane transporter activity | 133/17697 | 3.07$e$-05 | 0.004 | 0.003 |
| | GO:0009055 | Electron transfer activity | 114/17697 | 1.66$e$-04 | 0.011 | 0.009 |
| | GO:0004129 | Cytochrome-c oxidase activity | 28/17697 | 2.69$e$-04 | 0.011 | 0.009 |
| | GO:0015002 | Heme-copper terminal oxidase activity | 28/17697 | 2.69$e$-04 | 0.011 | 0.009 |

items. CC was significantly enriched in mitochondrial inner membrane, protein complex, matrix, ribosome, organellar ribosome, and other entries. MF was significantly enriched in the structural components of the ribosome, proton transmembrane transporter activity, cytochrome-c oxidase activity, electron transfer activity, heme-copper terminal oxidase activity, and other items. KEGG was sig-

nificantly enriched in oxidative phosphorylation, thermogenesis, ribosome, Parkinson's disease, cardiac muscle contraction, and other pathways. A pathview diagram with differentially expressed genes colored for the arachidonic acid metabolism pathway is displayed in (Figures 12(e)–12(i)), including hsa00190, hsa03010, hsa04260, hsa04714, and hsa05012.

TABLE 4: KEGG enrichment analysis of coexpressed genes based on the LUAD oxidative stress-related prognostic risk model.

| Ontology | ID | Description | BgRatio | $p$ value | $p$.adjust | $Q$ value |
|---|---|---|---|---|---|---|
| KEGG | hsa00190 | Oxidative phosphorylation | 133/8076 | $4.47e$-08 | $3.35e$-06 | $2.82e$-06 |
| | hsa04714 | Thermogenesis | 231/8076 | $4.98e$-07 | $1.87e$-05 | $1.57e$-05 |
| | hsa03010 | Ribosome | 158/8076 | $2.71e$-05 | $6.77e$-04 | $5.70e$-04 |
| | hsa05012 | Parkinson disease | 249/8076 | $6.90e$-05 | 0.001 | 0.001 |
| | hsa04260 | Cardiac muscle contraction | 87/8076 | $1.25e$-04 | 0.002 | 0.001 |

*3.8. Gene Set Enrichment Analysis of Oxidative Stress Gene Model in LUAD.* In order to additionally explore the molecular processes engaged in the identification of high-risk populations with poor prognosis of LUAD in the oxidative stress prognostic model composed of MAP3K19 and NTSR1, GSEA is a computational method utilized to ascertain if a predefined set of genes shows statistical differences between two biological states, which is typically utilized to estimate expression in a dataset sample. Changes in the pathway as well as biological process activity were analyzed [23]. To investigate the variations in biological mechanisms between the two groups of samples, employing the gene expression profile dataset, we obtained the reference gene set "c2.cp.kegg.v7.4.entrez.gmt" from the MSigDB database, utilizing the R package "clusterProfiler." The GSEA method is included in enrichment analysis and visualization of datasets. $p$ value < 0.05 was considered statistically significant (Figure 13 and Table 5).

*3.9. Pan-Cancer Analysis of Immune Cells and Immune Infiltration Based on MAP3K19 and NTSR1.* The infiltration abundance of immune cells was analyzed by CIBERSORT, and the correlation heatmap in pan-cancer showed that MAP3K19 (Figure 14(a)) and NTSR1 (Figure 14(b)) were significantly correlated with immune infiltration in more types of tumors. Based on TCGA LUAD transcription profile and CIBERSORT, we derived the proportions of 22 tumor-infiltrating immune cells. In our study, the use of RNA-seq data as well as global proteomic data inferred MAP3K19 and NTSR1 for pan-cancer tumor purity, immune score, and stromal score. From this, significantly correlated immune infiltration scores were identified, and we analyzed the scores for stromal cell levels, tumor purity, and immune cell infiltration levels in cancer tissues calculated with ESTIMATE expression was substantially positively linked to immune score, ESTIMATE score, and stromal score, respectively. The results suggest that NTSR1 (Figures 14(c)–14(e)) and MAP3K19 (Figures 14(f)–14(h)) were closely related to stromal cell level, tumor purity, and immune cell infiltration level in the lung tumor microenvironment and may affected the prognosis of LUAD by changing the tumor immune microenvironment.

*3.10. Correlation Analysis of Immune Cell Infiltration.* It was found that MAP3K19 was positively correlated with T cell CD4 memory resting and B cell memory (Figures 15(a) and 15(b)) and negatively correlated with mast cells activated and dendritic cells activated (Figures 15(c) and 15(d)). NTSR1 was positively correlated with neutrophils,

macrophages M0, and T cell gamma delta (Figures 15(e)–15(g)) and negatively linked to mast cells resting (Figure 15(h)). In our study, the use of RNA-seq data as well as global proteomic data inferred MAP3K19 and NTSR1 for pan-cancer tumor purity, immune score, and stromal score. From this, significantly correlated immune infiltration scores were identified, and we analyzed the scores for stromal cell levels, tumor purity, and immune cell infiltration levels in cancer tissues calculated with ESTIMATE expression which was substantially positively linked to immune score, ESTIMATE score, and stromal score, respectively. The results suggest that NTSR1 (Figures 14(c)–14(e)) and MAP3K19 (Figures 14(f)–14(h)) were closely related to stromal cell level, tumor purity, and immune cell infiltration level in the lung tumor microenvironment and may affected the prognosis of LUAD by changing the tumor immune microenvironment.

## 4. Discussion

Globally, lung cancer is the major cause of cancer-related mortalities, whereas LUAD is the most prevalent histological subtype of the disease. About two-thirds of LUAD have activated oncogenes. Most oncogene mutations often activate downstream signaling pathways through oxidative stress pathways and states, ultimately leading to lethal malignancies, including LUAD [24, 29, 30]. Molecularly targeted therapy significantly improves survival in patients with treatment-targeted lesions compared to conventional chemotherapy [31, 32]. However, the clinical efficacy of targeted drugs has been improved due to the lack of appropriate small molecules to bind major tumor-causing gene mutations.

For the survival of both normal cells and cancerous cells, redox hemostasis is fundamental. Nonetheless, numerous malignancies have increased levels of reactive oxygen species (ROS) and exhibit signs of chronic oxidative stress as a result of oncogenic injury, metabolic malformations, hypoxia, and proteotoxic stress [33]. The increased ROS at the sublethal level is implicated in the enhancement of tumor development via triggering mutations and altering cell signaling [34]. Because traditional cytotoxic agents additionally influence normal tissues, targeted approaches that induce catastrophic oxidative stress selectively in malignant cells would avail a better therapeutic window [35].

The LUAD cohort of TCGA availed both the expression and clinical data in this study. Among oxidative stress-related genes, we did differential expression analysis as well as univariate Cox analysis to screen 32 prognostic DEGs

(a)



(b)



(c)



(d)

Figure 12: Continued.

(e)



(f)



(g)

Figure 12: Continued.

(h)



(i)
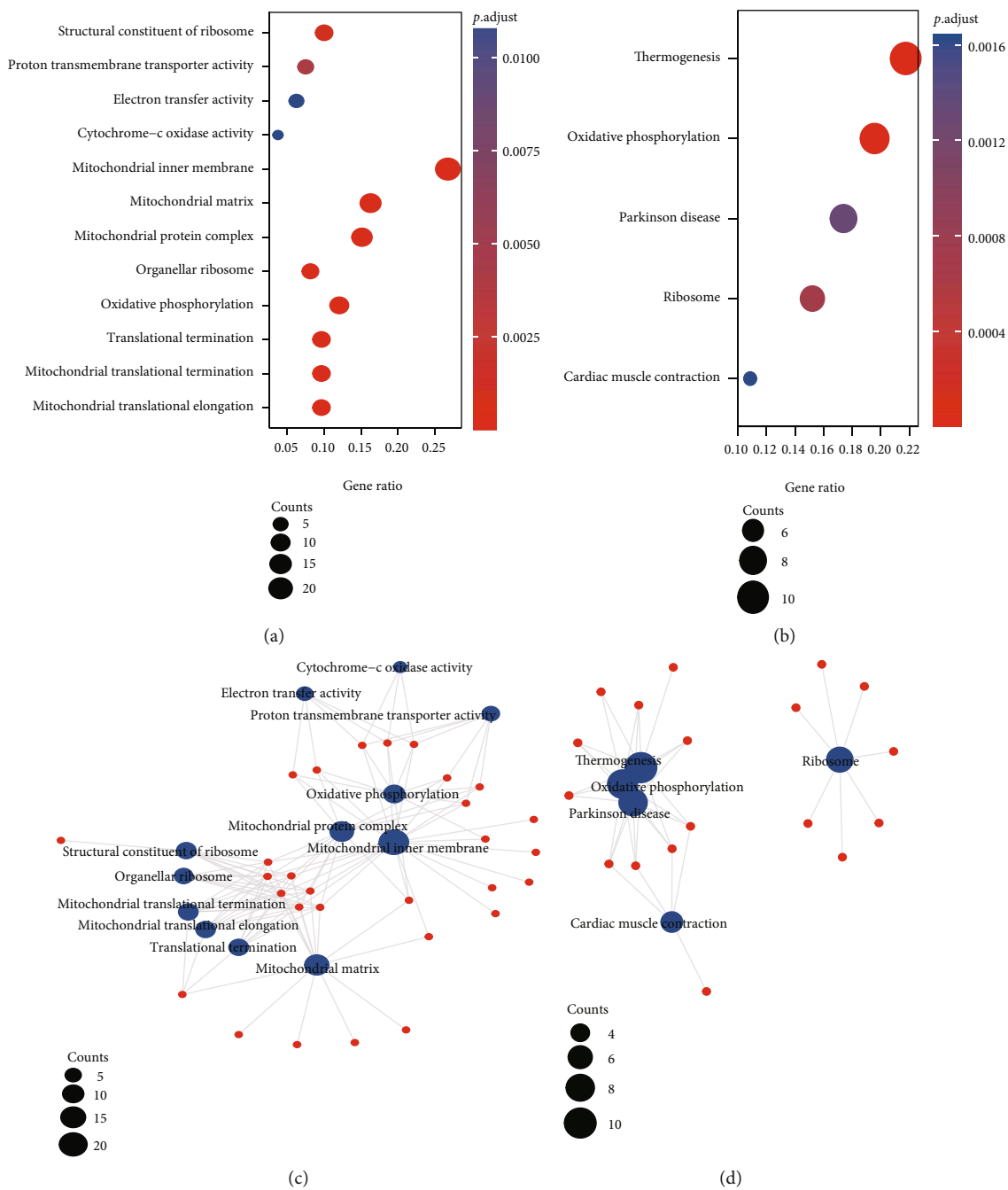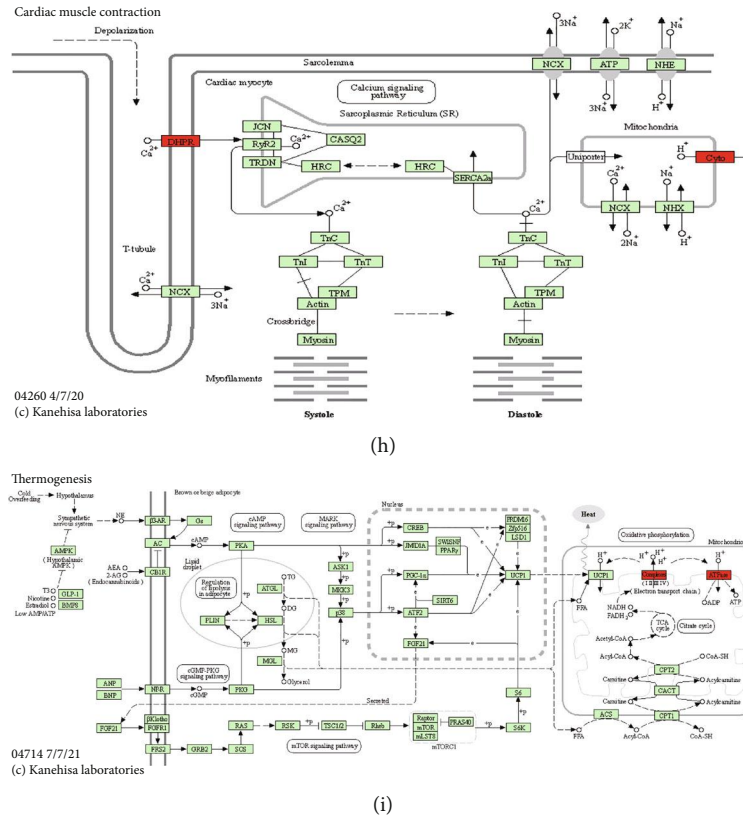
FIGURE 12: GO and KEGG enrichment analysis. (a–d) GO (a, c) and KEGG (b, d) enrichment analysis of 92 differentially expressed genes associated with LUAD prognosis. Bubble and bar graphs show the results of GO and KEGG enrichment analysis; (e–i) pathview diagrams with DEGs colored for expression are shown for hsa00190, hsa03010, hsa04260, hsa04714, and hsa05012, respectively. The figure shows term with $p$.adjust < 0.05. The length of the bars in the histogram denotes the amount of gene enrichment, the color denotes the significance, and the significance level increases gradually from blue to red.

from 147 differentially expressed oxidative stress-related genes and utilized Lasso-penalized Cox regression analysis constructed 2 gene markers associated with prognosis. At the same time, by constructing a PPI network, we analyzed the distribution of 18 differentially oxidative stress hub genes in transport vesicles, endoplasmic reticulum lumen, etc., affecting blood pressure, hormone secretion, nucleotide second messengers, and other functions to regulate oxidative stress effect. To construct and validate the oxidative stress gene signature affecting the diagnosis and prognosis of LUAD, we randomly divided 445 patients with complete clinical informatics data in TCGA-LUAD into the training set (2/3 of the total, $n = 296$) and the test set (1/3 of the total, $n = 149$); the grouping is normally distributed. We first performed a univariate Cox regression analysis of gene expression premised on the expression data of 147 oxidative stress-related genes in the training and test set, correspondingly. Thirty-five genes with prognostic values coexisting in both sets were identified. Then, we performed LASSO-Cox proportional hazards regression based on 35 prognostic-related differential oxidative stress genes in the training and validation set, correspondingly, and found that MAP3K19 and NTSR1 in both sets showed a better prediction of poor prognosis in LUAD ability. Therefore, the prognostic independence of 2-gene signature prognostic factors,

MAP3K19 and NTSR1, was further analyzed in combination with clinical features. A nomogram was constructed to effectively integrate prognostic variables and validated with calibration. Next, the difference analysis of MAP3K19 and NTSR1 in TCGA-LUAD and the correlation analysis of clinical variables indicated that both the high expression of NTSR1 and the low expression of MAP3K19 had a better diagnostic performance for the diagnosis of LUAD and the identification of poor prognosis. At the same time, MAP3K19 and NTSR1 were found to be significantly correlated with clinical variables such as LUAD pathological stage and TNM grade. Prognosis-related differential oxidative stress gene signatures are independent prognostic factors in patients with LUAD.

The importance of tumor immune activity on tumorigenesis and development, as well as individual variation at the gene level, has attracted more and more researchers to focus on the significance of differential genes that may be useful in distinguishing pRCC patients with heterogeneous responses and predicting prognosis potentially meaningful. The ESTIMATE analysis of pan-cancer species analysis of MAP3K19 and NTSR1 in TCGA database systematically recorded the abundance of 22 tumor-infiltrating immune compartments in LUAD samples through the CIBERSORT algorithm and integrated it with the MAP3K19 and NTSR1
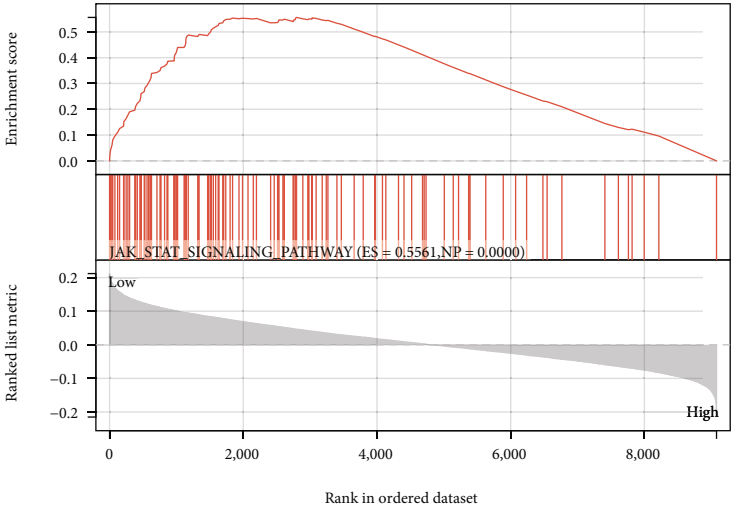
(a)
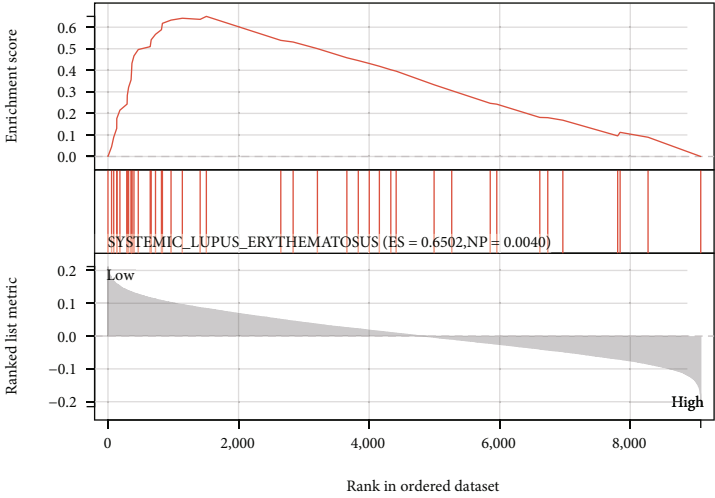


(b)
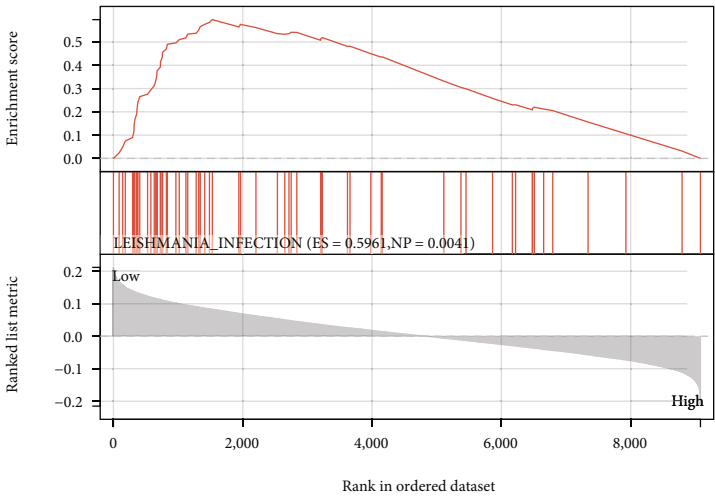


(c)

Figure 13: Continued.
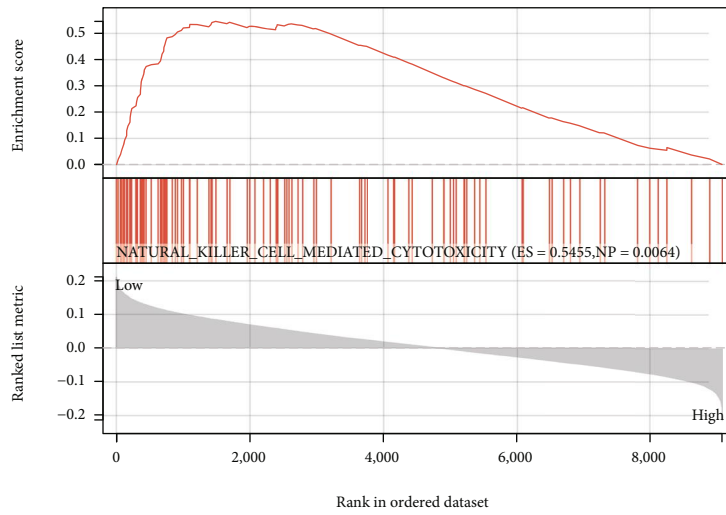
(d)
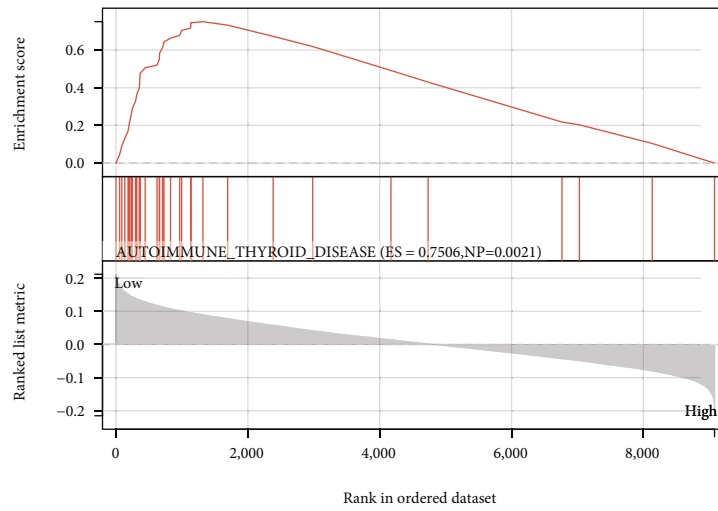


(e)
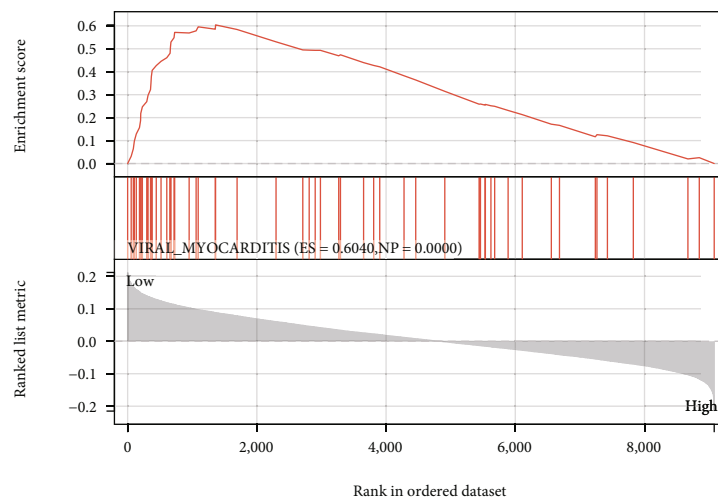


(f)

FIGURE 13: Continued.
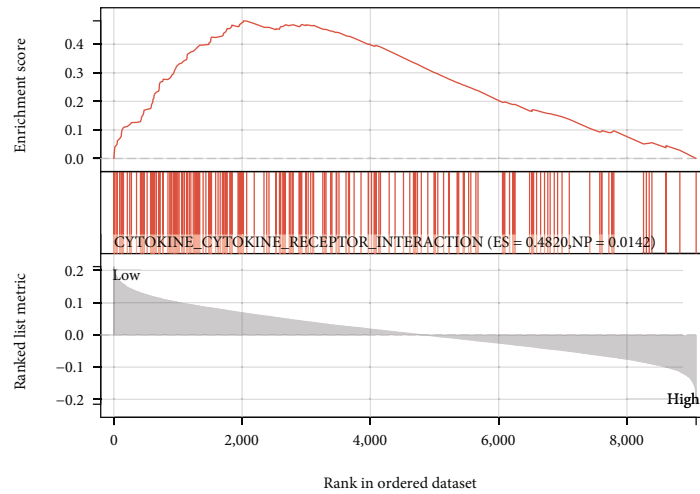
(g)
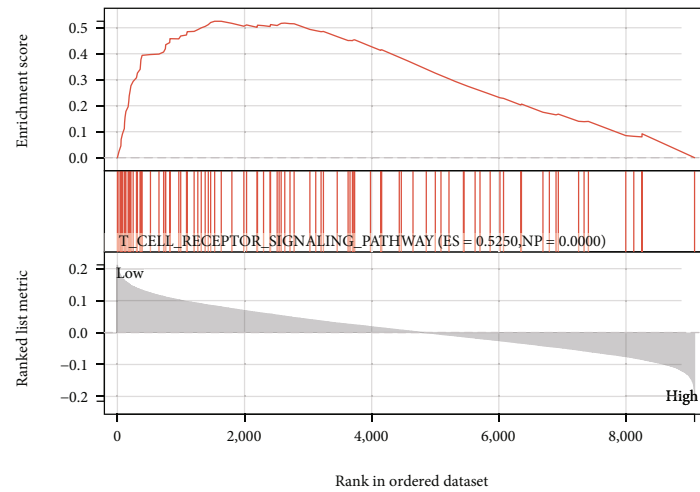


(h)
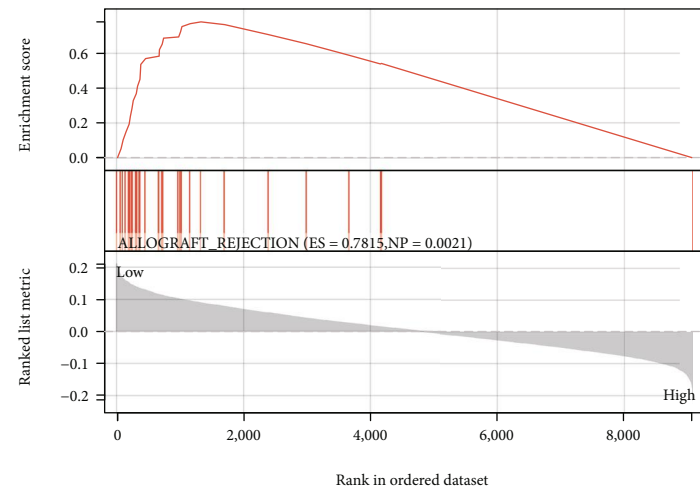


(i)

FIGURE 13: Gene set enrichment analysis of oxidative stress prognostic gene signature in LUAD. (a) JAK_STAT_SIGNALING_PATHWAY (ES = 0.5561, NP = 0.0000); (b) SYSTEMIC_LUPUS_ERYTHEMATOSUS (ES = 0.6502, NP = 0.0040); (c) LEISHMANIA_INFECTION (ES = 0.5961, NP = 0.0041); (d) NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY (NP = 0.0064); (e) AUTOIMMUNE_THYROID_DISEASE (ES = 0.7506, NP = 0.0021); (f) VIRAL_MYOCARDITIS (ES = 0.6040, NP = 0.0000); (g) CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION (ES = 0.4820, NP = 0.0142); (h) T_CELL_RECEPTOR_SIGNALING_PATHWAY (ES = 0.5250, NP = 0.0000); (i) ALLOGRAFT_REJECTION (ES = 0.7815, NP = 0.0021).

TABLE 5: Results of gene set enrichment analysis of gene signatures in prognostic models of oxidative stress.

| Term | ES | NES | $p$ value | FDR |
|------|------|------|------|------|
| VIRAL_MYOCARDITIS | 0.604 | 2.0023 | 0.0117 | 0.04 |
| AUTOIMMUNE_THYROID_DISEASE | 0.7506 | 2.0069 | 0.0021 | 0.0149 |
| ASTHMA | 0.7919 | 2.0269 | 0.002 | 0.0168 |
| NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY | 0.5455 | 1.9577 | 0.0064 | 0.0191 |
| JAK_STAT_SIGNALING_PATHWAY | 0.5561 | 2.0488 | 0.0259 | 0.023 |
| LEISHMANIA_INFECTION | 0.5961 | 1.8795 | 0.0041 | 0.0266 |
| ALLOGRAFT_REJECTION | 0.7815 | 1.9236 | 0.0021 | 0.0269 |
| T_CELL_RECEPTOR_SIGNALING_PATHWAY | 0.525 | 1.8848 | 0.0288 | 0.0149 |
| SYSTEMIC_LUPUS_ERYTHEMATOSUS | 0.6502 | 1.897 | 0.004 | 0.0293 |
| CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION | 0.482 | 1.8142 | 0.0142 | 0.046 |
| TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY | 0.4956 | 1.8005 | 0.0151 | 0.0487 |

molecular profiles to analyze the degree of immune infiltration in LUAD. Subsequently, further immune score, stromal score, and ESTIMATE score in the pan-cancer TME revealed the potential roles of MAP3K19 and NTSR1 in regulating stromal/immune scores and gene expression in tumors. This suggests that the oxidative stress prognostic model composed of MAP3K19 and NTSR1 may be involved in the molecular mechanism in the identification of high-risk populations with poor prognosis of LUAD.

Studies have shown that MAP3K19 level is elevated in COPD and bronchoalveolar lavage macrophages from patients with IPF. At the level of target gene transcription or protein synthesis, molecular studies have confirmed that MAP3K19 inhibitors are linked to pirfenidone or nintedanib. At the same time, MAP3K19 significantly attenuated bleomycin-induced pulmonary fibrosis [36] and is a central mediator of cigarette smoke-induced pulmonary inflammation and lower airway destruction [37]. In studies on lung cancer, targeting MAP3K19 has been reported to prevent human lung myofibroblast activation in vitro and in a humanized SCID model of idiopathic pulmonary fibrosis [38]. On the other hand, it can also phosphorylate MAP2K, thereby activating ERK as well as JNK kinases and increasing the KRAS mutant lung cancer cells' viability [39]. The mechanism of action of NTSR1 in various tumors has also been reported many times [40]. Effects, underlying mechanisms, and clinical roles of NTSR1 on gastric adenocarcinoma cell proliferation and invasion. Interfering with NTSR1 expression exhibits anti-invasive effects through the Jun/miR-494/SOCS6 axis in glioblastoma cells [41]. The mechanism of action of NTSR1 in various tumors has also been reported many times. Effects, underlying mechanisms, and clinical roles of NTSR1 on gastric adenocarcinoma cell proliferation and invasion. Interfering with NTSR1 expression causes anti-invasive effects through the Jun/miR-494/SOCS6 axis in glioblastoma cells. NTSR1 and Wnt/$\beta$-catenin enhance tumor growth in glioblastoma [42]. Furthermore, NTSR1 methylation is linked to the lateral and noninvasive progression of colorectal tumors, whereas lowered levels of methylation might enhance the malignant potential via activation of NTSR1 [43].

In conclusion, this study constructed and validated a 2-gene oxidative stress-related LUAD prognostic model, MAP3K19 and NTSR1, which were significantly correlated with clinical variables (including LUAD pathological stage and TNM grade) and significantly affected the infiltration of immune cells in the tumor microenvironment (TME) of LUAD. And they are all involved in the process of oxidative stress and the energy metabolism network of ROS. The results of enrichment analysis based on the biological functions of GO, KEGG, and GSEA and pathway signaling patterns suggest that the molecular expression of MAP3K19 and NTSR1 and other immune cells help in the process of oxidative stress, even though conclusive evidence is still unavailable. Studies have shown that during apoptosis, immune cells are attracted and aggregated by a set of signals that enhance programmed cell death [44, 45]. In terms of bioinformatics, various studies have revealed a possible link between tumor and immune infiltration [46, 47]. Additionally, to palliative targeted therapy, monotherapy with new immunotherapies, such as immune checkpoint inhibitors (ICIs), has also demonstrated quite successful outcomes in some individuals with advanced LUAD [48]. In this study, through risk group-based immune annotation analysis, we found that macrophages, Tregs, and other types of immune cells and costimulation of immune-related roles were significantly enriched in both cohorts, suggesting that there may be potential regulatory mechanisms.

This study found some limitations. In a bioinformatics study, the weakness of the absence of experimental as well as clinical validation remains, and the utilization of alternative cutoff criteria, statistical methodologies, or analytical tools may provide varied results. Furthermore, building a prognostic model by focusing on a single marker may result in the deletion of numerous other potential prognostic genes. In conclusion, we created a novel 2-gene signature associated with oxidative stress that was shown to be an independent prognostic predictor of OS in LUAD. Through functional annotation analysis, the gene signature was associated with tumor immunity; nonetheless, its underlying process is not clear and needs to be explored further.
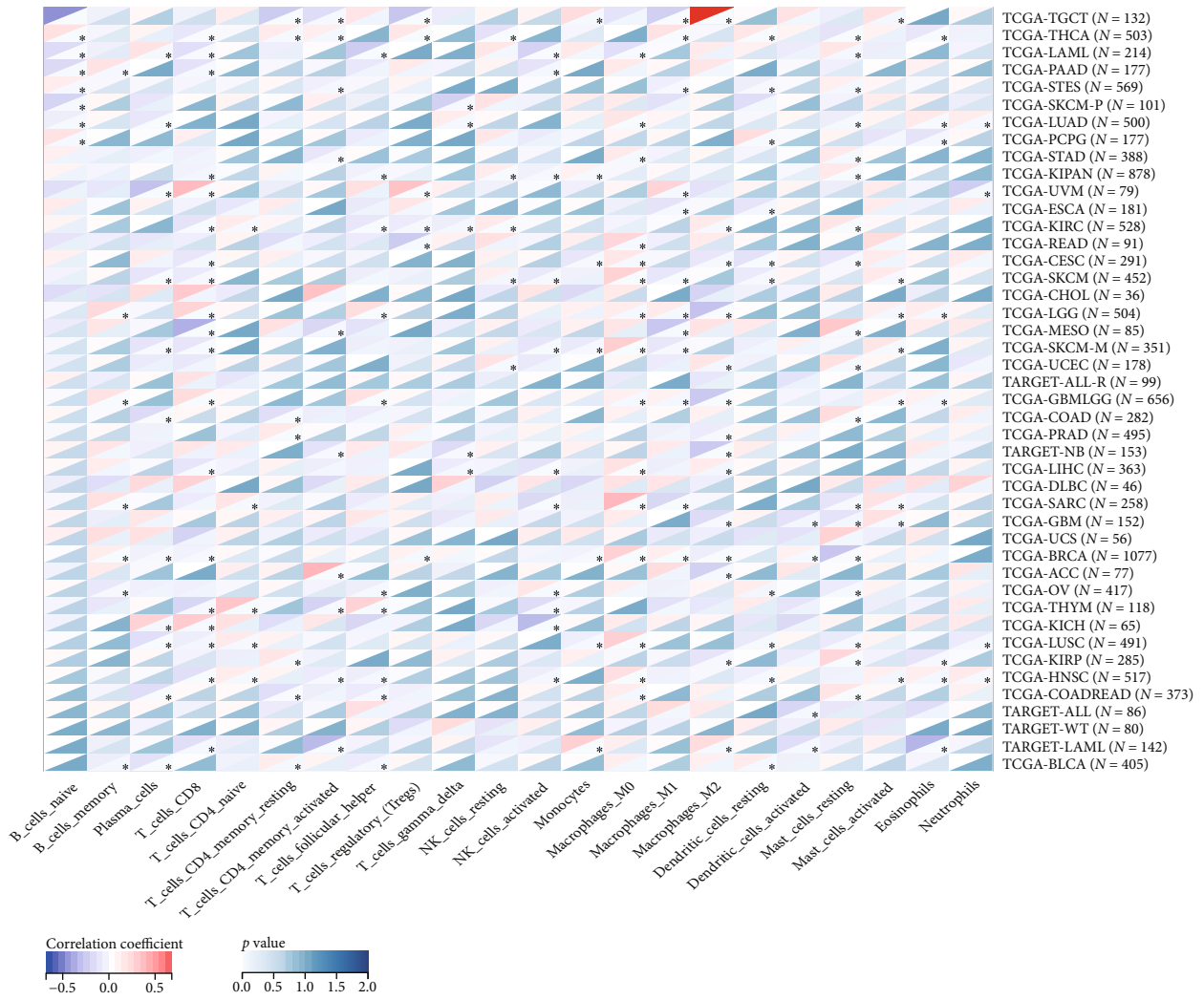
(a)

FIGURE 14: Continued.
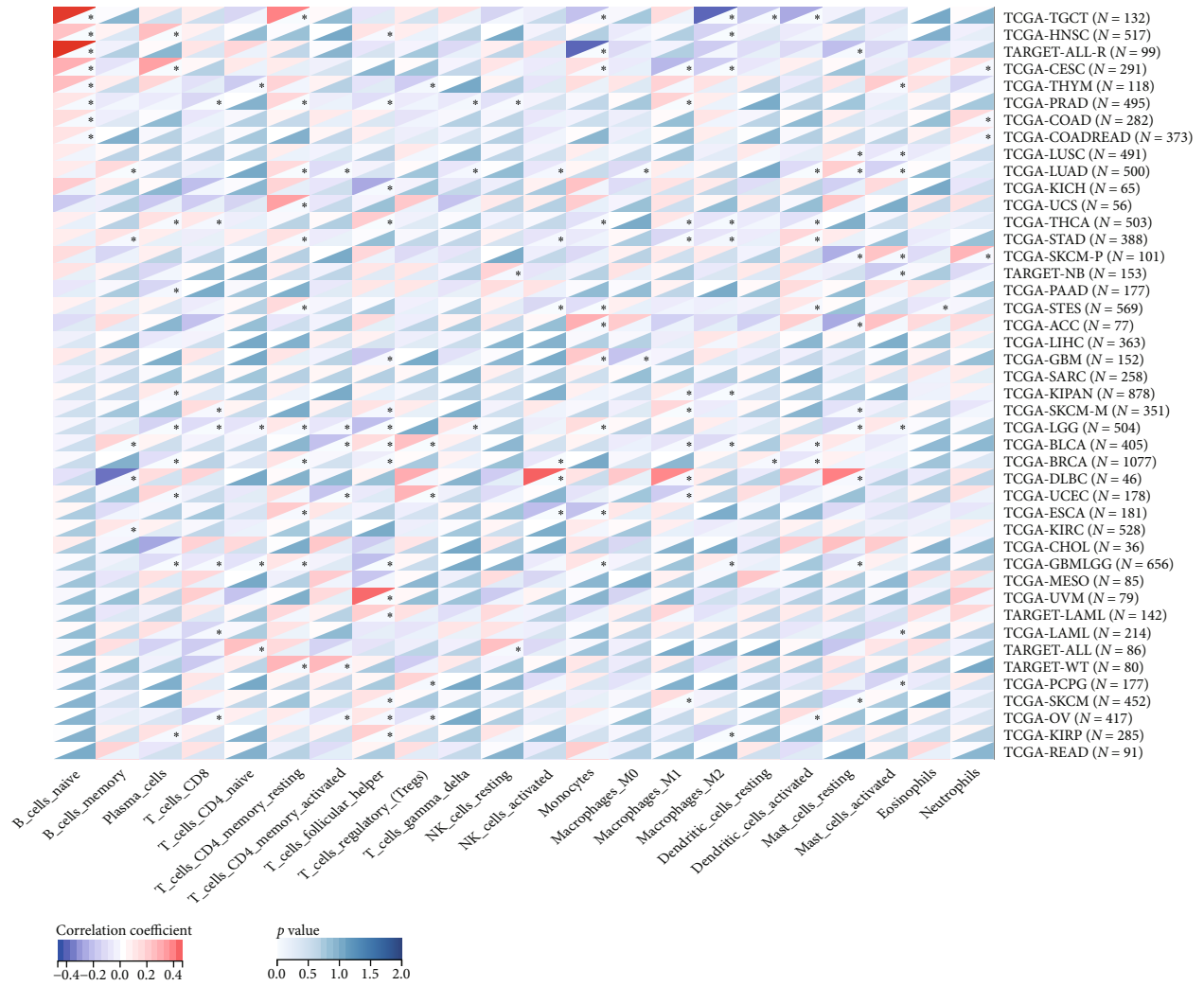
(b)
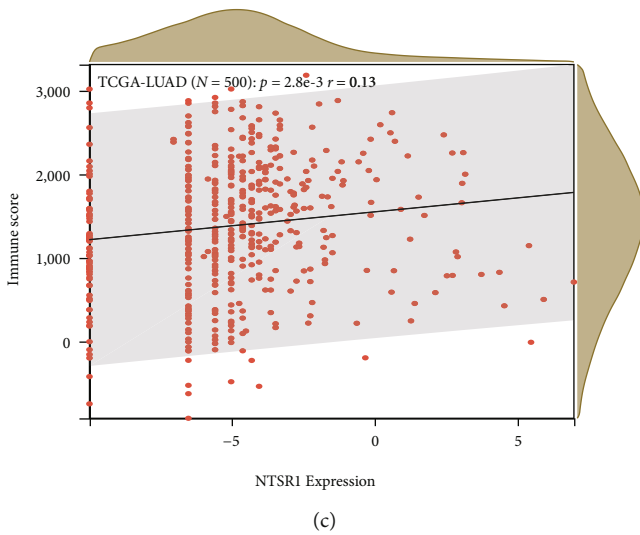


(c)
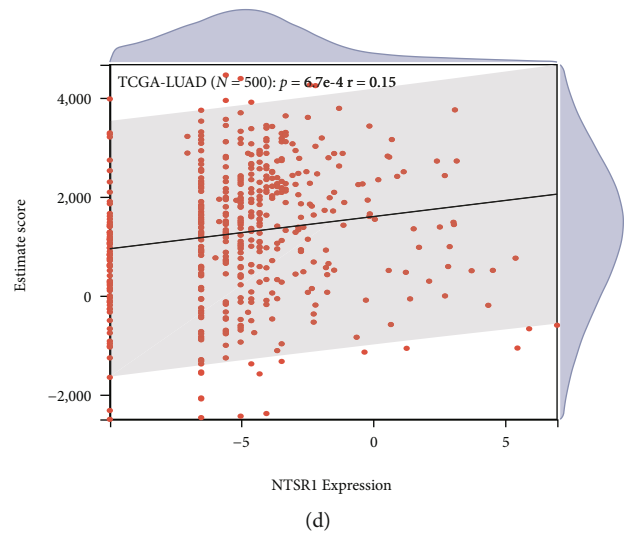


(d)

FIGURE 14: Continued.

(e)



(f)



(g)



(h)
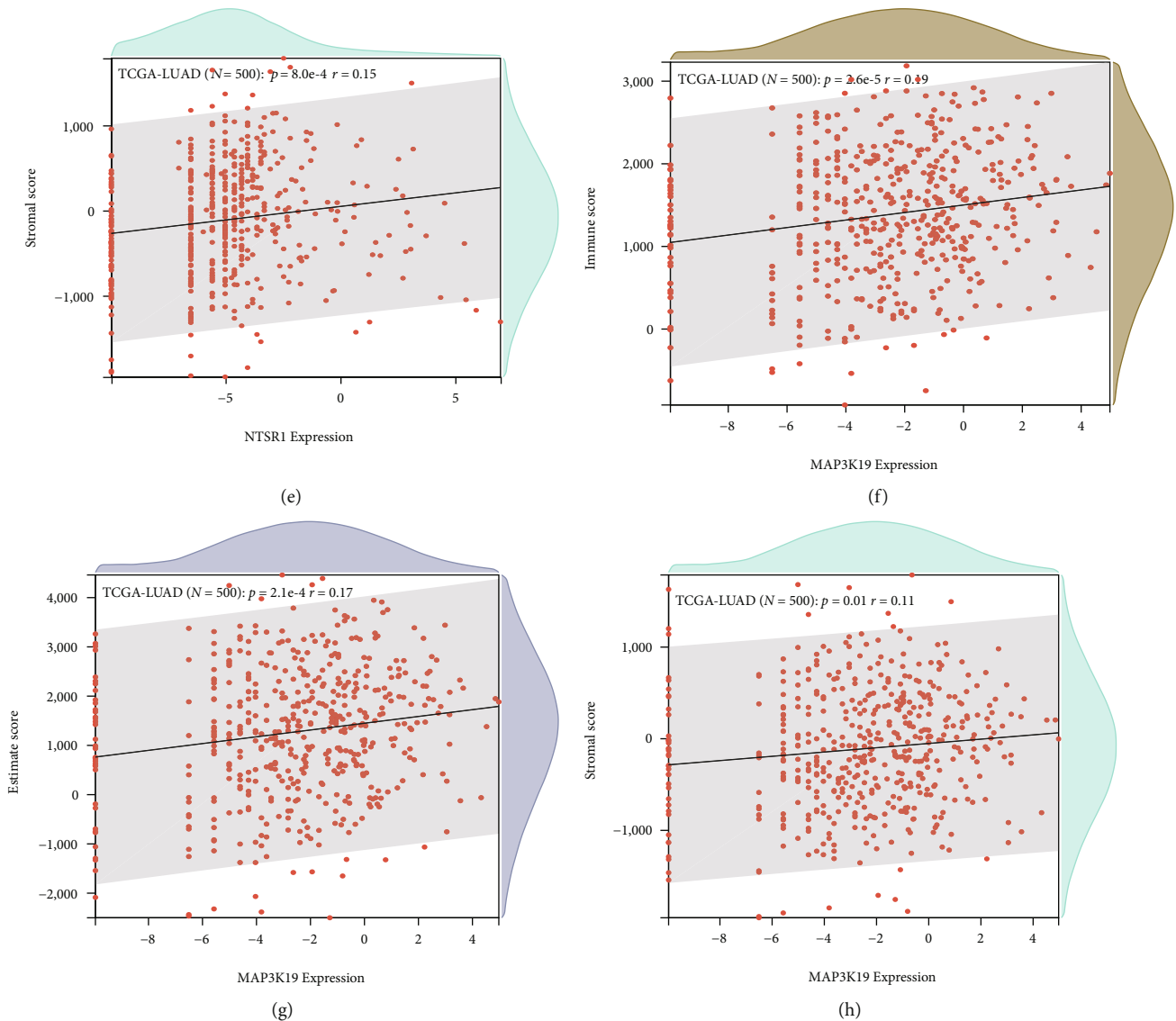
FIGURE 14: Immune cell infiltration analysis. (a, b) CIBERSORT analysis of MAP3K19 (a) and NTSR1 (b) heatmaps related to the infiltration abundance of immune cells in pan-cancer; (c–h) ESTIMATE analyzed the scores of tumor purity, stromal cell level, and immune cell infiltration level in tumor tissue and showed NTSR1 (c–e) and MAP3K19 (f–h) expression with immune score, ESTIMATE score, and stromal score, respectively, with a scatterplot positively correlated.

The TCGA-LUAD cohort availed both the clinical data and the expression data in this study. Among oxidative stress-related genes, we did differential expression analysis and univariate Cox analysis in order to screen 32 prognostic DEGs from 147 differentially expressed oxidative stress-related genes and utilized Lasso-penalized Cox regression analysis constructed 2 gene markers associated with prognosis. At the same time, by constructing a PPI network, we analyzed the distribution of 18 differentially oxidative stress hub genes in transport vesicles, endoplasmic reticulum lumen, etc., affecting blood pressure, hormone secretion, nucleotide second messengers, and other functions to regulate oxidative stress effect. To construct and validate the oxidative stress gene signature affecting the diagnosis and prognosis of LUAD, we randomly divided 445 patients with complete clinical informatics data in TCGA-LUAD into the training set (2/3 of the total, $n = 296$) and the test set (1/3 of the total, $n = 149$); the grouping is normally distributed. We first performed a univariate Cox regression analysis of gene expression premised on the expression data of 147 oxidative stress-related genes in the training and test sets, respectively. Thirty-five genes with prognostic values coexisting in both sets were identified. Then, we performed LASSO-Cox proportional hazards regression based on 35 prognostic-related differential oxidative stress genes in the training set as well as the validation set, respectively, and found that MAP3K19 and NTSR1 in the two sets showed a better prediction of poor prognosis in LUAD ability. Therefore, the prognostic independence of 2-gene signature prognostic factors, MAP3K19 and NTSR1, was further analyzed in
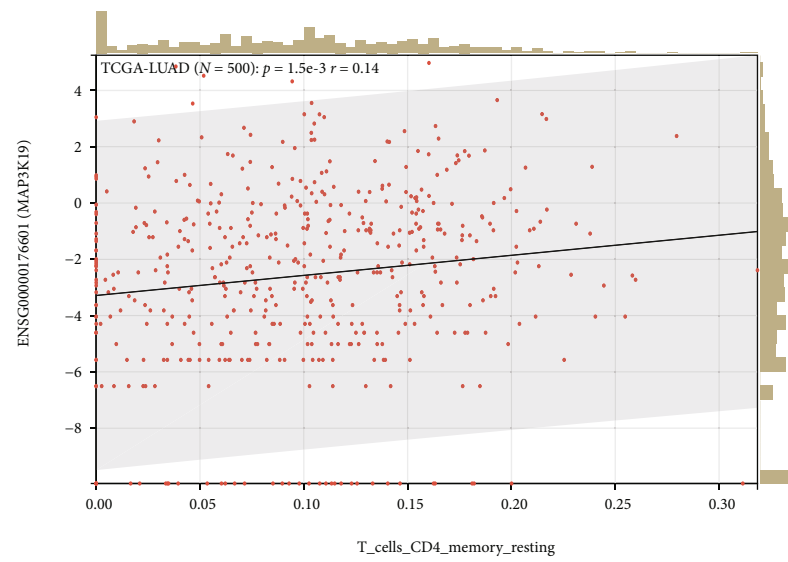
(a)
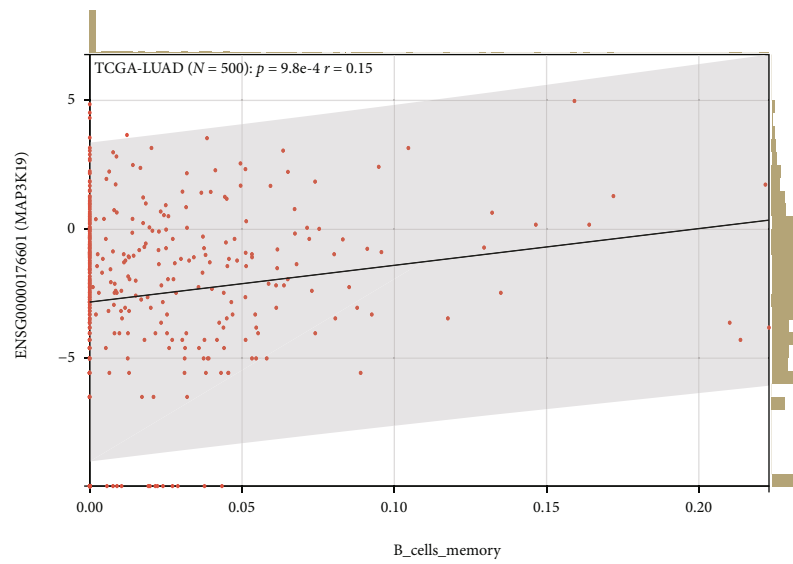


(b)

Figure 15: Continued.

(c)



(d)
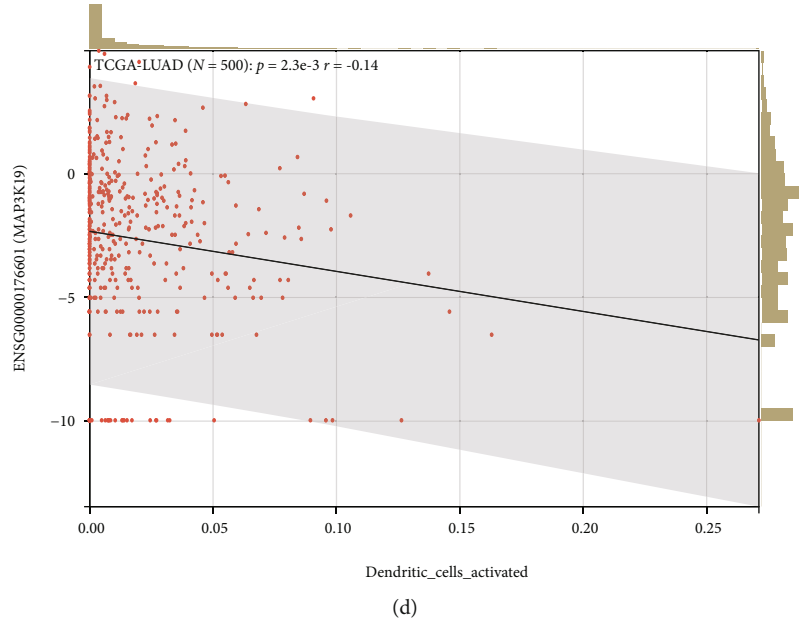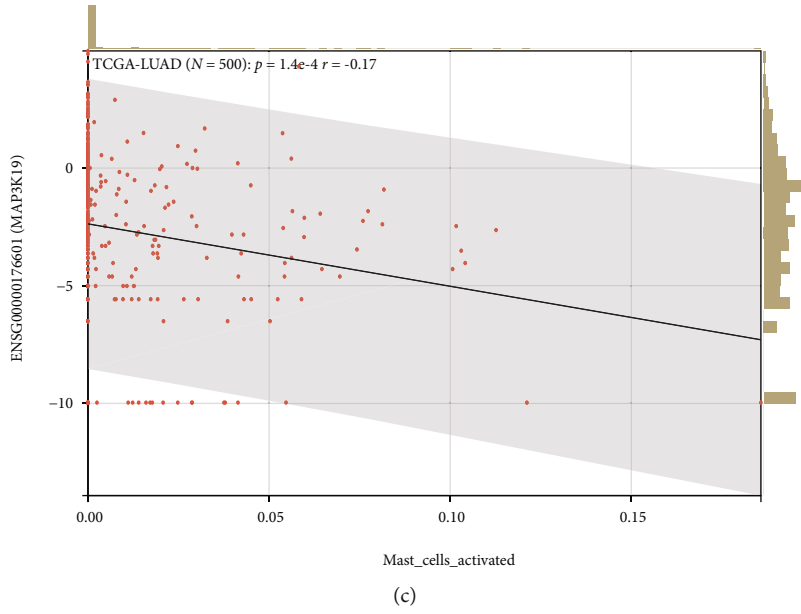
FIGURE 15: Continued.
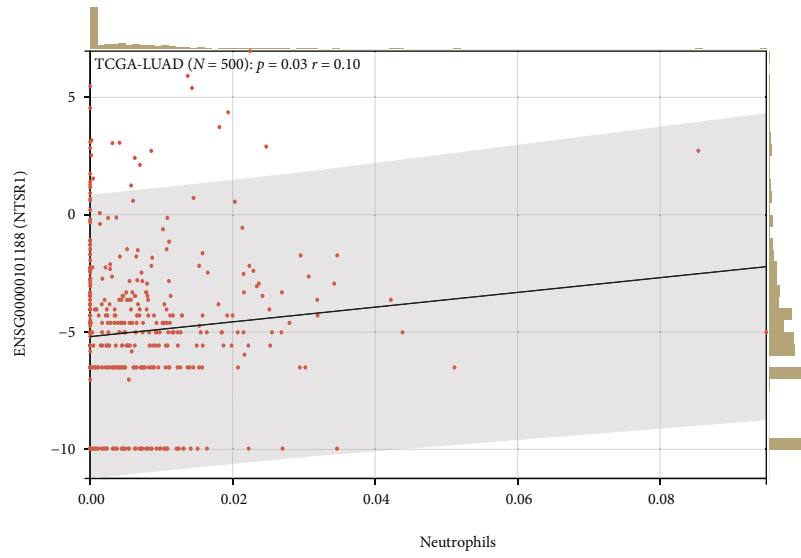
(e)
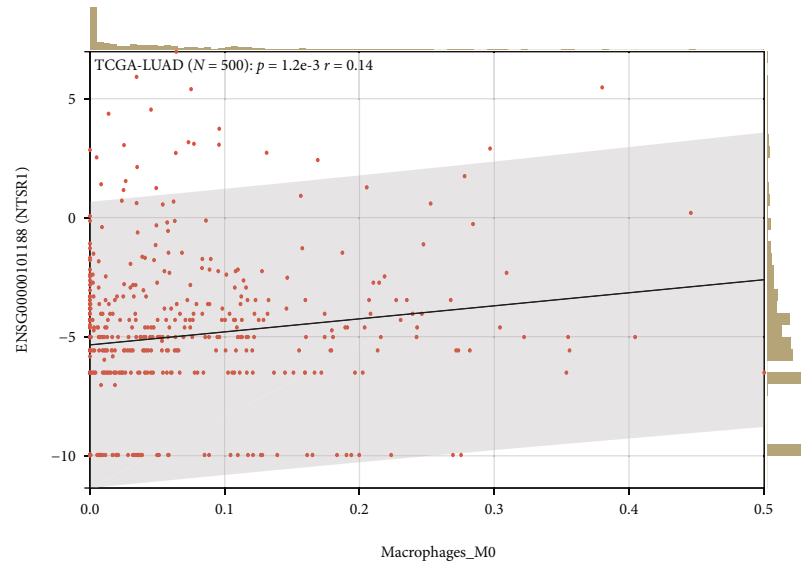


(f)

FIGURE 15: Continued.

(g)



(h)
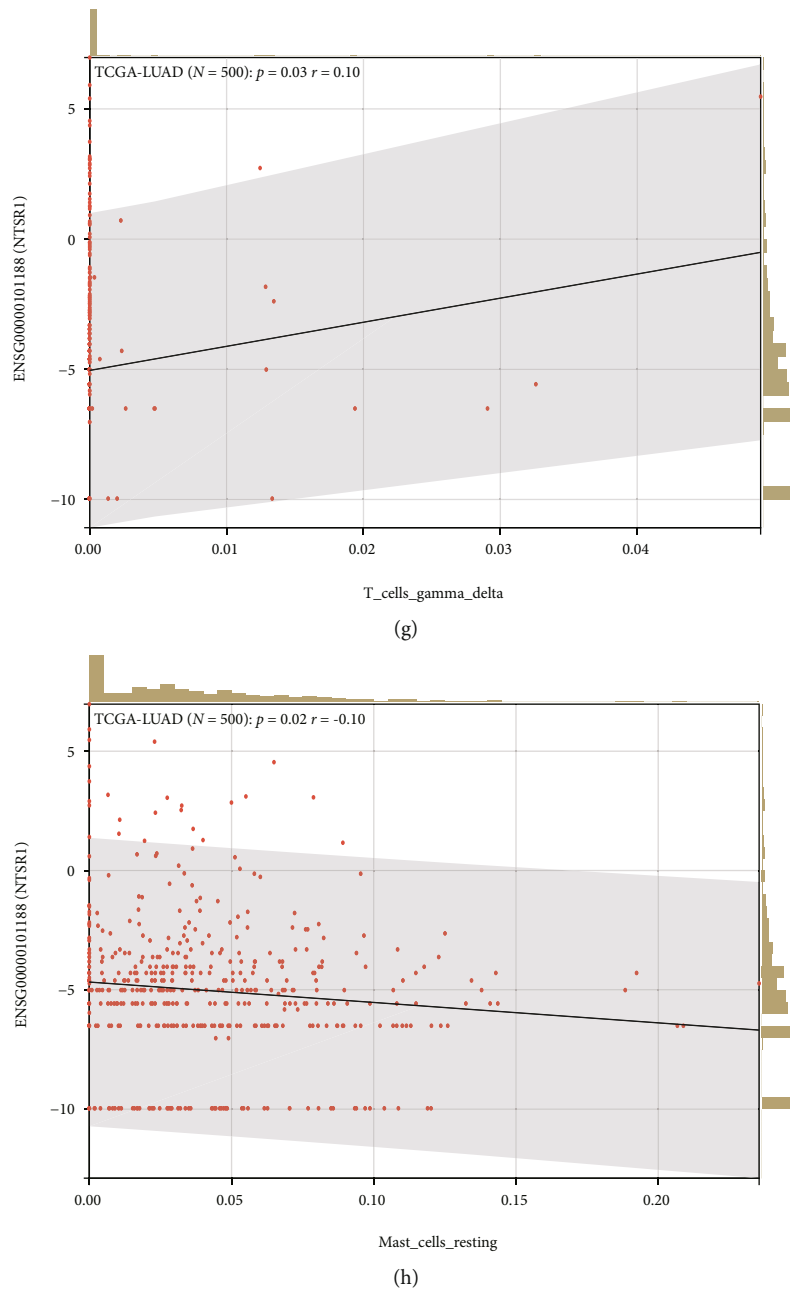
Figure 15: Correlation analysis of immune cell infiltration. (a, b) MAP3K19 was positively correlated with T cells CD4 memory resting and B cells memory; (c, d) MAP3K19 was negatively correlated with mast cells activated and dendritic cells activated; (e–g) NTSR1 was positively correlated with neutrophils, macrophages M0, and T cells gamma delta; (h) NTSR1 was negatively linked to mast cells resting.

combination with clinical features. A nomogram was constructed to effectively integrate prognostic variables and validated with calibration. Next, the difference analysis of MAP3K19 and NTSR1 in TCGA-LUAD and the correlation analysis of clinical variables indicated that both the high expression of NTSR1 and the low expression of MAP3K19 had a better diagnostic performance for the diagnosis of LUAD and the identification of poor prognosis. At the same time, MAP3K19 and NTSR1 were found to be significantly correlated with clinical variables such as LUAD pathological stage and TNM grade. Prognosis-related differential oxida-

tive stress gene signatures are independent prognostic factors in patients with LUAD. ESTIMATE analysis results of pan-cancer species analysis of MAP3K19 and NTSR1 in TCGA database. The abundances of 22 tumor-infiltrating immune compartments in LUAD samples were systematically recorded by the CIBERSORT algorithm and integrated with MAP3K19 and NTSR1 molecular profiles to analyze the degree of immune infiltration in LUAD. Subsequently, further immune score, stromal score, and ESTIMATE score in the pan-cancer TME revealed the potential roles of MAP3K19 and NTSR1 in regulating stromal/immune scores

and gene expression in tumors. This suggests that the oxidative stress prognostic model composed of MAP3K19 and NTSR1 may be involved in the molecular mechanism in the identification of high-risk populations with poor prognosis of LUAD.

## 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 conflicts of interest.

## Authors' Contributions

Xing Huang and Zhichai Lu conducted statistical analyses of the data and prepared the draft manuscript. Min He, Chen Chen, and Hanlin Ding edited the manuscript. Yipeng Feng and Yidan Sun provide critical comments to the manuscript. All authors checked and proofread the final version of the manuscript.

## Acknowledgments

## References

[1] J. W. Neal and G. W. Sledge, "Successes, toxicities and challenges in solid tumours," *Nature Reviews Clinical Oncology*, vol. 11, pp. 627-628, 2014.

[2] R. Herbst, D. Morgensztern, and C. Boshoff, "The biology and management of non-small cell lung cancer," *Nature*, vol. 553, pp. 446–454, 2018.

[3] T. S. Bowen, G. Schuler, and V. Adams, "Skeletal muscle wasting in cachexia and sarcopenia: molecular pathophysiology and impact of exercise training," *Journal of Cachexia, Sarcopenia and Muscle*, vol. 6, no. 3, pp. 197–207, 2015.

[4] T. Cheng, M. Xu, H. Zhang et al., "KLHDC8A expression in association with macrophage infiltration and oxidative stress predicts unfavorable prognosis for glioma," *Oxidative Medicine and Cellular Longevity*, vol. 2022, Article ID 2694377, 14 pages, 2022.

[5] W. Wang, S. Ren, Z. Wang, C. Zhang, and J. Huang, "Increased expression of TTC21A in lung adenocarcinoma infers favorable prognosis and high immune infiltrating level," *International Immunopharmacology*, vol. 78, article 106077, 2020.

[6] D. R. Camidge, W. Pao, and L. V. Sequist, "Acquired resistance to TKIs in solid tumours: learning from lung cancer," *Nature Reviews Clinical Oncology*, vol. 11, pp. 473–481, 2014.

[7] C. Chong and P. Jänne, "The quest to overcome resistance to EGFR-targeted therapies in cancer," *Nature Medicine*, vol. 19, pp. 1389–1400, 2013.

[8] J. J. Lin, G. J. Riely, and A. T. Shaw, "Targeting ALK: precision medicine takes on drug resistance," *Cancer Discovery*, vol. 7, no. 2, pp. 137–155, 2017.

[9] Q. Yao, X. Ge, Z. Lu, J. Shi, J. Shen, and J. Chen, "HAUS augmin-like complex subunit 1 influences tumour microenvironment and prognostic outcomes in glioma," *Journal of Oncology*, vol. 2022, Article ID 8027686, 19 pages, 2022.

[10] J. R. Lieffers, O. F. Bathe, K. Fassbender, M. Winget, and V. E. Baracos, "Sarcopenia is associated with postoperative infection and delayed recovery from colorectal cancer resection surgery," *British Journal of Cancer*, vol. 107, no. 6, pp. 931–936, 2012.

[11] W. Wang, Z. Lu, M. Wang et al., "The cuproptosis-related signature associated with the tumor environment and prognosis of patients with glioma," *Frontiers in Immunology*, vol. 13, article 998236, 2022.

[12] Y. Hamaguchi, T. Kaido, S. Okumura et al., "Proposal for new diagnostic criteria for low skeletal muscle mass based on computed tomography imaging in Asian adults," *Nutrition*, vol. 32, no. 11-12, pp. 1200–1205, 2016.

[13] S. T. McSorley, D. H. Black, P. G. Horgan, and D. C. McMillan, "The relationship between tumour stage, systemic inflammation, body composition and survival in patients with colorectal cancer," *Clinical Nutrition*, vol. 37, no. 4, pp. 1279–1285, 2018.

[14] H. W. Jung, J. W. Kim, J. Y. Kim et al., "Effect of muscle mass on toxicity and survival in patients with colon cancer undergoing adjuvant chemotherapy," *Supportive Care in Cancer*, vol. 23, no. 3, pp. 687–694, 2015.

[15] J. M. Kyriakis and J. Avruch, "Mammalian MAPK Signal Transduction Pathways Activated by Stress and Inflammation: A 10-Year Update," *Physiological Reviews*, vol. 92, no. 2, pp. 689–737, 2012.

[16] G. Malietzis, N. Johns, H. O. Al-Hassi et al., "Low muscularity and myosteatosis is related to the host systemic inflammatory response in patients undergoing surgery for colorectal cancer," *Annals of Surgery*, vol. 263, no. 2, pp. 320–325, 2016.

[17] B. K. Pedersen and M. A. Febbraio, "Muscles, exercise and obesity: skeletal muscle as a secretory organ," *Nature Reviews Endocrinology*, vol. 8, no. 8, pp. 457–465, 2012.

[18] A. S. Tolmacheva and G. A. Nevinsky, "Essential Protective Role of Catalytically Active Antibodies (Abzymes) with Redox Antioxidant Functions in Animals and Humans," vol. 23, no. 7, Article ID 3898, 2022.

[19] M. Diaz and P. Casali, "Somatic immunoglobulin hypermutation," *Current Opinion in Immunology*, vol. 14, no. 2, pp. 235–240, 2002.

[20] Team RC, *R: a language and environment for statistical computing*, R Foundation for Statistical Computing VAfhwR-po, 2005.

[21] M. C. Burger, C. Zhang, P. N. Harter et al., "CAR-engineered NK cells for the treatment of glioblastoma: turning innate effectors into precision tools for cancer immunotherapy," *Frontiers in Immunology*, vol. 10, p. 2683, 2019.

[22] A. J. Cruz-Jentoft, G. Bahat, J. Bauer et al., "Sarcopenia: revised European consensus on definition and diagnosis," *Age and Ageing*, vol. 48, no. 1, pp. 16–31, 2019.

[23] T. J. Osterman, M. Terry, and R. S. Miller, "Improving cancer data interoperability: the promise of the minimal common," *JCO Clinical Cancer Informatics*, vol. 4, pp. 993–1001, 2020.

[24] Y. Sun, Y. Zhang, S. Ren et al., "Low expression of RGL4 is associated with a poor prognosis and immune infiltration in lung adenocarcinoma patients," *International Immunopharmacology*, vol. 83, article 106454, 2020.

[25] A. K. M. Kassambara and P. Biecek, "survminer: drawing survival curves using 'ggplot2'. R package version 0.4.8," 2020, https://CRAN.R-project.org/package=survminer.

[26] P. Blanche, J.-F. Dartigues, and H. Jacqmin-Gadda, "Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks," *Statistics in Medicine*, vol. 32, no. 30, pp. 5381–5397, 2013.

[27] G. Bano, C. Trevisan, S. Carraro et al., "Inflammation and sarcopenia: a systematic review and *meta*-analysis," *Maturitas*, vol. 96, pp. 10–15, 2017.

[28] Z. Z. Fulop, S. Gurzu, T. Bara et al., "Lymph node ratio, an independent prognostic factor for patients with stage II-III rectal carcinoma," *Pathology, Research and Practice*, vol. 215, no. 6, article 152384, 2019.

[29] D. K. Simanshu, D. V. Nissley, and F. McCormick, "RAS proteins and their regulators in human disease," *Cell*, vol. 170, no. 1, pp. 17–33, 2017.

[30] Z. Wang, X. Ge, J. Shi, B. Lu, X. Zhang, and J. Huang, "*SPTSSA* is a prognostic marker for glioblastoma associated with tumor-infiltrating immune cells and oxidative stress," *Oxidative Medicine and Cellular Longevity*, vol. 2022, Article ID 6711085, 17 pages, 2022.

[31] M. G. Kris, B. E. Johnson, L. D. Berry et al., "Using multiplexed assays of oncogenic drivers in lung cancers to select targeted," *JAMA*, vol. 311, no. 19, pp. 1998–2006, 2014.

[32] J. Rotow and T. G. Bivona, "Understanding and targeting resistance mechanisms in NSCLC," *Nature Reviews Cancer*, vol. 17, no. 11, pp. 637–658, 2017.

[33] L. Neckers and P. Workman, "Hsp90 molecular chaperone inhibitors: are we there yet?," *Clinical Cancer Research*, vol. 18, no. 1, pp. 64–76, 2012.

[34] G. M. DeNicola, F. A. Karreth, T. J. Humpton et al., "Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis," *Nature*, vol. 475, no. 7354, pp. 106–109, 2011.

[35] C. Villa, A. Vasiljevic, M. L. Jaffrain-Rea et al., "A standardised diagnostic approach to pituitary neuroendocrine tumours (PitNETs): a European Pituitary Pathology Group (EPPG) proposal," *Virchows Archiv*, vol. 475, no. 6, pp. 687–692, 2019.

[36] S. A. Boehme, K. Franz-Bacon, D. N. DiTirro, T. W. Ly, and K. B. Bacon, "MAP3K19 is a novel regulator of TGF-$\beta$ signaling that impacts bleomycin-induced lung injury and pulmonary fibrosis," *PLoS One*, vol. 11, no. 5, article e0154874, 2016.

[37] S. A. Boehme, K. Franz-Bacon, J. Ludka, D. N. DiTirro, T. W. Ly, and K. B. Bacon, "MAP3K19 Is Overexpressed in COPD and Is a Central Mediator of Cigarette Smoke-Induced Pulmonary Inflammation and Lower Airway Destruction," *PLoS One*, vol. 11, no. 12, article e0167169, 2016.

[38] I. C. Jones, M. S. Espindola, R. Narayanan et al., "Targeting MAP3K19 prevents human lung myofibroblast activation both *in vitro* and in a humanized SCID model of idiopathic pulmonary fibrosis," *Scientific Reports*, vol. 9, article 19796, 2019.

[39] V. T. Hoang, K. Nyswaner, P. Torres-Ayuso, and J. Brognard, "The protein kinase MAP3K19 phosphorylates MAP2Ks and thereby activates ERK and JNK kinases and increases viability of KRAS-mutant lung cancer cells," *Journal of Biological Chemistry*, vol. 295, no. 25, pp. 8470–8479, 2020.

[40] D. A. Levine and The Cancer Genome Atlas Research Network, "Integrated genomic characterization of endometrial carcinoma," *Nature*, vol. 497, no. 7447, pp. 67–73, 2013.

[41] Q. Ou-yang, X. He, A. Yang, B. Li, and M. Xu, "Interference with NTSR1 Expression Exerts an Anti-Invasion Effect via the Jun/miR-494/SOCS6 Axis of Glioblastoma Cells," *Cellular Physiology and Biochemistry*, vol. 49, no. 6, pp. 2382–2395, 2018.

[42] R. A. Brooks, G. F. Fleming, R. R. Lastra et al., "Current recommendations and recent progress in endometrial cancer," *CA: a Cancer Journal for Clinicians*, vol. 69, no. 4, pp. 258–279, 2019.

[43] S. Sud, J. Holmes, M. Eblan, R. Chen, and E. Jones, "Clinical characteristics associated with racial disparities in endometrial cancer outcomes: a surveillance, epidemiology and end results analysis," *Gynecologic Oncology*, vol. 148, no. 2, pp. 349–356, 2018.

[44] M. R. Elliott and K. S. Ravichandran, "The dynamics of apoptotic cell clearance," *Developmental Cell*, vol. 38, no. 2, pp. 147–160, 2016.

[45] E. A. Ross, A. Devitt, and J. R. Johnson, "Macrophages: the good, the bad, and the gluttony," *Frontiers in Immunology*, vol. 12, article 708186, 2021.

[46] Y. Wang, F. Ren, Z. Song, X. Wang, and X. Ma, "Multiomics profile and prognostic gene signature of m6A regulators in uterine corpus endometrial carcinoma," *Journal of Cancer*, vol. 11, no. 21, pp. 6390–6401, 2020.

[47] S. Zuo, M. Wei, S. Wang, J. Dong, and J. Wei, "Pan-cancer analysis of immune cell infiltration identifies a prognostic Immune-Cell Characteristic Score (ICCS) in Lung Adenocarcinoma," *Frontiers in Immunology*, vol. 11, article 1218, 2020.

[48] J. Li, L. Zhu, Y. Shi, J. Liu, L. Lin, and X. Chen, "m6A demethylase FTO promotes hepatocellular carcinoma tumorigenesis via mediating PKM2 demethylation," *American Journal of Translational Research*, vol. 11, no. 9, pp. 6084–6092, 2019.

*Research Article*

# Prognostic Assessment of Oxidative Stress-Related Genes in Colorectal Cancer and New Insights into Tumor Immunity

**Zilu Chen** [1] **Kun Mei** [2] **Yao Xiao** [1] **Yan Xiong** [1] **Wei Long** [1] **Qin Wang** [1] **Jiang Zhong** [3] **Dongmei Di** [2] **Yunxi Ge** [3] **Yi Luo** [4,5] **Ziyun Li** [1,6] **Yan Huang** [3] **Renjun Gu** [1,7] and **Bin Wang** [2]

[1]*Nanjing University of Chinese Medicine, Nanjing 210023, China*
[2]*Department of Cardiothoracic Surgery, The Third Affiliated Hospital of Soochow University, Changzhou 213003, China*
[3]*Department of Ultrasound, Nanjing Hospital of Chinese Medicine Affiliated to Nanjing University of Chinese Medicine, Nanjing 210001, China*
[4]*Department of Oncology, Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing University of Chinese Medicine, Nanjing, Jiangsu 210028, China*
[5]*Department of Oncology, Jiangsu Province Hospital on Integration of Chinese and Western Medicine, Nanjing, Jiangsu 210028, China*
[6]*School of Acupuncture and Tuina, School of Regimen and Rehabilitation, Nanjing University of Chinese Medicine, Nanjing 210023, China*
[7]*School of Chinese Medicine & School of Integrated Chinese and Western Medicine, Nanjing University of Chinese Medicine, Nanjing 210023, China*

Correspondence should be addressed to Yan Huang; jacob6666@163.com, Renjun Gu; renjungu@hotmail.com, and Bin Wang; colin_iverson@163.com

Oxidative stress is crucial to the biology of tumors. Oxidative stress' potential predictive significance in colorectal cancer (CRC) has not been studied; nevertheless here, we developed a forecasting model based on oxidative stress to forecast the result of CRC survival and enhance clinical judgment. The training set was chosen from the transcriptomes of 177 CRC patients in GSE17536. For validation, 65 samples of colon cancer from GSE29621 were utilized. For the purpose of choosing prognostic genes, the expression of oxidative stress-related genes (OXEGs) was found. Prognostic risk models were built using multivariate Cox regression analysis, univariate Cox regression analysis, and LASSO regression analysis. The outcomes of the western blot and transcriptome sequencing tests were finally confirmed. ATF4, CARS2, CRP, GPX1, IL1B, MAPK8, MRPL44, MTFMT, NOS1, OSGIN2, SOD2, AARS2, and FOXO3 were among the 14 OXEGs used to build prognostic characteristics. Patients with CRC were categorized into low-risk and high-risk groups according on their median risk scores. Cox regression analysis using single and multiple variables revealed that OXEG-related signals were independent risk factors for CRC. Additionally, the validation outcomes from western blotting and transcriptome sequencing demonstrated that OXEGs were differently expressed. Using 14 OXEGs, our work creates a predictive signature that may be applied to the creation of new prognostic models and the identification of possible medication candidates for the treatment of CRC.

## 1. Introduction

Colorectal cancer (CRC) is a common malignant tumor of the gastrointestinal tract [1]. It is the second most lethal malignancy in China, after lung cancer, with an incidence of roughly 40,800 persons. The fourth malignant tumor accounts for around 195,600 annual deaths [2]. Currently, colorectal cancer patients can receive treatment through

| | *p* value | Hazard ratio | |
|---|---|---|---|
| AARS2 | <0.001 | 0.126 (0.037–0.422) | |
| ATF4 | 0.018 | 2.623 (1.176–5.849) | |
| CARS2 | 0.023 | 7.555 (1.316–43.384) | |
| CRP | 0.004 | 15.745 (2.394–103.571) | |
| CYBA | 0.024 | 2.160 (1.106–4.220) | |
| FOX03 | 0.001 | 0.384 (0.217–0.681) | |
| GPX1 | <0.001 | 3.863 (1.734–8.607) | |
| IL1B | 0.007 | 1.440 (1.104–1.879) | |
| MAPK8 | 0.026 | 3.219 (1.153–8.983) | |
| MRPL44 | <0.001 | 3.251 (1.617–6.537) | |
| MTFMT | 0.007 | 1.784 (1.175–2.708) | |
| NOS1 | 0.019 | 5.694 (1.330–24.379) | |
| OSGIN2 | 0.022 | 3.490 (1.196–10.183) | |
| SOD2 | 0.029 | 1.350 (1.032–1.766) | |
| TXN | 0.011 | 2.017 (1.176–3.457) | |



(a)



(b)

FIGURE 1: Continued.

(c)

FIGURE 1: Construction of the prognostic model in GEO-CRC. (a) According to the results of univariate Cox regression analysis, a total of 15 genes were identified as prognostic genes; (b) LASSO coefficient profiles of the prognostic genes; and (c) turning optimal parameter (lambda) screening in the LASSO model.
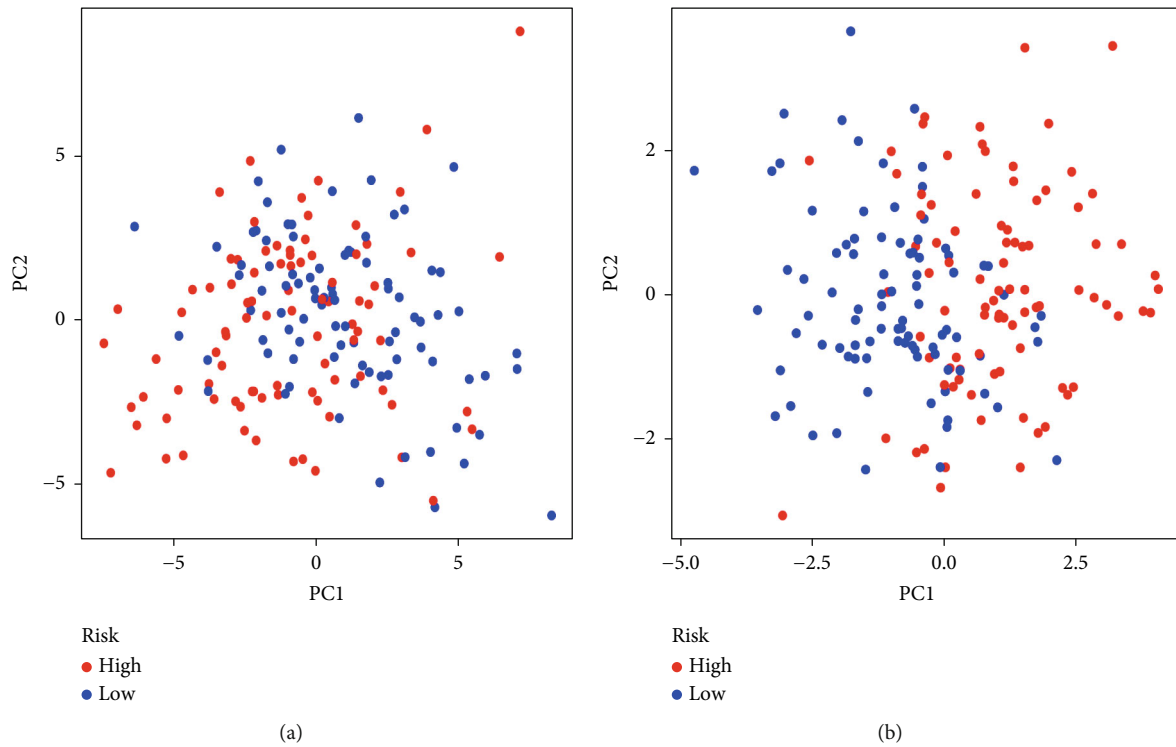


FIGURE 2: Principal component analysis. (a) PCA analysis was performed in high- and low-risk groups. (b) PCA analysis of prognosis model.

(a)



(b)



(c)

Figure 3: Continued.

AUC at 1 years: 0.946
AUC at 3 years: 0.684
AUC at 5 years: 0.724

(d)

| | $p$ value | Hazard ratio |
|---|---|---|
| Age | 0.076 | 1.015 (0.998–1.031) |
| Gender | 0.845 | 0.962 (0.653–1.418) |
| Grade | 0.202 | 1.327 (0.860–2.047) |
| Stage | 0.878 | 0.983 (0.791–1.222) |
| Risk score | <0.001 | 3.199 (2.386–4.291) |

(e)

| | $p$ value | Hazard ratio |
|---|---|---|
| Age | 0.164 | 1.013 (0.995–1.031) |
| Gender | 0.431 | 1.178 (0.783–1.773) |
| Grade | 0.519 | 1.150 (0.751–1.761) |
| Stage | 0.835 | 0.975 (0.768–1.238) |
| Risk score | <0.001 | 3.179 (2.369–4.266) |

(f)

Risk, AUC = 0.804
Age, AUC = 0.535
Gender, AUC = 0.493
Grade, AUC = 0.573
Stage, AUC = 0.456

(g)

Figure 3: Continued.

FIGURE 3: Evaluation of the prognostic signature in GEO-CRC. (a) Kaplan-Meier (K-M) curve of overall survival for the training set; (b) K-M curves of overall survival for the validation set; (c) ROC curves of 1-year, 3-year, and 5-year survival rates for the training set; (d) ROC curves of 1-year, 3-year, and 5-year survival rates for the validation set; (e, f) Cox regression analysis of risk scores and other clinical characteristics (age, gender, grade, stage); (g) ROC curve of clinical characteristics for the training set; and (h) ROC curve of clinical characteristics for the validation set.

surgery, endoscopic procedures, radiotherapy, chemotherapy, and immunotherapy [3], but the overall survival rate of colorectal cancer patients has not dramatically enhanced [4]. On account of the high invasiveness and paucity of awareness of early physical examination, diagnosis is often made once symptoms have advanced or the disease has metastasized, which poses significant challenges for the prognosis and course of treatment.

Tumor development consists of a multitude of complex variables. The oxidative stress factor plays a significant role in numerous stages of tumor advancement, including the transformation of normal cells into tumor cells, proliferation, tumor angiogenesis, and metastasis [5–8]. The term "tumor of oxidative stress" refers to an improper regulation mechanism of oxidative signaling and oxidative damage of macromolecules caused by an imbalance in the body's oxidation and antioxidation system of mutual limitation [9].

The effect of oxidative stress mainly involves reactive nitrogen species (RNS) and reactive oxygen species (ROS) as an outcome of chemical reactions; the resulting effect is often seen as a double-edged sword, with great controversy over the tumor-promoting and tumor-suppressing effects [6, 10], the specific effect of ROS levels on tumor cells themselves, [11] sensitivity and lack of oxygen, and the tumor microenvironment of regulatory factors. When the content of ROS is insufficient to break the balance between oxidative and antioxidant systems in the tumor growth environment, it can participate in the regulation of epithelial-mesenchymal transition (EMT), tumor angiogenesis, and other processes by activating PI3K/Akt and NF-κB signaling pathways [12–14]. ROS can also promote the metastasis and

proliferation of tumor cells [15]. At the same time, low levels of ROS can cause occasional DNA base mismatch and DNA damage in the body, which will be repaired immediately. However, when a large amount of ROS causes the amassing of DNA damage in tumor cells to the extent that it cannot be repaired, conventional base excision repair [16] and nucleotide excision repair cannot remove the damaged DNA. Thus, the tumor cells have to undergo programmed cell death.

Plasma medicine is an emerging academic area that combines clinical medicine, physics, and life sciences. Cold atmospheric plasma-activated medium (PAM), which is dependent on plasma production of active substances transferred to the medium, has a wide range of uses [17]. Its anticancer effect is generally considered to be through ROS and RNS [18–20]. Therefore, PAM was selected as an oxidative stress source to verify the relationship between oxidative stress and tumor prognosis.

In recent years, the relationship between the accumulation of ROS and immunotherapy has been more and more frequently mentioned. It is particularly important to clarify the technique of the body's immune feedback to a tumor and the escape of tumor cells from immune effect during oxidative stress [21, 22]. ROS production also affects the anticancer effects of CD8+ and CD4+ T cells in the tumor microenvironment [23, 24]. Many clinical trials and successful checkpoint immunotherapy instances have shown the crucial role that the cellular immune system plays in the treatment of cancer.

In the process of oxidative stress and colorectal cancer research, early prediction models and prognostic gene screening are particularly important. In this research, based

(a)



(b)



(c)



(d)
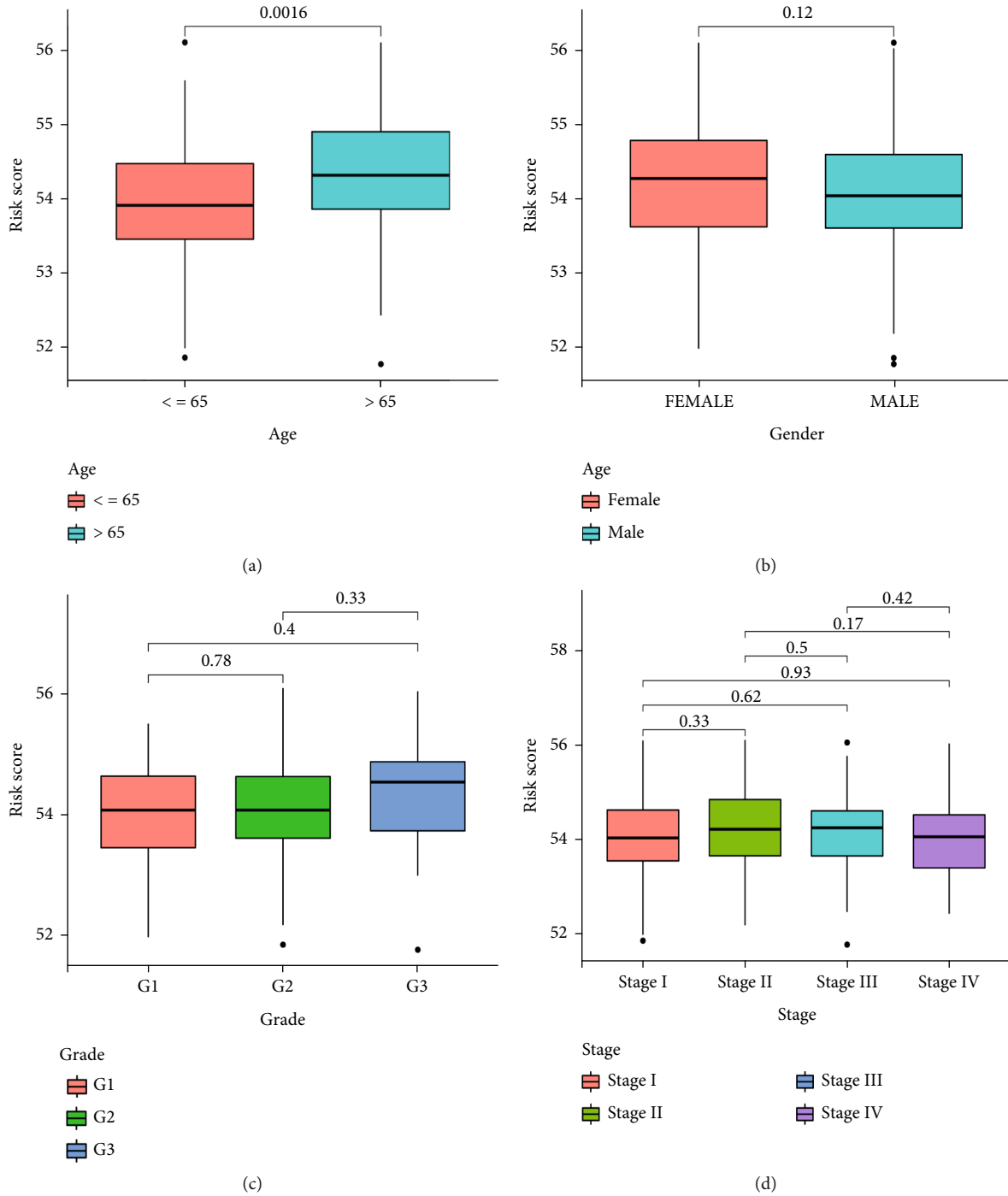
FIGURE 4: Relationship between the risk groups and clinical features in GEO-CRC. (a) Correlation between risk score and age; (b) correlation between risk score and gender; (c) correlation between risk score and grade; and (d) correlation between risk score and stage.

on information from Gene Expression Omnibus (GEO), related genes that may affect the prognosis of CRC were studied.

## 2. Materials and Methods

*2.1. Data Sources and Processing.* The CRC cohort's transcriptional dataset with aligning clinical data were downloaded from the Gene Expression Omnibus (GEO)

database. After comprehensive screening, data sets GSE17536 and GSE29621 were chosen for this study, in which 177 patients with symptoms of colorectal cancer were included in GSE17536 as a training set, and 65 patients with colorectal cancer were included in GSE29621 as the validation set.

*2.2. Screening for Oxidative Stress-Related Genes.* Eighty genes related to oxidative stress were retrieved from
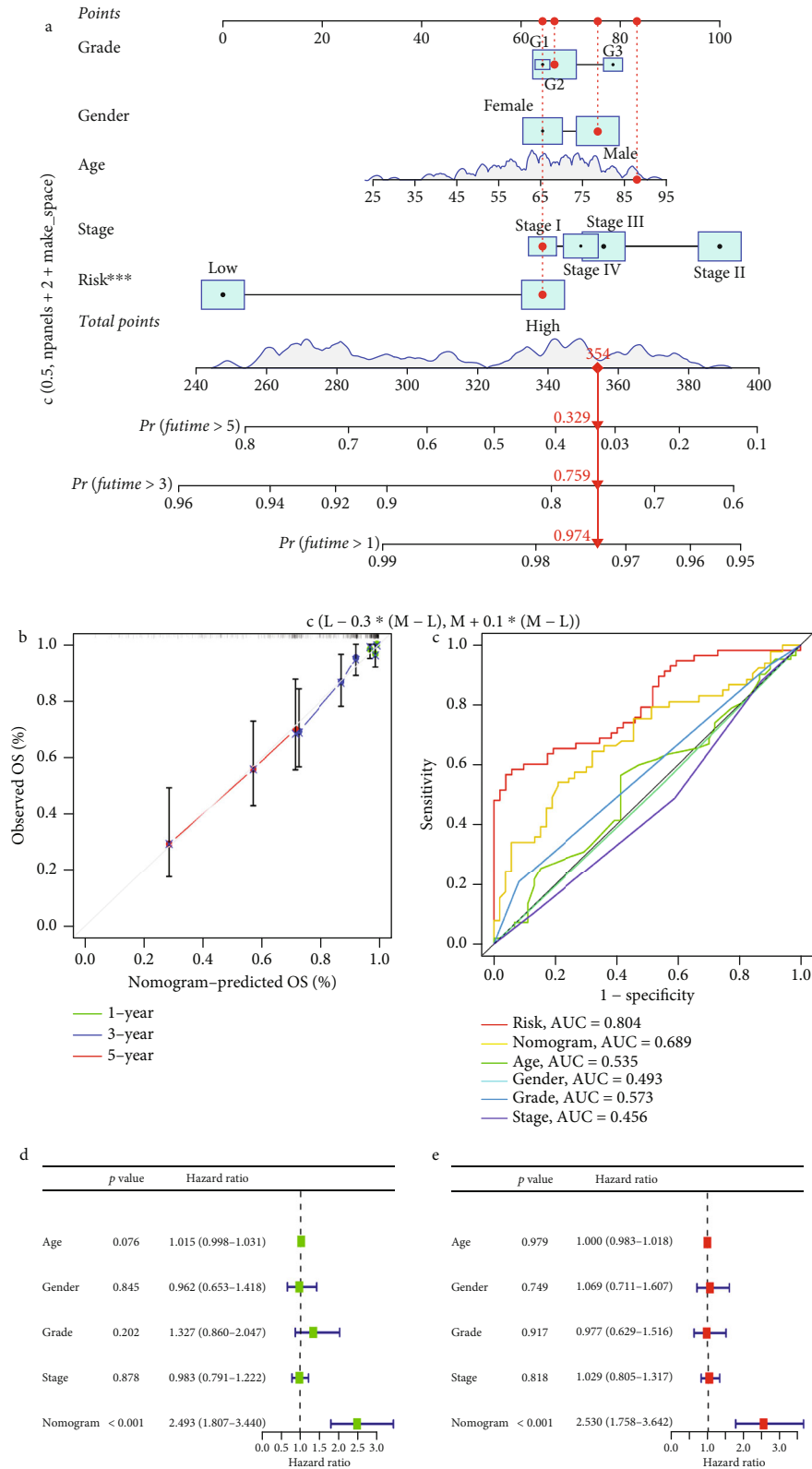
FIGURE 5: Nomogram prediction model and evaluation. (a) Nomogram of age, sex, stage, stage, and risk score for predicting 1-, 3-, and 5-year survival. (b) 1-, 3-, and 5-year calibration curves for the GEO dataset; (c) ROC curve of nomogram; (d) Cox regression analysis of risk score and other clinical characteristics of nomogram (age, gender, grade, stage).

(a)



(b)

Figure 6: Continued.

(c)
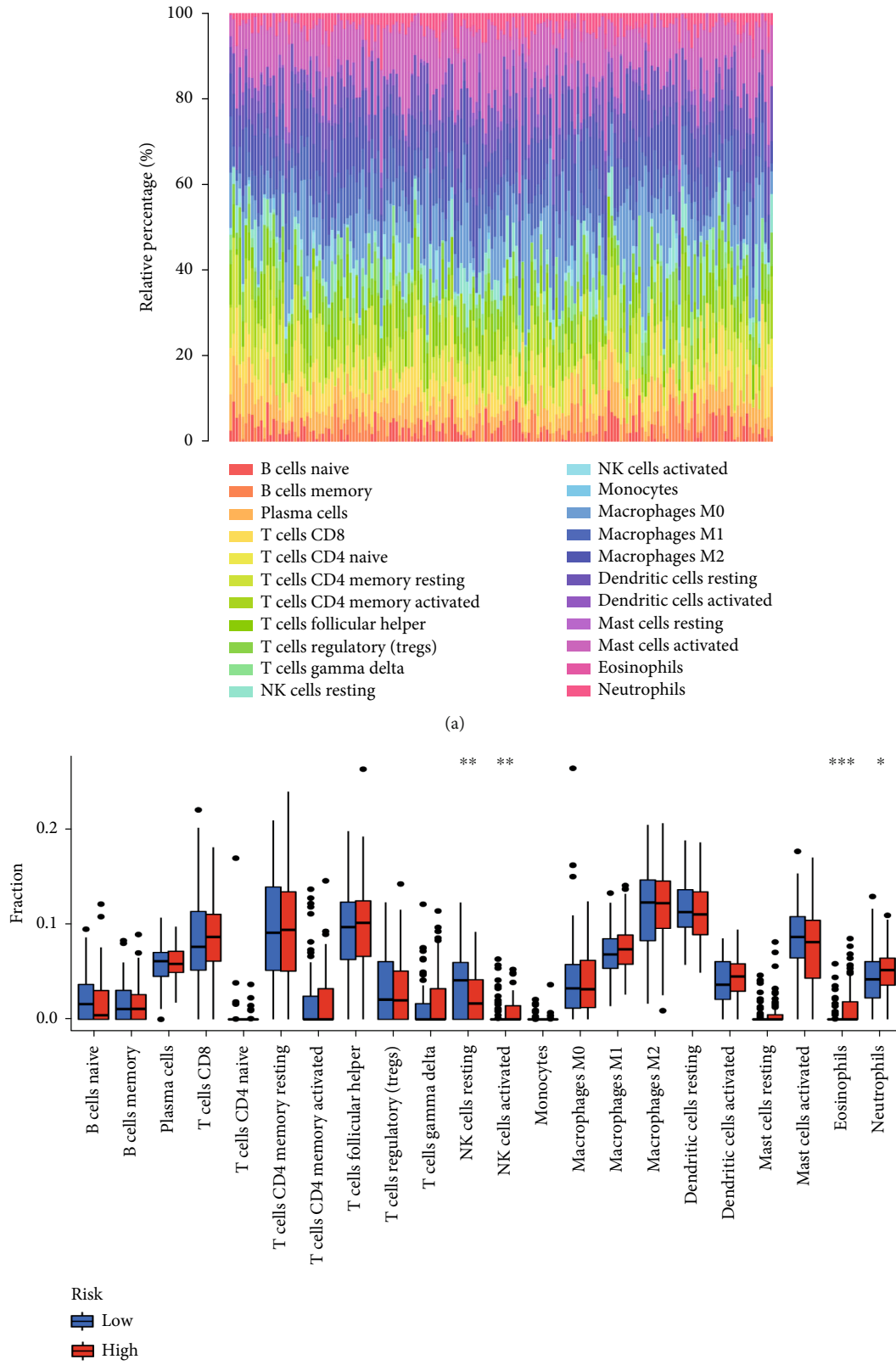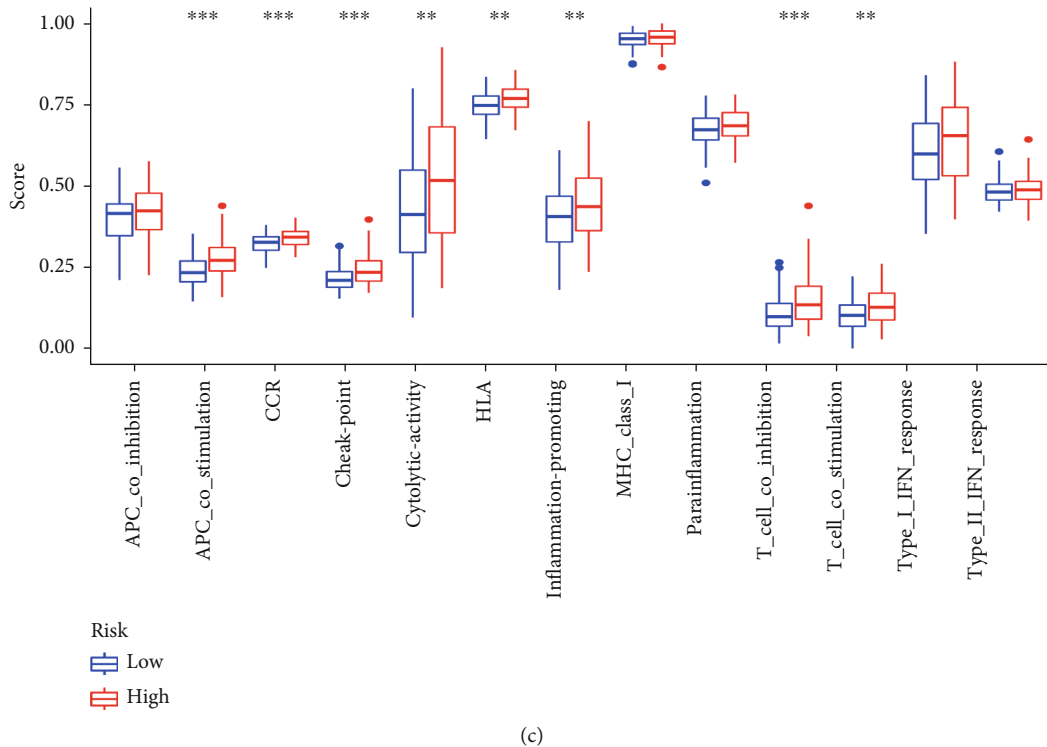
Figure 6: Immunoinfiltration analysis. (a) Tumor infiltrating immune cell distribution map; (b) correlation between risk score model and tumor infiltrating immune cells; (c) correlation between risk score model and tumor infiltrating immune cell function. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

GeneCards (https://www.genecards.org/). The cut-offs were set as relevance score > 20 (Supplement 1). Subsequently, 77 expressed genes of oxidative stress were identified by Venn diagram package (1.7.1).

2.3. Establishment and Assessment of Prognostic Risk Score Model. Univariate Cox regression screening was performed for prognostic OXEGs. The LASSO algorithm was implemented to find the value of the minimum error of cross validation and obtain the best prognostic gene of the model. Finally, stable OXEGs were constructed as the final prognostic model. Kaplan-Meier curves were utilized to establish prognostic differences between groups, and ROC curves were used to calculate the 1-, 3-, and 5-year survival of patients. Finally, the correlation between patients' low-risk and high-risk groups and clinical information was calculated.

2.4. Nomogram Prognosis Prediction Model Establishment. We used the "RMS" package in version R (4.2.0) to plot the lipopograph model in combination with patient's age, sex, grade, stage, and risk score. To show the accord between the actual survival probabilities at 1, 3, and 5 years and those predicted by the nomogram, calibration curves were developed. Finally, the model was validated using the ROC curve, multivariate Cox regression, and univariate Cox regression.

2.5. Correlation between Low-Risk and High-Risk Groups for Immune Cell Infiltration. CIBERSORT R package was used

to establish the amount of tumor-infiltrating immune cells in CRC tumor samples. Finally, we examined the functional differences of tumor immune cells through "reshape2," "GSVA," and "GSEABase" software package in R version (4.2.0).

2.6. Gene Enrichment Examination between High-Risk and Low-Risk Groups. In order to show the influence of potential biological pathways in the differential expression of OXEGs, r-packet clusterProfiler was utilized for gene ontology (GO) enrichment examination and KEGG pathway examination. The molecular signatures database (MSigDB) used with the Gene Set Enrichment Analysis (GSEA) software version (holdings) was C2 (C2. Cp. Kegg. 7.5.1. Entrez. GMT) used to evaluate the link between biological and genetic traits. The significance level was set to $P < 0.05$, and the number of arbitrary sample permutations was set to 1000.

2.7. Transcriptome Sequencing Validation of Colorectal Cancer. Three CRC tissues from the Anorectal Department, Nanjing Hospital of Traditional Chinese Medicine were collected and matched with normal tissues for transcriptome sequencing. Patients received no neoadjuvant chemotherapy or radiotherapy. Consent was obtained from the study participants prior to study commencement. Clinicopathological characteristics of these patients were also collected. The collected tissues were frozen in liquid nitrogen. The Declaration of Helsinki, the World Medical Association's code of ethics for human experimentation, was followed throughout the

(a)



(b)
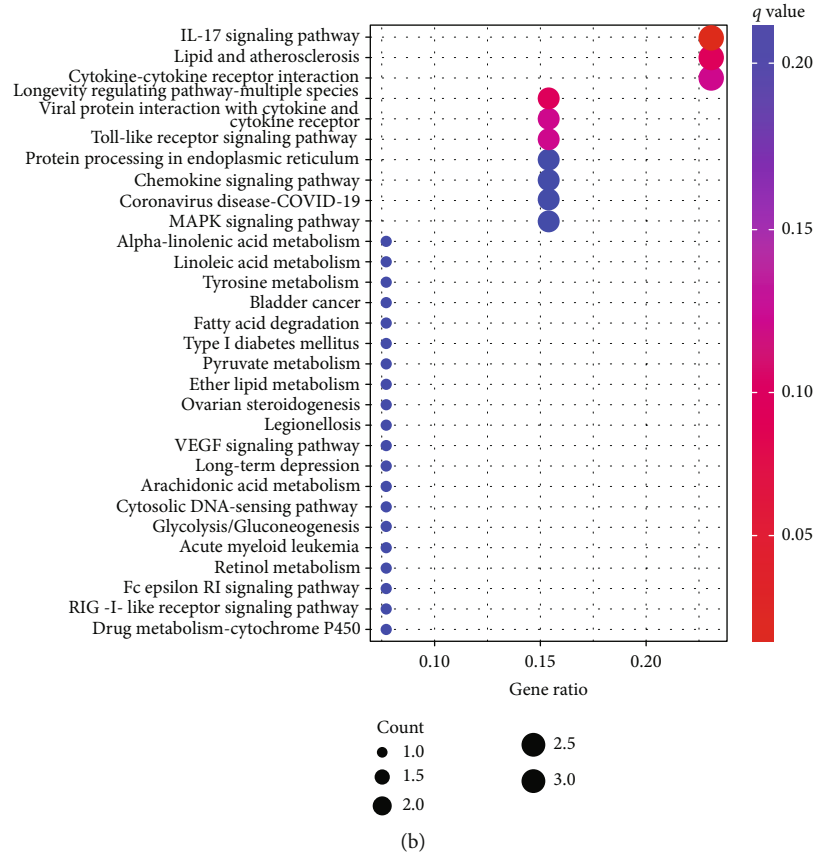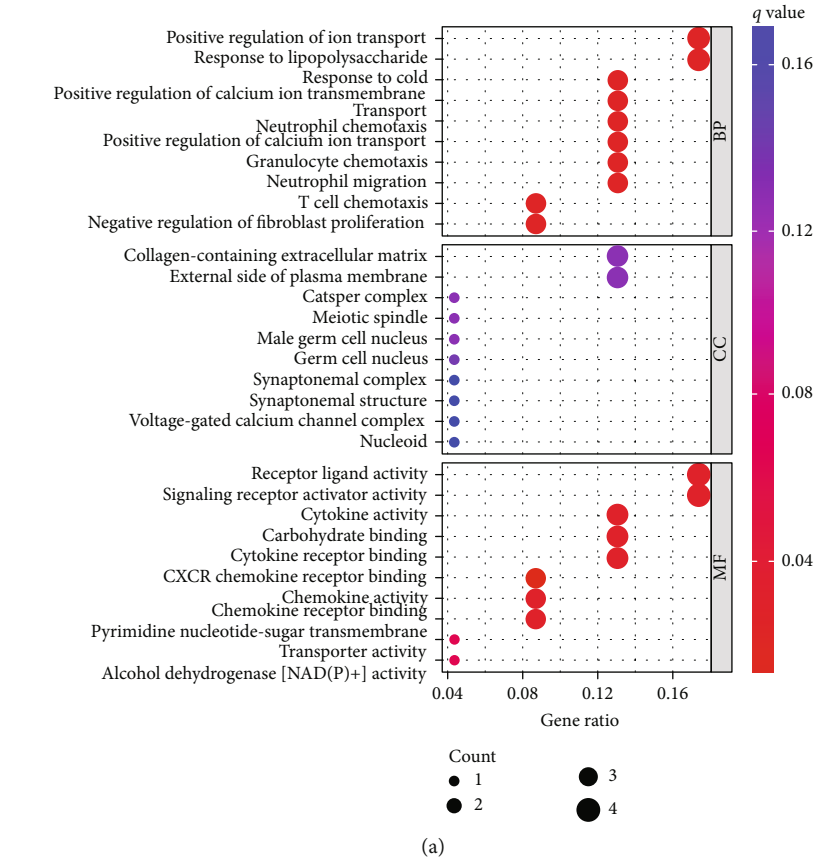
Figure 7: GO and KEGG enrichment analyses. (a) GO enrichment analysis; (b) KEGG enrichment analysis.
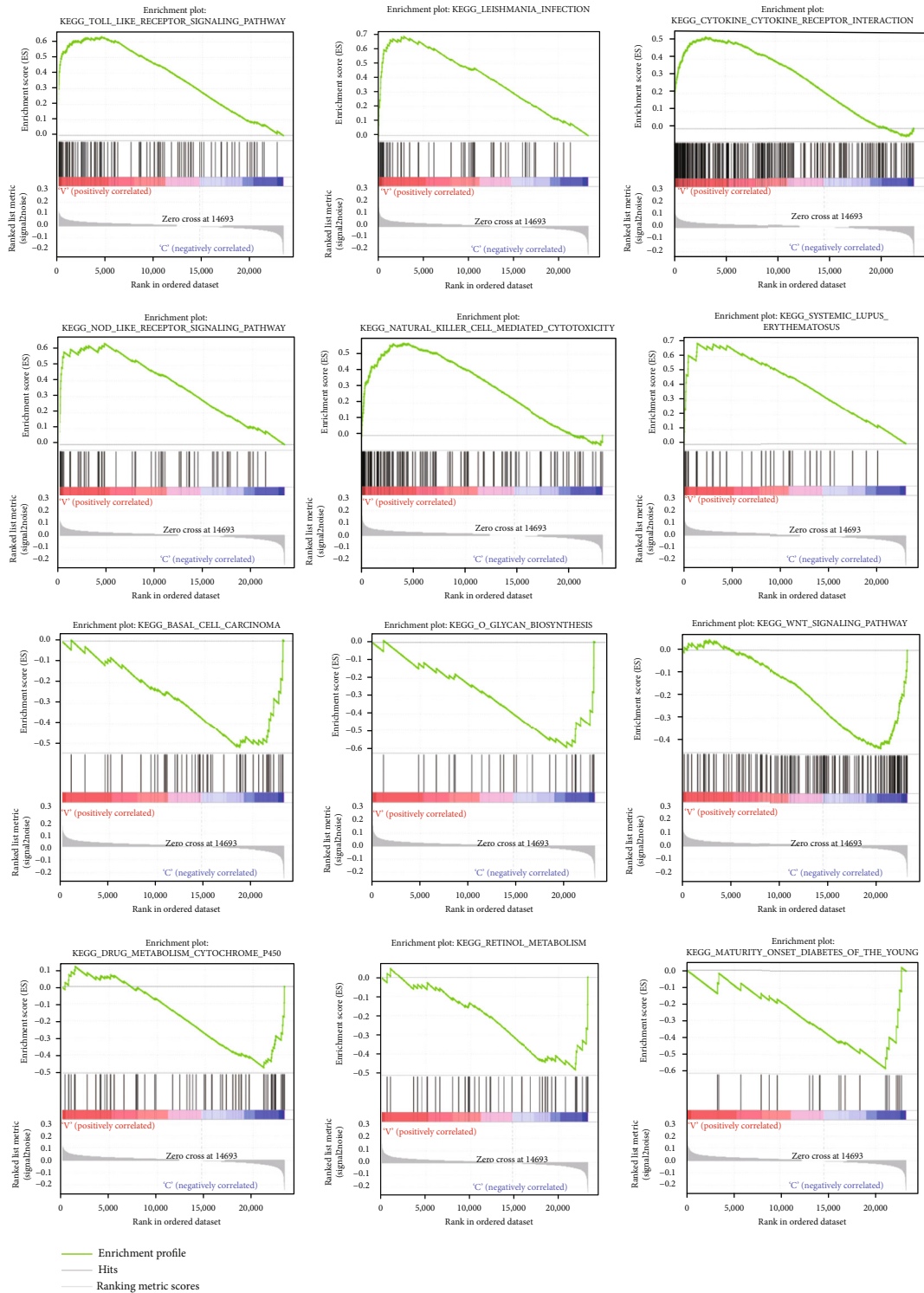
FIGURE 8: GSEA enrichment analysis.

performance of this study. The ethics committee of the Nanjing Hospital of Traditional Chinese Medicine had to approve the study. Informed consent was signed by all patients whose tissue samples were collected before the study began.

*2.8. Cell Line Culture and Treatment.* The CRC cell line SW480 was grown in 1640 media with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin, and 5% carbon dioxide in a humid environment at $37^\circ C$. All materials for cell culture were purchased from Gibco, USA. Cell lines within 10 generations were selected to reduce the influence of passage on experimental results. The particular plasma jet was designed by Nanjing Tech University. PAM is made by plasma jet spraying PBS solution at a distance of 5 mm. In

(a)



(b)

(c)
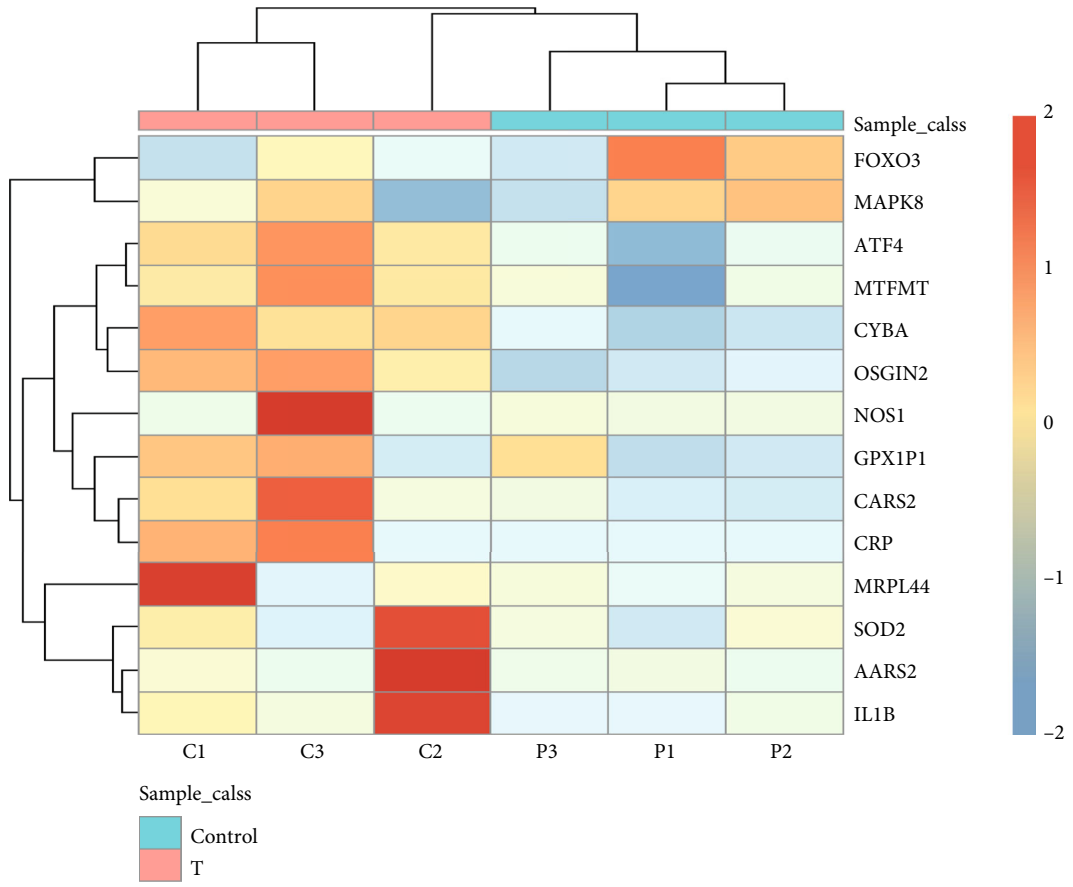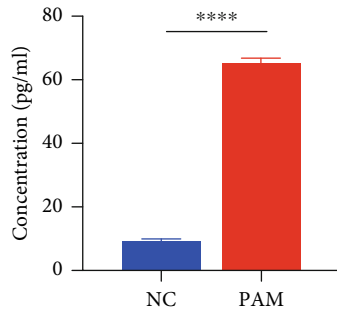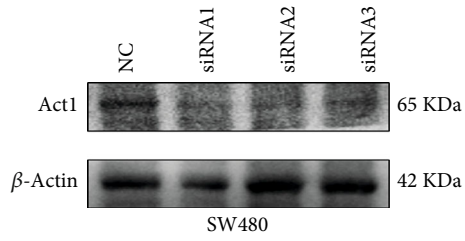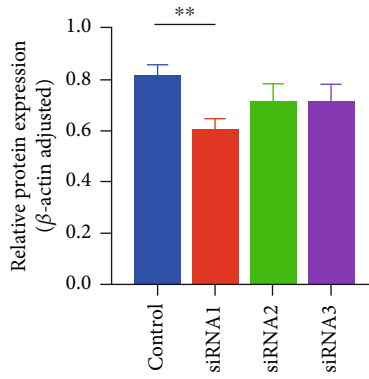
FIGURE 9: Continued.
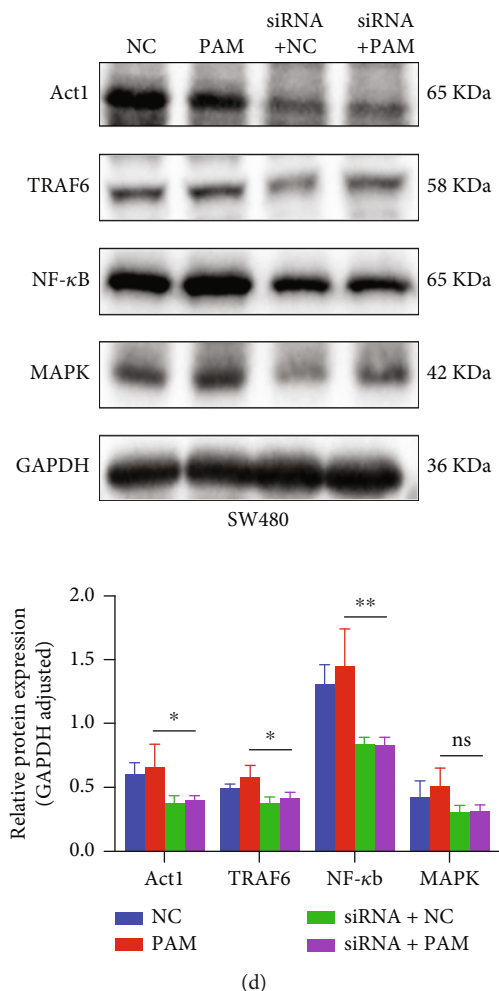
FIGURE 9: Validation of prognostic genes and enrichment pathways. (a) Prognostic gene sequencing heat map; (b) the increased expression of IL-17 in SW480 cell line was verified by ELISA; (c) validation of ACT1 knockdown in SW480 cell line; (d) validation of IL-17 pathway proteins. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

the control group, 200 μl of 50% PAM solution was added to the medium and grown for 24 hours.

### 2.9. Enzyme-Linked Immunosorbent Assay (ELISA).

Each group's culture medium supernatant was collected. Each batch of cells' supernatant was examined using ELISA kits from Quanzhou Ruixin Biological Technology Co., Ltd. in Quanzhou, China, to measure the amount of IL-17A components present.

### 2.10. Cell Transfection.

The SW480 ACT1-knockdown cells were transfected using LipofectamineTM 2000 Transfection Reagent (11668019). The Supplement 2 document includes the siRNA sequence screening. Six-well plates containing SW80 cells were planted with 5,105 cells per well. 250 l of Opti-MEM was added to two EP tubes once the cells had acquired a confluency of 60–70%. Then, 5 μg siRNA was added to one tube and 5 μl Lipofectamine 2000 to the other tube. After mixing, the tubes were left at noncold normal temperature for 5 min and then, the liquid of the two EP tubes was mixed gently and placed on an ultraclean table

for 20 min, followed by the incubator for 6 h before changing to complete medium and continuing to culture. The fresh medium was replaced after 24 hours, and 200 μl of 50% PAM solution was added to the medium and cultured for 24 hours.

### 2.11. Western Blotting.

RIPA lysis buffer (Epizyme Biomedical Technology, Shanghai, China) was used to extract total proteins in SW480 on ice (with 1% protease and phosphatase inhibitors). Before the samples were differentiated using 12% SDS-PAGE and conveyed to the polyvinylidene difluoride (PVDF) membranes, the total protein content was evaluated using the BCA Protein Assay Kit (TransGen Biotech, Beijing, China). After 15 minutes of blocking with Quick-Block Blocking Buffer from Beyotime Biotechnology in Shanghai, China, membranes were grown with various diluted primary antibodies overnight at 4°C. Following three TBST washes, the membranes were incubated for 1 hour with a second antibody that had been diluted. Proteintech (Wuhan, China) provided the ACT1 (26692-1-AP) primary antibody, while Abmart Technology provided the TRAF6

(T55175S), NF-B (p65) (T55034S), and MAPK P38 (T40075S) primary antibodies (Shanghai, China). The secondary antibody, goat antirabbit IgG (H + L) HRP (BL003A), was bought from Biosharp (Hefei, China).

*2.12. Statistical Analysis.* GraphPad Prism 8.0 was used to conduct all statistical analyses (GraphPad Software, San Diego, CA, USA). The Student $t$-test was used to compare the variations between the means of the two groups. The mean and standard deviation for all statistical data were displayed (SD). Statistics were identified as significant when $P < 0.05$.

# 3. Results

*3.1. Prognostic Risk Signature Construction of OXEGs.* Based on univariate Cox regression analysis, 15 genes were recognized as prognostic (Figure 1(a)). Finally, we screened 14 features of OSDEGs analysis after LASSO analysis and multivariate Cox regression analysis (Figures 1(b) and 1(c)). AARS2 and FOXO3 were the protective factors in the prognostic model, while ATF4, CARS2, CRP, CYBA, GPX1, IL1B, MAPK8, MRPL44, MTFMT, NOS1, OSGIN2, and SOD2 were considered as risk factors in the prognostic model. The risk score for each CRC patient in GSE17536 was assessed using the following equation: Risk score $= (-1.74 \times \text{AARS2 expression}) + (0.36 \times \text{ATF4 expression}) + (2.08 \times \text{CARS2 expression}) + (2.98 \times \text{CRP expression}) + (0.40 \times \text{CYBA expression}) + (-0.61 \times \text{FOXO3 expression}) + (0.68 \times \text{GPX1 expression}) + (0.06 \times \text{IL1B expression}) + (0.52 \times \text{MAPK8 expression}) + (0.47 \times \text{MRPL44 expression}) + (0.09 \times \text{MTFMT expression}) + (1.43 \times \text{NOS1 expression}) + (1.04 \times \text{OSGIN2 expression}) + (0.14 \times \text{SOD2 expression})$. Finally, CRC patients in GSE17536 were categorized as the high-risk and the low-risk groups based on the median risk score. We performed PCA analysis on both groups (Figures 2(a) and 2(b)) and found that the prognostic model genes could effectively identify between the low-risk and high-risk categories.

*3.2. Evaluation of the Prognostic Performance of the OXEGs Signature.* The low-risk group had a finer prognosis and a longer surviving time, whereas the high-risk group had an inferior prognosis and a shorter survival time, as shown by the Kaplan-Meier survival curve (Figure 3(a)). Additionally, we used the external validation dataset GSE29621 to confirm the prognostic risk profile's accuracy, and we found consistent variation in overall survival (OS) between the high-risk and the low-risk groups. (Figure 3(b)). Area under the curve (AUC) measurements for the 1-, 3-, and 5-year survival rates were 0.900, 0.781, and 0.804, respectively, while these measurements for the 1-, 3-, and 5-year AUC of the validation set were 0.946, 0.684, and 0.724, respectively. This data demonstrated that our prognostic prediction had good sensitivity and specificity (Figures 3(c) and 3(d)). Subsequently, we showed from Cox regression analysis that the prognostic risk models associated with OXEGs are independent predictors of CRC prognosis (Figures 3(e) and 3(f)). We further predicted the relationship between clinical fea-

tures and prognostic value by ROC (Figure 3(g)), and the AUC of the model was 0.804, while the AUC of age, gender, stage, and grade were 0.535, 0.493, 0.573, and 0.456, respectively, demonstrating the model's excellent sensitivity and specificity. The ROC curve of the validation set signified that the AUC value of the model was 0.724, which was higher than other clinical features (Figure 3(h)). Through clinical correlation analysis (Figures 4(a)–4(d)), we concluded that age was substantially different between the groups at high-and low-risk ($P < 0.05$), while gender, grade, and stage were not statistically significant.

*3.3. Nomogram Construction and Evaluation.* We constructed a nomogram based on age, sex, grade, stage, and risk scores for 14 OXEGs to foretell the 1-, 3-, and 5-year survival in CRC patients (Figure 5(a)). As can be seen from the calibration curve (Figure 5(b)), the survival prediction for year 5 is in good agreement with the actual value. In the ROC curve (Figure 5(c)), the value of AUC was 0.689, indicating that the model had high accuracy. The rosette model was later demonstrated to be an independent determinant of CRC prognosis by univariate and multivariate Cox regression analysis. (Figures 5(d) and 5(e)).

*3.4. Association between Tumor Immune Cell Infiltration and Risk Score.* Calculations were made to determine the variation in immune cells between the low-risk group and the high-risk group using the CIBERSORT method to get the waterfall diagram of immune cells in the tumor group (Figures 6(a) and 6(b)). As seen in the image, the high-risk group had elevated levels of neutrophils, eosinophils, and activated NK cells ($P < 0.05$), but the low-risk group had considerably higher levels of resting NK cells infiltrating their tissues. APC costimulation, inflammation-promoting, CCR, T cell costimulation, cytolytic activity, HLA, T cell coinhibition, and checkpoint were all more significant in the high-risk group, according to the analyses of immune cell function (Figure 6(c)).

*3.5. Gene Enrichment Analysis.* Twenty-seven differential genes were identified in the low-risk and high-risk categories (Supplement 3). We examined the GO and KEGG pathways of the differential genes and obtained a total of 68 KEGG signaling pathways and 832 significantly enriched biological processes (Supplements 4 and 5). The biological processes were enriched for positive control of ion transport, response to lipopolysaccharide, response to cold, positive control of calcium ion transmembrane transport, neutrophil chemotaxis, positive management of calcium ion transport, granulocyte migration, and T cell chemotaxis. In the cellular component, collagen-regulation of fibroblast extracellular matrix, external side of plasma membrane, CatSper complex, meiotic spindle, and male germ cell nucleus were enriched. In addition, we demonstrated significant enrichment of receptor ligand activity, signaling receptor activator activity, cytokine activity, carbohydrate binding, cytokine receptor binding, CXCR chemokine receptor binding, and chemokine activity in molecular function (Figure 7(a)). The IL-17 signaling pathway, lipid and atherosclerosis, cytokine-

cytokine receptor interaction, longevity regulating pathway-multiple species, viral protein interaction with cytokine and cytokine receptor, and Toll-like receptor signaling pathway were the pathways that KEGG enrichment analysis showed to be enriched (Figure 7(b)). The enrichment disparities between the high-risk and the low-risk groups may be seen by GSEA. Systemic lupus erythematosus, basal cell carcinoma, O-glycan biosynthesis, Wnt signaling pathway, retinol metabolism, drug metabolism cytochrome p450, and maturity-onset diabetes of the young were the major areas of enrichment in the low-risk group. Cytokine receptor interaction, leishmania infection, NOD-like receptor signaling pathway, natural killer cell mediated cytotoxicity, and systemic lupus were the major areas of enrichment in the high-risk group (Figure 8).

*3.6. Validation of Prognostic Genes and Enrichment Pathways.* Sequencing results portrayed that prognostic genes were different in colorectal cancer tissues and adjacent tissues. As can be seen from Figure 9, FOXO3 and MAPK8 were highly expressed in adjacent tissues. ATF4, MTFMT, CYBA, OSGIN2, NOS1, GPX1P1, CARS2, CRP, MRPL44, SOD2, AARS2, and IL-1B are highly expressed in colorectal cancer tissues (Figure 9(a)). After ELISA confirmed that IL-17A was increased in the oxidative stimulation group (Figure 9(b)), we knocked down the IL-17A receptor ACT1 (Figure 9(c)) and performed oxidative stimulation again. Western blot results signify the expressions of downstream signaling factors TRAF6, NF-$\kappa$B, and MAPK signaling pathways were decreased to a certain extent after ACT1 knockdown (Figure 9(d)).

# 4. Discussion

As one of the prevalent malignant tumors in the world, colorectal cancer has a high fatality rate. The 5-year survival rate for individuals with advanced CRC is around 14%, and more than 50% of patients receive their diagnosis at an advanced stage [25, 26]. There is still an urgency for developing a prognostic model in order to give tailored prediction and precision medicine for CRC patients who are dealing with the therapeutic conundrum.

In the body, the content of ROS in normal cells and tumor cells is different, and the sensitivity to oxidative stress is also different. Oxidative DNA damage has been widely accepted as an important feature of the occurrence of malignant tumors. At present, researchers have found a variety of novel treatment methods based on ROS to restore chemoresistance and overcome radiotherapy resistance, enhance the efficacy of chemoradiotherapy, [27–31] and bring certain guiding significance to clinical treatment. At the same time, oxidative stress products in the tumor microenvironment also affect the immune response of the body. Therefore, we established and verified the prognostic design in relation to OXEGs to foretell the prognosis of CRC patients.

Using multivariate and univariate Cox regression analysis with the LASSO technique, we examined the differentially expressed genes related to oxidative stress in GEO. 14

OXEGs were screened (AARS2, FOXO3, ATF4, CARS2, CRP, CYBA, GPX1, IL1B, MAPK8, MRPL44, MTFMT, NOS1, OSGIN2, and SOD2) to create a prognostic risk model for prediction. The model we created has good specificity and sensitivity and ROC testing showed that its AUC was 0.804. Additionally, we discovered that age was a predictive factor that affected CRC patients on its own. AARS2 and FOXO3 are protective factors in the prognostic model of CRC, and the former regulates the proliferation of colorectal cancer by affecting mitochondrial respiration [32]. The latter is associated with morbidity and mortality in CRC [33] and can modulate its mediated SOX2 expression to affect cancer migration, invasion, and stem cell proliferation [34]. ATF4 regulates tumor autophagy in CRC and affects tumor survival [35]. CRP is related to systemic inflammation, but its effect on colorectal cancer is not clear. CYBA can induce familial colorectal cancer by interfering with the integrity of intestinal barrier [36]. IL-1B, as a member of the interleukin family, is closely related to tumor immunity [37]. SOD2 contributes to the chemical resistance of colorectal cancer [38], and according to research, MAPK8 can promote the progression of colorectal cancer [39]. However, the functional roles of CARS2, GPX1, MRPL44, MTFMT, NOS1, and OSGIN2 in CRC are still unknown. The sensitivity and accuracy of this model are further verified by the combined GEO dataset. Our findings suggest that higher risk is associated with poorer outcomes.

A total of 68 KEGG signaling pathways and 832 significantly enriched biological processes were obtained through enrichment analysis. The most interesting point was the significant enrichment of the IL-17 pathway. When ROS content increased, levels of the proinflammatory cytokine IL-17A increased through retardation of the PI3K/AKT/mTOR pathway and selective autophagy [40]. We also demonstrated by ELISA that IL-17A levels were significantly increased in SW480 cells after oxidative stimulation. Many studies have also proven the relationship between IL-17A level and REDOX environment [41, 42]. To this end, we performed western blot verification of oxidative stimulation after knocking down the receptor ACT1 of IL-17A [43] and found that the downstream factor TRAF6 and the relative content of corresponding pathways including NF-$\kappa$B and MAPK were affected (see mechanism diagram in the Supplementary Material). MAPK may not be statistically significant due to the interference of other related pathways, which will be discussed further.

PAM can increase the content of ROS in tumor cells, and the accumulation of a large number of active substances leads to tumor cell death. Many researchers have affirmed the effect of PAM in the treatment of tumors in vivo and in vitro. Nakamura et al. [44] showed that PAM inhibited the metastasis of ovarian cancer through in vivo and in vitro experiments, and Utsumi et al. proved that PAM also has a certain therapeutic effect on chemotherapy resistant ovarian cancer [45]. To expand on the selection of PAM concentration in this study, we will further select a wider range of concentrations to verify the difference between low concentration and high concentration groups.

We estimated the variation in immune infiltration between the two categories using the CIBERSORT method in order to study the link between immune cell infiltration and risk ratings. We identified significant increases in monocyte, activated NK cells, eosinophils, and neutrophil levels in the low-risk group. Moreover, the largely enriched IL-17A pathway is closely linked to immune cells. Bruno et al. attenuated the antifungal host immune response by using IL-17A inhibitors, which increased the incidence of Candida infection [46]. These results suggest that oxidative stress-related gene tags may influence immune cell infiltration and hence the efficacy of colorectal cancer immunotherapy.

## 5. Conclusions

In conclusion, using 14 OXEGs, we created a prognostic model for colorectal cancer under oxidative stress that has a high predictive value. This study provides the possibility for individuals with CRC to get individualized care.

## 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 state that the publishing of this work does not include any conflicts of interest.

## Authors' Contributions

Zilu Chen and Kun Mei contributed equally to this work and share first authorship.

## Acknowledgments

## Supplementary Materials

*Supplementary 1.* Supplement 1: 80 genes related to oxidative stress.

*Supplementary 2.* Supplement 2: Three ACT1 knockdown siRNA sequences were used for screening.

*Supplementary 3.* Supplement 3: Table 3: Twenty-seven differential genes were identified in the low-risk and high-risk groups.

*Supplementary 4.* Supplement 4: Table 4: GO enrichment analysis.

*Supplementary 5.* Supplement 5: Table 4: KEGG enrichment analysis.

## References

[1] N. Hilbertina, N. C. Siregar, M. Abdullah, and S. I. Wanandi, "Determination of cancer-associated fibroblast and stromal phenotypes as novel prognostic factors for colorectal carcinomas associated with tumor budding," *Journal of Natural Science, Biology and Medicine*, vol. 10, pp. S68–S72, 2019.

[2] R. Zheng, S. Zhang, H. Zeng et al., "Cancer incidence and mortality in China, 2016," *Journal of the National Cancer Center*, vol. 2, no. 1, pp. 1–9, 2022.

[3] Z. Yuan, S. Weng, C. Ye, H. Hu, S. Zhang, and Y. Yuan, "CSCO guidelines for colorectal cancer version 2022: updates and discussions," *Chinese Journal of Cancer Research*, vol. 34, no. 2, pp. 67–70, 2022.

[4] E. R. Tabriz, M. Ramezani, and A. Aledavood, "Role of dietary supplements in the recurrence of colorectal cancer: an issue for cancer survivors," *Evidence Based Care*, vol. 12, no. 1, pp. 63-64, 2022.

[5] J. D. Hayes, A. T. Dinkova-Kostova, and K. D. Tew, "Oxidative stress in cancer," *Cancer Cell*, vol. 38, no. 2, pp. 167–197, 2020.

[6] C. Lu, D. Zhou, Q. Wang et al., "Crosstalk of microRNAs and oxidative stress in the pathogenesis of cancer," *Oxidative Medicine and Cellular Longevity*, vol. 2020, Article ID 2415324, 13 pages, 2020.

[7] H. Sies and D. P. Jones, "Reactive oxygen species (ROS) as pleiotropic physiological signalling agents," *Nature Reviews. Molecular Cell Biology*, vol. 21, no. 7, pp. 363–383, 2020.

[8] A. Parascandolo and M. O. Laukkanen, "Carcinogenesis and reactive oxygen species signaling: interaction of the NADPH oxidase NOX1-5 and superoxide dismutase 1-3 signal transduction pathways," *Antioxidants & Redox Signaling*, vol. 30, no. 3, pp. 443–486, 2019.

[9] H. Q. Ju, J. F. Lin, T. Tian, D. Xie, and R. H. Xu, "NADPH homeostasis in cancer: functions, mechanisms and therapeutic implications," *Signal Transduction and Targeted Therapy*, vol. 5, no. 1, p. 231, 2020.

[10] Y. Wang, H. Qi, Y. Liu et al., "The double-edged roles of ROS in cancer prevention and therapy," *Theranostics*, vol. 11, no. 10, pp. 4839–4857, 2021.

[11] H. J. Forman and H. Zhang, "Targeting oxidative stress in disease: promise and limitations of antioxidant therapy," *Nature Reviews. Drug Discovery*, vol. 20, no. 9, pp. 689–709, 2021.

[12] X. Liu, F. Yun, L. Shi, Z. H. Li, N. R. Luo, and Y. F. Jia, "Roles of signaling pathways in the epithelial-mesenchymal transition in cancer," *Asian Pacific Journal of Cancer Prevention*, vol. 16, no. 15, pp. 6201–6206, 2015.

[13] K. Xiang, H. Yanshan, Z. Chunmei, G. Minmin, W. Yan, and Y. Xiaojia, "GP5 regulates epithelial–mesenchymal transition in breast cancer via the PI3K/AKT signaling pathway," *Experimental Biology and Medicine (Maywood, N.J.)*, vol. 247, no. 17, pp. 1501–1517, 2022.

[14] M. Hajipour, K. Mokhtari, M. Mahdevar et al., "Identification of a novel interplaying loop of PPAR γ and respective lncRNAs are involved in colorectal cancer progress," *International Journal of Biological Macromolecules*, vol. 219, pp. 779–787, 2022.

[15] C. H. Chang and S. Pauklin, "ROS and TGFβ: from pancreatic tumour growth to metastasis," *Journal of Experimental & Clinical Cancer Research*, vol. 40, no. 1, p. 152, 2021.

[16] J. L. Yang, W. Y. Chen, S. Mukda, Y. R. Yang, S. F. Sun, and S. D. Chen, "Oxidative DNA damage is concurrently repaired by base excision repair (BER) and apyrimidinic endonuclease 1 (APE1)-initiated nonhomologous end joining (NHEJ) in cortical neurons," *Neuropathology and Applied Neurobiology*, vol. 46, no. 4, pp. 375–390, 2020.

[17] J. I. Ikeda, H. Tanaka, K. Ishikawa, H. Sakakita, Y. Ikehara, and M. Hori, "Plasma-activated medium (PAM) kills human cancer-initiating cells," *Pathology International*, vol. 68, no. 1, pp. 23–30, 2018.

[18] H. Tanaka, M. Mizuno, Y. Katsumata et al., "Oxidative stress-dependent and -independent death of glioblastoma cells induced by non-thermal plasma-exposed solutions," *Scientific Reports*, vol. 9, no. 1, p. 13657, 2019.

[19] X. Li, H. Wang, J. Wang et al., "Emodin enhances cisplatin-induced cytotoxicity in human bladder cancer cells through ROS elevation and MRP1 downregulation," *BMC Cancer*, vol. 16, no. 1, p. 578, 2016.

[20] X. Yang, G. Chen, K. N. Yu et al., "Cold atmospheric plasma induces GSDME-dependent pyroptotic signaling pathway via ROS generation in tumor cells," *Cell Death & Disease*, vol. 11, no. 4, p. 295, 2020.

[21] Z. Zhang, W. Q. Chen, S. Q. Zhang et al., "The human cathelicidin peptide LL-37 inhibits pancreatic cancer growth by suppressing autophagy and reprogramming of the tumor immune microenvironment," *Frontiers in Pharmacology*, vol. 13, article 906625, 2022.

[22] S. S. Xiong, "Prognosis, immunotherapy and drug therapy of oxidative stress-related genes in patients with hepatocellular carcinoma," *European Review for Medical and Pharmacological Sciences*, vol. 26, no. 14, pp. 5014–5032, 2022.

[23] H. Wang, J. Chen, G. Bai, W. Han, R. Guo, and N. Cui, "mTOR Modulates the Endoplasmic Reticulum Stress-Induced CD4+ T Cell Apoptosis Mediated by ROS in Septic Immunosuppression," *Mediators of Inflammation*, vol. 2022, Article ID 6077570, 13 pages, 2022.

[24] S. Renken, T. Nakajima, I. Magalhaes et al., "Targeting of Nrf2 improves antitumoral responses by human NK cells, TIL and CAR T cells during oxidative stress," *Journal for Immunotherapy of Cancer*, vol. 10, no. 6, article e004458, 2022.

[25] R. L. Siegel, K. D. Miller, A. Goding Sauer et al., "Colorectal cancer statistics, 2020," *CA: a Cancer Journal for Clinicians*, vol. 70, no. 3, pp. 145–164, 2020.

[26] Z. Li, T. She, H. Yang et al., "A novel tumor-homing TRAIL variant eradicates tumor xenografts of refractory colorectal cancer cells in combination with tumor cell-targeted photodynamic therapy," *Drug Delivery*, vol. 29, no. 1, pp. 1698–1711, 2022.

[27] L. Bin, Y. Yang, F. Wang et al., "Biodegradable silk fibroin nanocarriers to modulate hypoxia tumor microenvironment favoring enhanced chemotherapy," *Frontiers in Bioengineering and Biotechnology*, vol. 10, article 960501, 2022.

[28] T. G. Mhone, M. C. Chen, C. H. Kuo et al., "Daidzein synergizes with gefitinib to induce ROS/JNK/c-Jun activation and inhibit EGFR-STAT/AKT/ERK pathways to enhance lung adenocarcinoma cells chemosensitivity," *International Journal of Biological Sciences*, vol. 18, no. 9, pp. 3636–3652, 2022.

[29] J. A. Nickoloff, "Targeting replication stress response pathways to enhance genotoxic chemo- and radiotherapy," *Molecules*, vol. 27, no. 15, p. 4736, 2022.

[30] D. J. Hu, W. J. Shi, M. Yu, and L. Zhang, "High WDR34 mRNA expression as a potential prognostic biomarker in patients with breast cancer as determined by integrated bioinformatics analysis," *Oncology Letters*, vol. 18, no. 3, pp. 3177–3187, 2019.

[31] W. Shi, C. Li, T. Wartmann et al., "Sensory ion channel candidates inform on the clinical course of pancreatic cancer and present potential targets for repurposing of FDA-approved agents," *Journal of Personalized Medicine*, vol. 12, no. 3, p. 478, 2022.

[32] Z. Zhu, Q. Hou, B. Wang et al., "A novel mitochondria-related gene signature for controlling colon cancer cell mitochondrial respiration and proliferation," *Human Cell*, vol. 35, no. 4, pp. 1126–1139, 2022.

[33] C. Yu, A. M. Hodge, E. M. Wong et al., "Association of FOXO3 blood DNA methylation with cancer risk, cancer survival, and mortality," *Cell*, vol. 10, no. 12, p. 3384, 2021.

[34] T. Y. Chang, K. C. Lan, C. Y. Chiu, M. L. Sheu, and S. H. Liu, "ANGPTL1 attenuates cancer migration, invasion, and stemness through regulating FOXO3a-mediated SOX2 expression in colorectal cancer," *Clinical Science (London, England)*, vol. 136, no. 9, pp. 657–673, 2022.

[35] S. Han, L. Zhu, Y. Zhu et al., "Targeting ATF4-dependent prosurvival autophagy to synergize glutaminolysis inhibition," *Theranostics*, vol. 11, no. 17, pp. 8464–8479, 2021.

[36] L. Zhu, B. Miao, D. Dymerska et al., "Germline variants of CYBA and TRPM4 predispose to familial colorectal cancer," *Cancers (Basel)*, vol. 14, no. 3, p. 670, 2022.

[37] H. Imazeki, Y. Ogiwara, M. Kawamura, N. Boku, and C. Kudo-Saito, "CD11b(+) CTLA4(+) myeloid cells are a key driver of tumor evasion in colorectal cancer," *Journal for Immunotherapy of Cancer*, vol. 9, no. 7, article e002841, 2021.

[38] M. Paku, N. Haraguchi, M. Takeda et al., "SIRT3-mediated SOD2 and PGC-1α contribute to chemoresistance in colorectal cancer cells," *Annals of Surgical Oncology*, vol. 28, no. 8, pp. 4720–4732, 2021.

[39] X. Liu, L. Li, J. Bai et al., "Long noncoding RNA plasmacytoma variant translocation 1 promotes progression of colorectal cancer by sponging microRNA-152-3p and regulating E2F3/MAPK8 signaling," *Cancer Science*, vol. 113, no. 1, pp. 109–119, 2022.

[40] H. Gou, X. Chen, X. Zhu et al., "Sequestered SQSTM1/p62 crosstalk with Keap1/NRF2 axis in hPDLCs promotes oxidative stress injury induced by periodontitis," *Free Radical Biology & Medicine*, vol. 190, pp. 62–74, 2022.

[41] M. M. E. Tabaa, A. M. K. Fattah, M. Shaalan, E. Rashad, and N. A. el Mahdy, "Dapagliflozin mitigates ovalbumin-prompted airway inflammatory-oxidative successions and associated bronchospasm in a rat model of allergic asthma," *Expert Opinion on Therapeutic Targets*, vol. 26, no. 5, pp. 487–506, 2022.

[42] S. J. Yoo, H. R. Lee, J. Kim, I. S. Yoo, C. K. Park, and S. W. Kang, "Hypoxia-inducible factor-2 alpha regulates the migration of fibroblast-like synoviocytes via oxidative stress-induced CD70 expression in patients with rheumatoid arthritis," *International Journal of Molecular Sciences*, vol. 23, no. 4, p. 2342, 2022.

[43] N. C. Brembilla, L. Senra, and W. H. Boehncke, "The IL-17 family of cytokines in psoriasis: IL-17A and beyond," *Frontiers in Immunology*, vol. 9, p. 1682, 2018.

[44] K. Nakamura, Y. Peng, F. Utsumi et al., "Novel intraperitoneal treatment with non-thermal plasma-activated medium inhibits metastatic potential of ovarian cancer cells," *Scientific Reports*, vol. 7, no. 1, p. 6085, 2017.

[45] F. Utsumi, H. Kajiyama, K. Nakamura et al., "Effect of indirect nonequilibrium atmospheric pressure plasma on anti-proliferative activity against chronic chemo-resistant ovarian cancer cells in vitro and in vivo," *PLoS One*, vol. 8, no. 12, article e81576, 2013.

[46] M. Bruno, L. Davidson, H. J. P. M. Koenen et al., "Immunological Effects of Anti–IL-17/12/23 Therapy in Patients with Psoriasis Complicated by *Candida* Infections," *The Journal of Investigative Dermatology*, 2022.

*Research Article*

# Integrative Analysis Reveals the Potential Role and Prognostic Value of GOLM1 in Hepatocellular Carcinoma

**Yan Lin** [ID],[1] **Ziqin He,**[1] **Xing Gao** [ID],[1] **Lu Lu,**[1] **Cheng Lu,**[2] **Julu Huang,**[2] **Min Luo,**[1] **Jiazhou Ye** [ID],[2] **and Rong Liang** [ID][1]

[1]*Department of Digestive Oncology, Guangxi Medical University Cancer Hospital, Nanning, Guangxi 530021, China*
[2]*Department of Hepatobiliary Surgery, Guangxi Medical University Cancer Hospital, Nanning, Guangxi 530021, China*

Correspondence should be addressed to Jiazhou Ye; yejiazhou@gxmu.edu.cn and Rong Liang; liangrong@gxmu.edu.cn

Overexpression of Golgi membrane protein 1 (GOLM1) is closely associated with hepatocellular carcinoma (HCC) vascular invasion. How GOLM1 may be involved in angiogenesis in HCC remains unclear. We explored how GOLM1 promotes angiogenesis in HCC and potential prognostic value. Expression levels of GOLM1 in HCC patients and healthy controls were obtained from The Cancer Genome Atlas (TCGA). Differentially expressed genes (DEGs) between HCC patients and controls were compared. GOLM1 was knocked out in the HCC cell line, and RNA sequencing and DEG expression analysis were performed compared with control cells. Based on TCGA data and cell line RNA sequencing data, DEGs affected by a high expression of GOLM1 were identified. Subsequently, enrichment analysis was performed to explore the functions and pathways of the DEGs affected by a high expression of GOLM1. A relevant network analysis was built. Cox regression, genomic variance analysis scores, minimum absolute shrinkage and selection operator regression, and random forest regression models were applied to determine the best prognostic model and validated using the GSE54236 dataset from the Gene Expression Omnibus (GEO). We determined the effect of GOLM1 expression on immune cell infiltration in liver cancer. GOLM1 was overexpressed in HCC tissues compared with controls, and its level correlated with tumor purity and prognosis. 400 DEGs affected by highly expressed GOLM1 were identified in TCGA and cell line RNA sequencing data. Enrichment analysis revealed that these DEGs may be related to biological processes of oxidative stress and angiogenesis and involved in the VEGF signaling pathway and protein processing in endoplasmic reticulum. We predicted a comprehensive regulatory network in which GOLM1 activated VEGF signaling to promote HCC angiogenesis. GOLM1 may interact with E2F1 and IGF2BP3 to promote angiogenesis. GOLM1 overexpression was associated with greater immune cell infiltration. A random forest regression model was the best prognostic model. Our study reveals a potential molecular mechanism of GOLM1 in promoting HCC. We developed two prognostic models based on DEG associated with GOLM1 overexpression to help stratify HCC prognosis and improve individualized treatment.

## 1. Introduction

Liver cancer has the fourth highest mortality among malignant tumors worldwide [1]. Among the three pathological types of primary liver cancer, hepatocellular carcinoma (HCC) is the most common, accounting for about 90% of cases [2]. HCC is a typical vascular tumor, so angiogenesis plays an important role in its onset and progression [3]. Several angiogenesis pathways are abnormally activated in HCC

to support tumor development, including pathways involving vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF), and platelet-derived growth factor and their receptors, as well as pathways involving angiopoetin and Tie [4]. Targeted antiangiogenesis therapy has become one of the main strategies for treating HCC. Although a variety of antiangiogenic drugs are currently under development, only sorafenib and lenvatinib have been approved for the first-line treatment of advanced HCC

[2–4]. Therefore, in-depth studies on the mechanisms of HCC angiogenesis are needed to identify new targets for the development of effective antiangiogenic drugs.

We wondered whether Golgi membrane protein 1 (GOLM1), also known as GP73 or GOLPH2, may be a suitable therapeutic target in HCC. GOLM1 is expressed predominantly in epithelial cells [5] and can also be cleaved by proprotein convertase and secreted into the blood [6]. GOLM1 is overexpressed in a variety of malignancies including HCC, and its high expression correlates strongly with poor prognosis [7]. A previous multicenter study comparing serum GOLM1 and alpha-fetoprotein (AFP) in 4217 human subjects showed that GOLM1 had a sensitivity and specificity of 76.4% and 97.4%, respectively, for HCC, while AFP had a sensitivity and specificity of 58.2% and 85.3%, respectively [8]. Another meta-analysis involving 11 studies showed that GOLM1 was superior to AFP as a diagnostic marker in 5 studies, while the results were opposite or unclear in the remaining 6 studies [9]. Therefore, GOLM1 may even allow a more sensitive and specific diagnosis of liver cancer than AFP [8–10].

Our previous study found that serum GOLM1 levels were significantly higher in patients with HCC, and its sensitivity and specificity for diagnosing HCC were higher than those of AFP [11]. Further studies related to GOLM1 and drug resistance in HCC were conducted, and it was confirmed that GOLM1 promoted oxaliplatin resistance in human HCC cells [12]. In addition, it was also noted during the collection of clinical case data from HCC patients that GOLM1 elevated vascular invasion more significantly in HCC patients [13]. Other studies reported that GOLM1 promotes tumor metastasis by participating in the epithelial-mesenchymal transformation and recycling of epidermal growth factor receptor and receptor tyrosine kinases [14, 15]. Moreover, GOLM1 can enhance STAT3 phosphorylation by upregulating the epidermal growth factor receptor and then activating programmed death-ligand 1 transcriptional expression to inhibit immune responses [16]. Thus, evidence suggests that GOLM1 promotes the pathogenesis and progression of HCC through various mechanisms, whereas no detailed studies specifically addressing the GOLM1 gene and HCC angiogenesis have been reported both nationally and internationally.

To explore this possibility, we combined data from The Cancer Genome Atlas-Liver Hepatocellular Carcinoma (TCGA-LIHC) and experimental sequencing data from an HCC cell line. These findings may reveal new targets and strategies for targeted antiangiogenic therapy in HCC.

## 2. Material and Methods

*2.1. Data Collection.* Gene expression profiles were obtained from publicly available databases: The Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov), from which data on 371 HCC patients and 50 controls were extracted and Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo), from which the dataset GSE54236 [17] was extracted for 81 tumor tissues from 78 HCC patients and 80 controls from 54 consecutive patients with cirrhosis. GEO data were used as the validation dataset. The current study adheres to TCGA and GEO data access policies and publication guidelines. The Tumor Immune Estimation Resource database (TIMER, http://timer.cistrome.org/) was to identify the differential expression of GOLM1 in tumor and normal tissues of multiple cancer species.

*2.2. Immunohistochemistry (IHC).* Tumor and adjacent tissues from three patients with HCC were collected from the Guangxi Medical University Cancer Hospital, immediately fixed in 10% formaldehyde for 12 h, dehydrated, transparent, paraffin-embedded, and sectioned ($4\,\mu$m) for IHC. After dewaxing the paraffin sections into water, antigen repair was performed using sodium citrate buffer (pH 6.0) for 15 min at 95°C and washed 3 times with phosphate buffered saline (PBS). The slides were then incubated with 3% $H_2O_2$ for 30 min to block endogenous peroxidase activity and 5% bovine serum albumin (BSA) for 1 h at room temperature to block nonspecific binding sites and incubated with primary antibody GOLM1 (American, Proteintech, Cat No. 15126-1-AP) overnight at 4°C. After three washes with PBS, the sections were incubated with the corresponding horseradish peroxidase at 37°C. The sections were incubated with the corresponding horseradish peroxidase- (HRP-) coupled secondary antibody (China, Beijing, ZSGB-BIO, SP-9001) in a wet box for 1 h and washed three times with PBS. Diaminobenzidine (DAB) was incubated for 10 min for color development, and hematoxylin was incubated for 3 min for nuclear restaining. Finally, the gradient was dehydrated in ethanol, clear in xylene, and sealed in neutral gum. IHC images were acquired under a standard light microscope (Olympus, Tokyo, Japan) and analyzed using Image-Pro Plus software (Media Cybernetics, Rockville, MD, United States). This study was approved by the Ethics Committee of the Guangxi Medical University Cancer Hospital. All procedures involving human participants complied with the ethical standards of the research committee and its ethical standards. Informed consent was obtained from participants for all study procedures and sequencing protocols.

*2.3. Cell Cultures.* We purchased MHCC97H cells from the Shanghai Institutes for Biological Sciences of the Chinese Academy of Sciences (CAS) as an HCC in vitro model. After being thawed, resuscitated, and passaged, the cells were kept in RPMI 1640 medium containing 10% fetal bovine serum and cultured in 37°C, 5% $CO_2$ saturated humidity cell incubator.

*2.4. RNA Sequencing.* GOLM1 expression was silenced in MHCC97H cells using two small interfering RNAs (siRNAs). The siRNA sequences were: siRNA 1, 5′-agg-gaaacgtgcttggtaa-3′ and siRNA 2, 5′-gaatagaagaggtcaccaa-3′. Lentiviral vectors encoding short hairpin RNA (shRNA) were designed based on the siRNA sequences to knock down GOLM1 expression (GOLM1-KD) [12]. These vectors were constructed by Hanyin Co. (Shanghai, China). The lentiviruses expressing the negative control lentivirus (Vector) were also constructed by Hanyin Co. (Shanghai, China). Total RNA was isolated from GOLM1-KD MHCC97H cells

and control MHCC97H cells using TRIzol (Thermo Scientific, Uppsala, Sweden) and purified using the RNeasy kit (Qiagen, Valencia, CA). RNA-Seq libraries were constructed using the TruSeq Stranded mRNA-Seq Library Preparation Kit (Illumina, San Diego, California, USA). Samples were sequenced using the Illumina NovaSeq system, generating paired-end reads of 150 base pairs. Raw sequence reads were converted into fragments per exon kilobase per million mapped reads (FPKM) to quantify gene expression.

2.5. *Survival Analysis.* Expression profiles were normalized using the "voom" function in the limma package in R [18]. For analyzing survival as a function of GOLM1 expression, the best cut-off value was determined using the "surv_cut-point" function in the survminer package in R (http://www.rstudio.org). According to the best cut-off value, patients were divided into groups expressing low or high levels of GOLM1. Then, overall survival (OS) was compared between the two groups using the log rank test.

2.6. *Identification of Differentially Expressed Genes (DEGs).* The limma package in R was used to identify DEGs between HCC patients and controls in TCGA data, as well as mRNAs whose expression differed between GOLM1-KD and control cells. DEGs with expression differences showing $P$.adjust < 0.05 were considered significant and included in further analyses. DEGs that were up- or downregulated across both TCGA and RNA sequencing data were defined as DEGs affected by GOLM1-KD. DEGs whose expression was opposite to that of DEGs affected by GOLM1-KD were identified as genes associated with GOLM1 overexpression.

2.7. *Functional Enrichment Analysis.* To explore the biological processes and signaling pathways in which DEG associated with GOLM1 overexpression may be involved, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using the clusterProfiler package in R [19]. Results associated with $P <$ 0.05 were defined as significant.

2.8. *Constructing Regulatory Networks.* We used the RNAInter database (http://www.rna-society.org/rnainter) [20] to extract DEGs that interacted with GOLM1 ($P < 0.05$). The set screening criterion was a score > 0.5. In combination with KEGG pathway genes, the Pearson correlation test and the hypergeometric test were utilized using the expression profiles. Finally, a comprehensive regulatory network of GOLM1-KEGG correlation was obtained. Results associated with $P < 0.05$ were considered significant.

2.9. *Molecular Docking.* We then explored whether target proteins may be able to bind GOLM1. We downloaded the three-dimensional structures of GOLM1 and target proteins from the Protein Data Bank (http://www.rcsb.org). Molecular docking was performed using Hex 8.0.0 software [21], and the results were visualized with PyMol software [22]. For whether two molecules have binding ability between them, when the docking energy is less than 0 KJ/mol, it means that both have binding potential, and the smaller the energy, the higher the binding potential.

2.10. *Gene Set Variant Analysis (GSVA).* We extracted the mechanism genes for univariate Cox regression analysis to obtain the meaningful DEGs of univariate Cox regression and then calculated the GSVA score of these genes for individual sample using GSVA package in R [23].

2.11. *Least Absolute Shrinkage and Selection Operator (LASSO) Regression Models.* Cox regression was used to identify DEG associated with OS of HCC patients. The glmnet package in R [24] was used to integrate potentially prognostic DEGs into a binomial LASSO regression model. During LASSO regression, we retained potential predictors with nonzero coefficients in order to generate candidate DEGs. Areas under receiver operating characteristic curves (AUCs) were calculated using the pROC package in R [25]. Cox regression was used to construct a prognostic nomogram to predict two- and five-year OS of HCC patients in TCGA.

2.12. *Random Forest Algorithm.* Genes with prognostic value in the univariate Cox regression were obtained. Survival data were dimensionally reduced using a random forest survival algorithm [26], ranked based on factor importance, and then filtered for gene signatures. Forest plots were generated using the forestplot package in R.

2.13. *Immune Cell Infiltration.* The level of infiltration of different types of immune cells was assessed using CIBERSORT (https://cibersort.stanford.edu/) and ssGSEA in the GSVA routine. Immune cells indicated as 0 were excluded from the analysis. A set of marker genes for the immune cell types analyzed by ssGSEA was obtained from Bindea et al. [27]. TIMER 2.0 [28] was used to assess the levels of immune cell infiltration. We also evaluated potential correlations between candidate genes and immune cell types using Pearson correlation analysis, with significance defined as $P < 0.05$. |cor| > 0.2 was considered to indicate that a correlation existed.

2.14. *Data Analysis and Statistics.* All bioinformatics analyses in this study were performed based on the Bioinforcloud platform (http://www.bioinforcloud.org.cn).

# 3. Results

3.1. *GOLM1 Is Overexpressed in HCC and Strongly Associated with Poor Patient Prognosis.* The study flowchart is shown in Figure 1. We identified a total of 12040 DEGs between HCC and controls in TCGA data (Figure 2(a)). In TCGA specimens (371 HCC tissues and 50 healthy controls), we found that GOLM1 was abundantly expressed at different stages of HCC but weakly expressed in controls (Figures 2(b) and 2(c)). Similarly, immunohistochemistry analysis showed a higher expression of GOLM1 in tumor than in adjacent tissue (Figure 2(d)). Interestingly, HCC patients with high expression of GOLM1 had poorer OS than those with low expression (Figure 2(e)).

Furthermore, analysis of the TIMER database showed GOLM1 to be upregulated in several types of tumors (Figure 2(f)). The above results show that GOLM1 is highly expressed in HCC. In addition, GOLM1 expression is closely related to the poor prognosis of HCC.
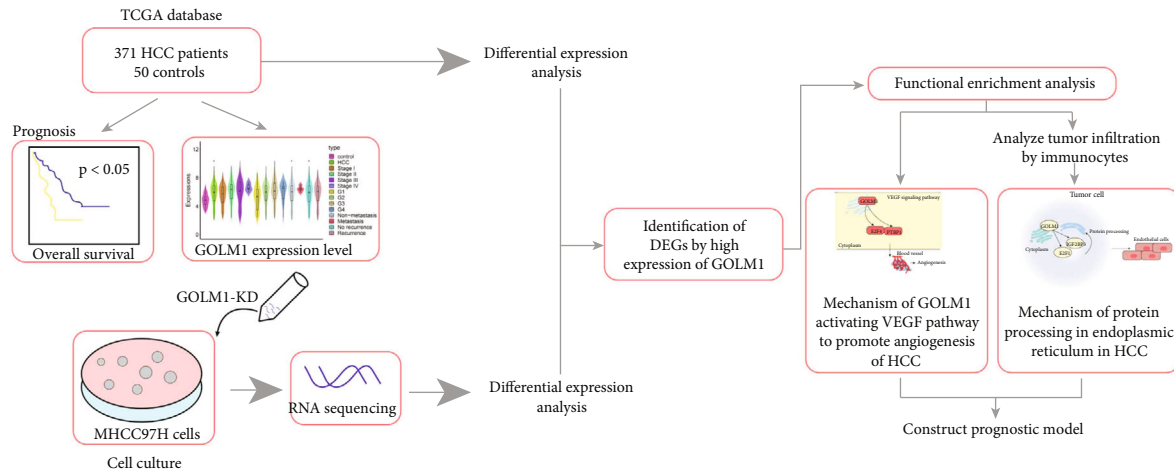
FIGURE 1: Study flowchart. The flow diagram of this study. DEGs: differentially expressed genes; HCC: hepatocellular carcinoma; KD: knocked down; TCGA: The Cancer Genome Atlas; VEFG: vascular endothelial growth factor.

### 3.2. Biological Functions of DEG Associated with GOLM1 Overexpression in HCC.
To identify DEG associated with GOLM1 overexpression, we performed differential expression analysis of the RNA sequencing data from our cell cultures. A total of 1363 DEGs between the GOLM1-KD MHCC97H cells and control MHCC97H cells groups were identified, comprising 744 upregulated and 619 downregulated genes (Figure 3(a)). Then, we analyzed overlapping DEGs found in 1363 DEGs cell cultures and the 14787 DEGs identified in TCGA. We found 737 overlapping genes, which were defined as DEGs affected by GOLM1-KD. In addition, 400 DEGs opposite to the DEGs affected by GOLM1-KD expression were identified as the specific DEGs associated with GOLM1 overexpression (Supplementary Table 1, Figure 3(b)).

These specific DEGs associated with GOLM1 overexpression were enriched for GO biological processes related to oxidative stress and angiogenesis: cellular response to oxidative stress, intrinsic apoptotic signaling pathway in response to oxidative stress, oxidative phosphorylation, and positive regulation of angiogenesis (Figure 3(c)). KEGG pathway analysis showed that these DEGs were involved mostly in protein processing in the endoplasmic reticulum, oxidative phosphorylation, apoptosis, and the VEGF signaling pathway (Figure 3(d)). Among them, the VEGF signaling pathway is activated in HCC and promotes angiogenesis [29], which attracted our attention (Figure 3(e)).

### 3.3. GOLM1 Activates the VEGF Signaling Pathway to Promote Angiogenesis in HCC.
To identify the regulatory network associated with GOLM1, pivot analysis was performed based on the RNAInter database (http://www.rna-society.org/rnainter) to find the genes interacting with GOLM1. The results identified 37 GOLM1 pivot genes with statistical significance, which indirectly regulated 12 KEGG pathways (Figure 4(a)). In particular, we identified six pivot genes in the VEGF signaling pathway: PTBP1, AR, CELF2, E2F4, DICER1, and CSTF2T. We also identified four pathway genes: HRAS, PTK2, PRKCB, and RAC2 (Figure 4(b)).

Thus, a comprehensive regulatory network of GOLM1, pivot genes, pathway genes, and the VEGF signaling pathway was constructed (Figure 4(c)). To further explore the regulation of target genes by GOLM1, we performed molecular docking analyses. The results showed that GOLM1 has the potential to bind E2F4 and PTBP1 (Figure 4(d)). Therefore, we postulate that GOLM1 targets E2F4 and PTBP1 to activate the VEGF signaling pathway, thereby promoting angiogenesis in HCC (Figure 4(e)).

### 3.4. Construction of a Prognostic Model for HCC.
In order to screen prognosis-related genes, mechanism genes in Figure 4(c) and 400 specific DEGs were extracted for univariate Cox regression analysis, and 52 DEGs associated with prognosis were obtained. To obtain the best prognostic model, 52 DEGs and GOLM1 were combined to construct four models based on GSVA, multifactor Cox regression, LASSO, or random forest regression. First, the GSVA scores of 53 prognostic genes were obtained by the GSVA model (Figure 5(a)). Then, 53 DEGs were subjected to multivariate Cox regression analysis to screen for independent prognostic factors and construct prognostic models, and finally, three prognostic genes were identified: HAVCR1, ETFDH, and MMP7 (Figure 5(b)). Similarly, LASSO regression analysis was performed on 53 DEGs to further remove redundant variables, and 14 genes were identified and used to construct prognostic models (Figure 5(c)). Finally, we used random forest regression models to identify 35 characteristic genes as the most relevant regulators of prognosis and constructed the corresponding models (Figure 5(d)).

To determine the best prognostic model, temporal ROC curves for median survival and survival at 1, 3, 5, and 8 years were plotted based on the risk scores of the four models (Figure 5(e)). The results showed that the random forest regression model had the optimal scoring efficacy. Subsequently, combined with the clinical characteristics of patients, we incorporated tumor distant metastasis, TNM stage, and random forest score models to construct 2-year and 5-year column line graph prediction models. The results showed that distant tumor metastasis, TNM stage, and

(a)



(b)



(c)

Figure 2: Continued.

(d)

(e)



(f)
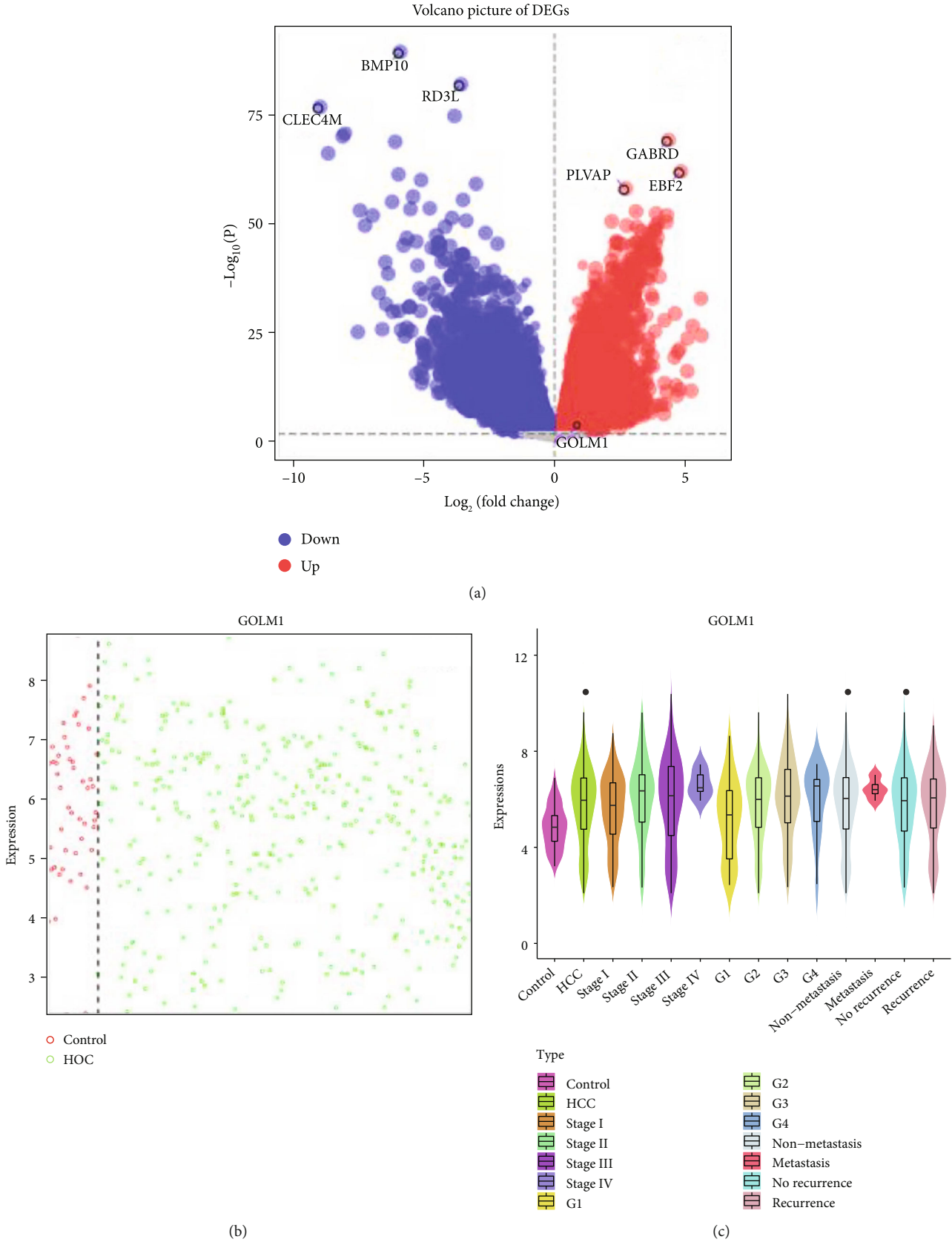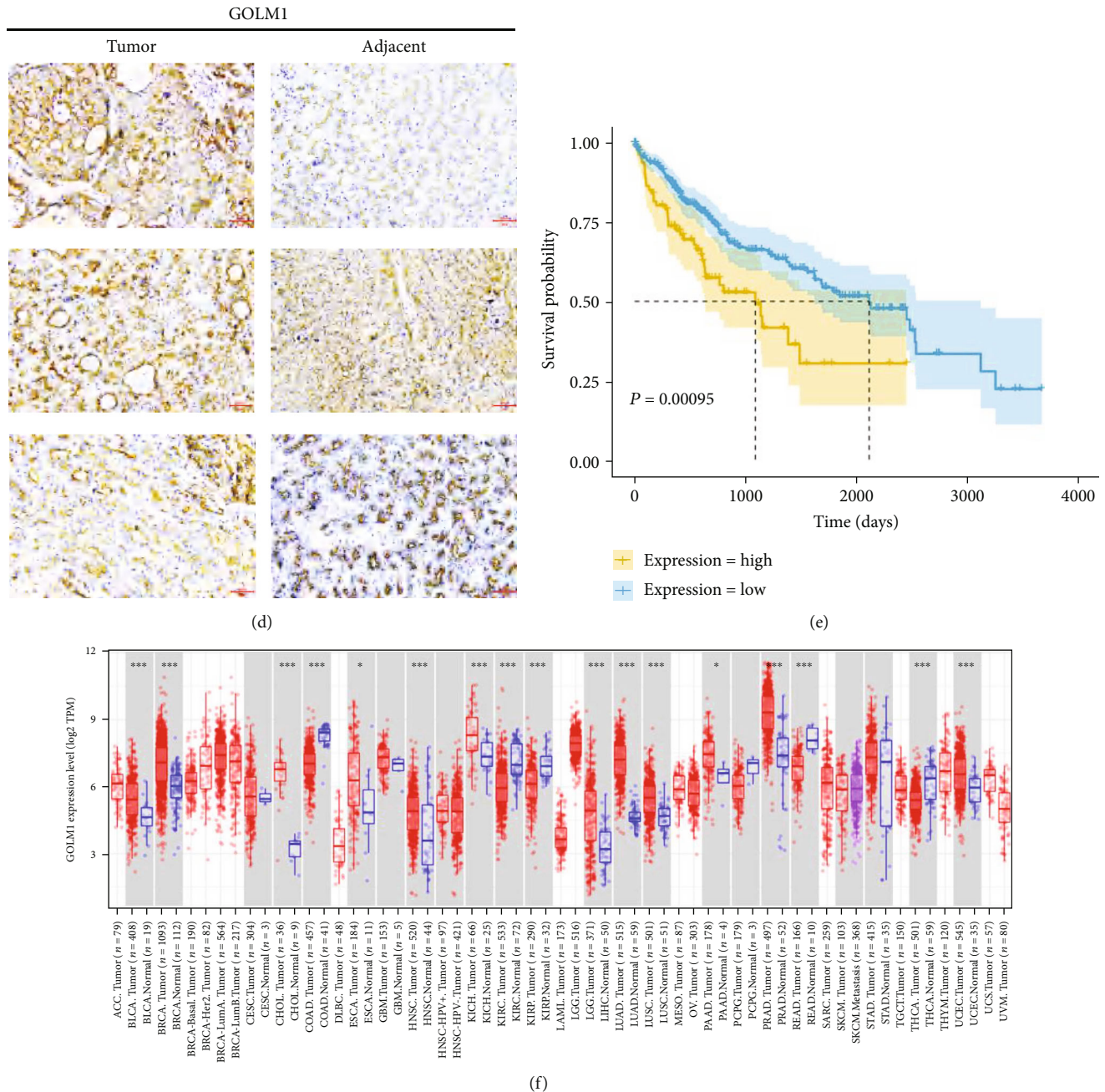
Figure 2: GOLM1 is overexpressed in hepatocellular carcinoma (HCC) and strongly associated with poor patient prognosis. (a) Differentially expressed genes (DEGs) between HCC patients and healthy controls from The Cancer Genome Atlas (TCGA). (b) GOLM1 expression level in 371 HCCs (green circle) and 50 adjacent normal tissues (red circle) from the TCGA. (c) GOLM1 expression at different HCC stages compare to the control group. (d) Immunohistochemistry against GOLM1 in tumor and adjacent tissues. (e) Kaplan-Meier curves of overall survival of patients in the high and low GOLM1 groups in TCGA. (f) GOLM1 expression in the Tumor Immune Estimation Resource database. DEGs: differentially expressed genes.

random forest score were independent prognostic factors for patients with HCC (Figure 5(f)). To validate the predictive value of the model, survival curves were used to demonstrate the OS and RFS curves of the clinical prognostic score model between samples of the high- and low-risk groups, and the results showed that patients in the high-risk group had significantly lower OS ($P < 0.0001$) and RFS ($P < 0.0001$) than those in the low-risk group, indicating that this clinical prognostic score model could effectively discriminate between

the high- and low-risk groups (Figure 5(g)). We confirmed this result in the validation dataset GSE54236 (additional Figures 1(a) and 1(b)).

3.5. Role of Protein Processing in the Endoplasmic Reticulum in HCC. We previously showed that DEGs associated with GOLM1 overexpression are predicted to participate in protein processing in the endoplasmic reticulum (Figure 3(d)) [30]. Multiple pathways, such as the endoplasmic

(a)

(b)

(c)
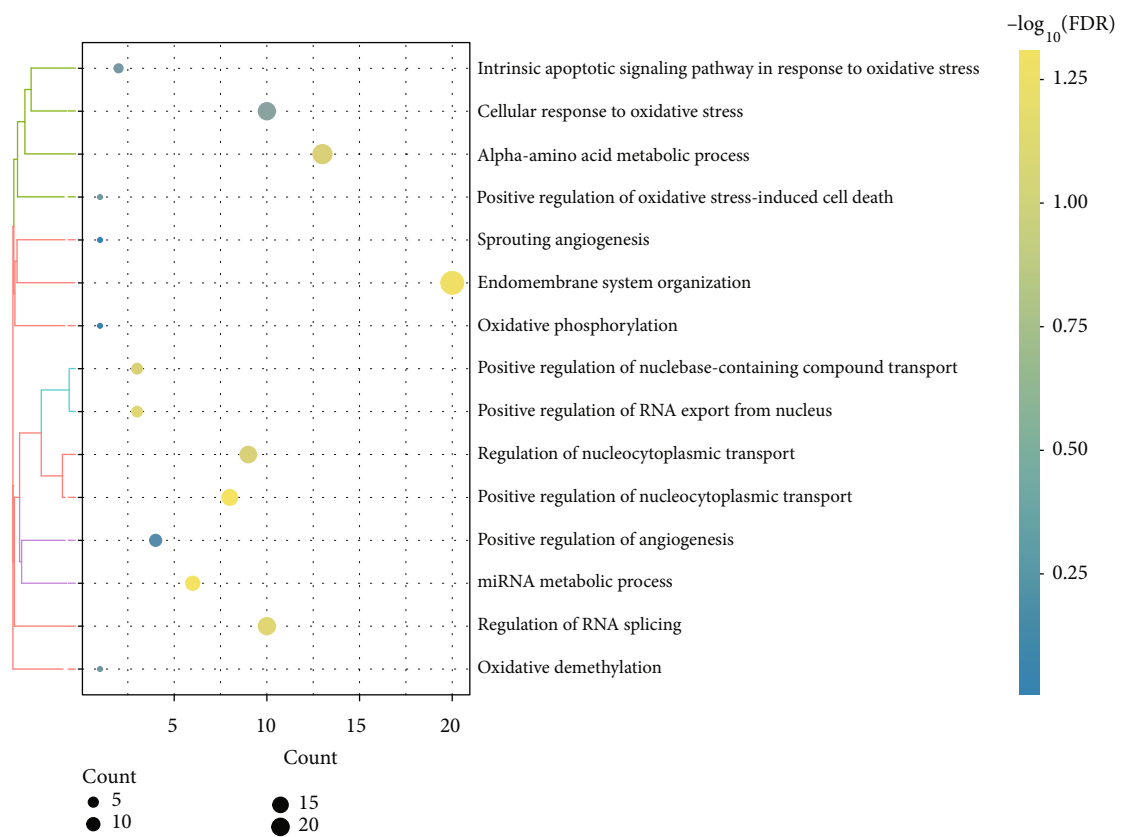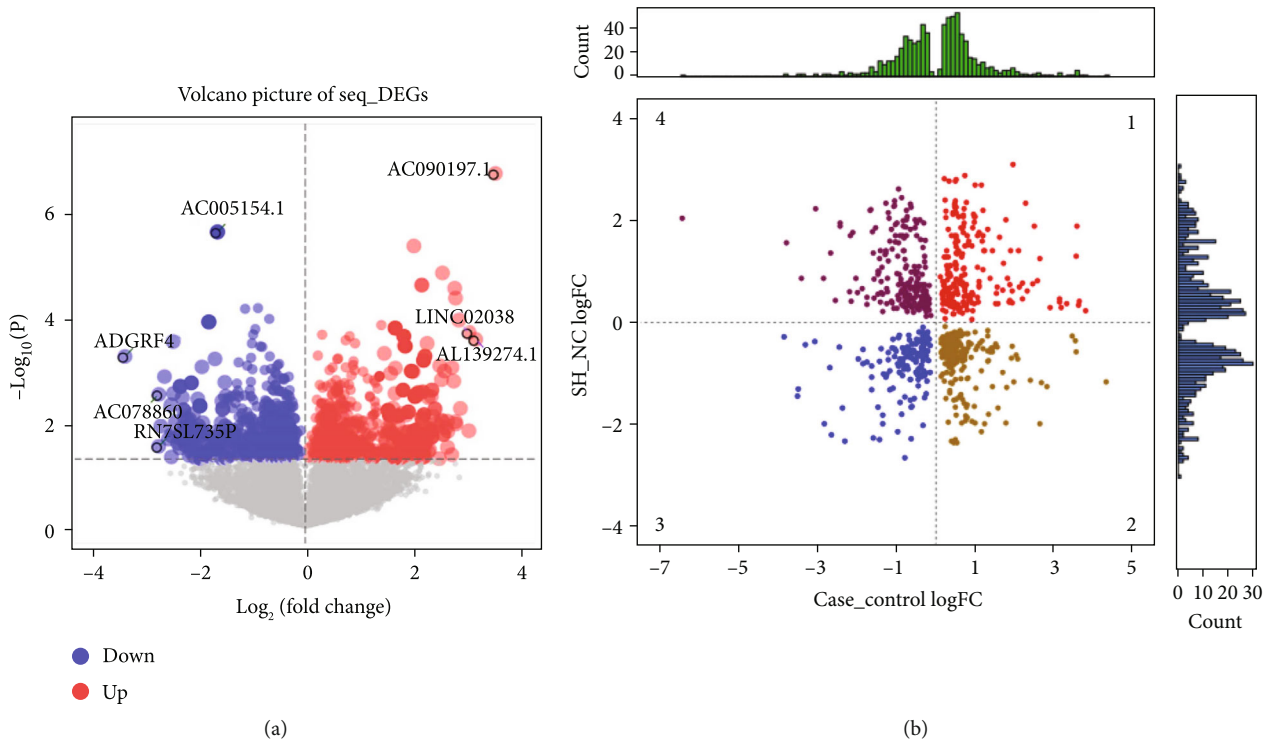
FIGURE 3: Continued.
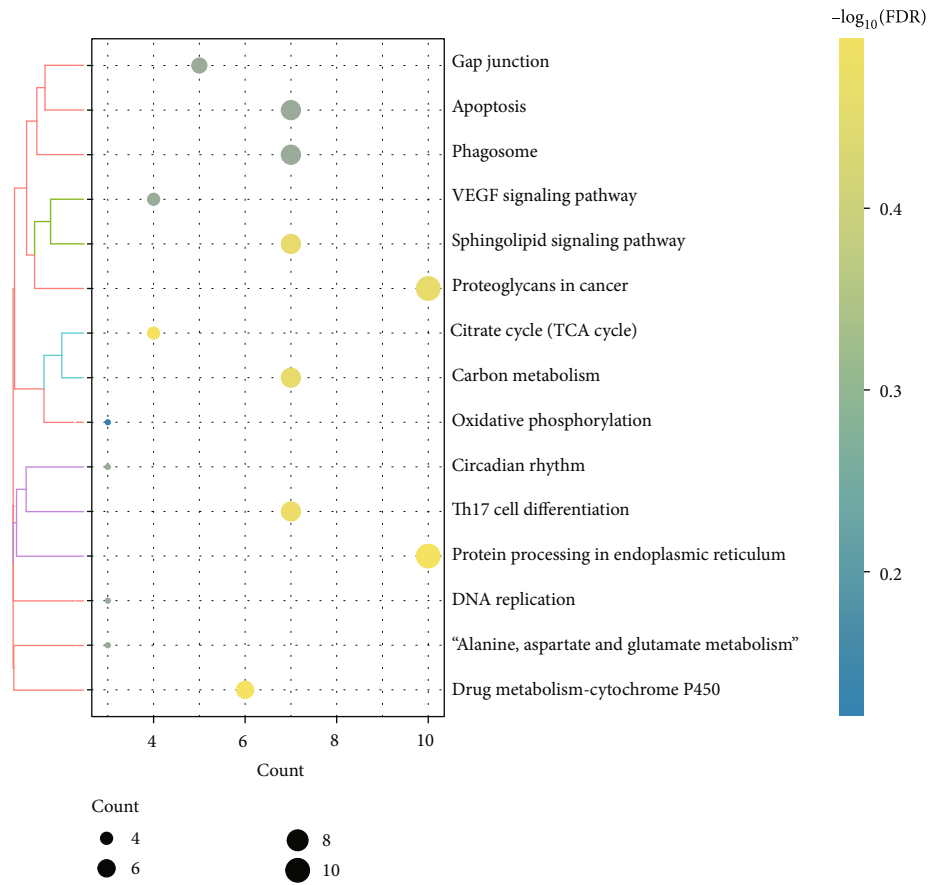
(d)
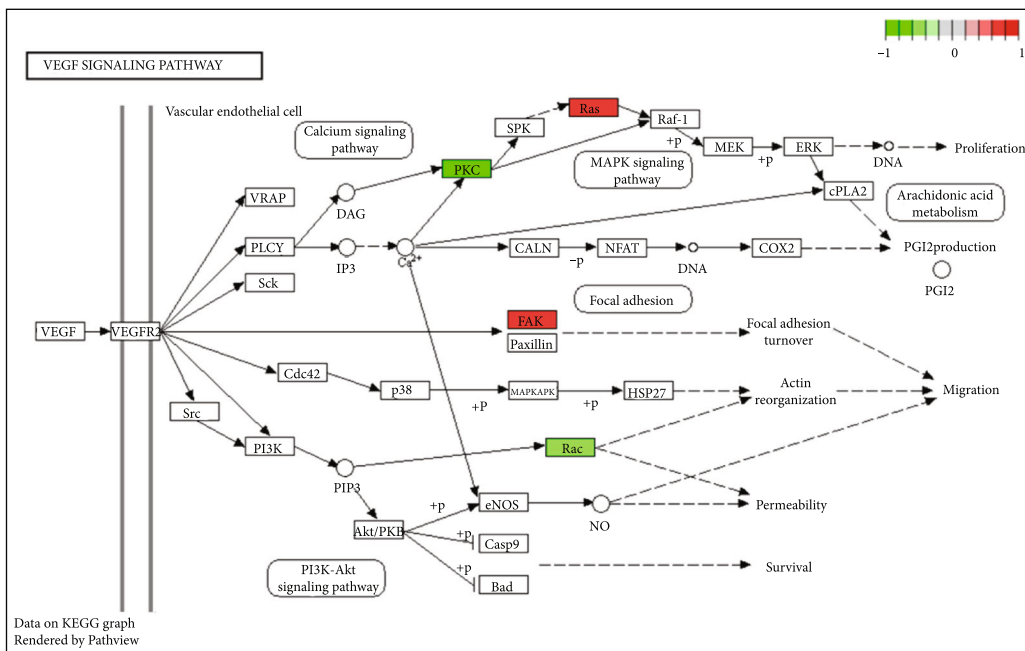


(e)

FIGURE 3: Identification and enrichment analysis of genes associated with high GOLM1 expression. (a) DEGs in GOLM1-knocked down (KD) MHCC97H cells and control MHCC97H cells. (b) DEGs between TCGA and RNA sequencing data. (c, d) GO and KEGG enrichment analyses of 400 specific DEGs associated with GOLM1 overexpression in HCC. (e) KEGG pathway annotations of the VEGF signaling pathway. Green indicates low expression, and red indicates high expression. DEGs: differentially expressed genes; GO: Gene Ontology; VEGF: vascular endothelial growth factor.

(a)

Figure 4: Continued.

(b)



(c)

Figure 4: Continued.

GOLM1                              Etotal: −657.85 kj/mol                              PTBP1

GOLM1                              Etotal: −544.67 kj/mol                              E2F4

(d)



(e)
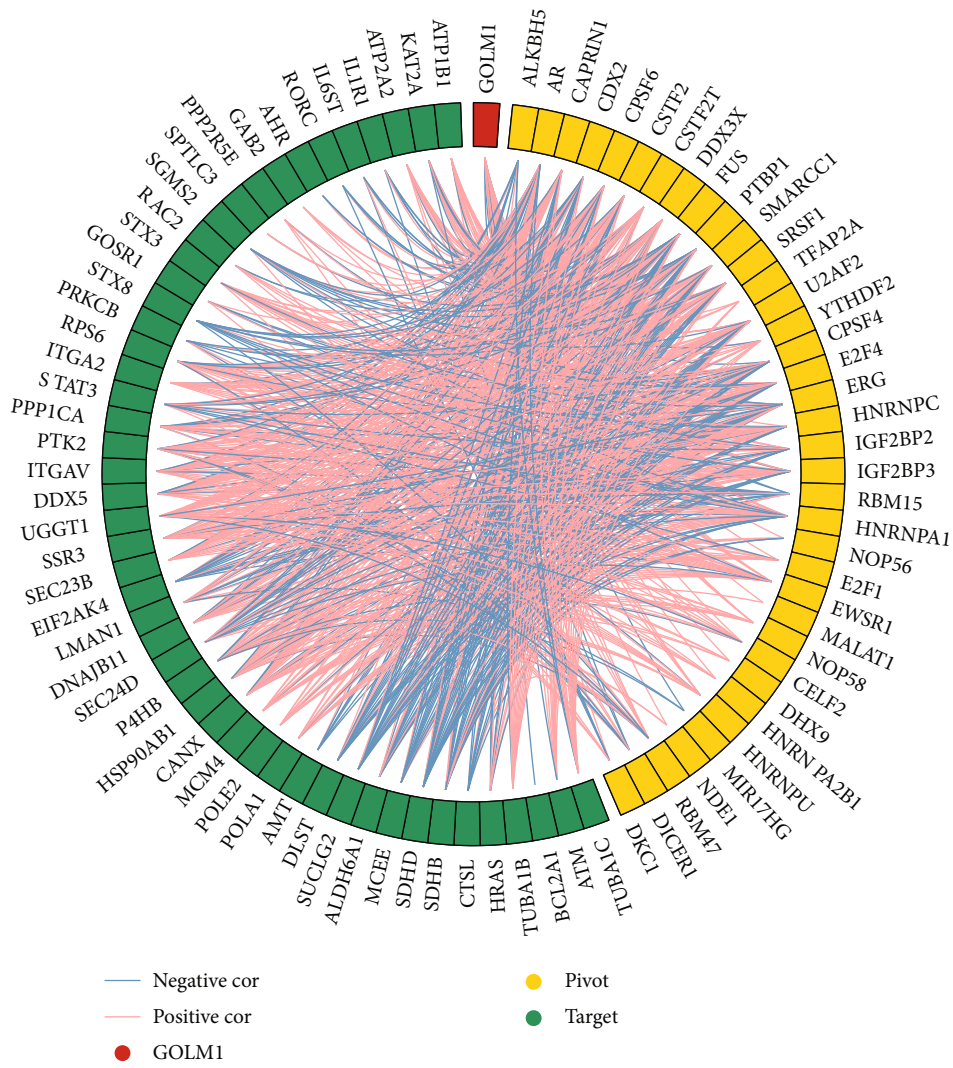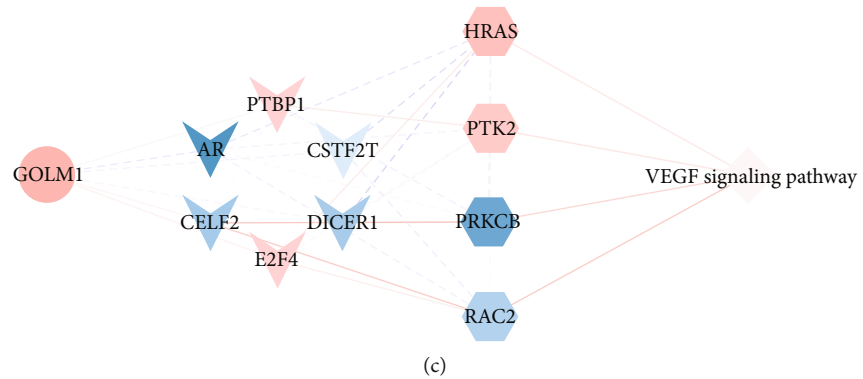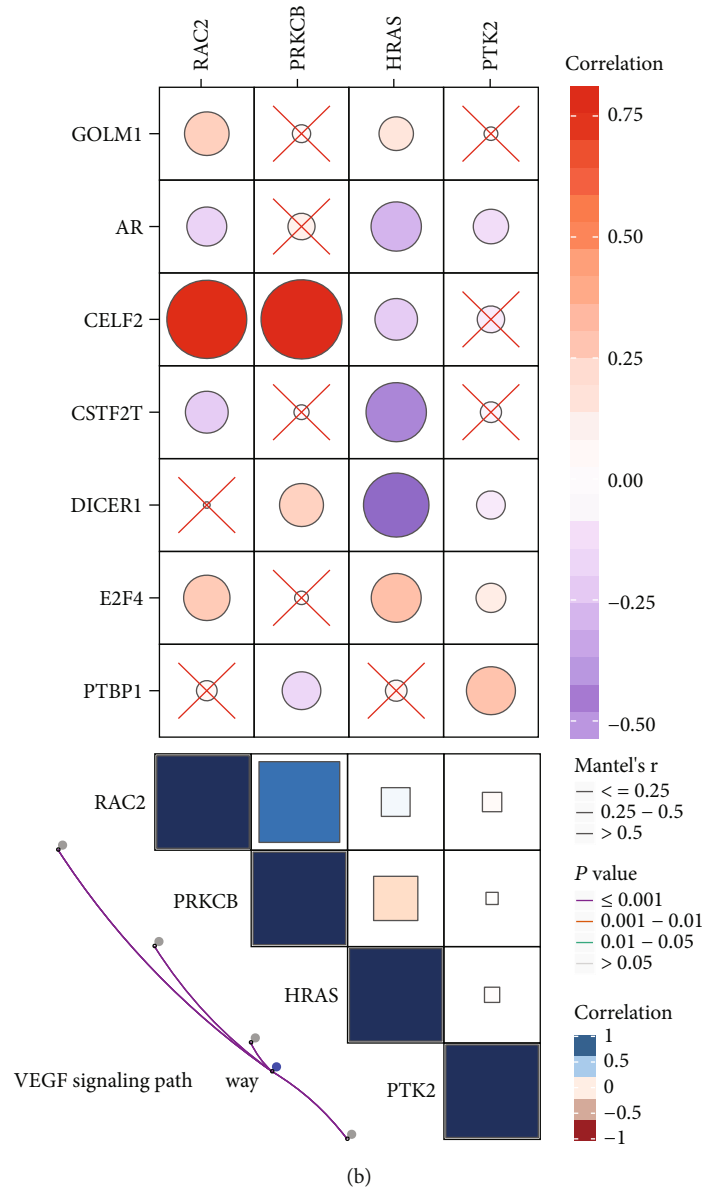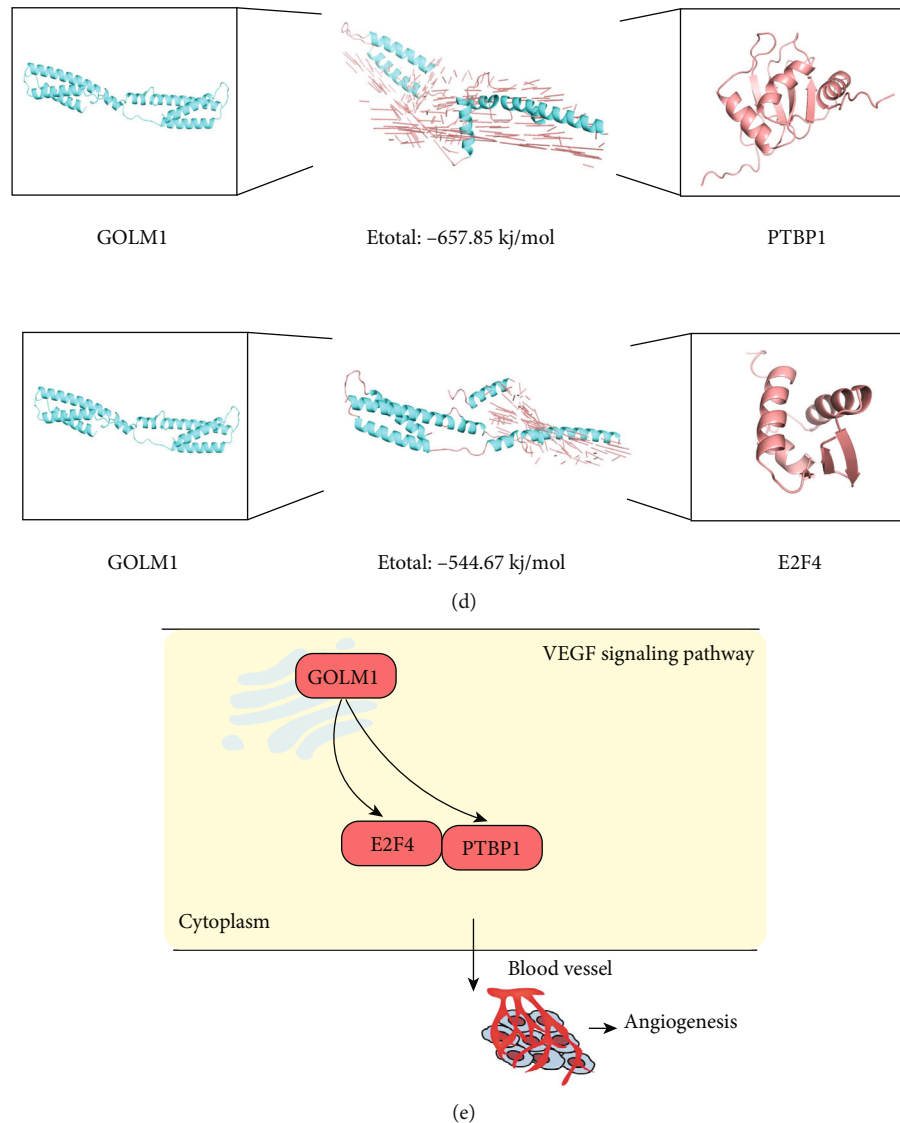
FIGURE 4: GOLM1 activates the vascular endothelial growth factor (VEGF) signaling pathway to promote angiogenesis in hepatocellular carcinoma. (a) Gene expression correlation network based on GOLM1 as well as pivot and target genes identified in the RNAInter database. (b) Pathway map interrelating GOLM1, pivot genes, pathway genes, and the VEGF signaling pathway. (c) Network view of GOLM1, pivot genes, target genes, and the VEGF pathway. Red represents upregulated gene expression, blue represents downregulated gene expression. (d) Molecular docking studies of PTBP1 and E2F4 with GOLM1. (e) Schematic of a potential mechanism by which GOLM1 promotes angiogenesis in HCC. VEFG: vascular endothelial growth factor.

reticulum-associated degradation pathway and the endoplasmic reticulum stress pathway are involved in protein processing in this organelle [30, 31]. Therefore, to explore the potential relevance of protein processing in the endoplasmic reticulum in HCC, we first constructed a comprehensive regulatory network including GOLM1, pivot genes, and pathway genes (Figure 6(a)). We then explored the potential of GOLM1 to bind other molecules in this network. The results predicted that GOLM1 could stably bind E2F1 and IGF2BP3 (Figure 6(b)).

Next, we extracted relevant genes from the comprehensive regulatory network for univariate Cox regression analysis, and 12 genes significantly associated with prognosis were identified. The 12 DEGs plus GOLM1 were combined to construct four models based on GSVA

(Figure 6(c)), multifactorial Cox regression (Figure 6(d)), LASSO (Figure 6(e)), and random forest (Figure 6(f)). The random forest regression model with 11 characteristic genes was the best prognostic model (Figure 6(g)). To validate its predictive value, we performed survival analysis according to the random forest risk score: patients with high scores had much worse OS and recurrence-free survival than those with low scores (Figure 6(h)). We confirmed this result in the validation dataset GSE54236 (additional Figures 2(a) and 2(b)).

3.6. Immune Infiltration in HCC. Studies have pointed out that protein processing in the endoplasmic reticulum plays a crucial role in immune responses [32, 33]. Using the TIMER 2.0 database, we showed that the 11 characteristic

(a)

Multivariate cox regression analysis



| Variable | Beta | P value | Hazard ratio (95% CI) |
|---|---|---|---|
| HAVCR1 | 0.22 | 0.002 | 1.245 (1.081–1.436) |
| ETFDH | 0.701 | 0.011 | 2.016 (1.175–3.457) |
| MMP7 | 0.163 | 0.020 | 1.177 (1.0.27–1.35) |

(b)



(c)

FIGURE 5: Continued.

(d)



GSVA_Score 1−Year AUC = 0.774
GSVA_Score Median−Year AUC = 0.779
GSVA_Score 3−Year AUC = 0.689
GSVA_Score 5−Year AUC = 0.587
GSVA_Score 8−Year AUC = 0.58

COX_Score 1−Year AUC = 0.687
COX_Score Median−Year AUC = 0.746
COX_Score 3−Year AUC = 0.7
COX_Score 5−Year AUC = 0.774
COX_Score 8−Year AUC = 0.643

Lasso_Score 1−Year AUC = 0.838
Lasso_Score Median−Year AUC = 0.86
Lasso_Score 3−Year AUC = 0.784
Lasso_Score 5−Year AUC = 0.756
Lasso_Score 8−Year AUC = 0.72

RandomForestSRC_Score 1−Year AUC = 0.889
RandomForestSRC_Score Median−Year AUC = 0.917
RandomForestSRC_Score 3−Year AUC = 0.882
RandomForestSRC_Score 5−Year AUC = 0.864
RandomForestSRC_Score 8−Year AUC = 0.72

(e)

Figure 5: Continued.

(f)



(g)

FIGURE 5: Establishment of the prognostic model for hepatocellular carcinoma (HCC) patients. (a) GSVA score of 53 prognostic genes. (b) Multivariate Cox regression analysis. (c) Establishment of the LASSO regression model. (d) Results of the random forest regression model for selecting prognostic characteristic genes. Thirty-five signature genes were selected by the random forest regression model. (e) The timeROC curve analysis of median survival and survival rates at one, three, five, and eight years for the above four models in TCGA. (f) Nomogram for predicting two- and five-year overall survival rates of HCC patients. The nomogram includes three variables: metastasis, N (presence or absence of lymphatic metastasis), and random forest risk score. (g) Performance validation of clinical prognostic models in the TCGA training sets. AUC: area under the receiver operating characteristic curve; GSVA: gene set enrichment analysis; LASSO: least absolute shrinkage and selection operator regression; RandomForestSRC: fast unified random forests for survival, regression, and classification.

genes correlated significantly with the abundance of neutrophils, endothelial cells, M2 macrophages, uncharacterized cells, and myeloid dendritic cells (Figures 7(a) and 7(b)). In addition, we analyzed the correlation between random forest risk score and immune checkpoints PDCD1, CD274, and CTLA4: the risk score correlated positively with expression of PDCD1 and CTLA4, but not with expression of CD274 (Figure 7(c)).

Based on the above results, we propose that GOLM1 may regulate protein processing in the endoplasmic reticulum by binding to E2F1 and IGF2BP3, thereby promoting the infiltration of endothelial cells and angiogenesis in HCC (Figure 7(d)).

## 4. Discussion

HCC is one of the cancers with higher incidence and mortality in the world [34]. Exploring how HCC occurs and progresses may help identify tumor markers, formulate effective treatment plans, and improve prognosis. HCC is a typical vascular tumor, and angiogenesis plays a key role in its growth [4]. However, vascular-related signaling pathways in HCC are still unclear, and new research is urgently needed to find new therapeutic targets. Our study shows that GOLM1 overexpression is closely related to vascular invasion of HCC. Therefore, exploring the relationship between GOLM1 and angiogenesis may help to identify new therapeutic targets in HCC.

(a)



GOLM1                      Etotal: -496.09 kj/mol                      E2F1

GOLM1                      Etotal: -616.56 kj/mol                      IGF2BP3

(b)



(c)

FIGURE 6: Continued.

Multivariate cox regression analysis

| Variable | Beta | P value | Hazard ratio (95% CI) |
|---|---|---|---|
| CSTF2 | 0.536 | 0.007 | 1.709 (1.16–2.517) |
| CAPRIN1 | 0.707 | 0.025 | 2.028 (1.091–3.769) |
| SSR3 | 0.527 | 0.046 | 1.693 (1.01–2.838) |



(d)



(e)



(f)
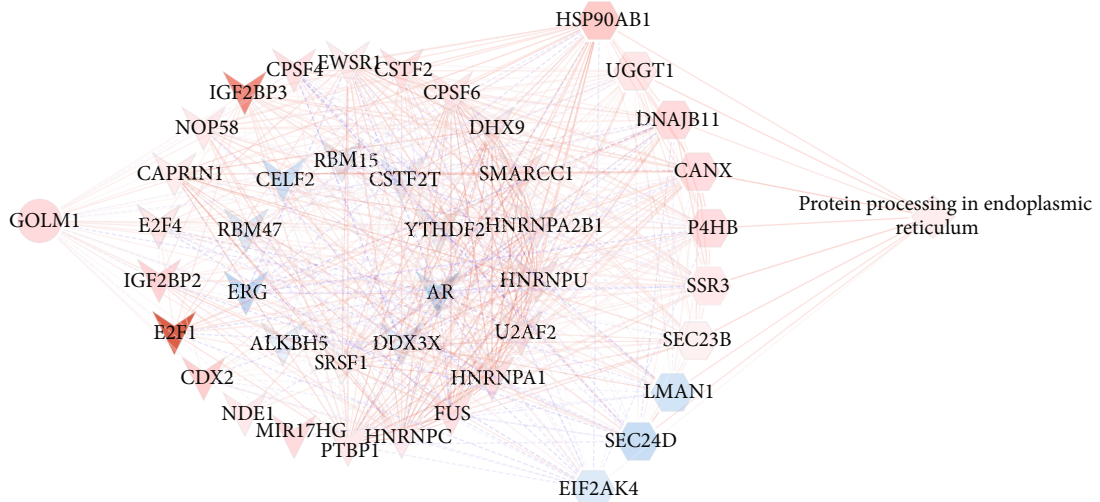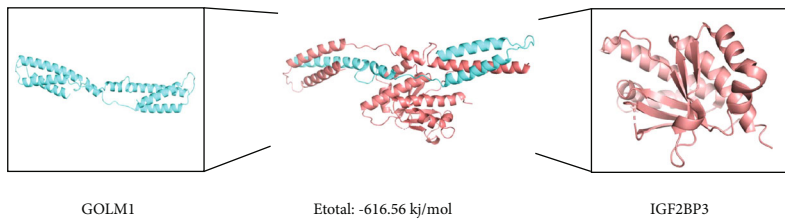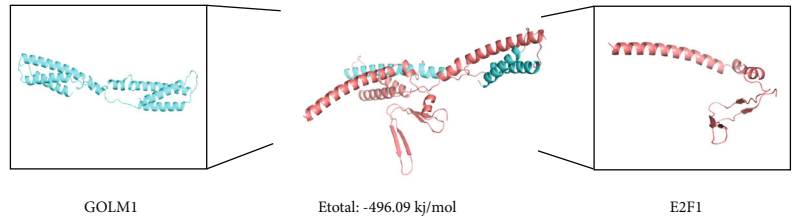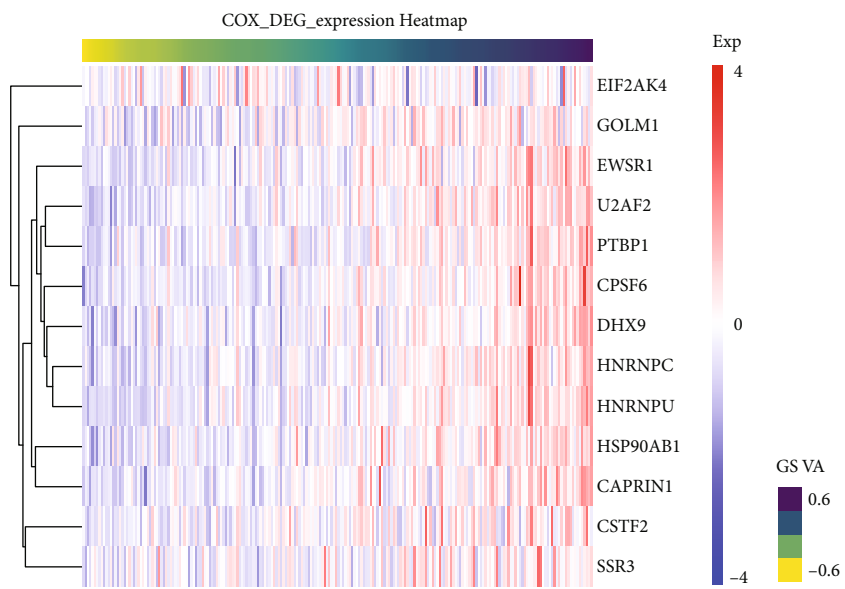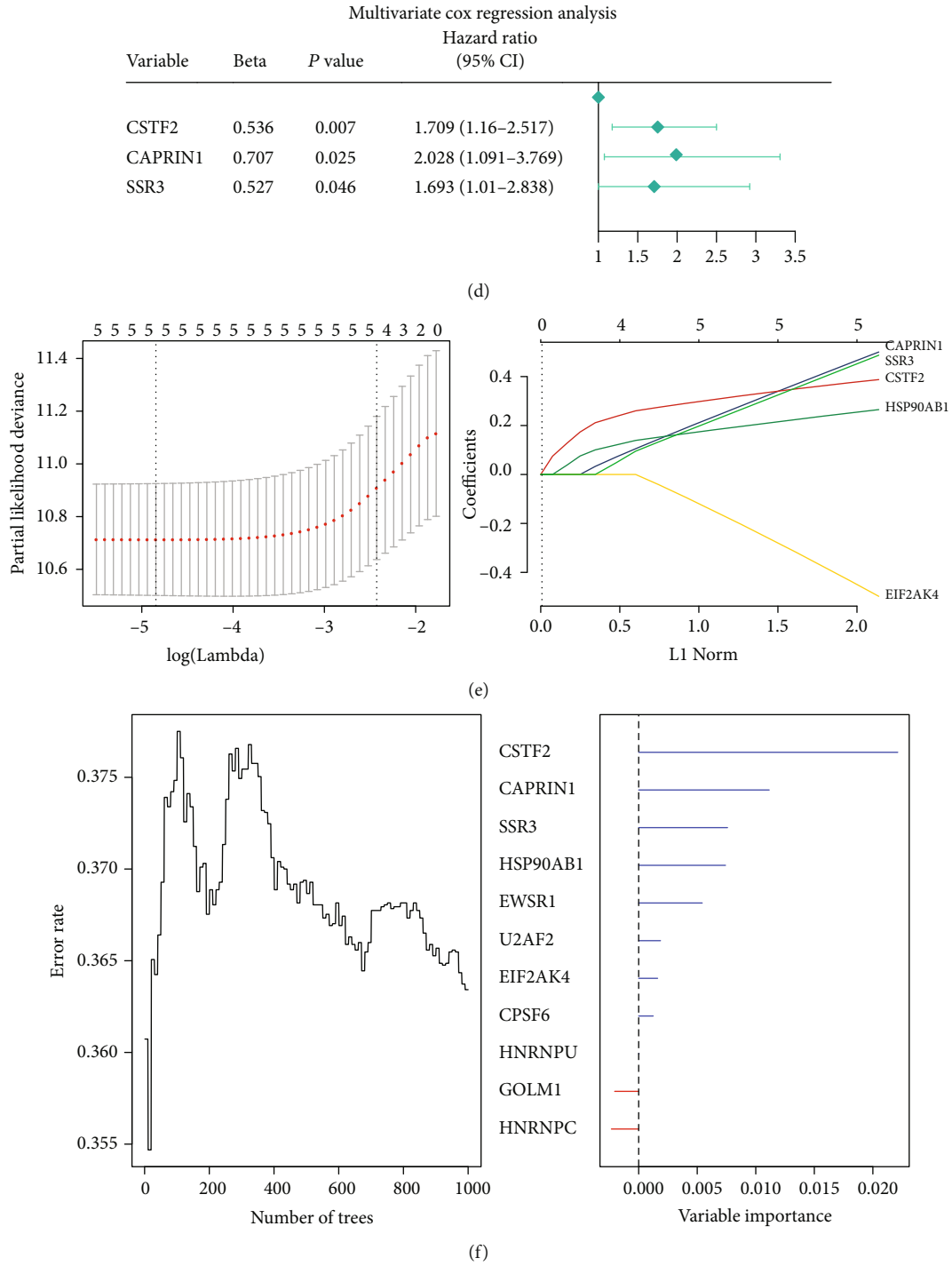
Figure 6: Continued.
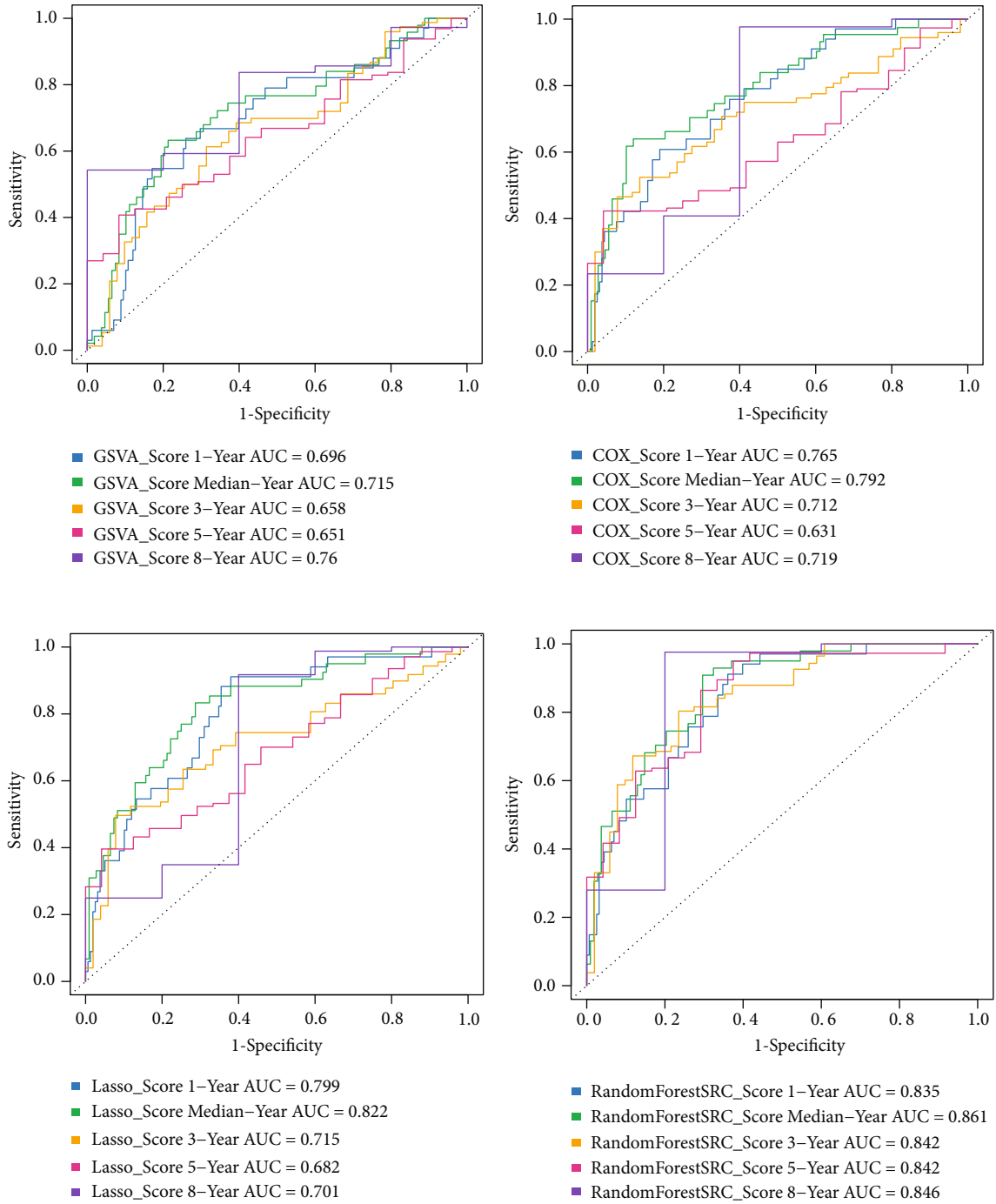
FIGURE 6: Continued.
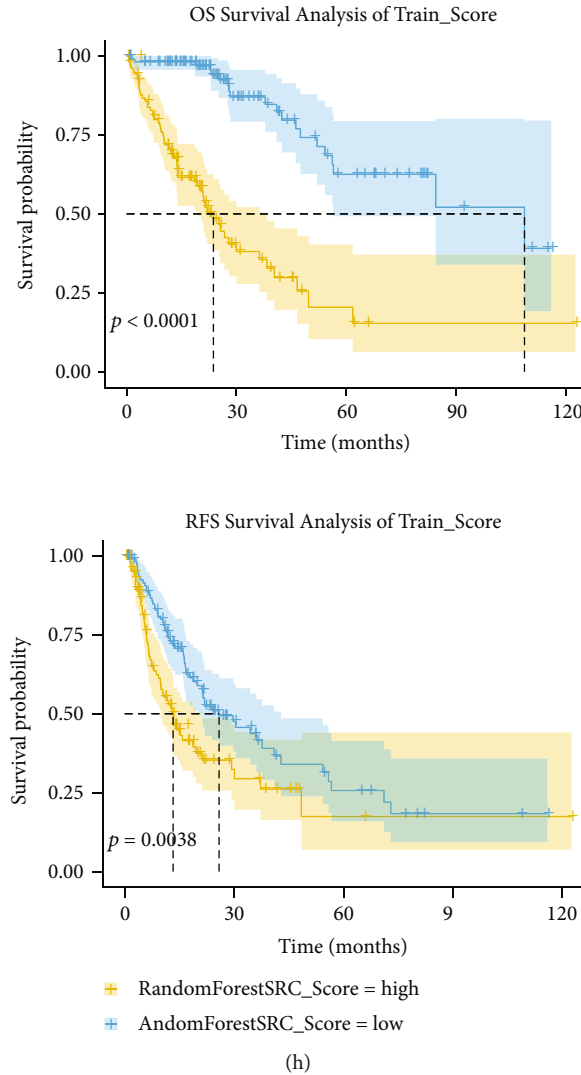
FIGURE 6: Role of GOLM1 in protein processing in the endoplasmic reticulum in hepatocellular carcinoma (HCC). (a) Regulatory network view of GOLM1, pivot genes, pathway genes, and the protein processing in the endoplasmic reticulum pathway. Red represents upregulated gene expression, and blue represents downregulated gene expression. (b) Molecular docking studies of E2F1 and IGF2BP3 with GOLM1. (c) GSVA score of the 13 prognostic genes. (d) Multivariate Cox regression analysis. (e) Establishment of the LASSO regression model. (f) The random forest regression model identified 11 signature genes associated with survival. (g) The timeROC curve analysis of median survival and survival rates at one, three, five, and eight years for the above four models in TCGA. (h) Performance validation of the optimal random forest regression model in the TCGA training set. AUC: area under the receiver operating characteristic curve; GSVA: gene set enrichment analysis; LASSO: least absolute shrinkage and selection operator regression; RandomForestSRC: fast unified random forests for survival, regression, and classification.

First, we used the TCGA database to explore GOLM1 expression in HCC. Consistent with previous studies, we found that GOLM1 was overexpressed in HCC compared with healthy controls and was closely associated with poor prognosis [35].

By analyzing sequencing data in both MHCC97H cells and TCGA database, we identified 400 specific DEGs associated with GOLM1 overexpression. GO analysis revealed that these genes are mainly involved in oxidative stress-related biological processes. KEGG pathway enrichment analysis showed that those genes were involved mainly in key pathways, such as the VEGF signaling pathway and protein processing in the endoplasmic reticulum. In view of the importance of angiogenesis in HCC, it is reasonable to hypothesize that GOLM1 promotes angiogenesis by activating VEGF signaling. The involvement of protein processing in the endoplasmic reticulum is also plausible, because endoplasmic reticulum stress has been implicated in HCC through its ability to promote tumor growth, metastasis, angiogenesis, and drug resistance [36]. In addition, increased oxidative stress is thought to be a recognized mechanism contributing to HCC [37]. Reactive oxygen species (ROS) are a source of oxidative stress generated in various organelles and stress pathways, such as mitochondria, peroxisomes, and endoplasmic reticulum [38]. Excessive ROS disrupts the integrity of proteins and lipids and may cause

(a)



(b)

Figure 7: Continued.

(c)
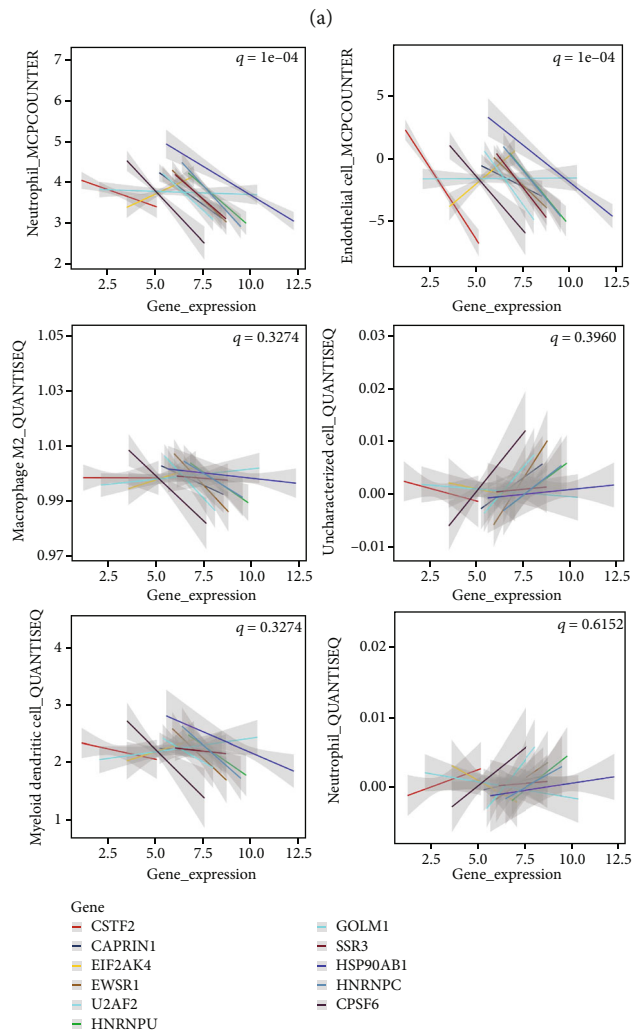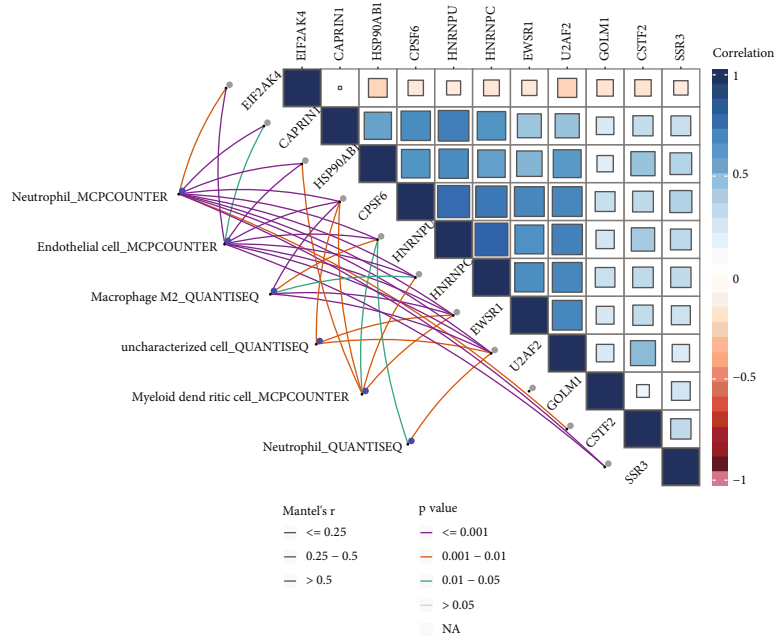


(d)
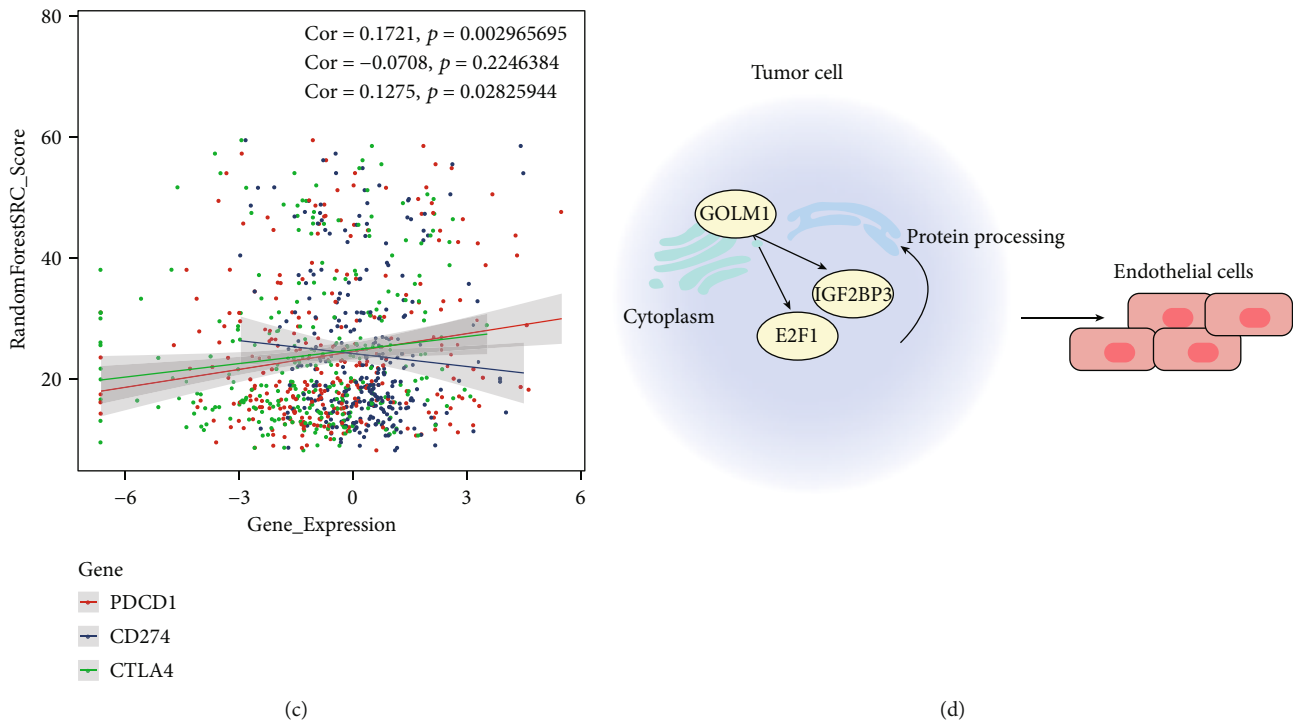
FIGURE 7: Immune infiltration in hepatocellular carcinoma (HCC) and relationship with GOLM1. (a, b) Correlation between 11 signature genes and immune cells. (c) Correlation between random forest regression model risk scores and immune checkpoints PDCD1, CD274, and CTLA4. (d) Schematic showing a potential mechanism by which GOLM1 may regulate protein processing in the endoplasmic reticulum. GOLM1 binds E2F1 and IGF2BP3 to promote tumor infiltration into the endothelium and affect angiogenesis in HCC.

genetic mutations inducing carcinogenesis [39]. The endoplasmic reticulum generates oxidative stress associated with endoplasmic reticulum stress when it contains endoplasmic reticulum redox protein 1 alpha and protein disulfide bond isomerase [40]. The endoplasmic reticulum oxidative stress triggers the release of hydrogen peroxide and calcium ions into the cytosol, further leading to increased mitochondrial oxidative stress and increased ROS. Hepatic oxidative stress leads to T cell tyrosine phosphatase (TCTPT) inactivation and promotes STAT3 signaling to drive HCC development [41].

Our results showed that GOLM1 might activate the VEGF pathway by binding to E2F4 and PTBP1. E2F4 is a novel tumor marker and well-established transcription factor that has been associated with HCC prognosis [42]. It is involved in the cell cycle, cell proliferation, resistance to apoptosis, and tumor progression [43, 44]. Recent studies reported that E2F4 overexpression is able to promote HCC cell proliferation by upregulating CDCA3 [45]. PTBP1 is an RNA binding protein that regulates RNA splicing and is involved in cellular processes such as the cell cycle, apoptosis, and immune activation [46]. PTBP1 regulates the alternative splicing of exon 10 in the Axl gene, allowing it to promote HCC cell invasion and metastasis [47]. E2F4 and PTBP1 have not previously been linked to HCC angiogenesis, so our results suggest that future studies should explore this possible link in detail.

Our results also suggest that GOLM1 may be involved in protein processing in the endoplasmic reticulum by binding to E2F1 and IGF2BP3, promoting endothelial cell infiltra-

tion. Endothelial cells are indispensably linked to angiogenesis, a complex, highly ordered process that is dependent on endothelial cells [48]. Our immunoinfiltration analysis found that endothelial cells were significantly infiltrated; however, the link between GOLM1 and endothelial cells is not known. E2F1 is a transcription factor involved mainly in the regulation of the cell cycle, cell proliferation, and apoptosis [49], and it is a key determinant of the survival of cells under endoplasmic reticulum stress [50]. IGF2BP3 is highly expressed in a variety of tumors including HCC, lung, and prostate cancers, and it helps maintain tumor cell growth, proliferation, invasion, and drug resistance through several oncogenic pathways [51]. IGF2BP3 can inhibit ZO-1 expression, enhancing the ability of HCC cells to invade [52]. Few studies have examined E2F1 and IGF2BP3 in HCC, so our results justify more detailed experiments into how they may contribute to disease onset and progression.

Based on the RNAInter database and VEGF signaling pathway, we identified 11 mechanism genes and then analyzed and verified these genes and 400 specific DEGs in TCGA and GEO databases. Finally, a new HCC prognostic model was constructed based on a random forest approach. Most of the 35 characteristic genes in the model have previously been linked to HCC. For example, low expression of UGP2 is associated with HCC progression [53]; RAN promotes the growth, migration, and invasion of HCC cells [54]; ITGAV is up-regulated in HCC and promotes tumor metastasis [55]; ETFDH is underexpressed in HCC and associated with poor OS [56]; and SERPINA3 mediates the upregulation of HNRNP-K transcriptional activity and

promotes the survival and proliferation of HCC cells [57]. Several genes in our model may be related to HCC but the potential connection requires further study. These genes include NDC1, TTBK1, PSMD11, FANCE, TRPM8, GSAP, SEPSECS, VMA21, PLA2G12A, SORD, UBTF, MCEE, and GOLM1.

We provide evidence that the random forest model is an independent prognostic factor for HCC, can be used to predict OS and recurrence-free survival, and can evaluate the prognosis of HCC patients. In order to improve the accuracy of the prognostic prediction, a nomogram was developed based on the random forest model as well as patient clinical characteristics. The OS nomogram also included the risk scores of metastasis, $N$ (presence or absence of lymphatic metastasis), and random forest score.

We also identified the optimal random forest regression model based on protein processing in the endoplasmic reticulum. The results of the Kaplan-Meier curve showed that the model can help identify the high-risk and low-risk HCC patients. Among the 11 characteristic genes identified, HNRNPC emerged as an independent prognostic factor for OS and disease-free survival in HCC patients, and it may be related to sorafenib treatment and anti-PD-1 immunotherapy response [58]. A recent study [59] suggested that the splicing regulator hnRNPU is a new transcriptional target of c-Myc in HCC. In that work, c-Myc upregulated hnRNPU, while hnRNPUSSR3 stabilized the c-Myc mRNA, thereby promoting c-Myc-driven HCC development. CPSF6, as an alternative polyadenylation factor, is an activator of pre-mRNA cleavage and polyadenylation processing [60]. CPSF6 is able to upregulate NQO1 to regulate HCC cell metabolism and thereby promote tumor development [61]. EWSR1 is strongly expressed in HCC, it is associated with histological grade and pathological T stage, and it is considered a novel tumor prognostic marker [62]. HSP90AB1 is also associated with HCC, and it may be involved in the progression from cirrhosis to HCC [63]. SSR3 is highly expressed in HCC and is associated with tumor size, TNM stage, differentiation grade, and poor prognosis [64]. CAPRIN1 is upregulated in HCC and can partially reverse the downregulation of c-MYC and CCND2 caused by miR-621 dysregulation, thereby promoting cell proliferation [65]. At present, the roles of EIF2AK4, U2AF2, and CSTF2 in HCC remain unclear. In this way, our results identify novel genes associated with HCC, and future investigation of these genes may provide new insights into the disease and its treatment.

In addition, various types of infiltrating immune cells have been described in the pathogenesis of HCC, and the potential role is not yet clear. To date, the main focus of cancer immunotherapy has been to interrupt immune checkpoints that inhibit antitumor lymphocytes. In addition to lymphocytes, the HCC milieu includes many other immune cell types, of which neutrophils are emerging as important contributors to the pathogenesis of hepatocellular carcinoma. A growing body of evidence supports neutrophils as key mediators of the immunosuppressive environment in which certain cancers develop and as drivers of tumor progression [66]. Little is known about the impact of endothelial cells on tumor cell behavior. In HCC patients, endothelial cells act as promoters of molecular crosstalk, enhancing HCC cell survival, migration, and invasion [67]. Recently, information about M2 macrophages promoting hepatocellular carcinoma metastasis revealed the mechanism of metastasis in HCC [68]. However the role of bone marrow dendritic cells in HCC is not known for the time being.

Although our study identified potential molecular mechanisms through which GOLM1 promotes HCC angiogenesis, it still has several limitations. First, our work was based mainly on bioinformatics predictions using previously published data from TCGA. Nevertheless, we validated our in silico findings using a GEO dataset and explored the role of GOLM1 in HCC using cell culture and RNA sequencing together. Second, the established nomogram model needs external validation. Since our prognostic model was constructed and validated using retrospective analysis of public databases, it should be confirmed in prospective studies. Future work should investigate, in vivo and in vitro, how GOLM1 promotes HCC angiogenesis.

## 5. Conclusions

Our study constructed HCC prognostic models based on DEG associated with GOLM1 overexpression, which may help to stratify HCC patients according to prognosis and to guide individualized treatment. Functional enrichment analysis of these genes led us to propose a mechanism by which GOLM1 promotes HCC angiogenesis. This may help develop effective treatments.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no competing interests.

## Authors' Contributions

Yan Lin and Ziqin He contributed equally to this work. Yan Lin and Ziqin He analyzed the data and drafted the manuscript. Jiazhou Ye and Rong Liang designed the study and revised the manuscript. Xing Gao, Lu Lu, Cheng Lu, Julu Huang and Min Luo produced and edited the chart. All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

## Acknowledgments

## Supplementary Materials

*Supplementary 1.* Figure S1. Validation of the random forest regression model in the GSE54236 validation set. (a) Kaplan-Meier curve of survival probability between the high-risk and low-risk groups in GSE54236. (b) The time-ROC curve analysis of the random forest regression model in GSE54236. AUC: area under the receiver operating characteristic curve; OS: overall survival; RFS: recurrence-free survival; RandomForestSRC: fast unified random forests for survival, regression, and classification.

*Supplementary 2.* Figure S2. Validation of the random forest regression model in the GSE54236 validation set. (a) Kaplan-Meier curves of survival probability between the high-risk and low-risk groups in GSE54236. (b) The time-ROC curve analysis of the random forest regression model in GSE54236. AUC: area under the receiver operating characteristic curve; OS: overall survival; RFS: recurrence-free survival; RandomForestSRC: fast unified random forests for survival, regression, and classification.

*Supplementary 3.* Supplementary Table 1. Details of the 400 specific DEGs associated with GOLM1 overexpression.

## References

[1] A. Villanueva, "Hepatocellular carcinoma," *The New England Journal of Medicine*, vol. 380, no. 15, pp. 1450–1462, 2019.

[2] J. M. Llovet, R. K. Kelley, A. Villanueva et al., "Hepatocellular carcinoma," *Nature Reviews Disease Primers*, vol. 7, no. 1, p. 6, 2021.

[3] M. Hilmi, C. Neuzillet, J. Calderaro, F. Lafdil, J. M. Pawlotsky, and B. Rousseau, "Angiogenesis and immune checkpoint inhibitors as therapies for hepatocellular carcinoma: current knowledge and future research directions," *Journal for Immunotherapy of Cancer*, vol. 7, no. 1, p. 333, 2019.

[4] M. A. Morse, W. Sun, R. Kim et al., "The role of angiogenesis in hepatocellular carcinoma," *Clinical Cancer Research*, vol. 25, no. 3, pp. 912–920, 2019.

[5] R. D. Kladney, G. A. Bulla, L. Guo et al., "GP73, a novel Golgi-localized protein upregulated by viral infection," *Gene*, vol. 249, no. 1-2, pp. 53–65, 2000.

[6] C. Bachert, C. Fimmel, and A. D. Linstedt, "Endosomal trafficking and proprotein convertase cleavage of cis Golgi protein GP73 produces marker for hepatocellular carcinoma," *Traffic*, vol. 8, no. 10, pp. 1415–1423, 2007.

[7] R. Liang, Z. Liu, X. Piao et al., "Research progress on GP73 in malignant tumors," *Oncotargets and Therapy*, vol. Volume 11, pp. 7417–7421, 2018.

[8] Y. Mao, H. Yang, H. Xu et al., "Golgi protein 73 (GOLPH2) is a valuable serum marker for hepatocellular carcinoma," *Gut*, vol. 59, no. 12, pp. 1687–1693, 2010.

[9] J. Yang, J. Li, W. Dai et al., "Golgi protein 73 as a biomarker for hepatocellular carcinoma: a diagnostic meta-analysis," *Experimental and Therapeutic Medicine*, vol. 9, no. 4, pp. 1413–1420, 2015.

[10] Y. Zhou, X. Yin, J. Ying, and B. Zhang, "Golgi protein 73 versus alpha-fetoprotein as a biomarker for hepatocellular carcinoma: a diagnostic meta-analysis," *BMC Cancer*, vol. 12, no. 1, p. 17, 2012.

[11] R. Liang and X. L. Luo, "Golgi protein-73 (GP73) in the diagnosis of primary liver cancer," *Guangxi Medical University*, 2012.

[12] J. Z. Ye, S. M. Yan, C. L. Yuan et al., "GP73 level determines chemotherapeutic resistance in human hepatocellular carcinoma cells," *Journal of Cancer*, vol. 9, no. 2, pp. 415–423, 2018.

[13] R. Liang, X. L. Luo, L. Y. Ge et al., "The diagnostic value of serum Golgi protein-73 in hepatocellular carcinoma," *Tumor*, vol. 32, no. 2, p. 115-8-4, 2012.

[14] Y. Liu, X. Zhang, S. Zhou et al., "Knockdown of Golgi phosphoprotein 73 blocks the trafficking of matrix metalloproteinase-2 in hepatocellular carcinoma cells and inhibits cell invasion," *Journal of Cellular and Molecular Medicine*, vol. 23, no. 4, pp. 2399–2409, 2019.

[15] Q. H. Ye, W. W. Zhu, J. B. Zhang et al., "GOLM1 modulates EGFR/RTK cell-surface recycling to drive hepatocellular carcinoma metastasis," *Cancer Cell*, vol. 30, no. 3, pp. 444–458, 2016.

[16] J. Yan, B. Zhou, L. Guo et al., "GOLM1 upregulates expression of PD-L1 through EGFR/STAT3 pathway in hepatocellular carcinoma," *American Journal of Cancer Research*, vol. 10, no. 11, pp. 3705–3720, 2020.

[17] E. Villa, R. Critelli, B. Lei et al., "Neoangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. Results from a prospective study," *Gut*, vol. 65, no. 5, pp. 861–869, 2016.

[18] M. E. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, article e47, 2015.

[19] 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.

[20] Y. Lin, T. Liu, T. Cui et al., "RNAInter in 2020: RNA interactome repository with increased coverage and annotation," *Nucleic Acids Research*, vol. 48, no. D1, pp. D189–D197, 2020.

[21] G. Macindoe, L. Mavridis, V. Venkatraman, M. D. Devignes, and D. W. Ritchie, "HexServer: an FFT-based protein docking server powered by graphics processors," *Nucleic Acids Research*, vol. 38, no. Web Server issue, pp. W445–W449, 2010.

[22] B. H. M. Mooers, "Shortcuts for faster image creation in PyMOL," *Protein Science*, vol. 29, no. 1, pp. 268–276, 2020.

[23] 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.

[24] R. Tibshirani, "The lasso method for variable selection in the Cox model," *Statistics in Medicine*, vol. 16, no. 4, pp. 385–395, 1997.

[25] X. Robin, N. Turck, A. Hainard et al., "pROC: an open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinformatics*, vol. 12, no. 1, p. 77, 2011.

[26] J. Wang, X. Chen, Y. Tian et al., "Six-gene signature for predicting survival in patients with head and neck squamous cell carcinoma," *Aging (Albany NY)*, vol. 12, no. 1, pp. 767–783, 2020.

[27] G. Bindea, B. Mlecnik, M. Tosolini et al., "Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer," *Immunity*, vol. 39, no. 4, pp. 782–795, 2013.

[28] T. Li, J. Fu, Z. Zeng et al., "TIMER2.0 for analysis of tumor-infiltrating immune cells," *Nucleic Acids Research*, vol. 48, no. W1, pp. W509–W514, 2020.

[29] H. H. Zhang, Y. Zhang, Y. N. Cheng et al., "Metformin incombination with curcumin inhibits the growth, metastasis, and angiogenesis of hepatocellular carcinoma in vitro and in vivo," *Molecular Carcinogenesis*, vol. 57, no. 1, pp. 44–56, 2018.

[30] D. S. Schwarz and M. D. Blower, "The endoplasmic reticulum: structure, function and response to cellular signaling," *Cellular and Molecular Life Sciences*, vol. 73, no. 1, pp. 79–94, 2016.

[31] S. A. Oakes and F. R. Papa, "The role of endoplasmic reticulum stress in human pathology," *Annual Review of Pathology*, vol. 10, no. 1, pp. 173–194, 2015.

[32] S. E. Bettigole and L. H. Glimcher, "Endoplasmic reticulum stress in immunity," *Annual Review of Immunology*, vol. 33, no. 1, pp. 107–138, 2015.

[33] J. S. So, "Roles of endoplasmic reticulum stress in immune responses," *Molecules and Cells*, vol. 41, no. 8, pp. 705–716, 2018.

[34] P. Bertuccio, F. Turati, G. Carioli et al., "Global trends and predictions in hepatocellular carcinoma mortality," *Journal of Hepatology*, vol. 67, no. 2, pp. 302–309, 2017.

[35] J. Yan, B. Zhou, H. Li, L. Guo, and Q. Ye, "Recent advances of GOLM1 in hepatocellular carcinoma," *Hepatic Oncology*, vol. 7, no. 2, article HEP22, 2020.

[36] J. Wu, S. Qiao, Y. Xiang et al., "Endoplasmic reticulum stress: multiple regulatory roles in hepatocellular carcinoma," *Biomedicine & Pharmacotherapy*, vol. 142, article 112005, 2021.

[37] F. Ciccarone, S. Castelli, and M. R. Ciriolo, "Oxidative stress-driven autophagy acROSs onset and therapeutic outcome in hepatocellular carcinoma," *Oxidative Medicine and Cellular Longevity*, vol. 2019, Article ID 6050123, 10 pages, 2019.

[38] T. Luangmonkong, S. Suriguga, H. A. M. Mutsaers, G. M. M. Groothuis, P. Olinga, and M. Boersema, "Targeting oxidative stress for the treatment of liver fibrosis," *Reviews of Physiology, Biochemistry and Pharmacology*, vol. 175, pp. 71–102, 2018.

[39] S. Reuter, S. C. Gupta, M. M. Chaturvedi, and B. B. Aggarwal, "Oxidative stress, inflammation, and cancer: how are they linked?," *Free Radical Biology and Medicine*, vol. 49, no. 11, pp. 1603–1616, 2010.

[40] H. M. Zeeshan, G. H. Lee, H. R. Kim, and H. J. Chae, "Endoplasmic reticulum stress and associated ROS," *International Journal of Molecular Sciences*, vol. 17, no. 3, p. 327, 2016.

[41] M. Grohmann, F. Wiede, G. T. Dodd et al., "Obesity drives STAT-1-dependent NASH and STAT-3-dependent HCC," *Cell*, vol. 175, no. 5, pp. 1289–1306.e20, 2018.

[42] Q. Zheng, Q. Fu, J. Xu, X. Gu, H. Zhou, and C. Zhi, "Transcription factor E2F4 is an indicator of poor prognosis and is related to immune infiltration in hepatocellular carcinoma," *Journal of Cancer*, vol. 12, no. 6, pp. 1792–1803, 2021.

[43] J. Hsu and J. Sage, "Novel functions for the transcription factor E2F4 in development and disease," *Cell Cycle*, vol. 15, no. 23, pp. 3183–3190, 2016.

[44] W. Ishioka, S. Oonuki, T. Iwadate, and K. I. Nihei, "Resorcinol alkyl glucosides as potent tyrosinase inhibitors," *Bioorganic & Medicinal Chemistry Letters*, vol. 29, no. 2, pp. 313–316, 2019.

[45] J. Liu, L. Xia, S. Wang et al., "E2F4 promotes the proliferation of hepatocellular carcinoma cells through upregulation of CDCA3," *Journal of Cancer*, vol. 12, no. 17, pp. 5173–5180, 2021.

[46] W. Zhu, B. L. Zhou, L. J. Rong et al., "Roles of PTBP1 in alternative splicing, glycolysis, and oncogensis," *Journal of Zhejiang University. Science. B*, vol. 21, no. 2, pp. 122–136, 2020.

[47] L. Shen, S. Lei, B. Zhang et al., "Skipping of exon 10 in Axlpre-mRNA regulated by PTBP1 mediates invasion and metastasis process of liver cancer cells," *Theranostics*, vol. 10, no. 13, pp. 5719–5735, 2020.

[48] S. Gao and C. T. Griffin, "RIPK3 modulates growth factor receptor expression in endothelial cells to support angiogenesis," *Angiogenesis*, vol. 24, no. 3, pp. 519–531, 2021.

[49] I. Shats, M. Deng, A. Davidovich et al., "Expression level is a key determinant of E2F1-mediated cell fate," *Cell Death and Differentiation*, vol. 24, no. 4, pp. 626–637, 2017.

[50] V. Pagliarini, P. Giglio, P. Bernardoni et al., "Downregulation of E2F1 during ER stress is required to induce apoptosis," *Journal of Cell Science*, vol. 128, no. 6, pp. 1166–1179, 2015.

[51] C. Mancarella and K. Scotlandi, "IGF2BP3 from physiology to cancer: novel discoveries, unsolved issues, and future perspectives," *Frontiers in Cell and Development Biology*, vol. 7, p. 363, 2020.

[52] Y. Gao, T. Luo, X. Ouyang, C. Zhu, J. Zhu, and X. Qin, "IGF2BP3 and miR191-5p synergistically increase HCC cell invasiveness by altering ZO-1 expression," *Oncology Letters*, vol. 20, no. 2, pp. 1423–1431, 2020.

[53] Q. Hu, S. Shen, J. Li et al., "Low UGP2 expression is associated with tumour progression and predicts poor prognosis in hepatocellular carcinoma," *Disease Markers*, vol. 2020, Article ID 3231273, 10 pages, 2020.

[54] L. Yao, Y. Zhou, Z. Sui et al., "RETRACTED: HBV-encoded miR-2 functions as an oncogene by downregulating TRIM35 but upregulating RAN in liver cancer cells," *eBioMedicine*, vol. 48, pp. 117–129, 2019.

[55] C. L. Kang, B. Qi, Q. Q. Cai et al., "LncRNA AY promotes hepatocellular carcinoma metastasis by stimulating ITGAV transcription," *Theranostics*, vol. 9, no. 15, pp. 4421–4436, 2019.

[56] Y. Wu, X. Zhang, R. Shen et al., "Expression and significance of ETFDH in hepatocellular carcinoma," *Pathology, Research and Practice*, vol. 215, no. 12, article 152702, 2019.

[57] E. Ko, J. S. Kim, J. W. Bae, J. Kim, S. G. Park, and G. Jung, "SERPINA3 is a key modulator of HNRNP-K transcriptional activity against oxidative stress in HCC," *Redox Biology*, vol. 24, article 101217, 2019.

[58] H. Jiang, G. Ning, Y. Wang, and W. Lv, "Identification of an m6A-related signature as biomarker for hepatocellular carcinoma prognosis and correlates with sorafenib and anti-PD-1 immunotherapy treatment response," *Disease Markers*, vol. 2021, Article ID 5576683, 15 pages, 2021.

[59] B. Zhang, H. Y. Wang, D. X. Zhao et al., "The splicing regulatory factor hnRNPU is a novel transcriptional target of c-Myc in hepatocellular carcinoma," *FEBS Letters*, vol. 595, no. 1, pp. 68–84, 2021.

[60] Y. Zhu, X. Wang, E. Forouzmand et al., "Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation," *Molecular Cell*, vol. 69, no. 1, pp. 62–74.e4, 2018.

[61] R. A. Rescorla, "Pavlovian conditioning. It's not what you think it is," *The American Psychologist*, vol. 43, no. 3, pp. 151–160, 1988.

[62] W. Jiang, T. Wu, X. Shi, and J. Xu, "Overexpression of EWSR1 (Ewing sarcoma breakpoint region 1/EWS RNA binding protein 1) predicts poor survival in patients with hepatocellular carcinoma," *Bioengineered*, vol. 12, no. 1, pp. 7941–7949, 2021.

[63] Y. Lin, R. Liang, J. Ye et al., "A twenty gene-based gene set variation score reflects the pathological progression from cirrhosis to hepatocellular carcinoma," *Aging (Albany NY)*, vol. 11, no. 23, pp. 11157–11169, 2019.

[64] S. Huang, W. Zhong, Z. Shi et al., "Overexpression of signal sequence receptor $\gamma$ predicts poor survival in patients with hepatocellular carcinoma," *Human Pathology*, vol. 81, pp. 47–54, 2018.

[65] Y. Zhang, W. You, H. Zhou et al., "Downregulated miR-621 promotes cell proliferation via targeting CAPRIN1 in hepatocellular carcinoma," *American Journal of Cancer Research*, vol. 8, no. 10, pp. 2116–2129, 2018.

[66] D. Geh, J. Leslie, R. Rumney, H. L. Reeves, T. G. Bird, and D. A. Mann, "Neutrophils as potential therapeutic targets in hepatocellular carcinoma," *Nature Reviews. Gastroenterology & Hepatology*, vol. 19, no. 4, pp. 257–273, 2022.

[67] Y. H. Wang, Y. Y. Dong, W. M. Wang et al., "Vascular endothelial cells facilitated HCC invasion and metastasis through the Akt and NF-$\kappa$B pathways induced by paracrine cytokines," *Journal of Experimental & Clinical Cancer Research*, vol. 32, no. 1, p. 51, 2013.

[68] J. Wu, W. Gao, Q. Tang et al., "Retracted: M2 macrophage–derived exosomes facilitate HCC metastasis by transferring $\alpha$M$\beta$2 integrin to tumor cells," *Hepatology*, vol. 73, no. 4, pp. 1365–1380, 2021.

*Research Article*

# Bulk and Single-Cell Transcriptome Analyses Revealed That the Pyroptosis of Glioma-Associated Macrophages Participates in Tumor Progression and Immunosuppression

**Lin Li,**[1] **Leyang Wu,**[1,2] **Xingpeng Yin,**[1] **Chenyang Li,**[1] **and Zichun Hua** ⬤[1,2,3]

[1]*The State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing, China*
[2]*Changzhou High-Tech Research Institute of Nanjing University and Jiangsu Target Pharma Laboratories Inc., Changzhou, China*
[3]*School of Biopharmacy, China Pharmaceutical University, Nanjing, China*

Correspondence should be addressed to Zichun Hua; zchua@nju.edu.cn

Glioma is the most common of all central nervous system (CNS) malignancies and is associated with a poor prognosis. Pyroptosis has been proven to be associated with the progression of multiple tumors and CNS diseases. However, the relationships between pyroptosis and clinical prognosis and immune cell infiltration are unclear in glioma. In this study, we conducted a comprehensive exploration of pyroptosis in glioma. First, prognosis-related genes were screened at each key regulatory locus in the pyroptosis pathway, and the prognostic ability and coexpression relationships of GSDMD and its upstream pathway genes NLRC4/CASP1/CASP4 were identified and well validated in multiple datasets. Tissue microarray-based immunohistochemistry results showed higher levels of NLRC4 and N-terminal GSDMD in high-grade gliomas, providing conclusive evidence of pyroptosis in gliomas. The robustness of the prognostic model based on these four genes was well validated in TCGA and CGGA cohorts. Bulk RNA-seq-based analysis showed that the group defined as the high-risk group according to the model showed activation of multiple inflammatory response pathways and impaired synaptic gene expression and had a higher infiltration of bone marrow-derived macrophages (BMDMs) and a hypersuppressed immune microenvironment. More importantly, three independent single-cell RNA-seq (scRNA-seq) datasets demonstrated that tumor-infiltrating macrophages, particularly BMDMs but not tissue-resident microglia, showed significant coexpression of the GSDMD and CASP genes, and BMDMs from high-grade gliomas accounted for a higher proportion of immune infiltrating cells and had higher expression of pyroptosis genes. Finally, we revealed the activation of pathways in response to LPS/bacteria and oxidative stress during BMDM development toward the pyroptosis cell fate by pseudotime trajectory analysis, suggesting potential BMDM pyroptosis initiators. The above results provide not only novel insights into the pathological mechanisms of glioma but also novel therapeutic targets for glioma, suggesting the potential application of pyroptosis inhibitors (e.g., disulfiram).

## 1. Introduction

Gliomas, which usually originate from glial cells or precursor cells and progress to astrocytomas, oligodendrogliomas, ventricular meningiomas, or oligodendroglial astrocytomas, account for approximately 80% of malignant tumors of the central nervous system (CNS) [1]. The World Health Organization classifies gliomas into 4 grades [2]. The 10-year survival rate for low-grade gliomas (grades I-II) is 47%, and the median survival time is 11.6 years, while the median overall survival time is 11.6 years, while the median overall survival

for grade IV gliomas is worse, at 15 months. With advances in targeted tumor treatment research and technology, there have been several breakthroughs in the identification of glioma molecular markers, such as isocitrate dehydrogenase (IDH) mutations [3] and O6-methylguanine-DNA methyltransferase (MGMT O6) promoter methylation [4]. However, these established markers are limited in their ability to elucidate the pathogenesis of glioma and are difficult to translate into targeted therapeutics. Therefore, it is urgent to explore new diagnostic assessment and prognostic analysis strategies

for glioma pathogenesis and progression mechanisms and to obtain novel therapeutic agents based on these mechanisms.

Pyroptosis is a proinflammatory mode of programmed cell death characterized by cell swelling and eventual rupture and the release of inflammatory contents following the perforation of the cell membrane, which is mediated by the N-terminal domain of the gasdermin protein [5, 6]. The gasdermin family has six members, including GSDMA, GSDMB, GSDMC, GSDMD, GSDME (DFNA5), and DFNB59. Except for DFNB59, the pyroptosis-mediating ability of all members has been well validated [7]. The inflammatory response caused by pyroptosis promotes immune cell infiltration to eliminate the pathogen or activate the tumor microenvironment [8–10]. However, excessive inflammatory responses not only damage normal cells but also reduce immune surveillance and the suppression of malignant cells, resulting in chronic inflammation and tumor immune escape [11]. As an important molecular marker of pyroptosis [9], IL1$\beta$ is generally released from the pore formed by oligomerized GSDMD [12] and has been shown to be closely associated with the formation of an immunosuppressive microenvironment in several tumor types [13–15]. Notably, several studies have shown that the proinflammatory cytokine IL1$\beta$ is significantly elevated in the serum of glioblastoma (GBM) patients and serves as a potential serum marker for this type of disease [16, 17]. In addition, a very recent study showed that monocyte-derived macrophages in gliomas secrete IL1$\beta$ in response to tumor cell induction, while Il1b knockdown significantly prolonged the survival time of primary glioma mice [18]. These studies suggest that aberrant pyroptosis signals may be present in glioma. In fact, pyroptosis has been demonstrated to be associated with the development of various peripheral inflammatory diseases and tumors, and several recent studies have reported that pyroptosis plays a critical role in the progression of CNS diseases, including Alzheimer's disease [19], multiple sclerosis [20], and stroke [21]. However, the relationship between pyroptosis and glioma, the most common primary tumor of the CNS, has rarely been reported. Here, we hypothesized that the progression of pyroptosis within gliomas could be used as a novel criterion for disease staging and prognostic assessment.

Although there have been several studies on prognosis-related pyroptosis genes in glioma [22, 23], they have been limited to establishing a prognostic model based on regression analysis while ignoring the correlation between pyroptosis and the tumor immune microenvironment. Moreover, these studies have tended to analyze all pyroptosis-related genes in general, but different gasdermin-mediated pyroptosis pathways are relatively independent from each other, and many genes are involved in other biological functions, such as apoptosis, so each pathway should be explored independently to reflect the situation of pyroptosis in glioma more accurately. The aim of this study was to identify the potential origin of pyroptosis activation signals in glioma through bioinformatics analysis of bulk RNA-seq data and single-cell RNA sequencing (scRNA-seq) data from glioma patients and to develop a risk score model based on markers of this signaling pathway to more effectively predict patient prognosis. This model will help to explore the relationship between pyroptosis and the

progression of glioma. In addition, we explored the potential link between pyroptosis and the immune microenvironment in glioma. We also used single-cell transcriptomics data to identify cell clusters in tumors with the ability to induce pyroptosis, and these clusters can provide targets for the development of novel therapies for glioma.

## 2. Materials and Methods

*2.1. Data Acquiring.* The bulk RNA-seq and clinical information of glioma patients were obtained from TCGA database (https://portal.gdc.cancer.gov/), CGGA database (http://www.cgga.org.cn/), and GEO database (https://www.ncbi.nlm.nih.gov/geo/). TCGA cohort contained 702 tumor samples, the CGGA cohort contained 325 tumor samples, the Bao dataset (GSE48865) [24] contained 274 tumor samples, and the Gravendeel dataset (GSE12907, GSE4271) [25] contained 276 tumor samples. TCGA and CGGA cohorts are used for candidate gene screening and prognostic model establishment and validation, while the other cohorts are used for candidate gene screening only. TCGA and CGGA datasets used for prognostic modeling screened samples according to the following criteria: (1) having WHO grade classification and ≥II; (2) having complete survival information, including overall survival and final events; and (3) having not received immune checkpoint blocker therapy. There were 597 samples in the filtered TCGA dataset and 306 samples in the CGGA dataset, and detailed clinical characteristics are summarized in Table 1.

The scRNA-seq expression profiles and cell annotation files of Cyril Neftel et al. (GSE131928) [26], which contained a total of 7930 cells from 28 patients, were obtained from the Single Cell Portal database (https://singlecell.broadinstitute.org/single_cell). The scRNA-seq expression profiles of Kai Yu et al. (GSE117891) [27], which contained 6148 cells from 13 patients, were obtained from the CGGA database. The scRNA-seq expression profiles of Andrew Venteicher et al. (GSE89567) [28], which contained 6341 cells from 10 patients, were obtained from the GEO database. A simplified workflow for the current study is depicted in Figure (1).

*2.2. Tumor Microenvironment Estimation.* Immune score, stromal score, and ESTIMATE score were calculated using the ESTIMATE R package [29]. CIBERSORT was used to predict the abundance of each type of cell infiltration in the tumor microenvironment (TME) [30]. Dysfunctional CD8+ T cell infiltration levels for TCGA_LGG and TCGA_GBM were obtained from the Tumor Immune Dysfunction and Exclusion (TIDE) portal (http://tide.dfci.harvard.edu/) [31]. Response to immune checkpoint blockade therapy in TCGA and CGGA cohorts was also predicted in the TIDE portal. In addition, we used single sample gene set enrichment analysis (ssGSEA) to predict immune inhibition scores and TGF-$\beta$ response scores (TBRS) in each sample based on the gene sets identified by Mariathasan et al. [32] (Supplementary Table. S1).

*2.3. Differentially Expressed Gene Analysis and Functional Annotation.* Analysis and functional annotation of

TABLE 1: Clinical characteristics of TCGA training cohort and CGGA validating cohort.

| Characteristics | TCGA dataset | | | | | CGGA dataset | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | High risk (N = 298) | Low risk (N = 299) | Total (N = 597) | p | FDR | High risk (N = 157) | Low risk (N = 156) | Total (N = 313) | p | FDR |
| Histology | | | | 4.40E − 38 | 2.60E − 37 | | | | 1.20E − 18 | 8.40E − 18 |
| A | 20 (3.35%) | 33 (5.53%) | 53 (8.88%) | | | 16 (5.11%) | 39 (12.46%) | 55 (17.57%) | | |
| AA | 54 (9.05%) | 56 (9.38%) | 110 (18.43%) | | | 33 (10.54%) | 24 (7.67%) | 57 (18.21%) | | |
| AO | 28 (4.69%) | 57 (9.55%) | 85 (14.24%) | | | 0 (0.0e + 0%) | 12 (3.83%) | 12 (3.83%) | | |
| AOA | 16 (2.68%) | 23 (3.85%) | 39 (6.53%) | | | — | — | — | | |
| GBM | 145 (24.29%) | 7 (1.17%) | 152 (25.46%) | | | 102 (32.59%) | 35 (11.18%) | 137 (43.77%) | | |
| O | 24 (4.02%) | 90 (15.08%) | 114 (19.10%) | | | 3 (0.96%) | 45 (14.38%) | 48 (15.34%) | | |
| OA | 11 (1.84%) | 33 (5.53%) | 44 (7.37%) | | | — | — | — | | |
| Grade | | | | 9.00E − 40 | 6.30E − 39 | | | | 7.70E − 16 | 4.60E − 15 |
| WHO II | 55 (9.21%) | 156 (26.13%) | 211 (35.34%) | | | 18 (5.75%) | 80 (25.56%) | 98 (31.31%) | | |
| WHO III | 98 (16.42%) | 136 (22.78%) | 234 (39.20%) | | | 34 (10.86%) | 40 (12.78%) | 74 (23.64%) | | |
| WHO IV | 145 (24.29%) | 7 (1.17%) | 152 (25.46%) | | | 102 (32.59%) | 35 (11.18%) | 137 (43.77%) | | |
| Gender | | | | 0.97 | 0.97 | | | | 0.53 | 0.6 |
| Female | 126 (21.11%) | 125 (20.94%) | 251 (42.04%) | | | 55 (17.57%) | 61 (19.49%) | 116 (37.06%) | | |
| Male | 172 (28.81%) | 174 (29.15%) | 346 (57.96%) | | | 102 (32.59%) | 95 (30.35%) | 197 (62.94%) | | |
| Age | | | | | | | | | | |
| Mean ± SD | 52.36 ± 15.86 | 42.40 ± 13.04 | 47.37 ± 15.34 | | | 47.03 ± 12.78 | 39.28 ± 9.71 | 43.17 ± 11.98 | | |
| OS | | | | | | | | | | |
| Mean ± SD | 524.51 ± 715.63 | 775.33 ± 915.69 | 650.13 ± 830.79 | | | 741.79 ± 905.26 | 2164.90 ± 1586.83 | 1451.07 ± 1472.63 | | |
| Censor | | | | 3.50E − 13 | 1.40E − 12 | | | | 1.30E − 10 | 5.10E − 10 |
| Alive | 168 (28.14%) | 251 (42.04%) | 419 (70.18%) | | | 21 (6.71%) | 74 (23.64%) | 95 (30.35%) | | |
| Dead | 130 (21.78%) | 48 (8.04%) | 178 (29.82%) | | | 136 (43.45%) | 82 (26.20%) | 218 (69.65%) | | |
| Temozolomide | | | | 5.90E − 02 | 0.2 | | | | 0.2 | 0.6 |
| No | 120 (20.10%) | 155 (25.96%) | 275 (46.06%) | | | 48 (15.34%) | 62 (19.81%) | 110 (35.14%) | | |
| Yes | 178 (29.82%) | 144 (24.12%) | 322 (53.94%) | | | 103 (32.91%) | 87 (27.80%) | 190 (60.70%) | | |
| Chemotherapy | | | | 0.14 | 0.28 | | | | 0.2 | 0.6 |
| No | 107 (17.92%) | 126 (21.11%) | 233 (39.03%) | | | 48 (15.34%) | 62 (19.81%) | 110 (35.14%) | | |
| Yes | 191 (31.99%) | 173 (28.98%) | 364 (60.97%) | | | 103 (32.91%) | 87 (27.80%) | 190 (60.70%) | | |
| IDH mutation | | | | 3.00E − 50 | 2.40E − 49 | | | | 1.20E − 23 | 2.80E − 22 |
| Mutant | 94 (15.75%) | 274 (45.90%) | 368 (61.64%) | | | 39 (12.46%) | 128 (40.89%) | 167 (53.35%) | | |
| Wild type | 199 (33.33%) | 24 (4.02%) | 223 (37.35%) | | | 118 (37.70%) | 27 (8.63%) | 145 (46.33%) | | |
| 1p19q codeletion | | | | 1.10E − 13 | 5.30E − 13 | | | | 1.50E − 13 | 9.20E − 13 |
| Codel | 33 (5.53%) | 114 (19.10%) | 147 (24.62%) | | | 4 (1.28%) | 58 (18.53%) | 62 (19.81%) | | |
| Noncodel | 261 (43.72%) | 184 (30.82%) | 445 (74.54%) | | | 148 (47.28%) | 95 (30.35%) | 243 (77.64%) | | |
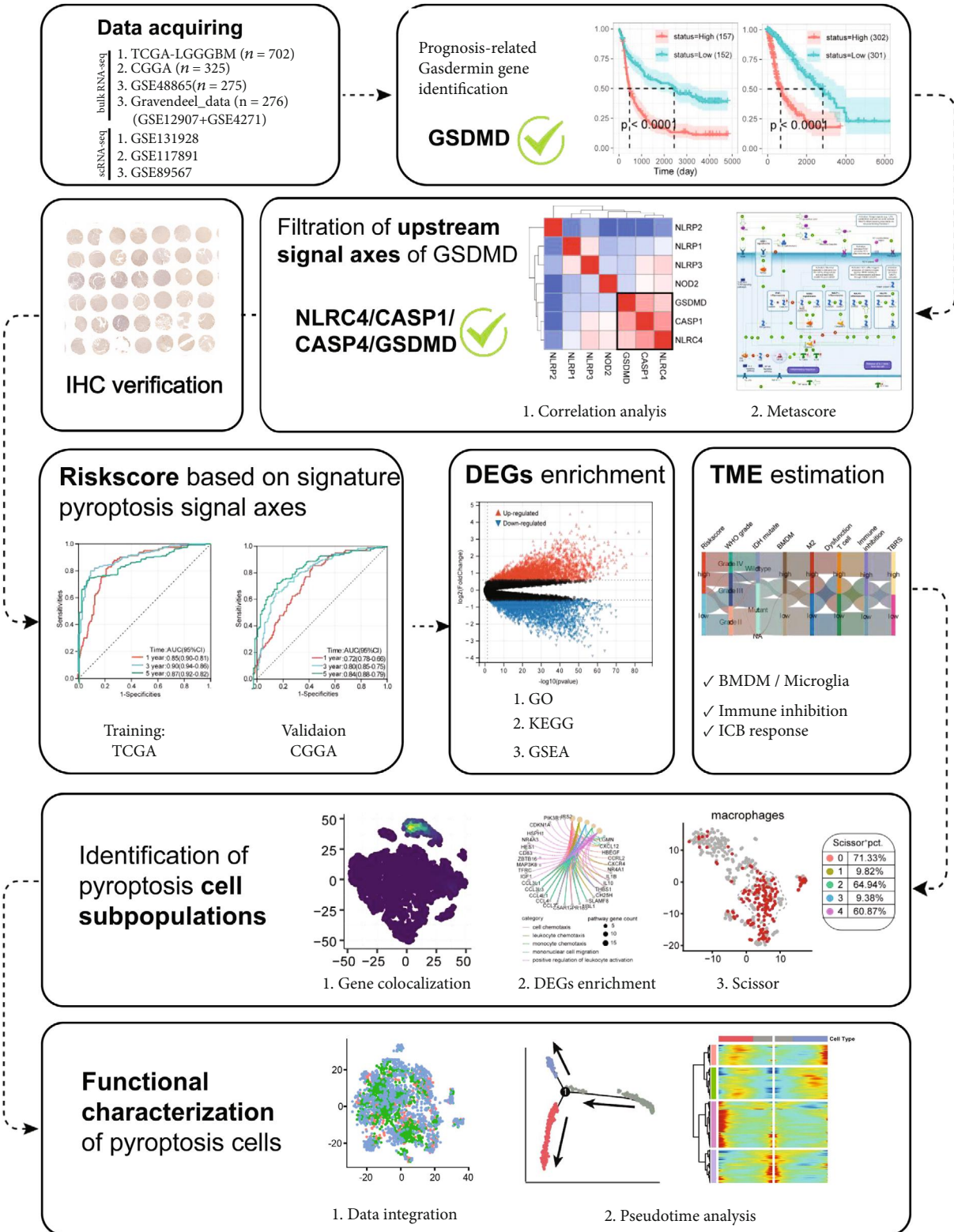
Figure 1: The complete research workflow. IHC: immunohistochemistry; DEGs: differentially expressed genes; GO: Gene Ontology analysis; KEGG: Kyoto Encyclopedia of Genes and Genomes; GSEA: gene set enrichment analysis; TME: tumor microenvironment; BMDM: bone marrow-derived macrophages; ICB: immune checkpoint blockade.

differentially expressed genes (DEGs) were performed using an empirical Bayesian approach by using the limma R package. Adjusted $p$ values less than 0.05 and absolute Log2 fold changes (log2FC) greater than 1.5 were considered DEGs and used for GO and KEGG functional annotation by the clusterProfiler R package.

*2.4. Visualization of the Pyroptosis Pathway.* To explore the GSDMD upstream signaling pathway, we divided TCGA cohort into the high- and low-expression groups by median GSDMD expression values and screened DEGs. The network building tool MetaCore™ version 5.4 (GeneGo) was used for pathway enrichment of differentially expressed genes as described previously [33], and the upstream signaling pathways of GSDMD were selected for visualization and analysis.

*2.5. Tissue Microarray and Immunohistochemistry (IHC).* Tissue microarrays purchased from Bioaitech were used for IHC. Each microarray contained 11 normal brain tissues, 7 grade I glioma samples, 32 grade II glioma samples, 22 grade III glioma samples, and 36 grade IV glioma samples. Anti-human GSDMD-N antibody (abcam, ab215203), anti-human NLRC4 antibody (abclonal, A13117), and anti-human PD1 antibody (servicebio, GB12338) were used for staining. The degree of IHC staining was reviewed and scored independently by two observers who were unaware of the clinical characteristics. The intensity of staining was scored according to the following criteria: cells with <25% staining were scored as (−, 1); cells with 25-49% staining were scored as (+, 2); cells with 50-74% staining were scored as (++, 3); and cells with 75-100% staining were scored as (+++, 4). The staining color was scored as negative light-yellow particle (1), brown-yellow particle (2), and brown particle (3). The final score was defined as the staining number score multiplied by the staining color score.

*2.6. Generation of Riskscore.* To establish a risk score that can assess the degree of activation of apoptotic pathways in individual patients, we performed multivariate Cox analysis on the screened highly conserved co-expressed gene cluster NLRC4/CASP1/CASP4/GSDMD using TCGA cohort as a training set. The Riskscore formula was constructed based on the coefficients of multivariate Cox analysis and validated for stability in the CGGA cohort. Kaplan–Meier curves were plotted to prove the prognostic value of the Riskscore, and log-rank tests were employed for analyzing statistical differences between the high- and low-risk groups. The accuracy of the Riskscore was assessed using receiver operating characteristic (ROC) curves. The independence of Riskscore was assessed using univariate and multifactorial Cox analyses.

*2.7. Copy Number Variation and Tumor Mutational Burden Analysis.* To determine copy number alteration events, we used the set of discrete copy number calls provided by GISTIC 2.0: homozygous deletion (−2); hemizygous deletion (−1); no-change (0); low-level gain (1); and high-level amplification (2). When more than half of the genes in the amplified or deleted peak region were high-level amplification (2) or homozygous deletion (−2), the copy number of the peak region is defined as changed. The oncoplot function in the maftools R package was used to visualize the general condition of the Mutation Annotation Format (MAF) of TCGA cohort in the form of a waterfall chart.

*2.8. Bone Marrow-Derived Macrophage (BMDM) and Microglia Estimation.* We used the single sample gene set enrichment analysis (ssGSEA) to predict BMDM and microglia infiltration scores in each sample based on the DEGs between microglia and BMDM demonstrated by Bowman et al. [34] as gene sets (Supplementary Table. S2). The gene set of DEGs in BMDM and microglia identified by Muller et al. [35] was used to validate the robustness of the above prediction. The same approach was used to predict the BMDM and microglia infiltration scores of macrophage subpopulations in the scRNA-seq dataset to distinguish BMDM and microglia at the single-cell level.

*2.9. scRNA-seq Data Processing.* The Seurat R package was used for scRNA-seq data processing as previously described [36]. Cells were removed if the number of expressed genes was less than 200 or more than 6,000, the UMI count was less than 1,000 and/or the percentage of mitochondrial genes was more than 0.1. The NormalizeData and ScaleData functions are used to normalize the matrix for subsequent cell clustering and dimensionality reduction. The first 2,000 highly variable genes identified by the FindVariableFeatures function were used in the RunPCA function for principal component analysis (PCA). The FindClusters function is used to cluster cells at a resolution of 0.5. RunTSNE is used to project cells into two dimensions and visualize them. The FindAllMarkers function was used to identify specific macrophage cluster DEGs compared to all other macrophage clusters. The harmony R package was used for integration and batch effect correction of expression profiles of BMDM and microglia from different datasets [37]. The monocle R package was used for performing differential expression and time-series analysis for single-cell expression experiments.

*2.10. Identifying Phenotype-Associated Subpopulations.* As previously described, the Scissor R package was used for phenotype-guided single-cell subpopulation identification [38]. Briefly, the Cyril Neftel single-cell expression matrix, TCGA bulk expression matrix, and phenotype of interest (overall survival in this study) were processed using Scissor. All cells can be divided into Scissor-positive (Scissor$^+$) cells and Scissor-negative (Scissor$^−$) cells, which are positively and negatively associated with the phenotype of interest, respectively.

*2.11. Statistical Analysis.* All statistical analyses were performed using R (4.1.2) software. Student's *t*-test (unpaired, two-tailed) was used to assess differences between two independent groups, and the Wilcoxon test is used for nonparametric tests between data that do not conform to a normal distribution. One-way analysis of variance (ANOVA) was used as a parametric method for data from more than two groups. The chi-square test was executed for the comparison of categorical variables between the high- and low-risk groups. The survivor and survminer R packages were used for survival analysis.

# 3. Results and Discussion

*3.1. GSDMD Significantly Correlated with the Progression and Overall Survival of Glioma.* Gasdermin proteins are the final executors of pyroptosis, and their expression level directly affects the possibility of pyroptosis occurring [39]. Considering that the gasdermin family contains five members that have been confirmed to mediate pyroptosis, we

(a)



(b)
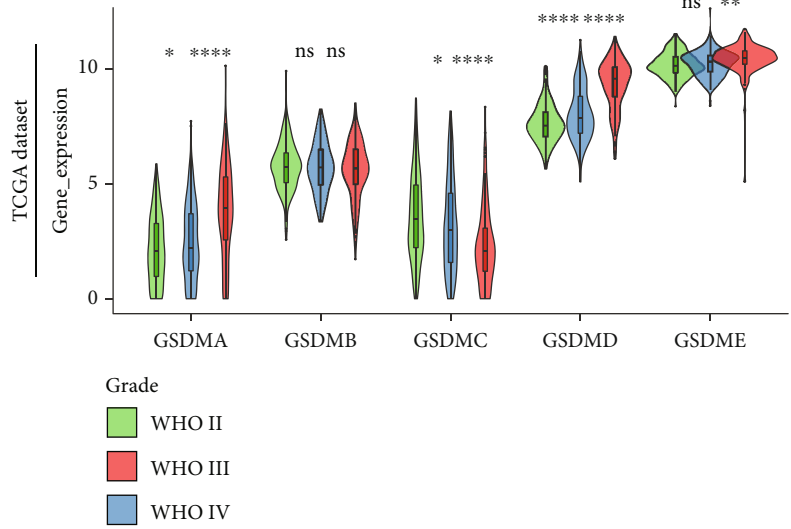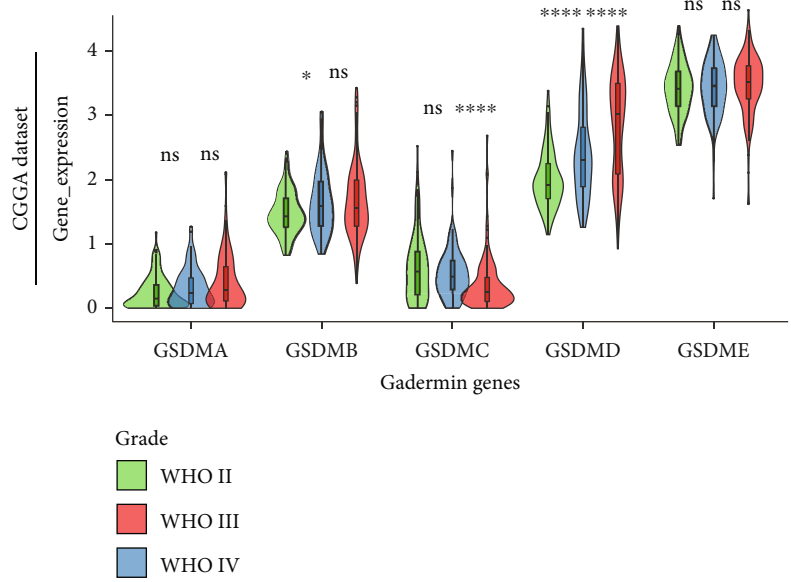
Figure 2: Continued.
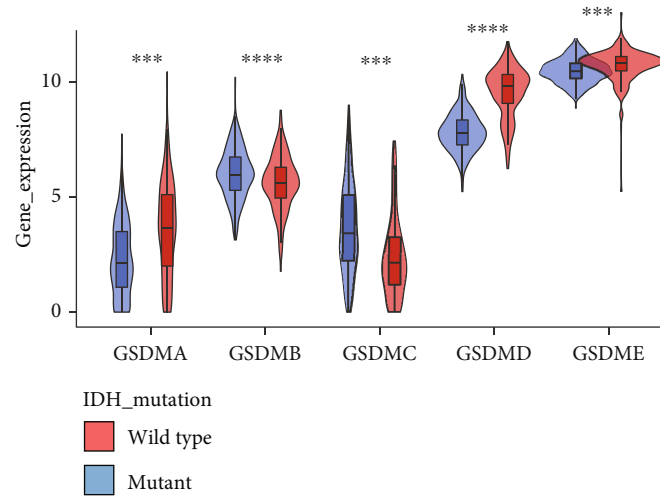
(c)
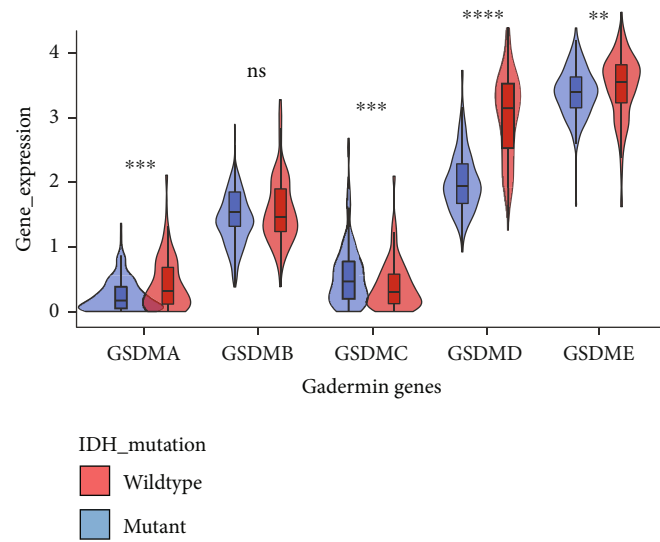


(d)
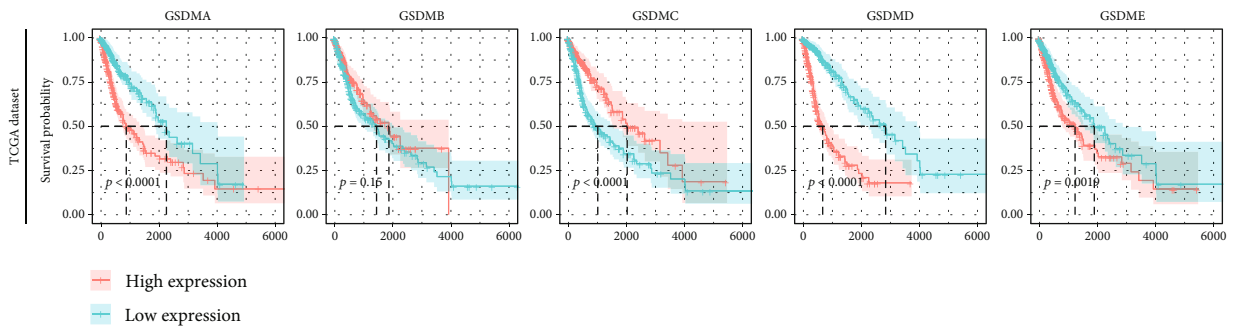


(e)

Figure 2: Continued.

(f)
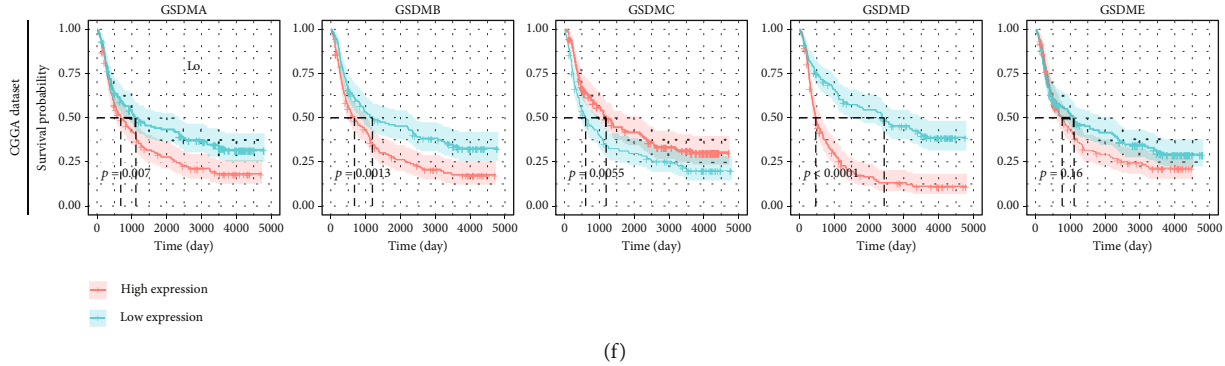
FIGURE 2: GSDMD expression was significantly associated with the progression and prognosis of glioma. (a, b) Expression levels of gasdermin family genes in patients with different WHO grades in TCGA and CGGA cohorts. (c, d) Expression levels of gasdermin family genes in patients with different IDH mutation phenotypes in TCGA and CGGA cohorts. (e, f) Kaplan–Meier plots for overall survival time (OS) of patients with different gasdermin family gene expression in TCGA and CGGA cohorts, using the median expression of each gene as a cutoff point to divide the high and low expression groups. Statistics were calculated using two-tailed, unpaired Student's $t$-test with Welch's correction in a-d. ns: not significant. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$.

examined the relationship between the different gasdermin genes and the progression or prognosis of glioma to investigate the most critical gene for the execution of pyroptosis in gliomas.

We first analyzed the expression of gasdermin family genes in different-grade gliomas in TCGA-LGGGBM and CGGA cohorts. The results showed that among the five pyroptosis genes of this family, only the expression of GSDMD showed a stable correlation with disease progression in both cohorts, and higher WHO grades corresponded to higher GSDMD expression (Figures 2(a) and 2(b)). This finding is consistent with the findings reported by Liu et al. that (1) GSDMD protein levels were elevated in clinical glioma tissue, accompanied by significant cleavage bands, and (2) GSDMD protein expression in GBM samples was higher than that in LGG samples [40]. IDH represents a major biomarker with diagnostic, prognostic, and predictive implications in glioma, and mutant phenotypes have a worse prognosis (Figure S1). The expression of multiple gasdermin genes in TCGA and CGGA cohorts was significantly different among the IDH phenotype groups (Figures 2(c) and 2(d)). However, in TCGA cohort, the expression of GSDMD (wild type vs. mutant) Log2FC = 1.784 was compared with the expression of gasdermin Log2FC in -0.823~0.899; in the CGGA cohort, the expression of GSDMD (wild type vs. mutant) Log2FC = 1.603 was compared with the expression of gasdermin Log2FC in -0.553~0.815. Therefore, among the gasdermin family members, GSDMD is the most differentially expressed gene among different IDH phenotypes of glioma.

To further confirm that GSDMD has a more significant indicative role in glioma than other gasdermin genes, we evaluated the overall survival time of each group based on the clinical information of patients from TCGA and CGGA datasets and gasdermin gene expression profiles, using the median expression of each gene as a cutoff point to divide the high and low expression groups (Figures 2(e) and 2(f)). The results showed that GSDMA, GSDMC, and GSDMD were significantly correlated with prognosis in both TCGA

and CGGA datasets and high expression of GSDMA and GSDMD corresponded to shorter survival, but GSDMC showed the opposite results. However, GSDMB and GSDME showed only limited prognostic relevance in a single dataset. Notably, in TCGA and CGGA cohorts, compared to the high expression group, the median survival time was prolonged 4.36-fold and 5.22-fold in the GSDMD low-expression group, while it was prolonged only 2.38-fold and 1.63-fold in the GSDMA low-expression group.

In conclusion, multiple gasdermin-mediated complex pyroptosis signaling networks may exist in gliomas. However, compared with other gasdermin genes, the pyroptosis triggered by GSDMD plays the most critical role in both the progression and the prognosis of glioma. Therefore, we performed subsequent data mining work around GSDMD.

*3.2. The NLRC4/CASP1/CASP4/GSDMD Pyroptosis Signaling Axis Can Be Used as a Prognostic Factor for Glioma.* The process of pyroptosis requires not only gasdermin expression but also upstream activation signals leading to gasdermin cleavage and N-terminal domain release, which are equally crucial. It has been well demonstrated that caspase-1, caspase-4, and caspase-5 are activators of GSDMD, capable of cleaving GSDMD at hGSDMD$_{276}$ and releasing GSDMD-NT, leading to pyroptosis [5]. CASP1, CASP4, and CASP5 were all significantly associated with overall survival in TCGA cohort (Figure 3(a)). Among them, CASP1 and CASP4 showed a particularly significant positive correlation with GSDMD ($R > 0.75$, $p < 10^{-16}$), while the coexpression of CASP5 was discrete (Figure 3(b)). The highly conserved coexpression relationships between CASP1/GSDMD and CASP4/GSDMD were validated in three other independent datasets (Figure S2), indicating CASP1/CASP4/GSDMD signaling axis activation in glioma.

Inflammasomes are the activator of multiple caspases, among which multiple inflammasomes such as NLRP1, NLRP3, and NLRC4 are directly involved in classical or nonclassical pyroptosis pathways and act as receptors of

(a)



(b)



(c)
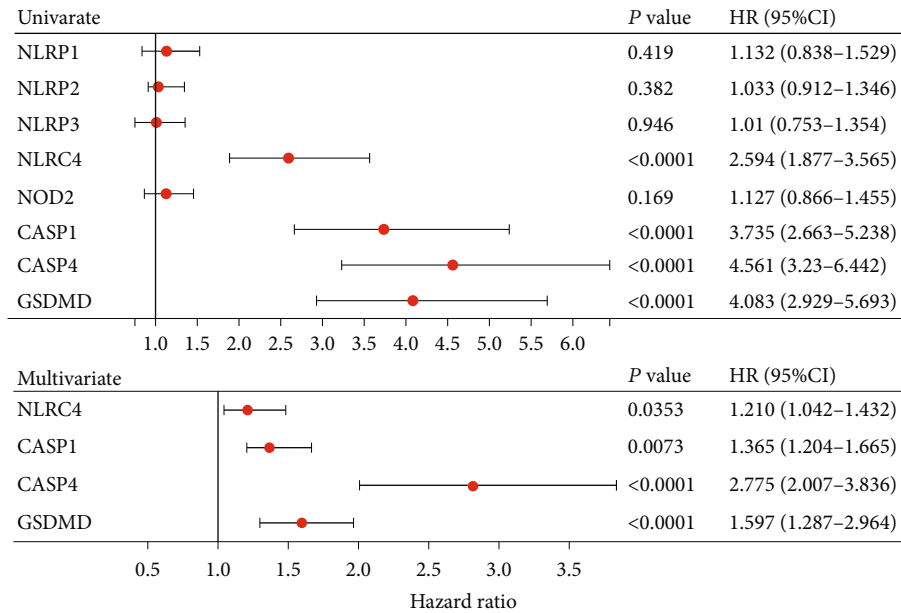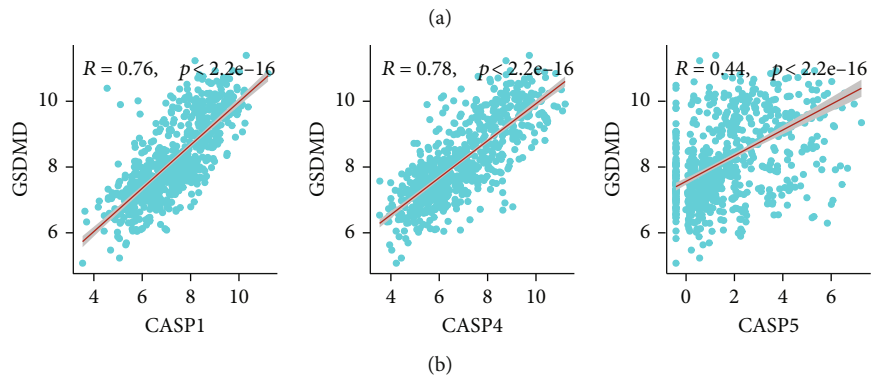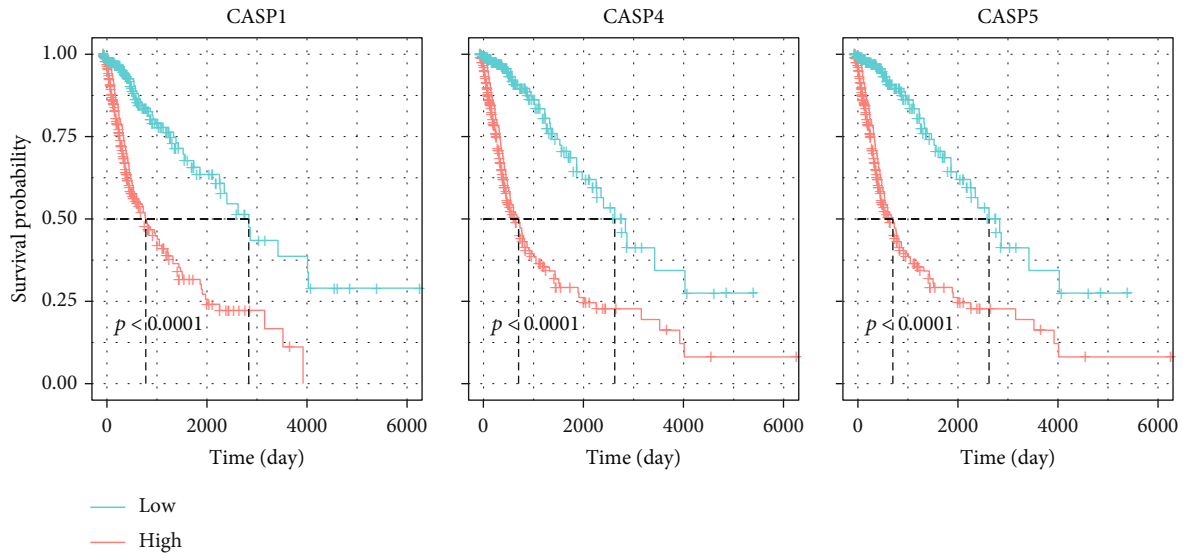
Figure 3: Continued.

(d)

FIGURE 3: Continued.

(e)



(f)



(g)

FIGURE 3: Continued.

(h)



(i)
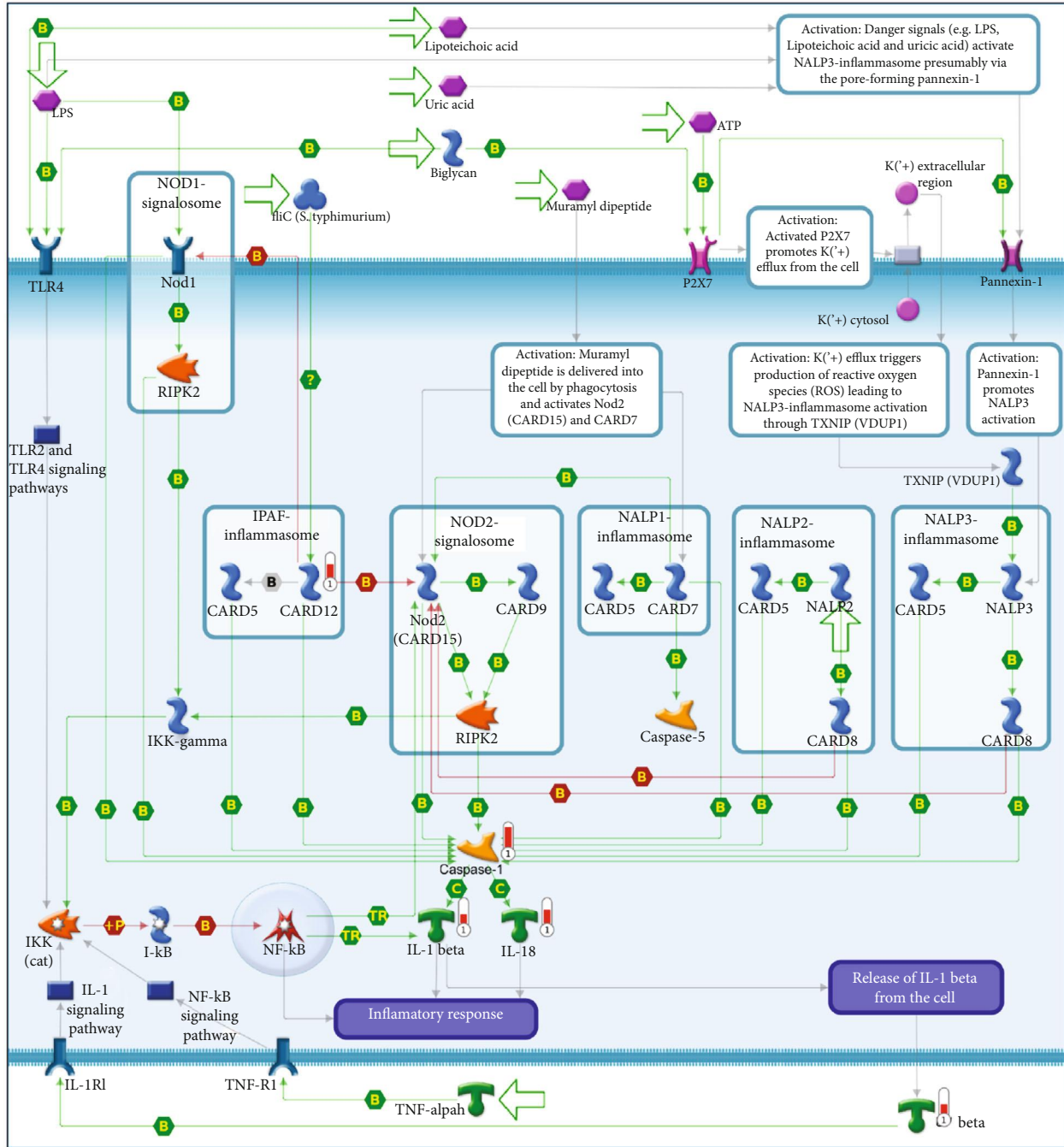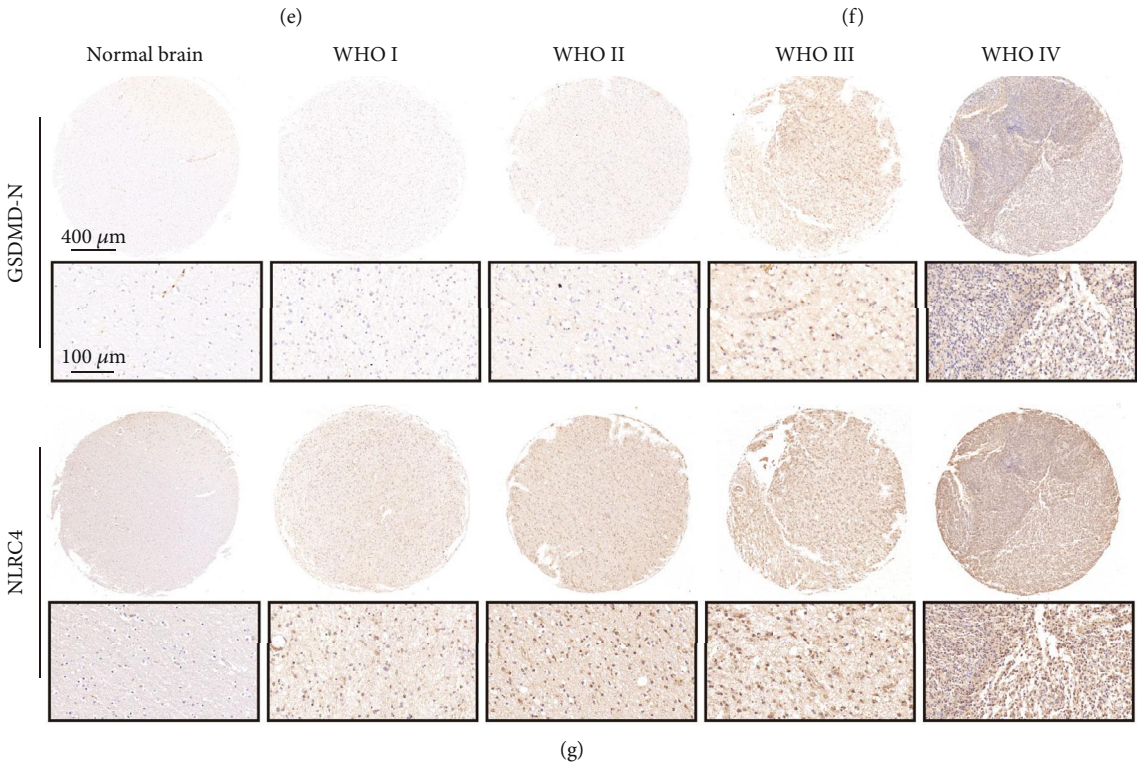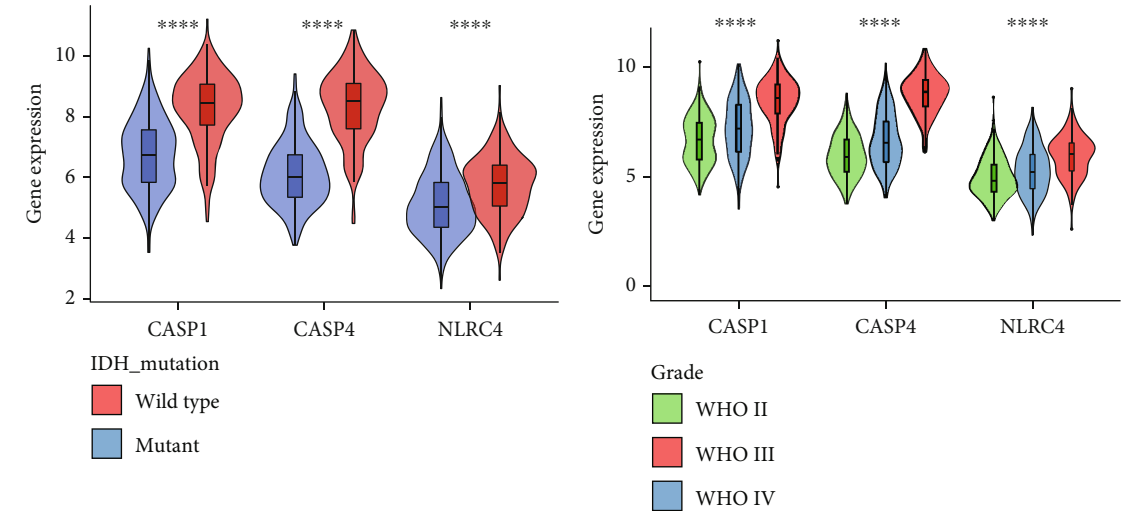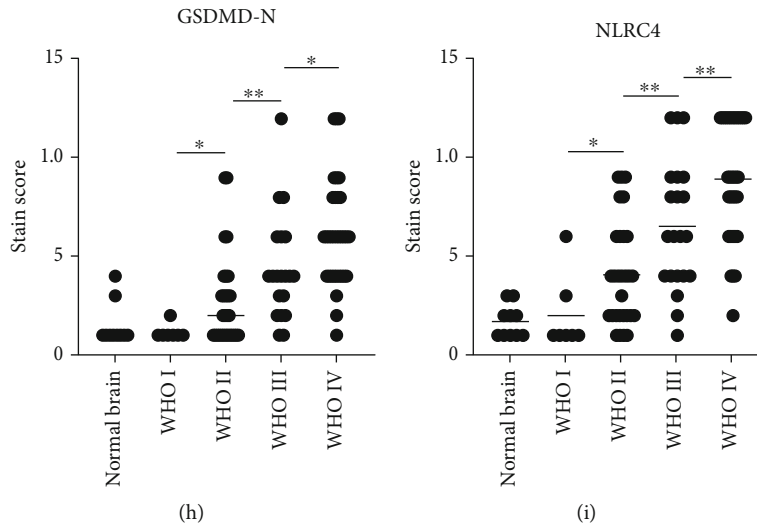
FIGURE 3: NLRC4/CASP1/CASP4/GSDMD signaling axis genes in TCGA cohort have similar expression patterns and jointly affect overall survival. (a) Kaplan–Meier plots for the OS of patients with different CASP gene expression in TCGA cohort, using the median expression of each gene as a cutoff point to divide the high- and low-expression groups. (b) Scatter plot of the correlation between the GSDMD gene and the expression of three caspase genes in TCGA cohort. The degree of correlation was examined using Spearman's coefficient. (c) Univariate Cox analysis of the effects of GSDMD, CASP4, and CASP1 and their upstream inflammasome gene expression on overall survival. Genes with $p < 0.05$ were selected for multivariate Cox analysis. Hazard ratios are presented as forest plots. (d) Pathway enrichment of DEGs between the GSDMD high- and low-expression groups using Metacore and visualization of CASP1 upstream gene hits. The red thermometer indicates the Log2FC of different genes. (e, f) Expression levels of CASP1, CASP4, and NLRC4 genes in patients with different WHO grades and IDH mutation phenotypes in TCGA cohort. (g) Representative sections of matched GSDMD-N and NLRC4 immunohistochemistry from normal brain tissue and different grades of glioma samples. (h, i) Semiquantitative results of GSDMD-N and NLRC4 staining levels in tissue microarrays. Statistics were calculated using one-way analysis of variance (ANOVA) in (f). Statistics were calculated using two-tailed, unpaired Student's $t$-test with Welch's correction in (e, h, i). $^{****}p < 0.0001$, $^{**}p < 0.01$, and $^{*}p < 0.05$.

pyroptosis [41]. Therefore, to explore the upstream activation signals of CASP genes, we selected representative NLRP1, NLRP2, NLRP3, NLRC4, and NOD2 inflammasome genes to examine their correlation with CASP1/CASP4/GSDMD. Univariate Cox regression analysis revealed that only NLRC4 of the CASP1 upstream inflammasome genes had a significant effect on overall survival (HR = 2.594, $p < 0.0001$), and multivariate Cox regression analysis demonstrated that NLRC4/CASP1/CASP4/GSDMD could jointly affect overall survival in TCGA cohort (Figure 3(c)). Among inflammasome genes, NLRC4 demonstrated its unique prognostic value in the CGGA cohort and Gravendeel dataset (Figure S3). Furthermore, we screened differentially expressed genes (DEGs) between the high and low GSDMD expression groups using the limma R package, performed pathway enrichment of DEGs, and analyzed GSDMD upstream gene hits using the network building tool MetaCore. The visualization results again demonstrated the unique association of the NLRC4 (CARD12 in the map) gene with GSDMD among inflammasome genes (Figure 3(d)). In addition, the expression of the CASP1, CASP4, and NLRC4 genes in TCGA cohort increased with disease progression and showed higher expression in the wild-type IDH group (Figures 3(e) and 3(f)), and the NLRC4/CASP1/CASP4/GSDMD coexpression relationship was verified in TCGA cohort and three other independent datasets (Figure S4). To further confirm pyroptosis in gliomas, we examined the protein levels of N-terminal GSDMD (GSDMD-N), which is

produced by cleavage of full-length GSDMD by caspase-1 and caspase-4 and is the most classic marker of pyroptosis, in gliomas of different disease grades and normal brain tissue. We also examined the protein levels of NLRC4. The results showed that GSDMD-N and NLRC4 were barely detectable in the normal brain tissue, while the levels of both GSDMD-N and NLRC4 increased in matched samples with increasing disease grade (Figure 3(g)). Semiquantitative analysis also showed significantly higher levels of GSDMD-N and NLRC4 in higher-grade glioma samples (Figures 3(h) and 3(i)), which provided conclusive evidence for GSDMD-mediated pyroptosis in gliomas.

To investigate in depth whether this signaling axis can be used as a valid guide for predicting patient prognosis, we constructed a scoring system based on the NLRC4/CASP1/CASP/GSDMD pyroptosis axis. We extracted the 4 genes with significant coefficients in the multivariate Cox analysis of TCGA dataset, used these data as the training set, and finally obtained the risk score formula: Riskscore = $1.02065 \times \text{expr}_{CASP4} + 0.46802 \times \text{expr}_{GSDMD} + 0.31184 \times \text{expr}_{CASP1} + 0.09562 \times \text{expr}_{NLRC4}$. The Riskscore for each patient was calculated, and the patients were divided into the high- and low-risk groups according to the median Riskscore (Figures 4(a) and 4(b)). There were significant differences in the histological classification, WHO grade, IDH mutation phenotype, and chromosome 1p19q codeletion phenotype between the two groups, while chemotherapy and temozolomide acceptance

(a)



(b)

FIGURE 4: Continued.

(c)



(d)



(e)
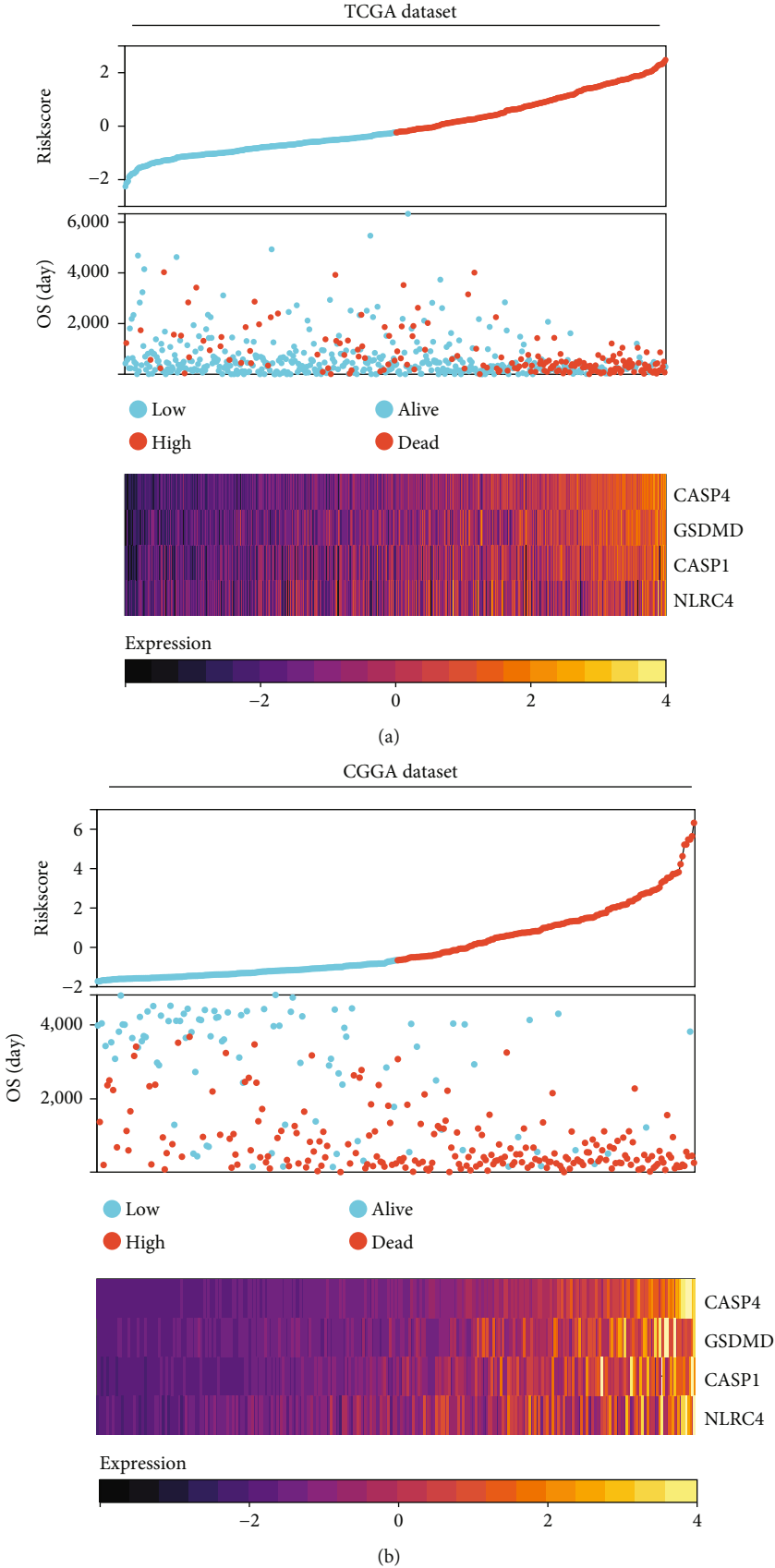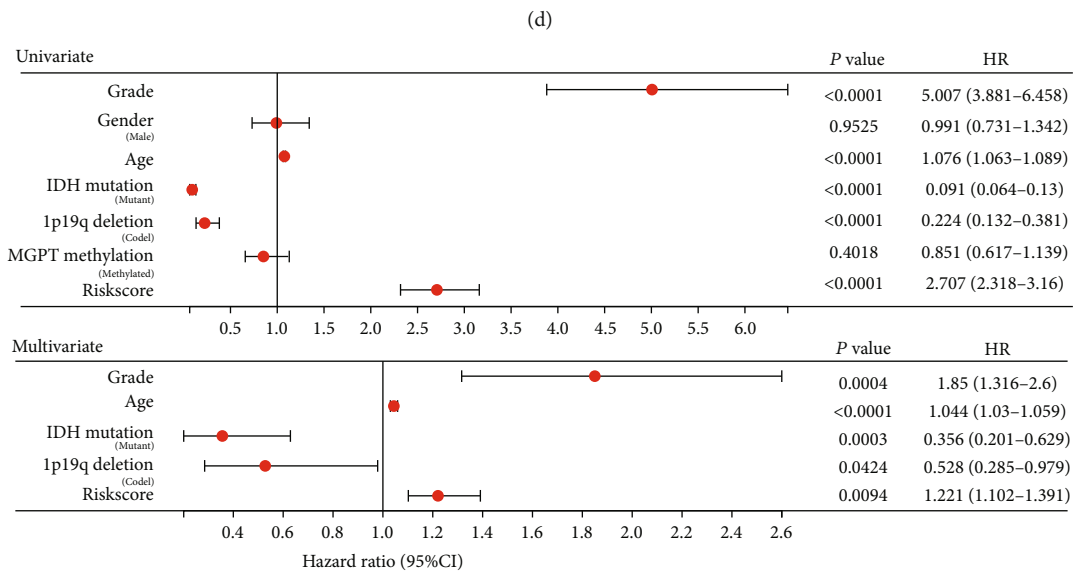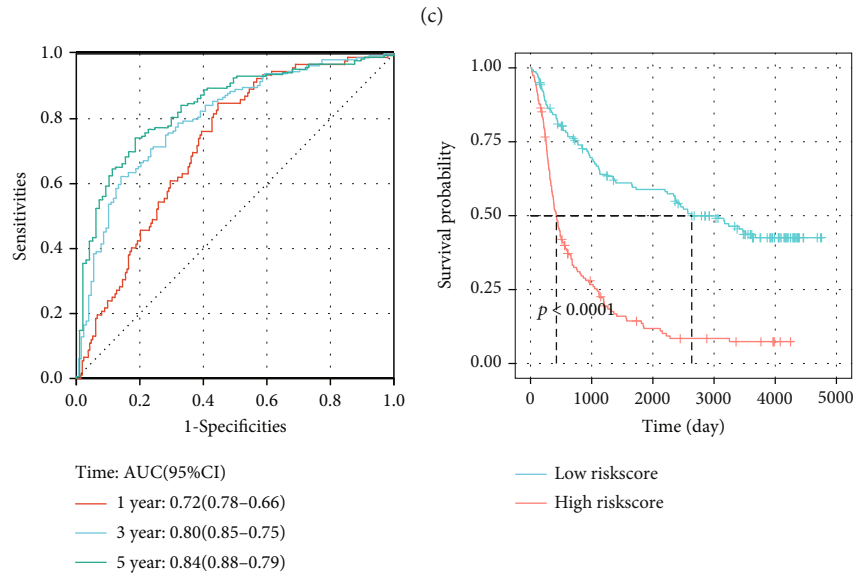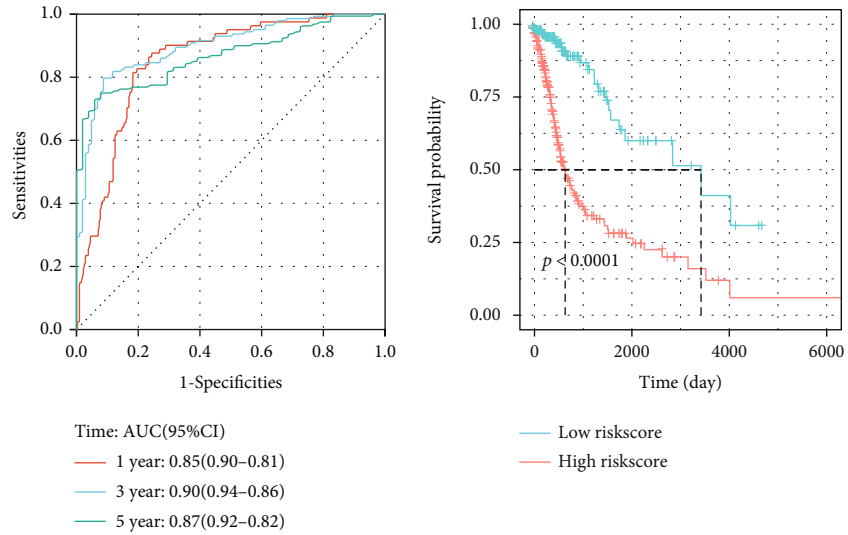
FIGURE 4: Continued.

(f)



(g)



(h)

FIGURE 4: Continued.

(i)



(j)
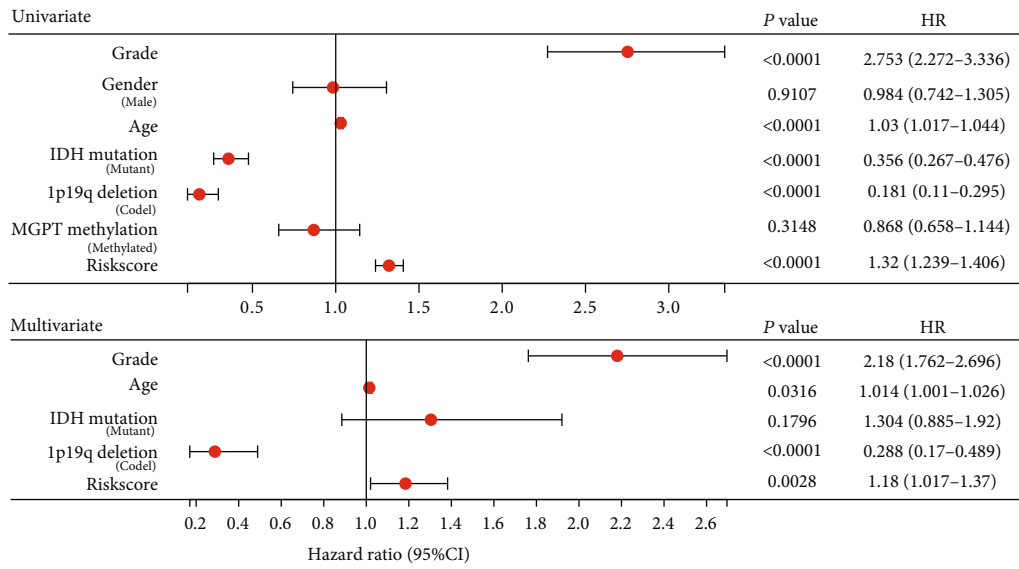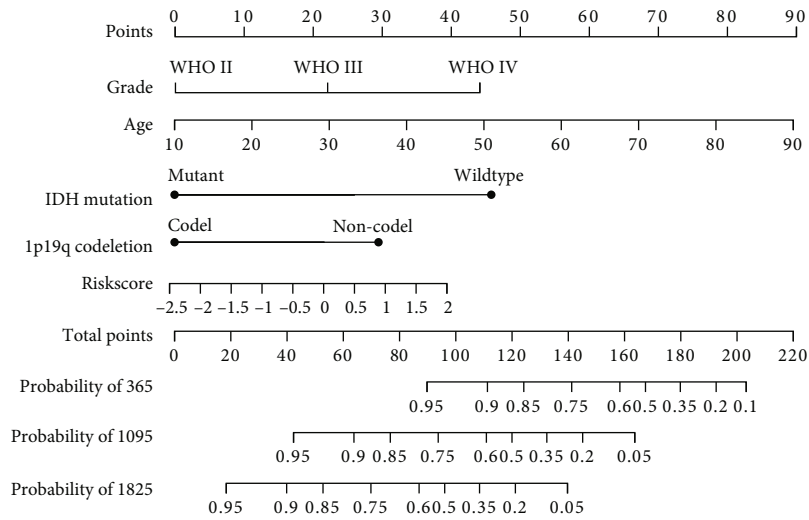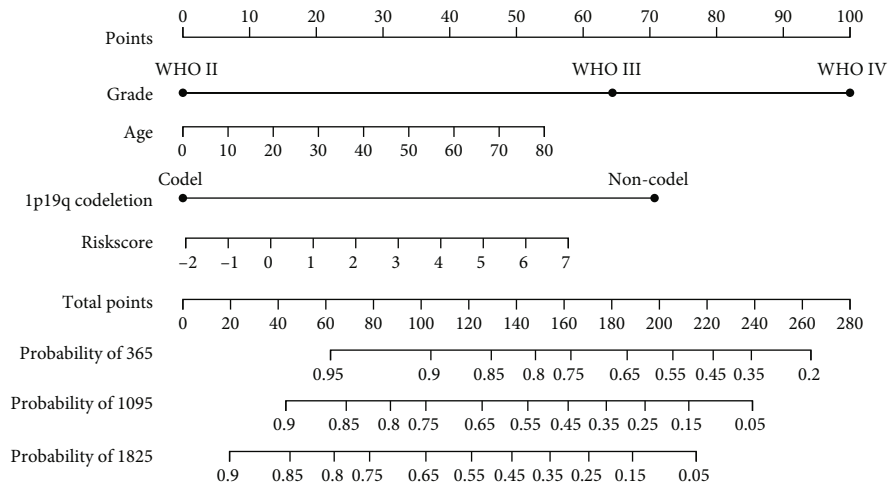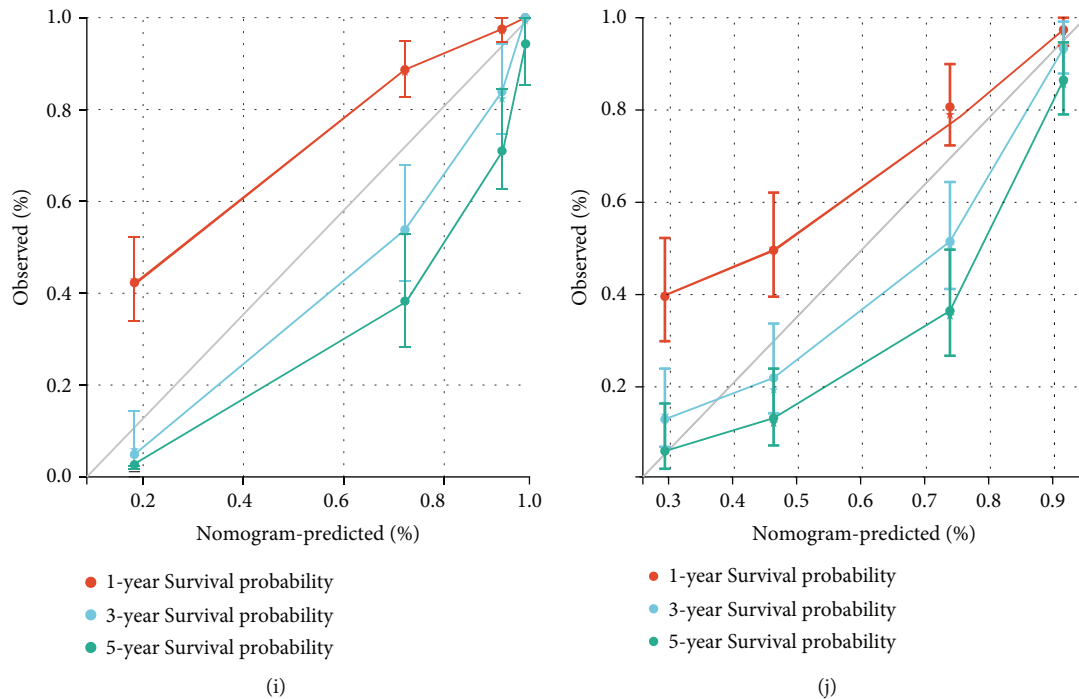
FIGURE 4: Stable prognostic model based on NLRC4/CASP1/CASP4/GSDMD pyroptosis signaling axis genes. (a, b) Relationship between risk score and overall survival of patients and expression levels of NLRC4, CASP1, CASP4, and GSDMD genes in TCGA training cohort and CGGA validation cohort. (c, d) Time-dependent ROC analysis and Kaplan–Meier analysis in TCGA training cohort and CGGA validation cohort to assess the prognostic value of the Riskscore, using the median Riskscore as a cutoff point to divide the high- and low-risk groups. (e, f) Univariate Cox analysis and multivariate Cox analysis in TCGA training cohort and CGGA validation cohort. Hazard ratios are presented as forest plots. (g, h) The nomogram for predicting the proportion of patients with 1-, 3-, and 5-year overall survival in TCGA and CGGA cohorts. (i, j) The calibration curves for the prediction of 1-, 3-, and 5-year overall survival in TCGA cohort and CGGA cohort.

differences were not significant (Table. 1). Time-dependent ROC and Kaplan–Meier curves were used to assess the prognostic ability of the four pyroptosis-associated genetic signatures. The results showed that the high-pyroptosis-risk group defined by the four signature genes had significantly shorter OS in TCGA training cohort and the CGGA external validation cohort (Figures 4(c) and 4(d)). The AUC (area under the ROC curve) was 0.85, 0.90, and 0.87 for the 1-year, 3-year, and 5-year OS in the training cohort and 0.72, 0.80, and 0.84 in the CGGA cohort, respectively. The ROC curves showed a similar prognostic value of our established prognostic model and the previously established 10-pyroptosis-gene prognostic model [22] and golden standard WHO grading system for predicting OS at 1, 3, and 5 years in TCGA and CGGA cohorts (Figure S5). Moreover, univariate and multifactorial Cox regression analyses revealed that the Riskscore could be used as a valid independent prognostic factor, as well as disease grade, age, IDH mutation status, and 1p19q codeletion status (Figures 4(e) and 4(f)). Nomograms based on the results of multivariate Cox regression analysis were used for scoring to assess the accuracy of the model. To correctly predict the 1-, 3-, and 5-year OS, we created a nomogram4 that included the WHO grade, age, IDH mutation status, 1p19q codeletion status, and the Riskscore (Figures 4(g) and 4(h)). The calibration curve study revealed agreement between the

patients' anticipated and observed 1-, 3-, and 5-year OS rates in both TCGA and CGGA cohorts (Figures 4(i) and 4(j)).

### 3.3. Differential Gene, Tumor Mutational Burden, and Drug Prediction Analysis Based on the Four-Pyroptosis-Gene Prognostic Model.

The distinct prognosis of the high- and low-risk groups defined by the four pyroptosis genes drove us to further explore the functional enrichment of the differential genes between the high- and low-risk groups and thus speculate on the potential mechanisms of pyroptosis involved in glioma disease progression. We analyzed differentially expressed genes between the high- and low-risk groups in TCGA dataset using the Limma R package. We screened DEGs with |log 2-fold change| >1.5 and adjusted $p$ < 0.05 and obtained a total of 498 upregulated genes and 518 downregulated genes (Figures 5(a) and 5(b)). Principal component analysis (PCA) showed that the high-risk group distinctly clustered apart from the low-risk group, revealing significant differences in expression profiles between the two groups (Figure 5(c)).

The GO enrichment results showed that upregulated DEGs were mainly involved in various immune response-related biological processes, such as cellular immune response, cellular defense response, and response to cytokines (Figure 5(d)). The KEGG enrichment results indicated that upregulated DEGs were mainly associated with
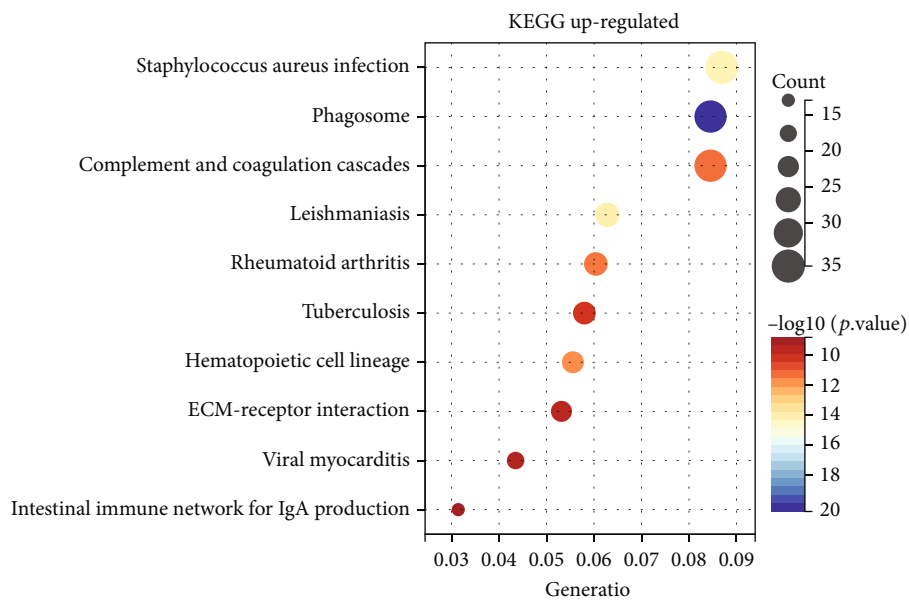
(a)



(b)



(c)

FIGURE 5: Continued.

(d)



(e)

Figure 5: Continued.
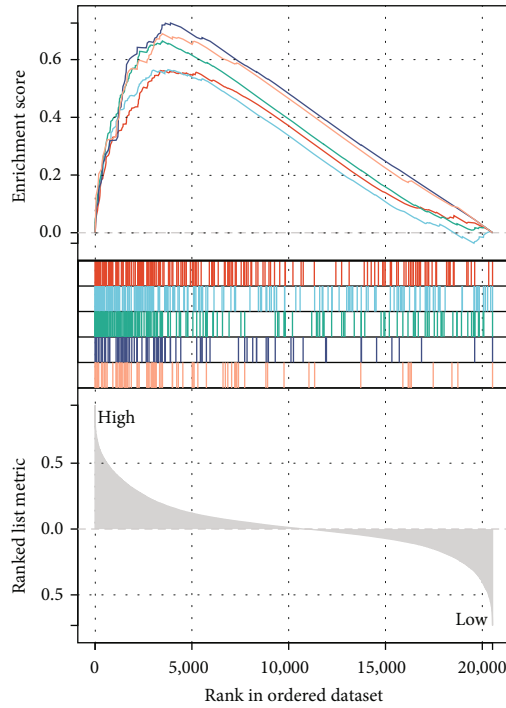
(f)
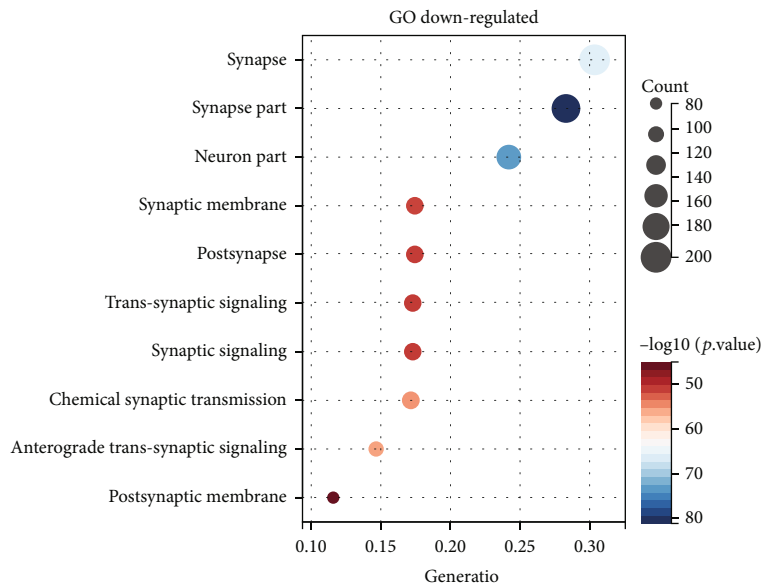


(g)

FIGURE 5: Upregulated DEGs were associated with inflammatory responses, and downregulated DEGs were associated with synaptic stabilization. (a) Heatmap of DEG expression profiles in TCGA cohort with clinical phenotypes, including the disease grade, IDH phenotype, Riskscore, and risk group. (b) Volcano map of DEGs between the two risk groups in TCGA cohort. (c) Principal component analysis (PCA) of all patients in TCGA cohort based on all gene expression profiles between risk groups. (d, e) Bubble plots of GO and KEGG enrichment analysis for upregulated DEGs. (f) HALLMARK terms were used for GSEA enrichment analysis based on the Riskscore of all patients in TCGA cohort, with the enriched terms represented by different color curves. (g) Bubble plot of GO enrichment analysis for downregulated DEGs.

inflammatory response signals, such as multiple bacterial and viral infections and autoimmune diseases such as rheumatoid arthritis (Figure 5(e)). Gene set enrichment analysis (GSEA) also showed the activation of various proinflammatory signaling pathways in the high-risk group, including IL2-STAT5, IL6-STAT3, and IFN-$\alpha$ responses (Figure 5(f)). Pyroptosis is a proinflammatory cell death mode in which large amounts of inflammatory substances are released during cell death and trigger an inflammatory response, which coincided with the activation of multiple aberrant immune response pathways present in the high-pyroptosis-risk group. In addition, the top ten GO terms enriched by downregulated DEGs all had a strong link with synapse formation, stabilization, and signal transduction (Figure 5(g)). Clinically, patients with high-grade glioma tend to develop degenerative diseases such as memory loss and cognitive impairment, and it has been demonstrated that impaired cognitive function is associated with shorter survival in glioblastoma patients [42]. Furthermore, pyroptosis is strongly associated with Alzheimer's disease progression, and GSDMD serves as an important marker for AD [19]. This is consistent with our analysis that there is an association between high-grade gliomas, corresponding to a high risk of pyroptosis, and neurodegenerative diseases. Taken together, these results indicate that a high degree of CASP1/CASP4/NLRC4/GSDMD pyroptosis is accompanied by the activation of proinflammatory signaling pathways in the brain and is closely associated with impaired establishment and stability of neuronal synapses.

We also explored differences in copy number variation (CNV) and tumor mutational burden (TMB) between the high-risk and low-risk groups. A significantly higher proportion of samples in the high-risk group had CNV. We screened for the genes that differed most significantly between the high-risk and low-risk groups, including high-level amplified genes and homozygous deletion genes. Interestingly, several interferon alpha (IFNA) family genes in the high-risk group were homozygously deleted (Figure S6a), and the activation of the inflammasome was previously reported to have an antagonistic effect on the type I interferon response in macrophages [43]. However, a correlation of other genes with pyroptosis could not be identified in previous studies. In addition, the CNV in the NLRC4, CASP1, CASP4, and GSDMD genes did not vary significantly between the high- and low-risk groups (Figure S6b). The low-risk group had an IDH1 mutation rate of 86.95%, and the majority of these samples also had mutations in the ATRX and CIC genes, which are characteristic of LGGs, such as oligodendrogliomas [44] (Figure S6c). In contrast, more EGFR, TTN, and PTEN mutations, which are usually characteristic of GBM [45], were observed in the high-risk group (Figure S6d).

We also performed a preliminary drug sensitivity analysis. The drug sensitivity data and expression profile data for glioma cell lines were obtained from Genomics of Drug Sensitivity in Cancer (GDSC) and the Cancer Therapeutics Response Portal (CTRP). Multiple drug candidates were screened by correlation analysis of the expression levels of the four pyroptosis genes of the cell lines with the IC50 of different drug treatments in the GDSC database (Figure S7a, b) and the

drug sensitivity (1-(AUC/30)) in the CTRP database (Figure S7c, d). However, the sensitivity to each of these drug candidates can only be correlated with the expression of one of the pyroptosis genes, so the combination is more appropriate for this pyroptosis target.

### 3.4. Increased Infiltration of BMDMs and the Immunosuppressive Microenvironment in the High-Risk Group.

We demonstrated that the high-risk group was associated with multiple inflammatory response signaling pathways (Figures 5(e)–5(g)), and we speculated that this may be associated with the altered infiltration of immune cells caused by the NLRC4/CASP1/CASP4/GSDMD pyroptosis axis. The ESTIMATE R package was used to predict the stromal score, immune score, and ESTIMATE score (stromal score + immune score), and the results showed that the high-risk group had a higher immune score and stromal score, which represented a higher degree of immune infiltration and tumor malignancy (Figure 6(a)). CIBERSORT was used to predict immune cell infiltration in TCGA cohort, and the most abundant immune cells in gliomas were M2 macrophages, which were further increased in the high-risk group. (Figure 6(b)). Previous studies have shown that macrophages in gliomas, especially those with the M2 phenotype, play an important role in the formation of the immunosuppressive microenvironment and tumor progression [46, 47]. Since brain macrophages can be divided into bone marrow-derived macrophages (BMDMs) and tissue-resident microglia and function differently, we used the DEGs between microglia and BMDMs demonstrated by Bowman et al. (Supplementary Table. S2) [34] as gene sets and assessed the microglia and BMDM infiltration in each sample by ssGSEA (single sample GSEA). BMDM infiltration differed remarkably between the high- and low-risk groups, while microglia did not change significantly (Figure 6(c)). Notably, the BMDM infiltration score had a significant positive correlation with M2 macrophage infiltration ($R = 0.56, p < 0.0001$) (Figure 6(d)) and the Riskscore ($R = 0.69, p < 0.0001$) (Figure 6(e)), which could not be observed in microglia. In addition, high infiltrations of M2 macrophages and BMDMs were strongly associated with poor prognosis, whereas microglia were of opposite and limited predictive value ($p = 0.029$) (Figure 6(f)). To verify the robustness of the association of high BMDM infiltration with poor prognosis and the positive correlation between BMDM infiltration and the Riskscore, we further used the microglia and BMDM differentially expressed genes from the study of Muller et al. (Supplementary Table. S3) to predict microglial and BMDM infiltration scores [35]. The results once again demonstrated the excellent prognostic value of BMDM infiltration, but not microglial infiltration (Figure S8b), and the robust positive correlation between BMDM infiltration and the Riskscore ($R = 0.75, p < 0.0001$) or M2 macrophage infiltration ($R = 0.62, p < 0.0001$) (Figure S8c, d).

Dysfunctional CD8+ T cell infiltration predicted by Tumor Immune Dysfunction and Exclusion (TIDE) was higher in the high-risk group (Figure 6(g)). ssGSEA based on the immune checkpoint gene set and the TGF-$\beta$ response score (TBRS) gene set associated with the anti-PD1

(a)



(b)

FIGURE 6: Continued.

(c)

(d)

(e)



(f)



(g)                                                        (h)                                                        (i)
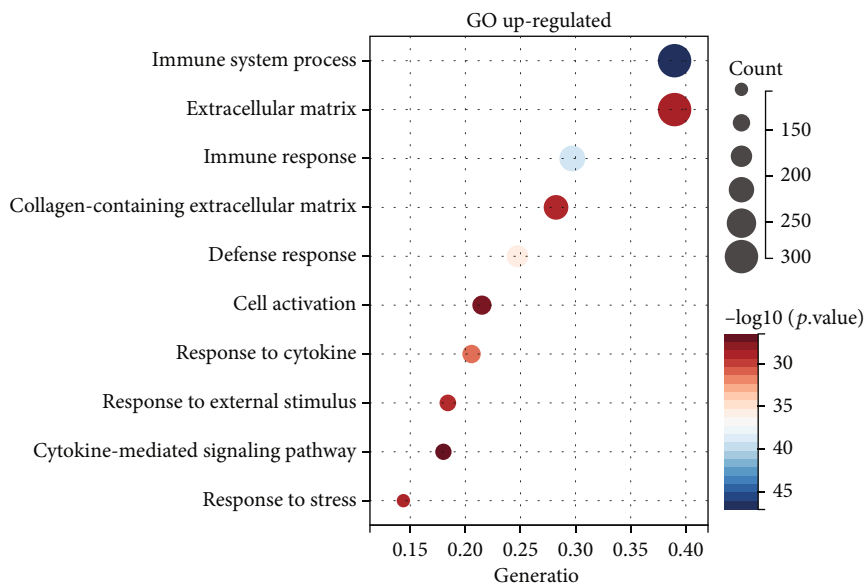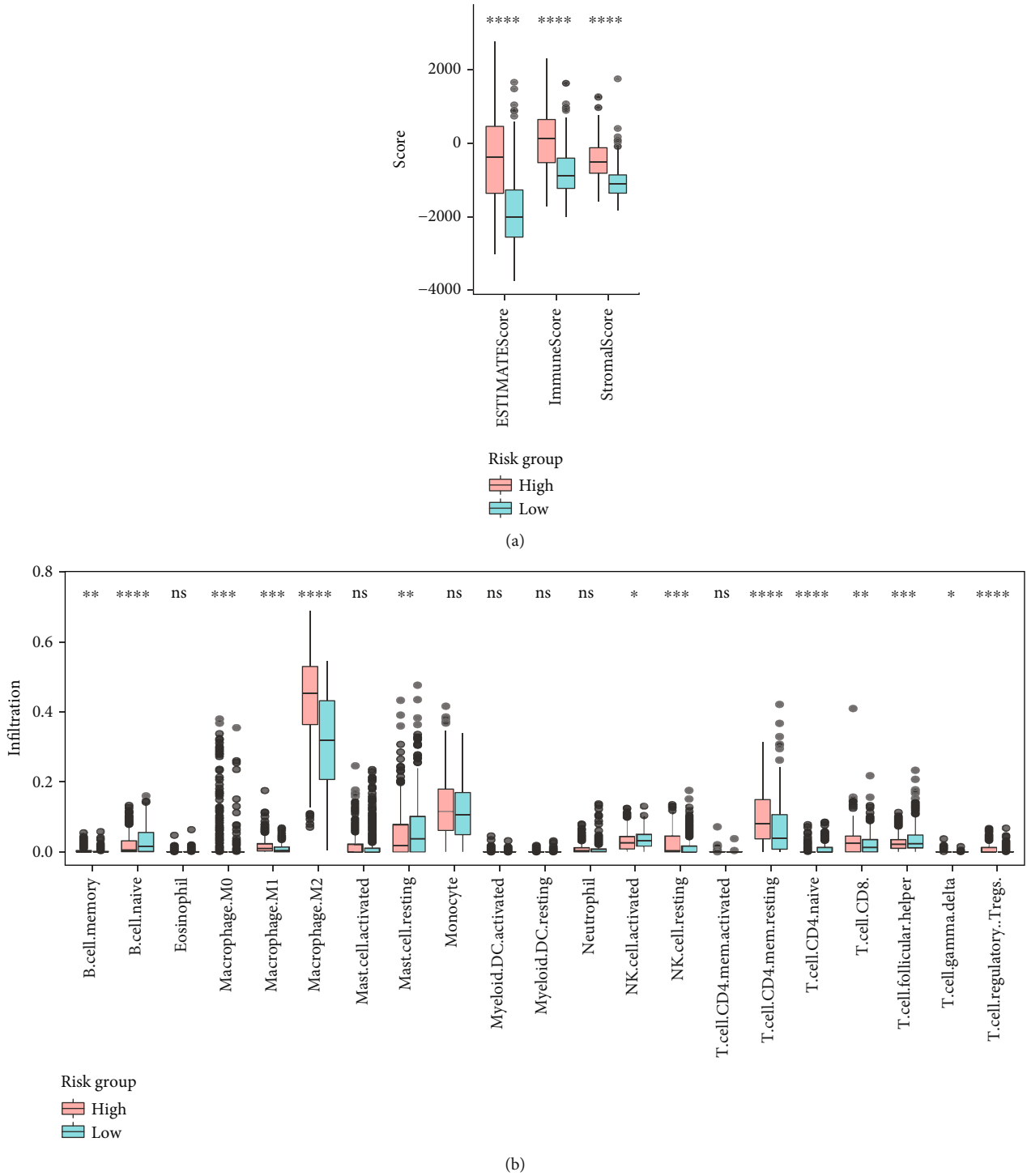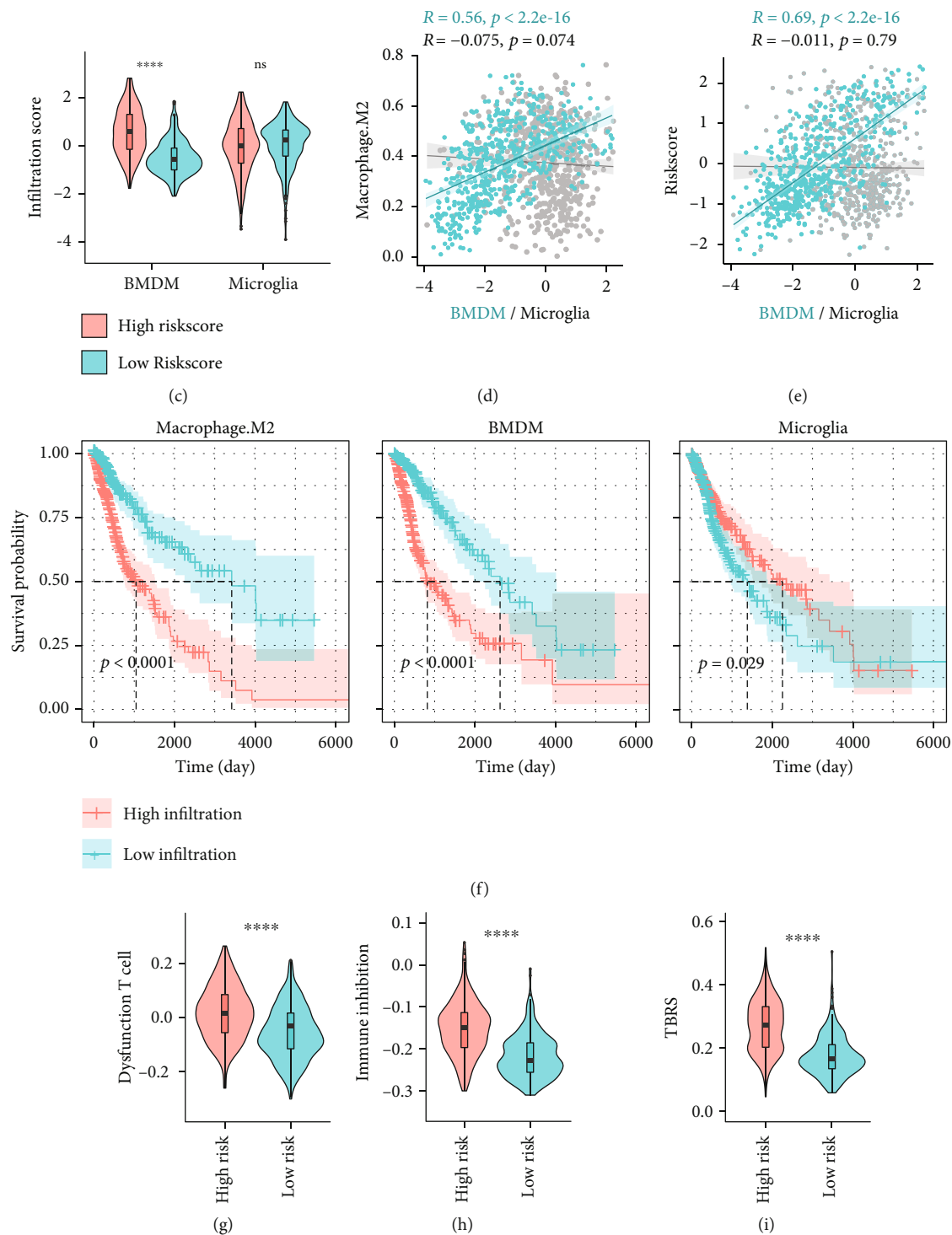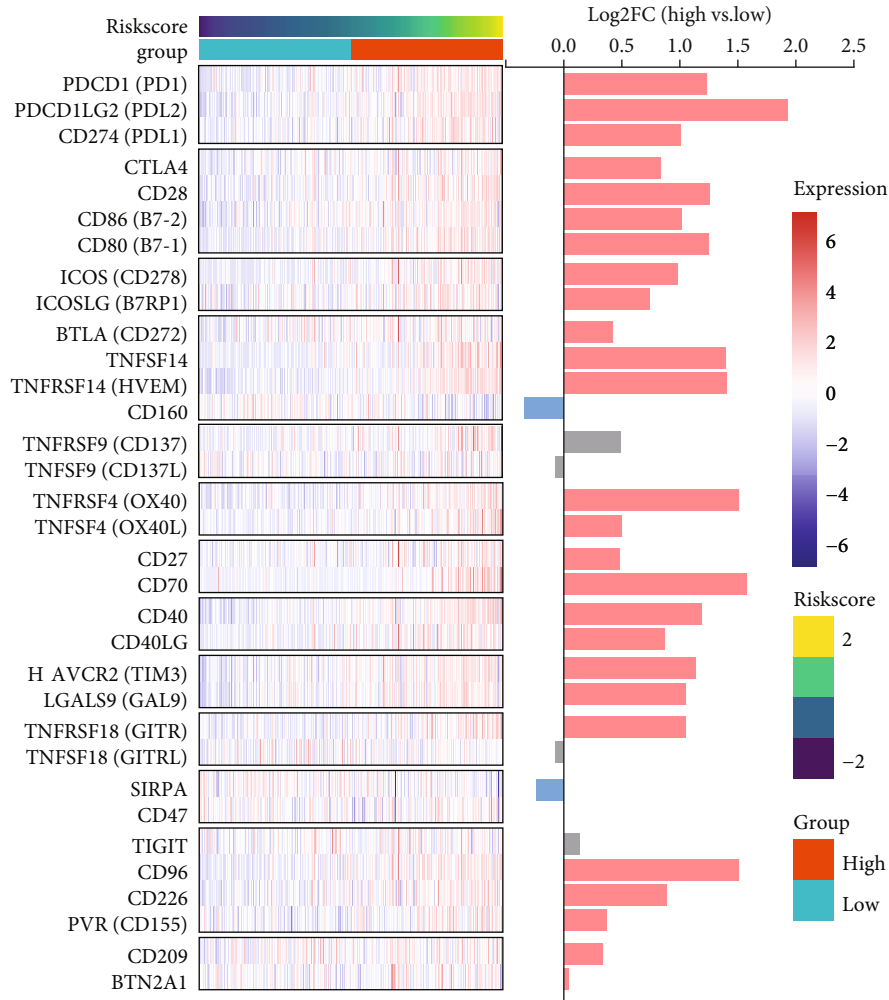
Figure 6: Continued.

(j)
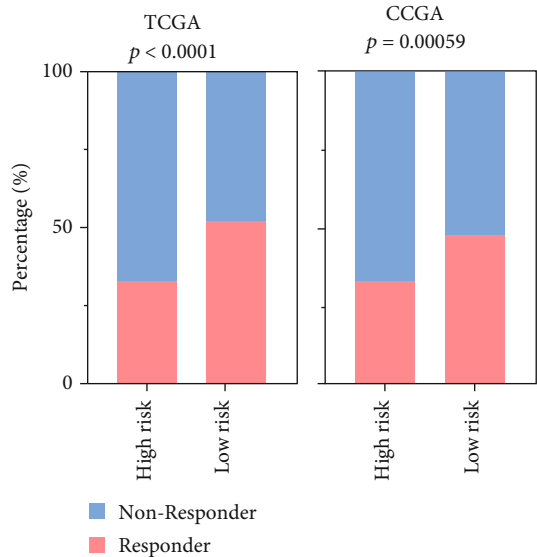


(k)

Figure 6: Continued.

(l)



(m)
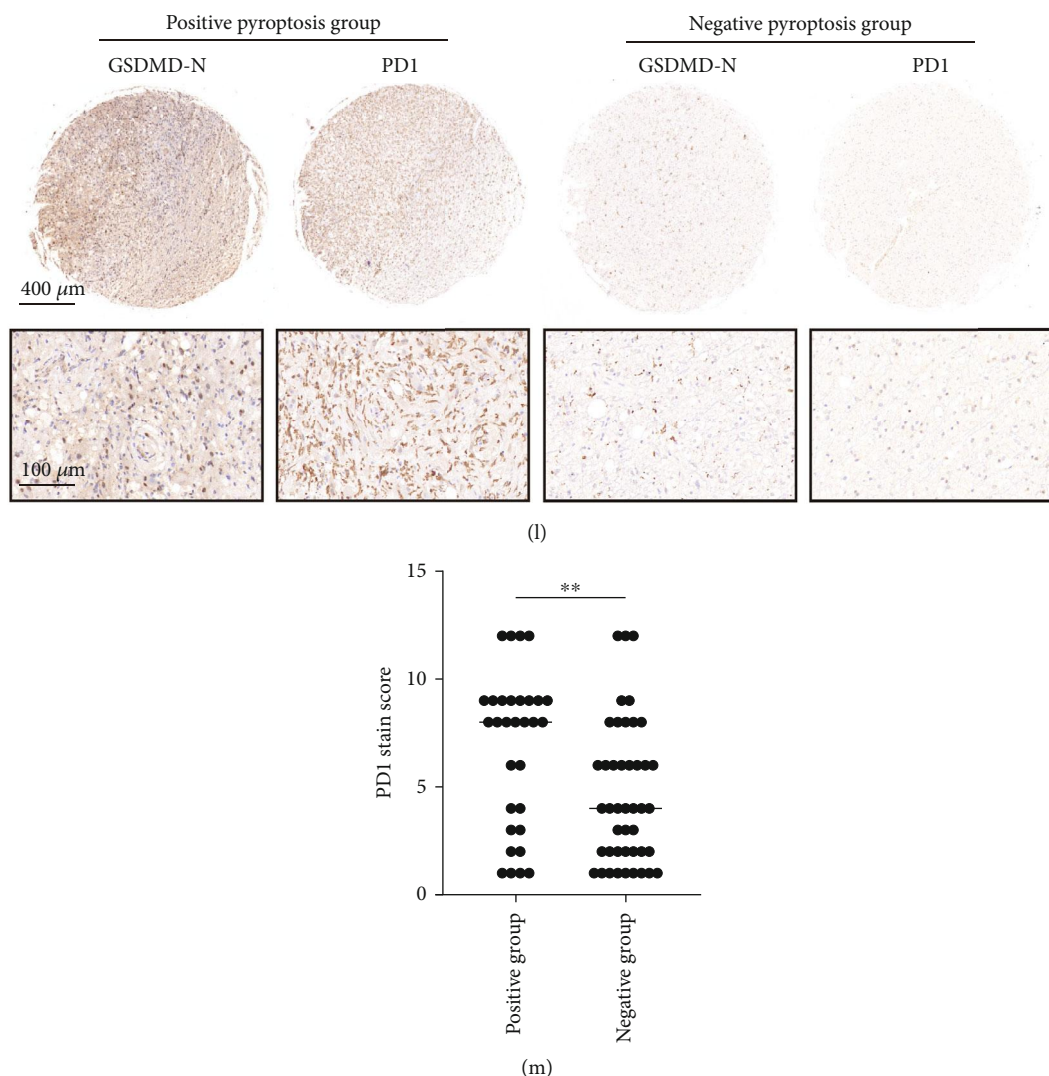
FIGURE 6: The Riskscore, defined by the pyroptosis axis, was associated with the immunosuppressive tumor microenvironment. (a) Stromal score, immune score, and ESTIMATE score of TCGA cohort predicted based on ESTIMATE R package. (b) Immune cell infiltration in TCGA cohort predicted by CIBERSORT. (c) ssGSEA based on macrophage and microglial DEGs to calculate the BMDM infiltration score and microglial infiltration score for each sample. (d) Scatter plots of the correlation between the infiltration score of BMDMs or microglia and the infiltration of M2-type macrophages. Spearman's coefficient was used to evaluate the degree of correlation. (e) Scatter plots of the correlation between the infiltration fraction of BMDMs or microglia and the Riskscore. Spearman's coefficient was used to evaluate the degree of correlation. (f) Kaplan–Meier curves for the correlation between M2 macrophage, BMDM, and microglial infiltration and overall survival time in TCGA cohort, using the median infiltration score as a cutoff point to divide the high- and low-infiltration groups. (g) Dysfunctional CD8+ T cell infiltration score predicted by Tumor Immune Dysfunction and Exclusion (TIDE) in TCGA cohort. (h, i) Immunosuppression scores (h) and TGF-$\beta$ response score (TBRS) (i) calculated for each sample based on ssGSEA with different characteristic gene sets. (j) Heatmap of the expression of immune checkpoint genes in TCGA cohort, aligned by immune checkpoint gene pairing and displaying the Log2FC (high-risk vs. low-risk) of the corresponding gene expression on the $y$-axis. The results for $p > 0.05$ are shown in gray. (k) The ICB responses of TCGA (left) and CGGA cohorts (right) based on the TIDE prediction results are presented in the stacked histograms. (l) Representative sections of matched GSDMD-N and PD1 immunohistochemistry from high pyroptosis group (GSDMD-N score > 5) and low pyroptosis group (GSDMD-N score < 5). (m) Semiquantitative results of PD1 staining levels in tissue microarray. Statistics were calculated using two-tailed, unpaired Student's $t$-test with Welch's correction in (a–c) and (g–i). Statistics were calculated using the Chi-squared test in (k). ns: not significant. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$.

treatment response identified by Sanjeev et al. (Supplementary Table. S1) [32] was used to evaluate the immunosuppression score (Figure 6(h)) and TBRS (Figure 6(i)) in each sample, which were also higher in the high-risk group, representing a hypersuppressed immune microenvironment

and disappointing anti-PD1 treatment response rate. Due to the lack of open access to immune checkpoint blockade (ICB) glioma therapy cohorts, we used TIDE to predict the response to ICB therapy in TCGA and CGGA cohorts. Response rates were significantly lower in the high-risk

group than in the low-risk group in both cohorts (Figure 6(k)), consistent with previous results of a higher TBRS in the high-risk group. In addition, the landscape analysis of immune checkpoint receptor and ligand genes demonstrated a positive correlation between gene expression and the Riskscore for the majority of immune checkpoints (Figure 6(j)).

In addition, we divided the samples into the positive pyroptosis group (GSDMD-N score > 5) and the negative pyroptosis group (GSDMD-N score < 5) based on the staining results of GSDMD-N of tissue microarray and compared the PD1 immunohistochemical staining levels in the two groups. The results showed that the PD1 level in the positive pyroptosis group was significantly higher than that in the negative pyroptosis group (Figures 6(l) and 6(m)), indicating that pyroptosis in the microenvironment of glioma was accompanied by deepening immunosuppression. Therefore, it is hypothesized that inhibition of pyroptosis in glioma may facilitate the alleviation of the immunosuppressive microenvironment.

Although the Riskscore was previously demonstrated to be associated with proinflammatory signals (Figure 5), it is not contradictory to mediating the establishment of a suppressive immune microenvironment. Pyroptosis is a type of proinflammatory cell death, and prolonged infiltration of IL1$\beta$, IL2, and IL6 in the inflammatory environment induces Tregs [48] and promotes the invasion and proliferation of glioma stem cells (GSCs) [49]. Our study showed that the activation of the NLRC4/CASP1/CASP4/GSDMD pyroptosis axis was significantly and positively correlated with M2-type BMDM infiltration, suggesting that blocking pyroptosis in glioma may be a potential approach to reduce macrophage infiltration. Pyroptosis inhibitors have potential as adjuvant therapeutic agents in high-grade glioma, such as dimethyl fumarate, which has been approved by the FDA to reduce macrophage infiltration by inhibiting pyroptosis to achieve efficacy in the treatment of multiple sclerosis [20] and has been demonstrated to cross the blood–brain barrier [50].

### 3.5. The Activation of the NLRC4/CASP1/CASP4/GSDMD Signaling Axis Is Mainly Present in Peripheral BMDMs.

To elucidate the potential link between macrophage infiltration and the NLRC4/CASP1/CASP4/GSDMD pyroptosis signaling axis, we further explored the source cells of pyroptosis in gliomas at single-cell resolution. The scRNA-seq expression profiles and cell annotation files of Neftel et al. [26] were obtained from the Single Cell Portal database, containing a total of 7930 cells from 28 patients (Figure 7(a)). After data cleaning and cell type clustering, cells were classified into malignant cells, macrophages, T cells, oligodendrocytes, and astrocytes (Figures 7(b) and 7(c)). ssGSEA predicted the BMDM score and microglia score in the same way as presented previously and was used to differentiate between peripheral-derived BMDMs and tissue-resident microglia (Figure S9).

The coexpression of CASP and GSDMD is required for the occurrence of pyroptosis, while CASP4, GSDMD and NLRC4, CASP1, and GSDMD colocalization signals were located in the macrophage/microglia population (Figure 7(d)) and were significantly concentrated in the BMDM cluster (Figure 7(e)). The gene expression levels of NLRC4, CASP1, CASP4, and GSDMD were all higher in BMDMs than in microglia (Figure 7(f)), suggesting that BMDMs are more sensitive to pyroptosis or more prone to pyroptosis than other cells. The BMDM cluster was further divided into five subpopulations from cluster_0 to cluster_4, while CASP4+ GSDMD+ and NLRC4+ CASP1+ GSDMD+ cells were clearly concentrated in the cluster_0 cell cluster (Figure 7(g)). With Log2FC > 1.5 and adjusted $p$ < 0.05, 118 signature genes of the cluster_0 BMDM subpopulation were screened, including various chemokine genes, such as CCL3, CCL4, and CXCL12, and the proinflammatory cytokine gene IL1B (Figure 7(h)). There is no doubt that GO enrichment analysis hits BP terms of multiple cell chemotaxis and migration-related pathways (Figure 7(i)). Therefore, the positive cycle that accompanies the release of inflammatory molecules such as DAMPs and multiple chemokines during macrophage pyroptosis leads to more infiltration of peripheral monocytes and macrophages, which may trigger a worse prognosis. To use big data and survival information to aid in scRNA-seq data analysis, we used the Scissor R package to assess the relevance of single cells to patient overall survival. Briefly, using TCGA-LGGGBM dataset and the Cyril Neftel scRNA-seq dataset, Scissor was applied to distinguish cells with high expression of survival-related genes, where scissor-positive (Scissor+) cells were those associated with poor survival and scissor-negative (Scissor-) cells were those associated with good prognosis. As expected, the macrophage population was filled with a large number of cells associated with poor survival, and the NLRC4+ CASP1+ CASP4+ GSDMD+ BMDM concentrated cluster had a very high proportion of cells associated with a poor prognosis (71.33%) (Figure 7(j)).

Other glioma scRNA-seq datasets were used to validate our above findings. First, we obtained the dataset from the CGGA database from Yu et al., including 6148 cells from 13 patients [27]. The cells were classified into 6 cell types, including BMDMs and microglia, by downscaling analysis and cell-type identification (Figures 8(a) and 8(b)). BMDMs had significantly higher expression levels of GSDMD, CASP1, CASP4, and NLRC4 than microglia (Figure 8(c)). In addition, we calculated the PyropScore for each cell using ssGSEA based on the expression levels of the four genes. The PyropScore of BMDMs was significantly higher than that of microglia (Figure 8(d)). The NLRC4+ CASP1+ GSDMD+ and CASP4+ GSDMD+ cell populations were also significantly enriched in the BMDM cell cluster (Figure 8(e)). Consistent results were obtained based on the analysis of another independent dataset, GSE89567, which contained 6341 cells from 10 patients [28] (Figures 8(f)–8(j)).

### 3.6. The Activation of Pathways of Response to LPS/Bacteria and Oxidative Stress in Pyroptotic BMDMs.

To further explore the differences between tumor-infiltrating BMDMs and microglia and the potential triggers of BMDM pyroptosis, the expression profiles of all BMDMs and microglia were extracted from the above 3 datasets, and the Harmony R

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

FIGURE 7: Continued.

(i)      (j)
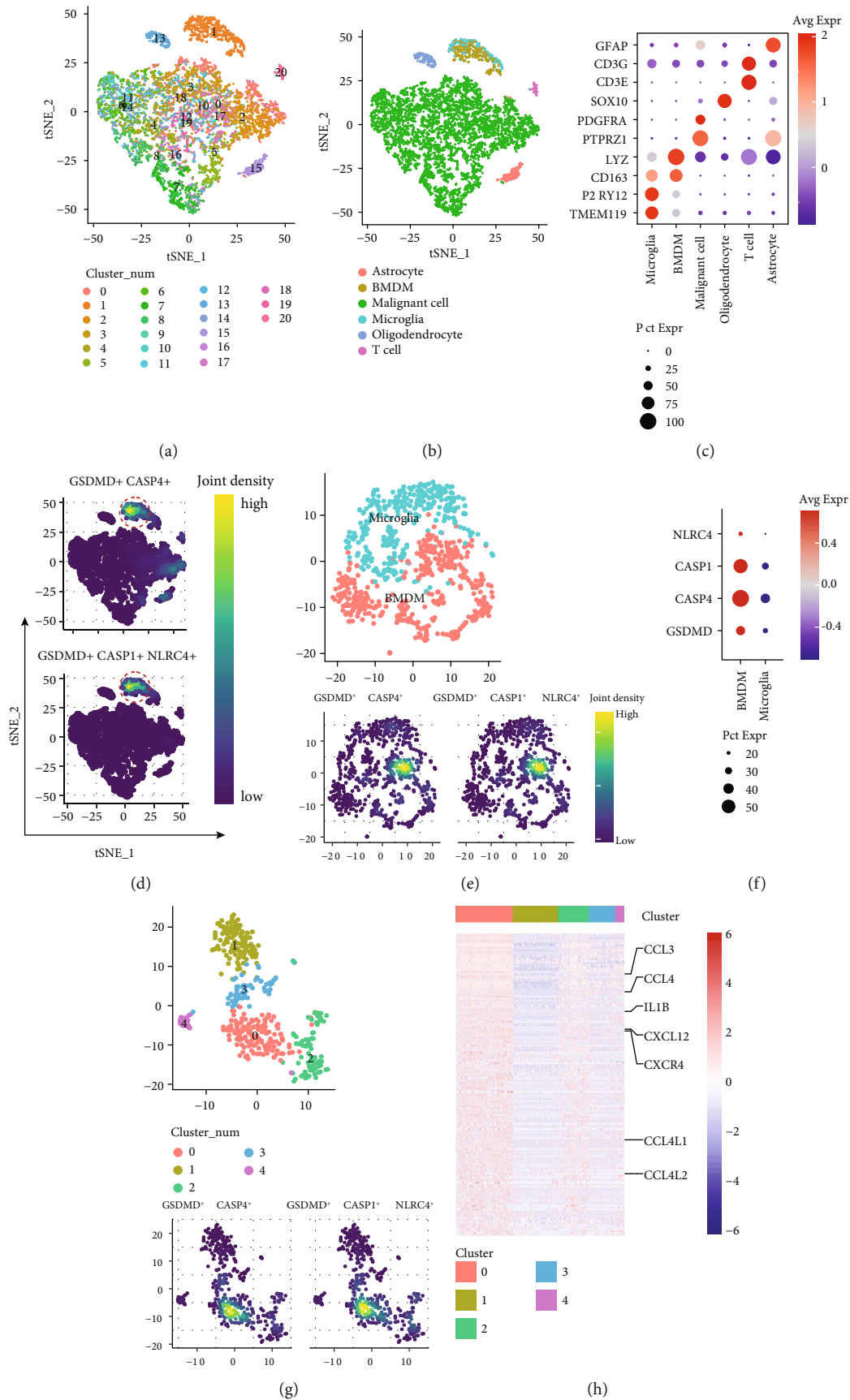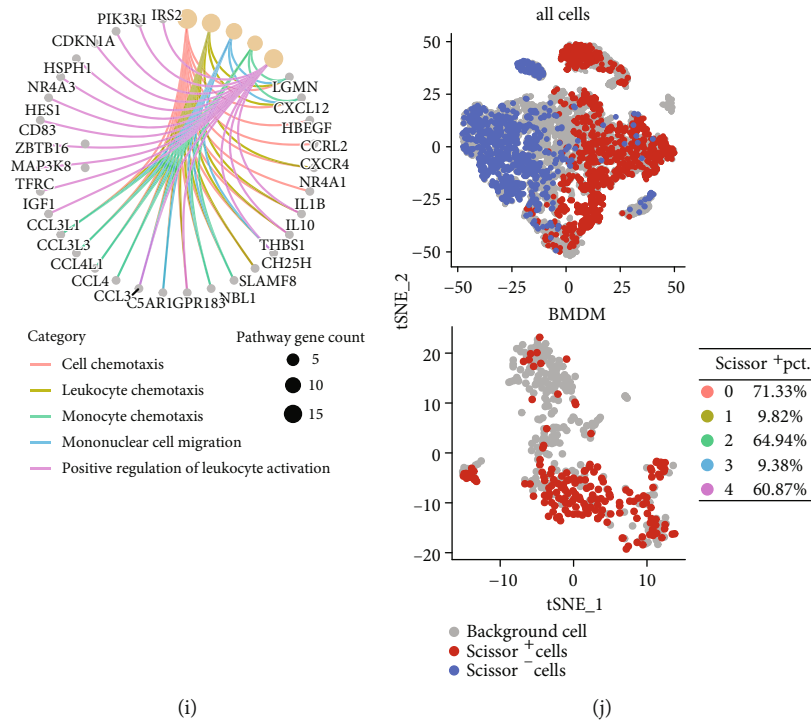
Figure 7: Single-cell transcriptomics revealed that the NLRC4/CASP1/CASP4/GSDMD pyroptosis axis was colocalized in peripheral-derived BMDMs. (a) *t*-distributed stochastic neighbor embedding (tSNE) plot of all single cells. (b) tSNE plot of all cells with cell-type annotations. (c) The signature gene expression matrix for cell cluster identification. (d) tSNE plot of CASP4+ GSDMD+ and NLRC4+ CASP1+ GSDMD+ cells in all cells. Red circles highlight the colocation cluster. (e) tSNE plots of microglia and peripheral-derived macrophages (BMDMs) (top) and distribution of CASP4+ GSDMD+ and NLRC4+ CASP1+ GSDMD+ cells in microglia and BMDMs (bottom). (f) Comparison of NLRC4, CASP1, CASP4, and GSDMD gene expression in BMDMs and microglia. (g) tSNE plots of the distribution of subpopulations of BMDMs (top) and distribution of CASP4$^+$ GSDMD+ and NLRC4+ CASP1+ GSDMD+ cells in BMDM subpopulations (bottom). (h) Heatmap of DEG expression used to distinguish macrophage clusters_0~clusters_4 and highlight genes of interest on the right. (i) GO enrichment analysis of DEGs of BMDM cluster_0. (j) tSNE plots with scissors prediction results of all cells (top) and BMDM (bottom). The percentage of scissor-positive cells in each subpopulation of BMDMs is shown on the right.

package was used for batch effect correction. A total of 1423 BMDMs and 1441 microglia from 36 patients were finally integrated (Figure 9(a)). Notably, BMDM infiltration gradually replaced most microglia as disease grade progressed (Figure 9(b)), and BMDMs from high-grade gliomas had significantly higher expression of CASP1, CASP4, GSDMD, and NLRC4 genes than low-grade gliomas, whereas microglia did not have this conserved relationship (Figure 9(c)). This result is consistent with the previous results of a positive correlation between the Riskscore and BMDM infiltration obtained based on bulk RNA-seq analysis (Figure 6(e)). We defined CASP1+ GSDMD+ and CASP4+ GSDMD+ cells as pyroptotic cells and observed a significantly higher proportion of pyroptotic cells in BMDMs than in microglia (Figure 9(d)). To explore the characteristic changes during BMDM development toward the pyroptosis cell fate, we explored the BMDM differentiation trajectory using the Monocle R package (Figure 9(e)). Notably, pseudotime analysis revealed two cell fates of the BMDM developmental trajectory in gliomas, one of which had a significantly higher proportion of pyroptotic cells (Figure 9(f)). By integrating pathway enrichment and trajectory information, we found that BMDM development

toward the pyroptosis cell fate was accompanied by the activation of the response pathway to lipopolysaccharide (LPS), bacteria and oxidative stress (Gene Cluster 1) (Figure 9(g)). LPS and bacterial infection (e.g., *Salmonella typhimurium*) are classic inducers of pyroptosis in macrophages [51], while oxidative stress has also been recently reported to lead to caspase1-GSDMD-mediated pyroptosis [52, 53]. Although the presence of LPS and bacterial infection in gliomas is less likely, aberrant activation of this response pathway and reactive oxygen species in the tumor microenvironment may lead to macrophage pyroptosis. In addition, we evaluated the activation of the LPS/bacteria response pathway and the oxidative stress response pathway in TCGA cohort using ssGSEA based on the corresponding genes in Gene Cluster 1. The results showed that the high-risk group had significantly higher LPS/bacteria response and oxidative stress response pathway scores (Figure 9(h)) and that high LPS/bacteria response pathway scores and oxidative stress pathway scores were associated with shorter overall survival (Figure 9(i)). The above results suggest that the activation of LPS/bacteria or oxidative stress pathways is associated with poor prognosis, suggesting that they may be triggers of BMDM pyroptosis,

(a)

(b)

(c)

(d)

(e)

(f)

Figure 8: Continued.

(g)

(h)

(i)

(j)
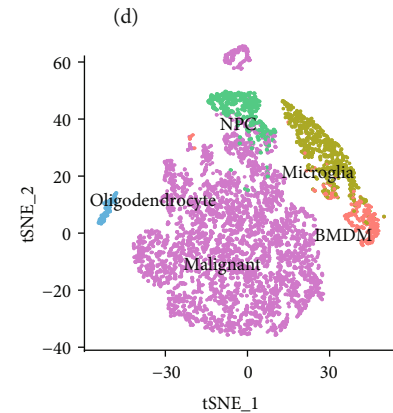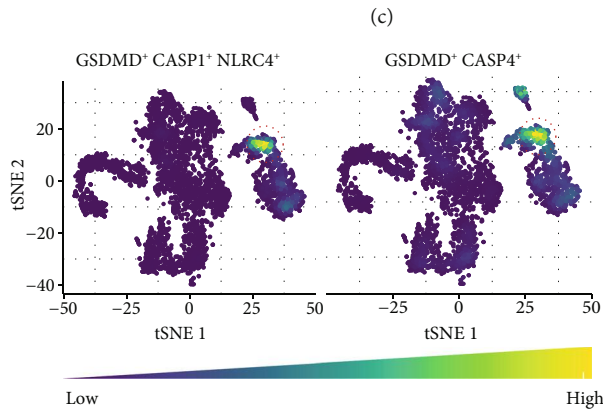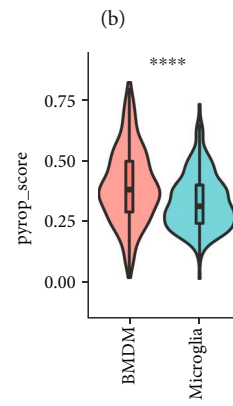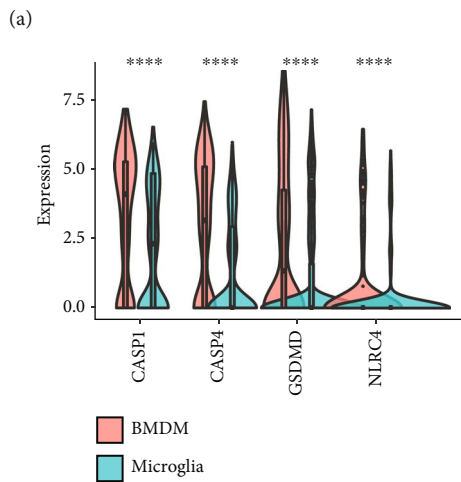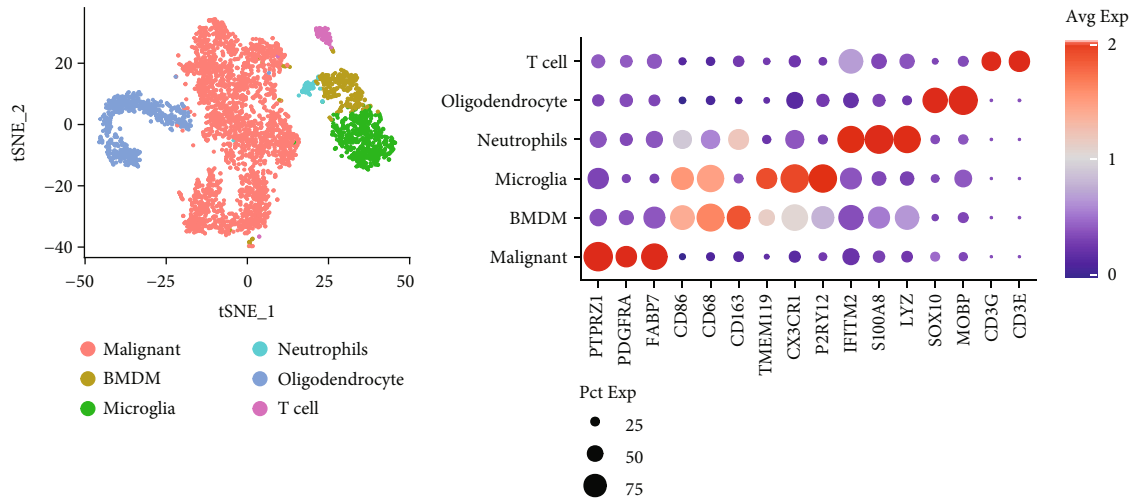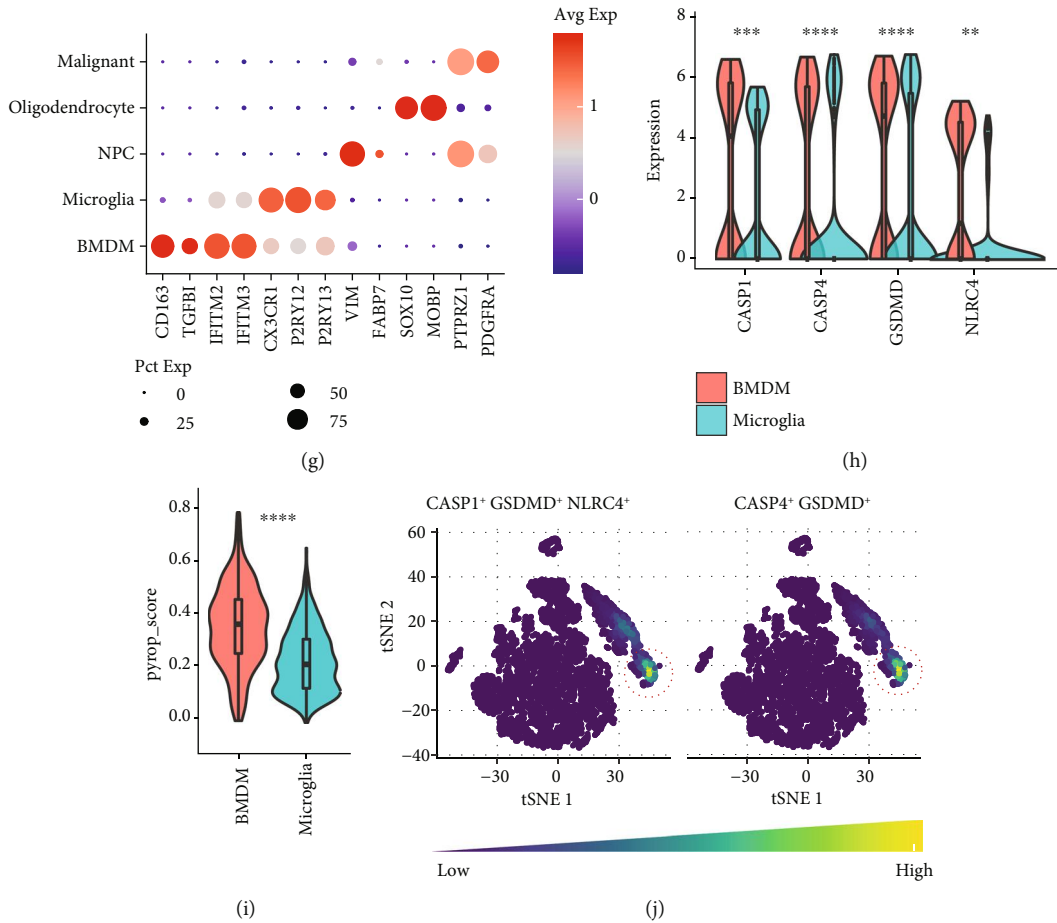
FIGURE 8: Analysis of other datasets confirmed the high expression of pyroptosis genes in BMDMs. (a–e) Visualization and analysis of the GSE117891 dataset. (a) $t$-distributed stochastic neighbor embedding (tSNE) plot of all single cells with cell type annotations. (b) The signature gene expression matrix for cell cluster identification. (c) Comparison of CASP1, CASP4, GSDMD, and NLRC4 expression in BMDMs and microglia. (d) Comparison of the PyropScore between BMDMs and microglia, which was calculated using ssGSEA based on the expression levels of the four genes. (e) The tSNE plots reveal NLRC4+ CASP1+ GSDMD+ cells and CASP4+ GSDMD+ cells in all cells. Red circles highlight the positive cell cluster. (f–j) Visualization and analysis of the GSE89567 dataset. All analysis and visualization methods are the same as those in (a–e). Statistics were calculated using two-tailed, unpaired Student's $t$-test with Welch's correction in (d, i). ns: not significant. $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$.

thereby affecting the tumor microenvironment and leading to tumor progression.

## 4. Discussion

Gliomas are the most common primary tumors of the central nervous system (CNS) and remain incurable, and a deeper understanding of their pathobiology is urgently needed [54]. The glioma tumor microenvironment has a large number of chemokines, cytokines, and growth factors. Despite the recruitment of a high abundance of infiltrating immune cells, such as microglia, peripheral macrophages, CD8+ T cells, CD4+ T cells, and Tregs, the chronic inflammatory environment leads to the establishment of a tumor immunosuppressive microenvironment, which ultimately promotes tumor development [55, 56].

Recently, an inflammatory cell death known as pyroptosis has emerged as an important mediator of the inflammatory response, and as research progresses, pyroptosis is being

proven to be closely associated with an increasing number of types of inflammatory diseases and tumors [10, 57, 58]. Although several recent studies have expanded on the involvement of gasdermin family genes in pyroptosis pathways [59–61], the GSDMD-mediated pyroptosis signaling pathway triggered by inflammasomes has been shown to be the pathway most associated with the formation of an immunosuppressive microenvironment in a variety of tumors [62]. For example, in pancreatic ductal adenocarcinoma (PDA) and head and neck squamous cell carcinoma (HNSCC), inflammasomes of tumor-associated macrophages activate caspase-1 and mediate the cleavage of GSDMD and the release of mature IL1$\beta$, resulting in the suppression of CD8+ T cells [13, 14]. Although inflammasome-mediated pyroptosis in glioma has not been reported, IL1$\beta$ has been shown to be a serum marker in glioblastoma [16, 17], which prompted us to explore potential pyroptosis pathways in glioma.

In this study, we conducted a comprehensive exploration of pyroptosis in glioma. We screened for prognosis-related

(a)

(b)



(c)



(d)                                                          (e)                                                          (f)

FIGURE 9: Continued.

(g)



(h)

(i)
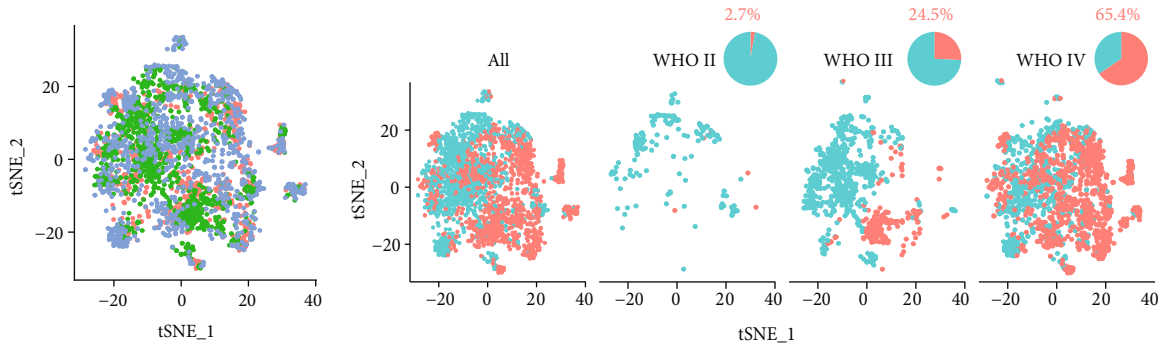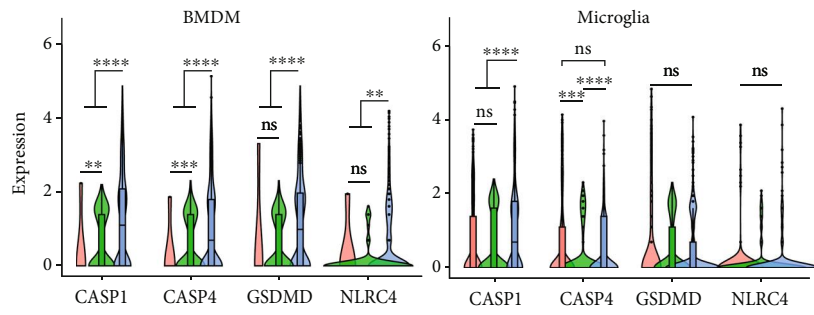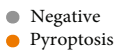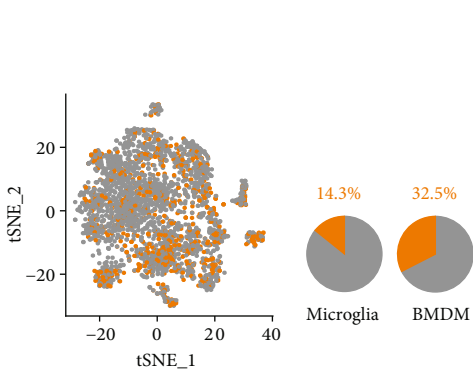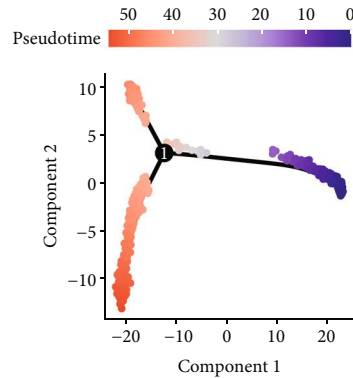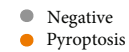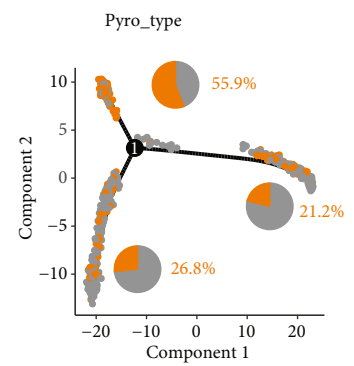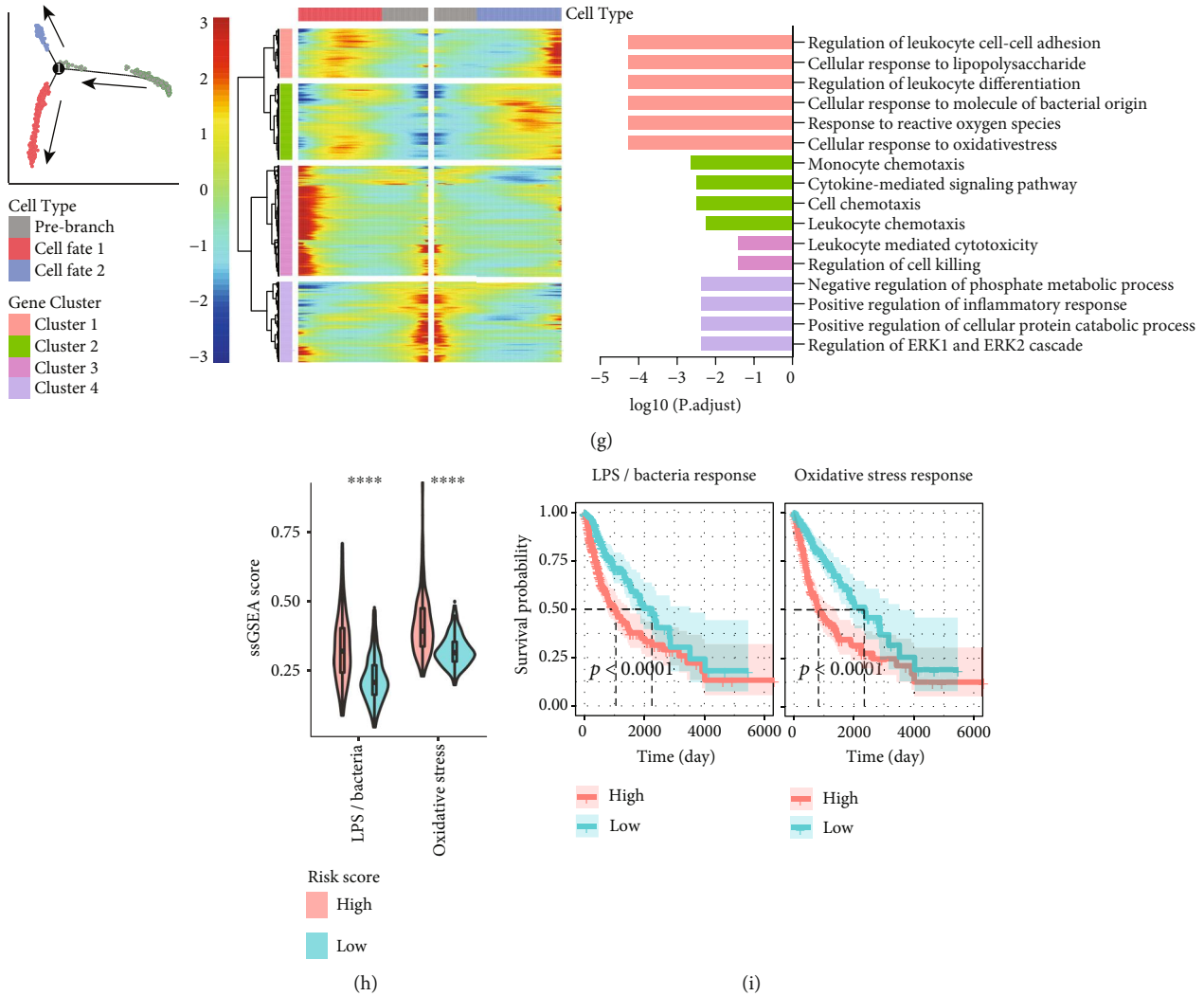
FIGURE 9: Activation of oxidative stress pathways in pyroptotic BMDMs was associated with poor prognosis. (a) BMDM and microglia identified in the GSE131928, GSE117891 and GSE117891 datasets were extracted and integrated using the harmony R package. (b) tSNE plots show the distribution of BMDMs and microglia in glioma samples of different disease grades after correction for batch effects. The pie charts show the percentage of BMDMs and microglia in each WHO grade of sample. (c) Comparison of CASP1, CASP4, GSDMD, and NLRC4 expression in BMDMs (left) and microglia (right) from samples of different WHO grades. (d) CASP1+ GSDMD+ and CASP4+ GSDMD+ cells were defined as pyroptotic cells, and the pyroptosis type of the cells was projected in the tSNE plot. The pie charts show the percentage of BMDMs and microglia in each WHO grade of sample. The pie chart shows the percentage of pyroptotic cells in BMDMs and microglia. (e, f) Trajectory of all BMDMs along pseudotime. The colors from blue to red represent the forward order of pseudotime. The pyroptosis type of the cells is projected on the trajectory. (g) Heatmap revealing the dynamic changes in gene expression during the differentiation process. From the middle to the left and to the right represent the process of changes in gene expression toward differentiation to the two cell fates. Differences in enriched pathways by GO between different phases (right panel). Genes were grouped into four clusters according to their expression patterns, and the results of GO enrichment analysis for each gene cluster are presented in different colors and shown below. (h, i) ssGSEA in TCGA cohort based on the LPS/bacterial response pathway and oxidative stress response pathway genes in Gene Cluster 1, comparing their differences between the high- and low-risk groups and their impact on overall survival. Statistics were calculated using the Wilcoxon test in (c). Statistics were calculated using two-tailed, unpaired Student's $t$-test with Welch's correction in (h). ns: not significant. $^{**}p < 0.01$, $^{***}p < 0.001$, and$^{****}p < 0.0001$.

genes at each key regulatory site of the pyroptosis pathway, obtained the NLRC4/CASP1/CASP4/GSDMD gene cluster, and developed a robust prognostic model based on this cluster. The differentially expressed genes that were upregulated in the high-risk group, defined by the expression of the four genes, were associated with the activation of multiple inflammatory response pathways and increased immune cell infiltration, which are typical results of pyroptosis. In addition,

we demonstrated an immunosuppressive microenvironment in the high-risk group using multiple methods, including TIDE, ssGSEA, and landscape analysis of immune checkpoint expression profiles. Moreover, the differentially expressed genes that were in the high-risk group were strongly associated with synaptic establishment and synaptic signaling, suggesting that pyroptosis can lead to synaptic impairment and neurodegenerative diseases and may

explain to some extent the clinical phenomenon of cognitive dysfunction associated with poor prognosis in glioblastoma patients [42]. Notably, the Riskscore was significantly positively correlated with BMDM infiltration, while the single-cell transcriptomics further demonstrated that NLRC4+ CASP1+ GSDMD+ and CASP4+ GSDMD+ cells were concentrated in a specific peripheral-derived BMDM cluster. Gene characteristics of this cell cluster were found to be associated with monocyte/leukocyte chemotaxis and the expression of IL1$\beta$, an important cytokine involved in the establishment of the immunosuppressive microenvironment, in several previous studies. More importantly, our analysis based on bulk RNA-seq datasets showed a significant positive correlation between pyroptosis gene expression and BMDM infiltration, and independent validation in multiple single-cell RNA-seq datasets further provided conclusive evidence that the four pyroptosis genes were highly expressed only in BMDMs and that pyroptosis gene expression levels were even higher in BMDMs from high-grade glioma samples. Finally, we revealed the activation of LPS/bacteria and oxidative stress response pathways during BMDM development toward the pyroptosis cell fate by pseudotime analysis, suggesting potential BMDM pyroptosis initiators. This is the first demonstration of a strong association between the pyroptosis signaling pathway and BMDM in glioma, providing novel insights into the pathological mechanisms of glioma.

Disulfiram (DSF), a recently demonstrated pyroptosis inhibitor [63], has been well validated in preclinical studies for the treatment of glioblastoma and has advanced to the clinical study phase as a novel adjuvant [64, 65]. In these studies, DSF was considered only as an acetaldehyde dehydrogenase (ALDH) inhibitor based on its classic function of treating alcohol addiction, but given its new status, the drug's function as an inhibitor of pyroptosis in glioma needs to be re-examined to guide the screening of suitable patients. In addition, our bulk RNA-seq-based analysis showed a significant positive correlation between the Riskscore and BMDM infiltration, and we further provided conclusive evidence in multiple independent single-cell RNA-seq datasets that four pyroptosis genes are highly expressed only in BMDMs. Several previous studies have provided solid evidence that the infiltration of BMDMs leads to tumor progression and the establishment of an immunosuppressive microenvironment [34, 66]. Thus, the model we developed can be used to predict the pyroptosis and BMDM infiltration levels in a patient's tumor microenvironment, thus assisting in the selection of candidate antipyroptosis drugs and antimacrophage drugs for the treatment of glioma. Therefore, this retrospective study is of great value, as it provides an in-depth exploration of glioma pathogenesis and its results suggest possibilities for drug development and repurposing based on the pyroptosis signaling pathway.

However, there are still some limitations: (1) the study was conducted based on retrospective data; thus, selection bias might be unavoidable, and (2) although we provided evidence based on bulk transcriptome, single-cell transcriptome and tissue microarray immunohistochemistry data demonstrating a strong relationship between glioma progression and pyroptosis, complex in vivo experiments, such as testing the rate of glioma tumorigenesis and immunosuppression of the tumor microenvironment in GSDMD-deficient mice, can provide more conclusive evidence for the value of pyroptosis as a drug target, which is a promising direction for subsequent studies.

## 5. Conclusions

Our study revealed a critical role of pyroptosis in maintaining immunosuppression in the tumor microenvironment and established a robust pyroptosis score as a prognostic biomarker. We further identified the pyroptosis BMDM cluster at single-cell resolution and preliminarily explored the trigger of BMDM pyroptosis, aberrant activation of pathways in response to LPS/bacteria and oxidative stress, providing potential targets for novel therapies against glioma, such as pyroptosis inhibitors and antimacrophage drugs.

## Abbreviations

CNS:        Central nervous system
ESTIMATE:   Estimation of STromal and immune cells in MAlignant tumor tissues using Expression data
TIDE:       Tumor Immune Dysfunction and Exclusion
ssGSEA:     Single-sample gene set enrichment analysis
TME:        Tumor microenvironment
scRNA-seq:  Single-cell RNA sequencing
BMDM:       Bone marrow-derived macrophages
GBM:        Glioblastoma
LGG:        Low-grade glioma
WHO:        The World Health Organization
IDH:        Isocitrate dehydrogenase
TCGA:       The Cancer Genome Atlas
CGGA:       Chinese Glioma Genome Atlas
GEO:        Gene-expression omnibus
DEGs:       Differentially expressed genes
Log2FC:     Log2 fold change
CNV:        Copy number variation
TMB:        Tumor mutational burden
ICB:        Immune checkpoint blockade
DSF:        Disulfiram
GO:         Gene Ontology analysis
KEGG:       Kyoto Encyclopedia of Genes and Genomes.

## Data Availability

The datasets analyzed during the current study are available in TCGA database (https://portal.gdc.cancer.gov/), CGGA database (http://www.cgga.org.cn/), GEO database (https://www.ncbi.nlm.nih.gov/geo/), and Single Cell Portal database (https://singlecell.broadinstitute.org/single_cell).

## Conflicts of Interest

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

## Authors' Contributions

## Acknowledgments

## Supplementary Materials

Supplementary Figure S1: patients with IDH mutations have a shorter overall survival. Supplementary Figure S2: CASP1 and CASP4 were positively correlated with GSDMD expression in different datasets. Supplementary Figure S3: the inflammasome gene NLRC4 was associated with prognosis in different datasets. Supplementary Figure S4: the NLRC4/CASP1/CASP4/GSDMD gene clusters have highly conserved positive correlation expression pattern. Supplementary Figure S5: performance of the prognostic model developed in this study compared with the established models. Supplementary Figure S6: CNV and TMB differences between the high-and low-risk groups. Supplementary Figure S7: drug sensitivity analysis based on the four pyroptosis genes. Supplementary Figure S8: BMDM infiltration has a robust correlation with glioma prognosis, Riskscore, and M2 macrophage infiltration. Supplementary Figure S9: macrophage clusters were clustered into BMDM and microglia subpopulations by ssGSEA. Supplementary Table S1: gene sets for ssGSEA to estimate tumor microenvironment inhibitory degree. Supplementary Table S2: gene sets from Robert B for ssGSEA to estimate BMDM and microglia infiltration. Supplementary Table S3: gene sets from Sören M for ssGSEA to estimate BMDM and microglia infiltration. *(Supplementary Materials)*

## References

[1] K. Yang, Z. Wu, H. Zhang et al., "Glioma targeted therapy: insight into future of molecular approaches," *Molecular Cancer*, vol. 21, no. 1, pp. 1–32, 2022.

[2] D. N. Louis, A. Perry, P. Wesseling et al., "The 2021 WHO classification of tumors of the central nervous system: a summary," *Neuro-Oncology*, vol. 23, no. 8, pp. 1231–1251, 2021.

[3] S. Han, Y. Liu, S. R. J. Cai et al., "IDH mutation in glioma: molecular mechanisms and potential therapeutic targets," *British Journal of Cancer*, vol. 122, no. 11, pp. 1580–1589, 2020.

[4] M. Weller, R. Stupp, G. Reifenberger et al., "MGMT promoter methylation in malignant gliomas: ready for personalized medicine?," *Nature Reviews. Neurology*, vol. 6, no. 1, pp. 39–51, 2010.

[5] J. Shi, Y. Zhao, K. Wang et al., "Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death," *Nature*, vol. 526, no. 7575, pp. 660–665, 2015.

[6] Y. Wang, W. Gao, X. Shi et al., "Chemotherapy drugs induce pyroptosis through caspase-3 cleavage of a gasdermin," *Nature*, vol. 547, no. 7661, p. 99, 2017.

[7] P. Yu, X. Zhang, N. Liu, L. Tang, C. Peng, and X. Chen, "Pyroptosis: mechanisms and diseases," *Signal Transduction and Targeted Therapy*, vol. 6, no. 1, pp. 1–21, 2021.

[8] I. Jorgensen, M. Rayamajhi, and E. A. Miao, "Programmed cell death as a defence against infection," *Nature Reviews Immunology*, vol. 17, no. 3, pp. 151–164, 2017.

[9] Y. Tan, Q. Chen, X. Li et al., "Pyroptosis: a new paradigm of cell death for fighting against cancer," *Journal of Experimental & Clinical Cancer Research*, vol. 40, no. 1, pp. 1–5, 2021.

[10] L. Wu, L. Li, S. Li et al., "Macrophage-mediated tumor-targeted delivery of engineered Salmonella typhimurium VNP20009 in anti-PD1 therapy against melanoma," *Acta Pharmaceutica Sinica B*, vol. 14, 2022.

[11] X. Xia, X. Wang, Z. Cheng et al., "The role of pyroptosis in cancer: pro-cancer or pro"host"?," *Cell Death & Disease*, vol. 10, no. 9, p. 650, 2019.

[12] C. L. Evavold, J. Ruan, Y. Tan, S. Xia, H. Wu, and J. C. Kagan, "The pore-forming protein Gasdermin D regulates Interleukin-1 secretion from living macrophages," *Immunity*, vol. 48, no. 1, p. 35, 2018.

[13] L. Chen, C. F. Huang, Y. C. Li et al., "Blockage of the NLRP3 inflammasome by MCC950 improves anti-tumor immune responses in head and neck squamous cell carcinoma," *Cellular and Molecular Life Sciences*, vol. 75, no. 11, pp. 2045–2058, 2018.

[14] D. Daley, V. R. Mani, N. Mohan et al., "NLRP3 signaling drives macrophage-induced adaptive immune suppression in pancreatic carcinoma," *The Journal of Experimental Medicine*, vol. 214, no. 6, pp. 1711–1724, 2017.

[15] I. Kaplanov, Y. Carmi, R. Kornetsky et al., "Blocking IL-1$\beta$ reverses the immunosuppression in mouse breast cancer and synergizes with anti-PD-1 for tumor abrogation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 4, pp. 1361–1369, 2019.

[16] Y. Gao, E. Zhang, B. Liu et al., "Integrated analysis identified core signal pathways and hypoxia characteristics of human glioblastoma," *Journal of Cellular and Molecular Medicine*, vol. 23, no. 9, pp. 6228–6237, 2019.

[17] L. Tarassishin, D. Casper, and S. C. Lee, "Aberrant expression of interleukin-1$\beta$ and inflammasome activation in human malignant gliomas," *PLoS One*, vol. 9, no. 7, article e103432, 2014.

[18] Z. Chen, B. Giotti, M. Kaluzova et al., "A paracrine circuit of IL-1$\beta$/IL-1R1 between myeloid and tumor cells drives glioblastoma progression," *bioRxiv*, 2022.

[19] H. Shen, C. Han, Y. Yang et al., "Pyroptosis executive protein GSDMD as a biomarker for diagnosis and identification of Alzheimer's disease," *Brain and Behavior*, vol. 11, no. 4, 2021.

[20] F. Humphries, L. Shmuel-Galia, N. Ketelut-Carneiro et al., "Succination inactivates gasdermin D and blocks pyroptosis," *Science*, vol. 369, no. 6511, p. 1633, 2020.

[21] H. Kim, J. S. Seo, S. Y. Lee et al., "AIM2 inflammasome contributes to brain injury and chronic post-stroke cognitive impairment in mice," *Brain, Behavior, and Immunity*, vol. 87, pp. 765–776, 2020.

[22] M. Zhang, Y. Cheng, Z. Xue, Q. Sun, and J. Zhang, "A novel pyroptosis-related gene signature predicts the prognosis of glioma through immune infiltration," *BMC Cancer*, vol. 21, no. 1, pp. 1–7, 2021.

[23] B. Chao, F. Jiang, H. Bai, P. Meng, L. Wang, and F. Wang, "Predicting the prognosis of glioma by pyroptosis-related signature," *Journal of Cellular and Molecular Medicine*, vol. 26, no. 1, pp. 133–143, 2022.

[24] Z. S. Bao, H. M. Chen, M. Y. Yang et al., "RNA-seq of 272 gliomas revealed a novel, recurrent PTPRZ1-MET fusion transcript in secondary glioblastomas," *Genome Research*, vol. 24, no. 11, pp. 1765–1773, 2014.

[25] L. A. Gravendeel, M. C. Kouwenhoven, O. Gevaert et al., "Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology," *Cancer Research*, vol. 70, p. 9065, 2010.

[26] C. Neftel, J. Laffy, M. G. Filbin et al., "An integrative model of cellular states, plasticity, and genetics for glioblastoma," *Cell*, vol. 178, no. 4, p. 835, 2019.

[27] K. Yu, Y. Q. Hu, F. Wu et al., "Surveying brain tumor heterogeneity by single-cell RNA-sequencing of multi-sector biopsies," *National Science Review*, vol. 7, no. 8, pp. 1306–1318, 2020.

[28] A. S. Venteicher, I. Tirosh, C. Hebert et al., "Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq," *Science*, vol. 355, no. 6332, 2017.

[29] K. Yoshihara, M. Shahmoradgoli, E. Martínez et al., "Inferring tumour purity and stromal and immune cell admixture from expression data," *Nature Communications*, vol. 4, p. 1, 2013.

[30] 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.

[31] P. Jiang, S. Gu, D. Pan et al., "Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response," *Nature Medicine*, vol. 24, no. 10, pp. 1550–1558, 2018.

[32] S. Mariathasan, S. J. Turley, D. Nickles et al., "TGF$\beta$ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells," *Nature*, vol. 554, no. 7693, pp. 544–548, 2018.

[33] L. Wu, F. Zhou, W. Xin et al., "MAGP2 induces tumor progression by enhancing uPAR-mediated cell proliferation," *Cellular Signalling*, vol. 91, article 110214, 2022.

[34] R. L. Bowman, F. Klemm, L. Akkari et al., "Macrophage ontogeny underlies differences in tumor-specific education in brain malignancies," *Cell Reports*, vol. 17, no. 9, pp. 2445–2459, 2016.

[35] S. Müller, G. Kohanbash, S. J. Liu et al., "Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment," *Genome Biology*, vol. 18, pp. 1–4, 2017.

[36] X. Li, M. Garg, T. Jia et al., "Single-cell analysis reveals the immune characteristics of myeloid cells and memory T cells in recovered COVID-19 patients with different severities," *Frontiers in Immunology*, vol. 12, 2022.

[37] I. Korsunsky, N. Millard, J. Fan et al., "Fast, sensitive and accurate integration of single-cell data with harmony," *Nature Methods*, vol. 16, no. 12, p. 1289, 2019.

[38] D. Sun, X. Guan, A. E. Moran et al., "Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data," *Nature Biotechnology*, vol. 40, pp. 527–538, 2021.

[39] W. T. He, H. Q. Wan, L. C. Hu et al., "Gasdermin D is an executor of pyroptosis and required for interleukin-1$\beta$ secretion," *Cell Research*, vol. 25, no. 12, pp. 1285–1298, 2015.

[40] J. Liu, L. Gao, X. Zhu et al., "Gasdermin D is a novel prognostic biomarker and relates to TMZ response in glioblastoma," *Cancers*, vol. 13, no. 22, p. 5620, 2021.

[41] G. Sollberger, G. E. Strittmatter, M. Garstkiewicz, J. Sand, and H. D. Beer, "Caspase-1: the inflammasome and beyond," *Innate Immunity*, vol. 20, no. 2, pp. 115–125, 2014.

[42] H. Bruhn, I. Blystad, P. Milos et al., "Initial cognitive impairment predicts shorter survival of patients with glioblastoma," *Acta Neurologica Scandinavica*, vol. 145, no. 1, pp. 94–101, 2022.

[43] T. P. Burke, P. Engström, R. A. Chavez, J. A. Fonbuena, R. E. Vance, and M. D. Welch, ""Inflammasome-mediated antagonism of type I interferon enhances Rickettsia pathogenesis," nature," *Microbiology*, vol. 5, no. 5, p. 688, 2020.

[44] S. Haase, M. B. Garcia-Fabiani, S. Carney et al., "Mutant ATRX: uncovering a new therapeutic target for glioma," *Expert Opinion on Therapeutic Targets*, vol. 22, no. 7, pp. 599–613, 2018.

[45] J. A. Benitez, J. Ma, M. Antonio et al., "PTEN regulates glioblastoma oncogenesis through chromatin-associated complexes of DAXX and histone H3.3," *Nature Communications*, vol. 8, pp. 1–5, 2017.

[46] H. Zhang, Y. B. Luo, W. T. Wu et al., "The molecular feature of macrophages in tumor immune microenvironment of glioma patients," *Computational and Structural Biotechnology Journal*, vol. 19, pp. 4603–4618, 2021.

[47] N. Zhang, H. Zhang, Z. Wang et al., "Immune infiltrating cells-derived risk signature based on large-scale analysis defines immune landscape and predicts immunotherapy responses in glioma tumor microenvironment," *Frontiers in Immunology*, vol. 12, 2021.

[48] T. Chinen, A. K. Kannan, A. G. Levine et al., "An essential role for the IL-2 receptor in $T_{reg}$ cell function," *Nature Immunology*, vol. 17, no. 11, pp. 1322–1333, 2016.

[49] S. M. Wang, H. Y. Lin, Y. L. Chen et al., "CCAAT/enhancer-binding protein delta regulates the stemness of glioma stem-like cells through activating PDGFA expression upon inflammatory stimulation," *Journal of Neuroinflammation*, vol. 16, pp. 1-2, 2019.

[50] R. Kunze, A. Urrutia, A. Hoffmann et al., "Dimethyl fumarate attenuates cerebral edema formation by protecting the blood-brain barrier integrity," *Experimental Neurology*, vol. 266, pp. 99–111, 2015.

[51] L. Wu, F. Bao, L. Li, X. Yin, and Z. Hua, "Bacterially mediated drug delivery and therapeutics: strategies and advancements," *Advanced Drug Delivery Reviews*, vol. 187, p. 114363, 2022.

[52] H. B. Li, Z. Y. Xia, Y. F. Chen, D. Qi, and H. Zheng, "Mechanism and therapies of oxidative stress-mediated cell death in ischemia reperfusion injury," *Oxidative Medicine and Cellular Longevity*, vol. 2018, article 2910643, p. 2, 2018.

[53] S. Wang, L. Y. Ji, L. Li, and J. M. Li, "Oxidative stress, autophagy and pyroptosis in the neovascularization of oxygen-induced retinopathy in mice," *Molecular Medicine Reports*, vol. 19, no. 2, pp. 927–934, 2019.

[54] M. E. Davis, "Glioblastoma: overview of disease and treatment," *Clinical Journal of Oncology Nursing*, vol. 20, no. 5, pp. 2–8, 2016.

[55] A. Gieryng, D. Pszczolkowska, K. A. Walentynowicz, W. D. Rajan, and B. Kaminska, "Immune microenvironment of gliomas," *Laboratory Investigation*, vol. 97, no. 5, pp. 498–518, 2017.

[56] H. Zhang, Z. Wang, Z. Dai et al., "Novel immune infiltrating cell signature based on cell pair algorithm is a prognostic marker in cancer," *Frontiers in Immunology*, vol. 12, p. 3486, 2021.

[57] X. Zheng, W. Chen, F. Gong, Y. Chen, and E. Chen, "The role and mechanism of pyroptosis and potential therapeutic targets in sepsis: a review," *Frontiers in Immunology*, vol. 12, p. 711939, 2021.

[58] P. Liu, Z. Zhang, and Y. Li, "Relevance of the pyroptosis-related inflammasome pathway in the pathogenesis of diabetic kidney disease," *Frontiers in Immunology*, vol. 12, 2021.

[59] W. Deng, Y. Bai, F. Deng et al., "Streptococcal pyrogenic exotoxin B cleaves GSDMA and triggers pyroptosis," *Nature*, vol. 602, no. 7897, pp. 496–502, 2022.

[60] J. Hou, R. Zhao, W. Xia et al., "PD-L1-mediated gasdermin C expression switches apoptosis to pyroptosis in cancer cells and facilitates tumour necrosis," *Nature Cell Biology*, vol. 22, no. 10, pp. 1264–1275, 2020.

[61] Z. Zhou, H. He, K. Wang et al., "Granzyme A from cytotoxic lymphocytes cleaves GSDMB to trigger pyroptosis in target cells," *Science*, vol. 368, no. 6494, 2020.

[62] B. E. Burdette, A. N. Esparza, H. Zhu, and S. Wang, "Gasdermin D in pyroptosis," *Acta Pharmaceutica Sinica B*, vol. 11, no. 9, pp. 2768–2782, 2021.

[63] J. J. Hu, X. Liu, S. Xia et al., "FDA-approved disulfiram inhibits pyroptosis by blocking gasdermin D pore formation," *Nature Immunology*, vol. 21, no. 7, pp. 736–745, 2020.

[64] C. C. Liu, C. L. Wu, M. X. Lin, C. I. Sze, and P. W. Gean, "Disulfiram sensitizes a therapeutic-resistant glioblastoma to the TGF-beta receptor inhibitor," *International Journal of Molecular Sciences*, vol. 22, no. 19, p. 10496, 2021.

[65] C. Lu, X. Y. Li, Y. Y. Ren, and X. Zhang, "Disulfiram: a novel repurposed drug for cancer therapy," *Cancer Chemotherapy and Pharmacology*, vol. 87, no. 2, pp. 159–172, 2021.

[66] L. Pinton, E. Masetto, M. Vettore et al., "The immune suppressive microenvironment of human gliomas depends on the accumulation of bone marrow-derived macrophages in the center of the lesion," *Journal for Immunotherapy of Cancer*, vol. 7, no. 1, p. 58, 2019.

*Research Article*

# *SPTSSA* Is a Prognostic Marker for Glioblastoma Associated with Tumor-Infiltrating Immune Cells and Oxidative Stress

**Ziheng Wang [ID],[1,2] Xinqi Ge,[1] Jinlong Shi,[2] Bing Lu,[1] Xiaojin Zhang,[1] and Jianfei Huang [ID][1]**

[1]*Department of Clinical Biobank & Institute of Oncology, Affiliated Hospital of Nantong University, Nantong, China 226000*
[2]*Department of Neurosurgery, Affiliated Hospital of Nantong University, China 226000*

Correspondence should be addressed to Jianfei Huang; jfhuang@ntu.edu.cn

*Background.* *SPTSSA* encodes the small subunit A of serine palmitoyltransferase. It catalyzes the formation of sphingoid long-chain base backbone of sphingolipids. Its role in glioma prognosis and tumor-infiltrating immune cells remains unclear. *Methods.* We analyzed *SPTSSA* expression and association with clinical prognosis using GEPIA and CGGA database. Then, GSEA was performed to identify relevant biological functions of *SPTSSA*. The correlations between *SPTSSA* expression and tumor immune infiltrates were investigated using CIBERSORT and TIMER. Finally, IHC and IF were performed to confirm the value of prognosis and the correlation with immune infiltration. *Results.* *SPTSSA* expression was significantly upregulated in diffuse glioma compared to normal tissues and associated with poor survival in GEPIA and CGGA database. Then, we identified biological processes and signaling pathways associated with *SPTSSA* expression. The result showed that *SPTSSA* enriched in the GO term like oxidative stress. Finally, we showed that *SPTSSA* expression was significantly associated with tumor-infiltrating immune cells and overall survival via IHC. *Conclusion.* These findings suggest that *SPTSSA* expression might be used as a prognostic biomarker for glioma and potential target for novel glioma therapy.

## 1. Introduction

In adults, in the central nervous system, the glioblastoma multiforme (GBM) is known to be the most prevalent form of malignancy. Representing almost 15% of all brain tumors, it has an incidence of 3.4 per 100,000 [1–3]. Eighty percent of GBMs is primary (de novo) GBMs and mainly occurs in older patients; the remaining GBMs are secondary GBMs derived from lower-grade astrocytoma or oligodedrogliomas and mainly occur in younger patients. GBM is an aggressive neoplasm; if untreated, patients have a median survival of 3 months [4]. The current standard treatments include surgical resection, chemotherapy with temozolomide, and radiotherapy [5]. With a median survival level of 12-14 months, the prognosis of GBM remains poor despite the advances in radiotherapy and surgery. Less than 5% of patients survive longer than 5 years after diagnosis [6–8]. Novel GBM treatments with improved clinical outcomes are urgently needed.

Cancer immunotherapy takes advantage of the body's own immune system to eradicate tumor cells [9, 10]. Current GBM immunotherapy approaches include checkpoint inhibitor treatment, adoptive cell therapy, dendritic-cell-based therapy, and peptide vaccination [11, 12]. Because the immune system plays a key role in the formation and establishment of tumors, a deep understanding of tumor microenvironment is essential to elucidate tumor-immune interactions and develop effective immunotherapy for GBM. Previous studies suggest that both tumor-associated macrophages (TAMs) and tumor-infiltrating neutrophils (TINs) could affect the treatment outcome and overall survival in GBM [13–16]. However, comprehensive analysis of various immune cell subtypes of GBM is lacking.

With the rapid development of various techniques for gene expression analysis and accumulation of large gene expression databases on clinical samples, bioinformatics analysis plays a significant role in screening and identification of candidate biomarkers for various diseases including cancers [17–19]. Bioinformatics not only provides data for identification of functionally differentially expressed genes (DEGs) for cancer diagnosis and prognosis but can also infer

the percentage of tumor-infiltrating immune cells from gene expression profiles [20, 21]. Finally, IHC result confirmed that *SPTSSA* was a novel biomarker associated with the infiltrating immune cells.

In the current study, we used bioinformatics analysis on GEPIA and CGGA databases and identified *SPTSSA* expression correlating with the prognosis of glioma patients. We further determined the correlation of *SPTSSA* expression with tumor-infiltrating immune cells using CIBERSORT, TIMER, and IHC. Our data provide rationale for future clinical and experimental studies of *SPTSSA* in GBM.

## 2. Material and Methods

*2.1. GEPIA Dataset Analysis.* We used Gene Expression Profiling Interactive Analysis (GEPIA) (http://gepia.cancer-pku.cn/), an interactive web server to identify cancer types that showed differential expression of *SPTSSA* gene between cancerous and normal tissues. Among cancer types that demonstrated differential expression of *SPTSSA* gene, we analyzed the association of *SPTSSA* expression and overall survival.

*2.2. Clinical Information and the CGGA mRNA Matrix.* In this study, the glioma samples obtained from the CGGA network (http://www.cgga.org.cn) numbering 1018 were included. The respective clinicopathological information and the informed consent for all these samples were obtained. The institutional review board of the Tiantan Hospital approved this study. To ascertain the differences in the *SPTSSA* expression and the survival value, an analysis was conducted initially. Besides, the mRNAseq_325 (Illumina HiSeq 2000 or 2500) and the mRNAseq_693 (Platform: Illumina HiSeq) datasets were also downloaded for further investigations. From the total number of 1,018 glioma samples comprising the two datasets, 693 samples were from mRNAseq_693, and 325 samples were from the mRNAseq_325. For batching and normalizing the two mRNA matrices, the limma and Sva packages were utilized. Table 1 enumerates the 749 completed clinical information contained in the clinicopathological characteristics of the patients received from the CGGA databases. The R software was used to conduct the gene expression and the survival analyses (version 3.6.2).

*2.3. Gene Set Enrichment Analysis (GSEA).* To determine the statistical significance of a previously outlined set of genes and the presence of consistent differences concerning two biological states, researchers utilize a computational method like GSEA [22, 23]. In this research, GSEA produced a preliminary list that classified the genes based on their association with the *SPTSSA* expression. Moreover, it elaborates on the remarkable differences between the survival of low- and high-*SPTSSA* groups.

In every analysis, gene set permutations were performed repetitively for 1000 times. We created a phenotype label based on the *SPTSSA*'s expression level. Furthermore, we used the normalized enrichment score (NES) and nominal $p$ value to categorize the enriched pathways in every phenotype [24]. Significantly enriched gene sets occurred at a discovery rate of $|NES| > 1$ and (FDR) $< 0.05$.

*2.4. Tumor-Infiltrating Immune Cell Analysis Using CIBERSORT.* CIBERSORT is a gene expression-based analytical tool for characterizing immune cell composition (http://cibersort.stanford.edu). Using CIBERSORT, we calculated the percentage of infiltrating immune cells in glioma tissues. Wilcox test was used to analyze the difference between *SPTSSA* high- and low-expression groups. R language survival package was used to determine the relationship between infiltrating immune cells and overall survival.

*2.5. TIMER Database Analysis.* Tumor Immune Estimation Resource (TIMER, http://cistrome.shinyapps.io/timer/) is a web server for comprehensive analysis of tumor-infiltrating immune cells. Using TIMER database, we further validated the correlation between *SPTSSA* expression and tumor-infiltrating immune cells.

*2.6. Sample Collection.* The *SPTSSA* expression analysis utilized IHC to assess the glioma tissues of 35 patients on paraffin-embedded, formalin-fixed slides. These samples were collected from the Affiliated Hospital of Nantong University from 2004 to 2014. Using these tissues, the TMA was constructed through the Tissue Microarray System (Quick-Ray, UT06, Unitma, Seoul, South Korea) based on the approach mentioned previously. The clinicopathological information obtained included the differentiation grade, histological type, age, and sex. All patients in the study signed and issued a written informed consent. The Affiliated Hospital of Nantong University's Human Research Ethics Committee approved the study protocol (2018-K020).

*2.7. Construction of the Tissue Microarray (TMA).* From the tissue areas containing >50% tumor, the pertinent regions were selected for the TMA construction from each block of the glioma tissue. Using the MTA-1 Manual Tissue Arrayer (Beecher Instruments, Sun Prairie, WI, USA), representative tumor cores measuring around 1 mm and two/three in number were transferred from the glioma tissue blocks to the recipient TMA blocks in each case. In this manner, the TMAs were constructed.

*2.8. Preparation of Monoclonal Antibody against SPTSSA.* Mab (monoclonal antibody) was prepared from female SPF (specific pathogen free) mice which firstly received 60 $\mu$g polypeptide (3.0 mg/mL) subcutaneous injections. Then, four mice received either four subcutaneous injection polypeptide. Indirect ELISA was performed to analyze the titer of IgG antibody. No. 4 mouse was chosen to perform cell fusion.

*2.9. Immunohistochemistry (IHC).* To quench the endogenous peroxidase, the TMA sections were incubated for 15 minutes with methanol and 3% $H_2O_2$ after being deparaffinized. By heating the sections in sodium citrate buffer (10 mmol/L, pH 6.0) for 3 minutes in a pressure cooker, the antigen was retrieved. Subsequently, for one hour, with the primary goat anti-*SPTSSA* antibody-diluted bovine

TABLE 1: Clinical information analysis based on CGGA database.

|  |  | Total (749) | Low expression (374) | High expression (375) | $\chi^2$ | $p$ |
|---|---|---|---|---|---|---|
| PRS_type | Primary | 502 | 267 | 235 | 6.691 | 0.035 |
|  | Recurrent | 222 | 95 | 127 |  |  |
|  | Secondary | 25 | 12 | 13 |  |  |
| Grade | WHO II | 218 | 127 | 91 | 16.663 | 0 |
|  | WHO III | 240 | 128 | 112 |  |  |
|  | WHO IV | 291 | 119 | 172 |  |  |
| Gender | Male | 307 | 167 | 140 | 4.147 | 0.042 |
|  | Female | 442 | 207 | 235 |  |  |
| Age | ≤41 | 342 | 169 | 173 | 0.068 | 0.795 |
|  | >41 | 407 | 205 | 202 |  |  |
| Radio_status | No | 124 | 74 | 50 | 5.644 | 0.018 |
|  | Yes | 625 | 300 | 325 |  |  |
| Chemo_status | No | 229 | 134 | 95 | 9.718 | 0.002 |
|  | Yes | 520 | 240 | 280 |  |  |
| IDH_mutation_status | Wildtype | 339 | 157 | 182 | 3.247 | 0.072 |
|  | Mutant | 410 | 217 | 193 |  |  |
| 1p19q_codeletion_status | Noncodel | 614 | 283 | 331 | 7.888 | 0.005 |
|  | Codel | 155 | 91 | 64 |  |  |

serum albumin, the tissue sections were incubated. The phosphate-buffered saline was used to wash the sections before being incubated for 15 minutes with horseradish peroxidase-conjugated donkey anti-goat antibody (Abcam) and washed further. Prior to the light counterstaining with hematoxylin, the sections were incubated for 15 minutes with diaminobenzidine solution (Kem-En-Tec Diagnostics, Taastrup, Denmark) to develop the color.

2.10. Immunofluorescence Staining. The glioma samples were cut in thin sections of 3 $\mu$m for immunostaining. Using the Fluorescence Kit (NEL 797001KT; PerkinElmer) and the Opal 8-color Fluorophore TSA, the multiplex immunofluorescence (IF) was done for CD83/CD56/CD20/CD68/CD66b/CD8/CD4/CD3 on the 3 $\mu$m formalin-fixed glioma tissue sections. Since combining four or more antibodies using IHC was challenging technically, instead of IHC, we performed the multiplexed IF.

2.11. Multispectral Analysis and Imaging. With an interactive image segmentation system, the Vectra 3 automated quantitative pathology imaging system, the image analysis, and acquisition were performed on the glioma slides. According to the intensity and the staining pattern on each selected image, the pathologist decided the fluorescent intensity count cutoff value for positivity for each marker of interest (CD83/CD56/CD20/CD68/CD66b/CD8/CD4/CD3). The extragerminal center area and the germinal center area were included for each image of the normal glioma tissue. The manual tissue segmentation function of the software was utilized for differentiating the two areas and analyzed independently.

2.12. Survival Analysis. Using SPTSSA medium expression level, glioma samples were stratified into high-SPTSSA expression and low-SPTSSA expression groups. Kaplan-Meier survival analysis was used to estimate the survival distributions. Evaluating the statistical significance required the log-rank test between stratified survival groups through the GraphPad Prism package. Then, we filtered the survival and gene expression data using Cox regression analyses at $p < 0.05$.

# 3. Results

3.1. SPTSSA Was Significantly Upregulated in Diffuse Glioma Compared to Normal Tissues and Associated with Poor Survival. We searched Gene Expression Profiling Interactive Analysis (GEPIA) database for SPTSSA expression in various tumor types. TCGA (http://tcga-data.nci.nih.gov/tcga/) and GTEx (http://commonfund.nih.gov/GTEx/) datasets from 33 tumor types were retrieved. The analysis of gene expression profile across all tumor samples and paired normal tissues indicated that seven tumor types (DLBC, GBM, LGG, LIHC, PAAD, TGCT, and THYM) showed significant higher SPTSSA expression in tumor tissues than in normal tissues (Figure 1(a)). Among these seven tumor types, we separated tumor cases into high-SPTSSA expression cases (>median expression level) and low-SPTSSA expression cases (≤median expression level) (Figure 1(b)). Log rank survival analysis indicated that high-SPTSSA expression in GBM and LGG was associated with poor survival (HR = 2.3, $p$(HR) = 7.3$e$ − 10, Logrank $p$ = 2.7$e$ − 10) (Figure 1(c)).
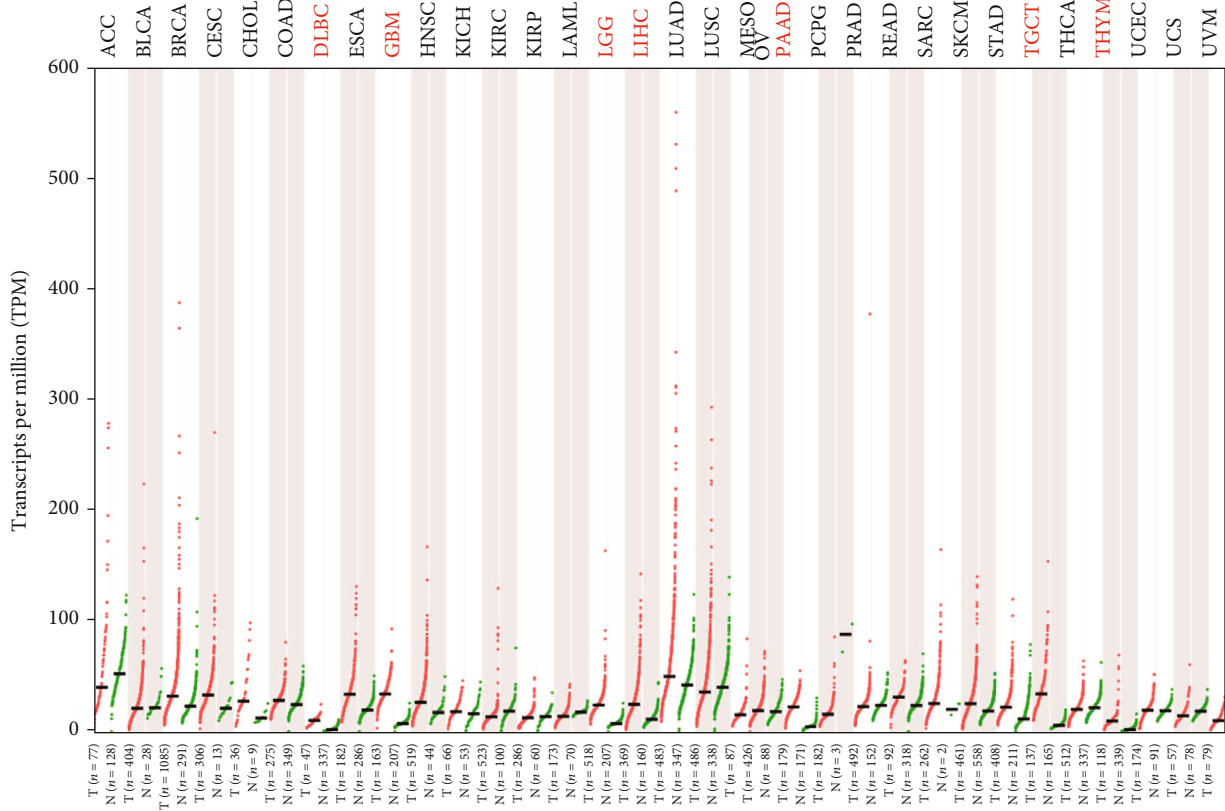
(a)
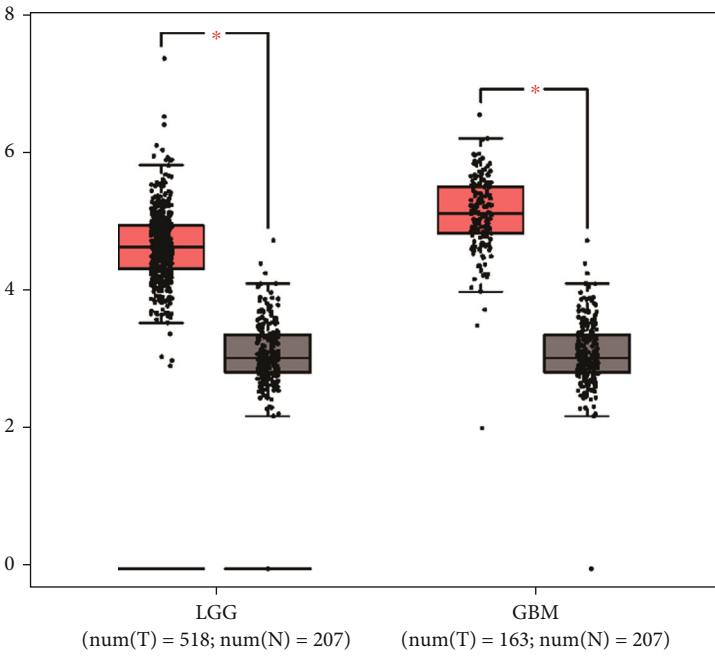


(b)
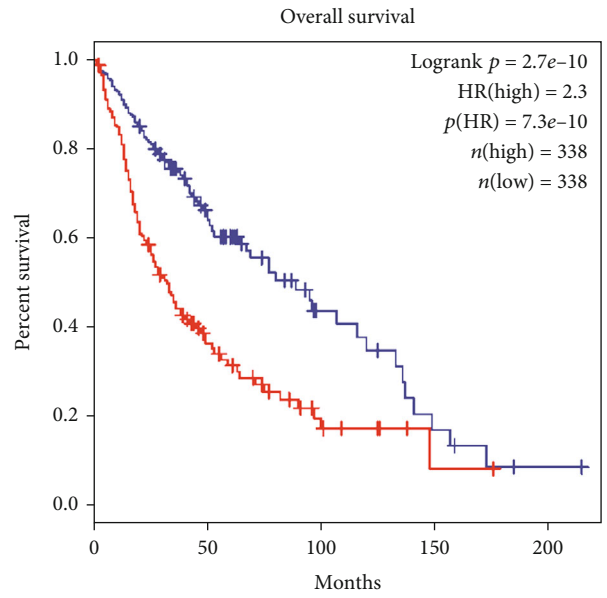


(c)

FIGURE 1: (a) Overview of expression of all cancers and normal tissues in GEPIA database. (b) Differences of SPTSSA in normal cells, low-grade glioma, and glioblastoma. (c) Grouped by median, high expression of SPTSSA was associated with poor survival.

### 3.2. Correlation of SPTSSA Expression with Clinical Characteristics.

To confirm our observation, we next analyzed SPTSSA expression using CGGA database. Log-rank test analysis showed that high-SPTSSA expression was significantly associated with poor survival ($p < 0.001$) (Figure 2(a)). In univariate analysis, PRS-type, histology, grade, age, chemotherapy, IDH-mutation, 1p19q-codeletion, and SPTSSA expression were all significantly associated with survival (Figure 2(b)). In multivariate analysis, PRS-type, grade, IDH-mutation, 1p19q-codeletion, and SPTSSA remained significantly associated with survival (Figure 2(c)). Moreover, a nomogram was carried out to investigate individualized survival probability (Supplementary Figure 1A), and calibration curve was carried out to demonstrate the accuracy of the nomogram in predicting prognosis at different time points (Supplementary Figure 1B). The cox analysis between SPTSSA and OS, PFI, DSS, and DFI was also carried out (Supplementary Figure 2A-2D). The results showed that SPTSSA correlated positively with OS, PFI, and DSS based on pan-cancer analysis. All the results indicate that SPTSSA could serve as an independent prognostic predictor for glioma.

### 3.3. Multifactorial Integrated Survival Analysis in CGGA Database.

The radiotherapy (Figure 3(a)), chemotherapy (Figure 3(b)), IDH1 genotypes (Figure 3(c)), and 1p19q status (Figure 3(d)) were added as variables in multifactorial analysis to further investigate the clinical value of SPTSSA. The correlation between the expression of SPTSSA and the survival rate with chemotherapy was analyzed subsequently. The poorest outcome was noticed in the case of the highest SPTSSA expression with chemotherapy (Figure 3(b), cherry), whereas the expression of SPTSSA without chemotherapy (Figure 3(b), purple) indicated the lowest. The role of the corresponding radiotherapy ($p < 0.0001$) SPTSSA as an important indicator was evidenced by the higher expression of SPTSSA (Figure 3(a), cherry and green) in the IDH1-mutant groups (Figure 3(a), purple and blue) revealing poor survival. The patients with high expression of SPTSSA (Figure 3(d), green) and noncodel 1p19q finally indicated the worst prognosis.

### 3.4. Correlations between SPTSSA and Immunotherapy, Immunotherapy Response Prediction.

Cancer immunotherapy is radically transforming cancer [25], and the use of immunotherapy in cancer treatment is on the rise [26]. Several studies have shown that immunotherapy and targeted therapy are effective in melanomas [27]. Thus, it is necessary to investigate the relationship between SPTSSA and immunotherapy and immunotherapy response prediction. In our study, we downloaded 109 samples from GSE91061, and the results showed that SPTSSA expression level has a positive correlation with anti-PD-1/CTLA-4 therapy both in LGG and GBM (Supplementary Figure 3A). And immunotherapy response prediction of AUC is 0.688 which reveals that SPTSSA can predict immunotherapy response of glioma patients (Supplementary Figure 3B). Furthermore, we investigated the relationship between SPTSSA and immune suppressive factors, immune promot-ing factors, MHC factors, chemokine, and receptors (Supplementary Figure 3C). The results showed that SPTSSA has a positive relationship with most immune suppressive factors, immune promoting factors, MHC factors, chemokine, and receptors which indicate that SPTSSA could have a good immune therapy towards glioma patients.

### 3.5. Connections between SPTSSA and Genomic Alteration.

The majority of cancers harbor at least one genomic alteration that could lead to potential treatment options, with 84% showing at least one. And tailored medicine is often based on specific genetic alterations that improve treatment outcomes [28]. Thus, exploring the relationship between SPTSSA and genomic alteration seems to be necessary. In our study, we found that SPTSSA is positively correlated with CALN1 in GBM, while no positive correlated genomic alteration found in LGG. And in GBM, the gain of SPTSSA genomic alteration located in 17q13.2, while the loss of SPTSSA genomic alteration located in 14q13.1 and 14q24.2 (Supplementary Figure 4A-4B).

### 3.6. Gene Set Enrichment Analysis (GSEA).

To distinguish the differentially activated signaling pathways in GBM, we performed the GSEA between low- and high-SPTSSA expression data sets. Under the MSigDB Collection's (c5.all.v7.1.symbols.gmt) enrichment analysis, the SPTSSA generated significant differences that were reported in the GSEA ($p < 0.05$). Our selection of the most highly enriched signaling pathways was based on their normalized enrichment scores ($|NES| > 1$) (Figure 4). The outcome revealed that negative regulation of response to oxidative stress, negative regulation of mitotic cell cycle, neuron death in response to oxidative stress, positive regulation of cellular catabolic process, and transcription factor complex were enriched in low expression phenotype (Figure 4(a)).

### 3.7. Connections between SPTSSA and TMB, MSI, and Immune Checkpoint.

There are studies which reveal that lower TMB values are associated with longer mean overall survival times, concluding that TMB is a marker of tumor malignancy [29]. Instability of microsatellites (MSI) results from mutations in DNA mismatch repair (MMR) genes, which fail to repair errors in DNA replication in repetitive sequences (microsatellites) [30]. There are studies found that MSI is associated with poor differentiation, proximal location, and failure of chemotherapy in colorectal cancers [31]. Thus, investigating the relationship between SPTSSA and TMB, MSI seem to be necessary. Regarding TMB, we found that SPTSSA expression level is correlated with TMB in GBM and LGG, while in MSI, we found no sense (Figures 5(a) and 5(b)).

Further, immune cells, checkpoint expression, and MSI status play a significant role in prognosis [32]. Thus, we further investigate the relationship between SPTSSA and immune checkpoint. And the results showed that SPTSSA correlated with most immune checkpoints (Figure 5(c)).

### 3.8. Associations between SPTSSA and DNA Methylation.

DNA methylation is mediated by DNA methyltransferase (DNMT) and is affected by the environment [33–35]. Thus,

(a)



(b)

FIGURE 2: Continued.

| | p value | Hazard ratio |
|---|---|---|
| SPTSSA | 0.026 | 1.147 (1.017–1.294) |
| PRS_type | <0.001 | 1.911 (1.625–2.247) |
| Histology | 0.078 | 0.672 (0.432–1.046) |
| Grade | <0.001 | 2.757 (2.018–3.768) |
| Gender | 0.545 | 1.061 (0.876–1.284) |
| Age | 0.038 | 1.234 (1.011–1.506) |
| Ratio | 0.280 | 0.861 (0.657–1.129) |
| Chemo | 0.002 | 0.686 (0.540–0.872) |
| IDH_mutation | <0.001 | 0.587 (0.466–0.739) |
| 1p19q_codeletion | <0.001 | 0.395 (0.283–0.551) |



(c)
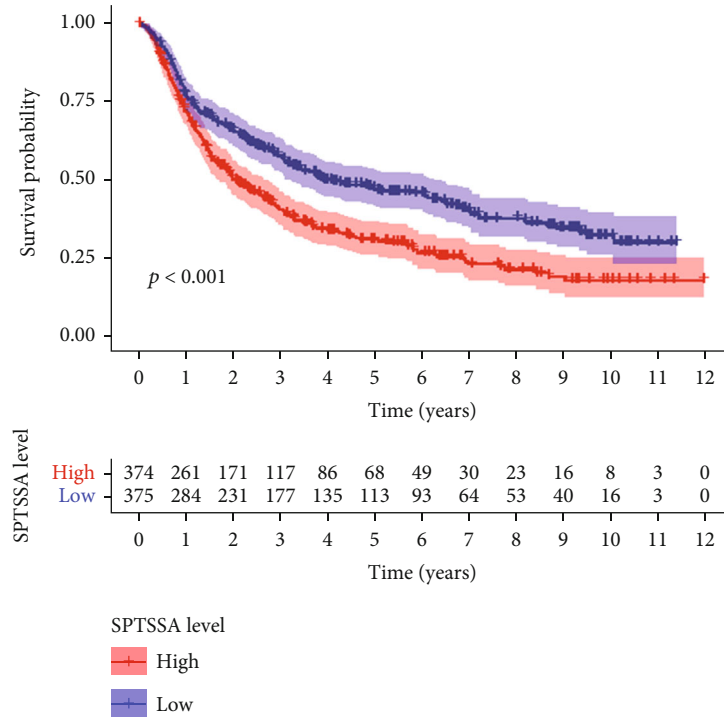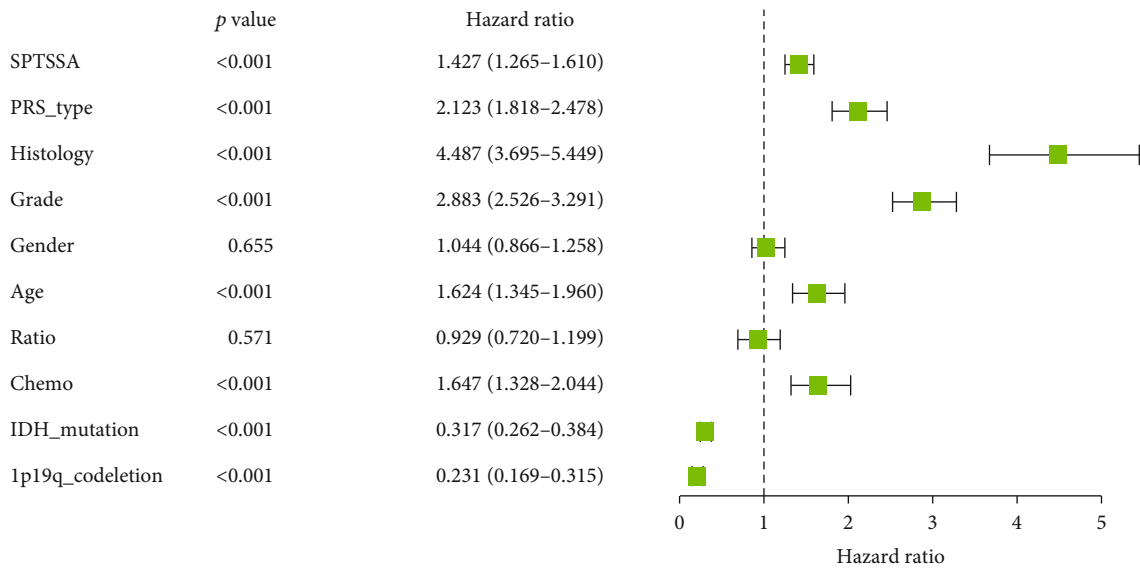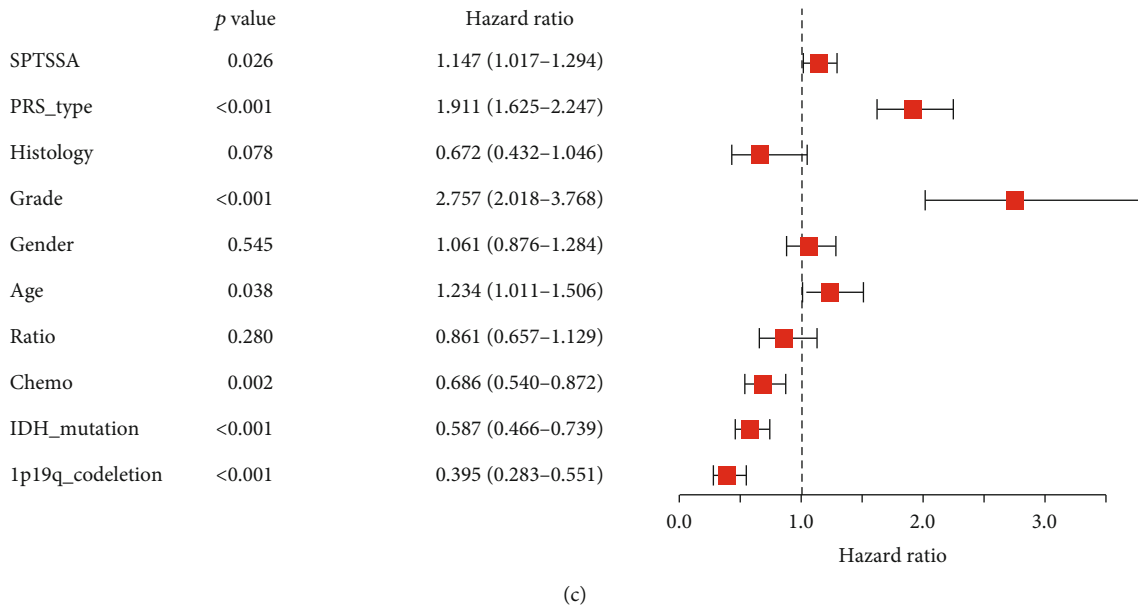
FIGURE 2: (a) Using CGGA database, grouped in median, high expression of SPTSSA predicts poor prognosis. (b, c) Univariate and multivariate Cox analyses indicated that SPTSSA was an independent predictor for OS.

we investigate the relationships between SPTSSA and four methytransferases (DNMT3B, DNMT3A, DNMT2, and DNMT1) (Supplementary Figure 5A). $p < 0.05$ and $R > 0.20$ indicated a significant and positive relationship, respectively. The results showed that SPTSSA has a positive correlation with four methytransferases in LGG, while in GBM, SPTSSA only has a positive correlation with DNMT3B, DNMT3A, and DNMT2.

### 3.9. Correlation of SPTSSA Expression and Infiltrating Immune Cells.
We used CIBERSORT analysis to evaluate the correlation of SPTSSA expression with tumor-infiltrating immune cells (TIICs). As Figure 6(b) illustrates, our results demonstrated two statistically significant associations. First, the ratio of monocytes, NK cells (activated), T follicular helper (Tfh) cells, naive CD4+ T cells, memory B cells, and naive B cells were substantially lower in tumor cases with high SPTSSA expression. Second, eosinophils, dendritic cells (activated and resting), and macrophages (M0) were considerably higher in tumor cases with high SPTSSA expression.

### 3.10. Correlation of Infiltrating Immune Cells and Overall Survival.
Using log-rank test survival analysis, we showed that high number of macrophage M0 cells ($p < 0.001$), T cells CD4 naive ($p < 0.001$), monocytes ($p < 0.001$), macrophages M2 ($p < 0.001$), dendritic cells activated ($p < 0.001$), T cells gamma delta ($p = 0.004$), T cells regulatory (Tregs) ($p = 0.009$), neutrophils ($p = 0.030$), and plasma cells ($p = 0.032$) were significantly associated with survival (Figure 6(a)).

### 3.11. Validation of Correlation between SPTSSA and Infiltrating Immune Cells.
Using the TIMER database (Figure 6(b)), the correlation between the infiltrating immune cells and the SPTSSA expression was analyzed further to confirm our observation. The SPTSSA expression

being significantly correlated with numbers of tumor-infiltrating dendritic cells, neutrophils, and macrophages was revealed clearly.

### 3.12. Immunohistochemistry (IHC).
To estimate the expression of SPTSSA in glioma tissues, we used IHC. Our findings revealed that in contrast to low-grade (I and II) gliomas, high-grade (III and IV) gliomas have highly expressed SPTSSA (Figure 7(a)). Log rank survival analysis indicated that high-SPTSSA expression in glioma was associated with poor survival (Logrank $p = 0.045$) (Figure 7(b)). In univariate analysis and multivariate analysis, SPTSSA is the only factor significantly associated with survival (Figure 7(c)).

### 3.13. Correlation Analysis between SPTSSA Expression and Immunosuppressive Markers.
To investigate the relationship between SPTSSA and the tumor immunology infiltrating cells, we focused on the correlations between SPTSSA and well-known immunosuppression-related genes. Analysis with TMA showed that SPTSSA had highly positive correlation with CD8, CD66b, and CD20 (Supplementary Figure 5B).

## 4. Discussion

In the current study, by employing various bioinformatics analysis tools, we identified that SPTSSA expression was upregulated in diffuse glioma and associated with poor survival. We characterized SPTSSA-related biological processes and signaling pathways. Finally, we provided evidence that SPTSSA expression was correlated with tumor immune infiltrates by CIBERSORT, TIMER, IHC, and IF.

GEPIA is a web-based interactive tool for mining RNA sequencing data on TCGA and the GTEx databases [17, 18]. It covers over 45,000 genes. GEPIA allows experimental

(a)



(b)

FIGURE 3: Continued.

(c)
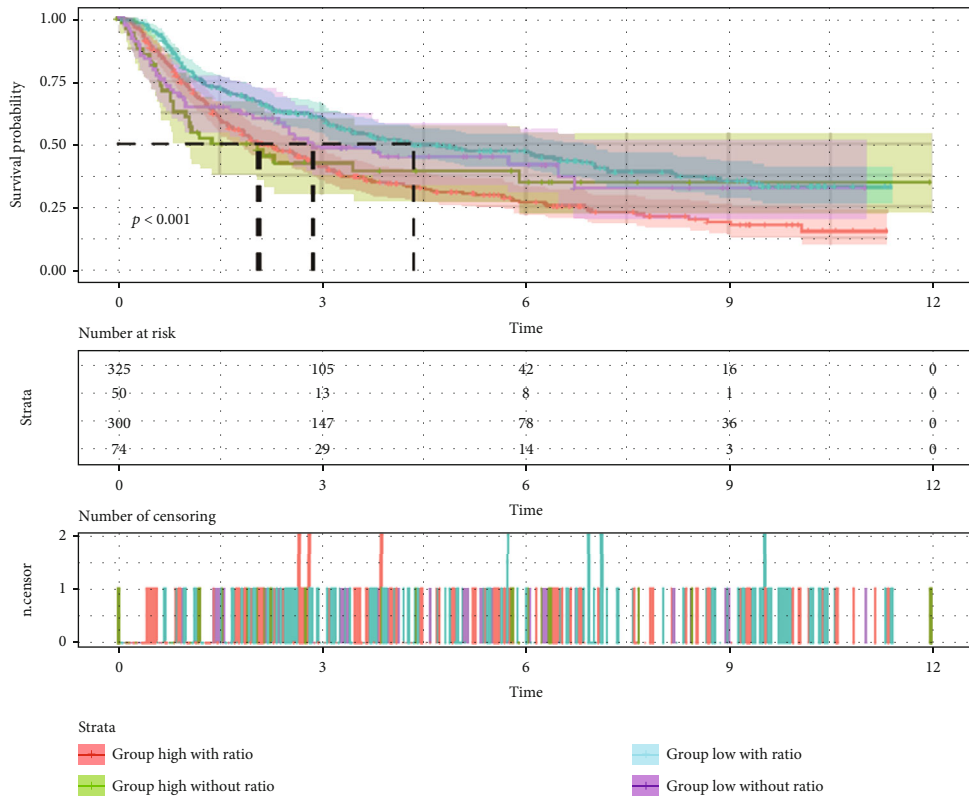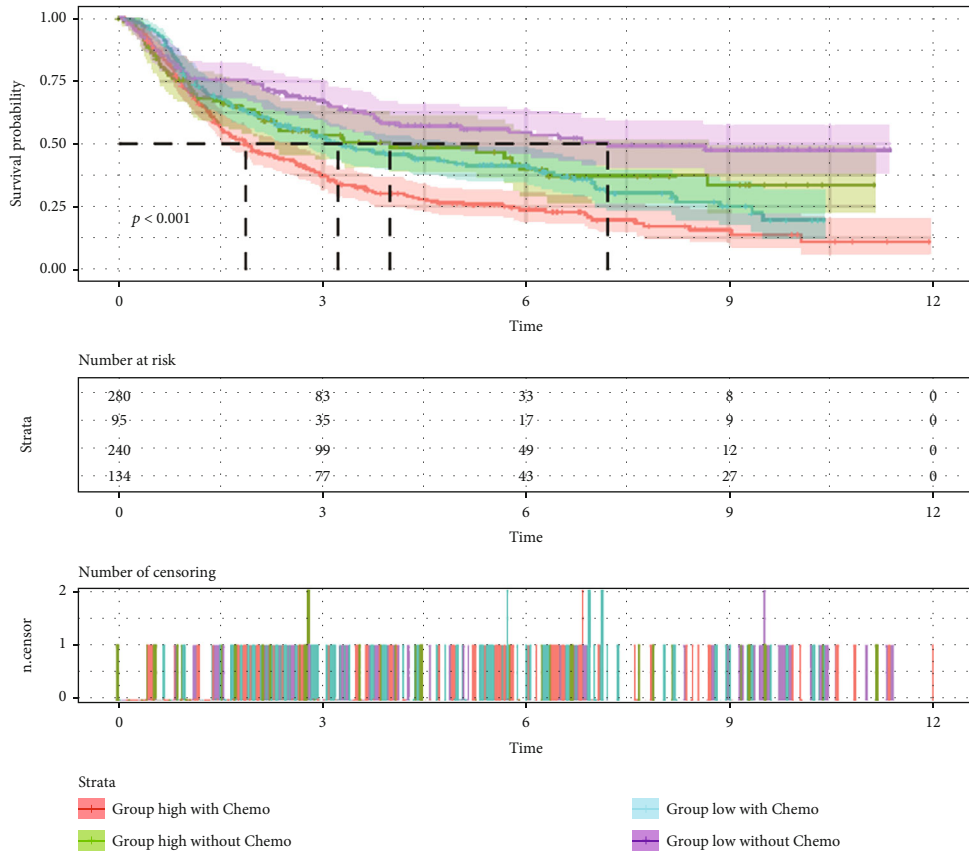


(d)
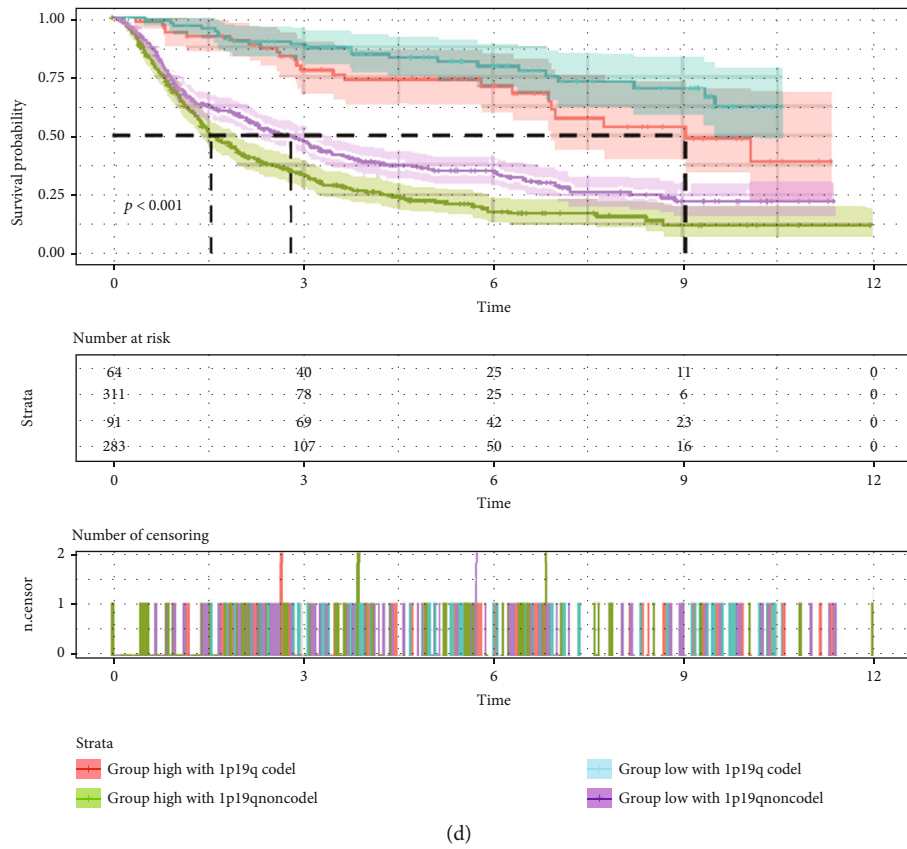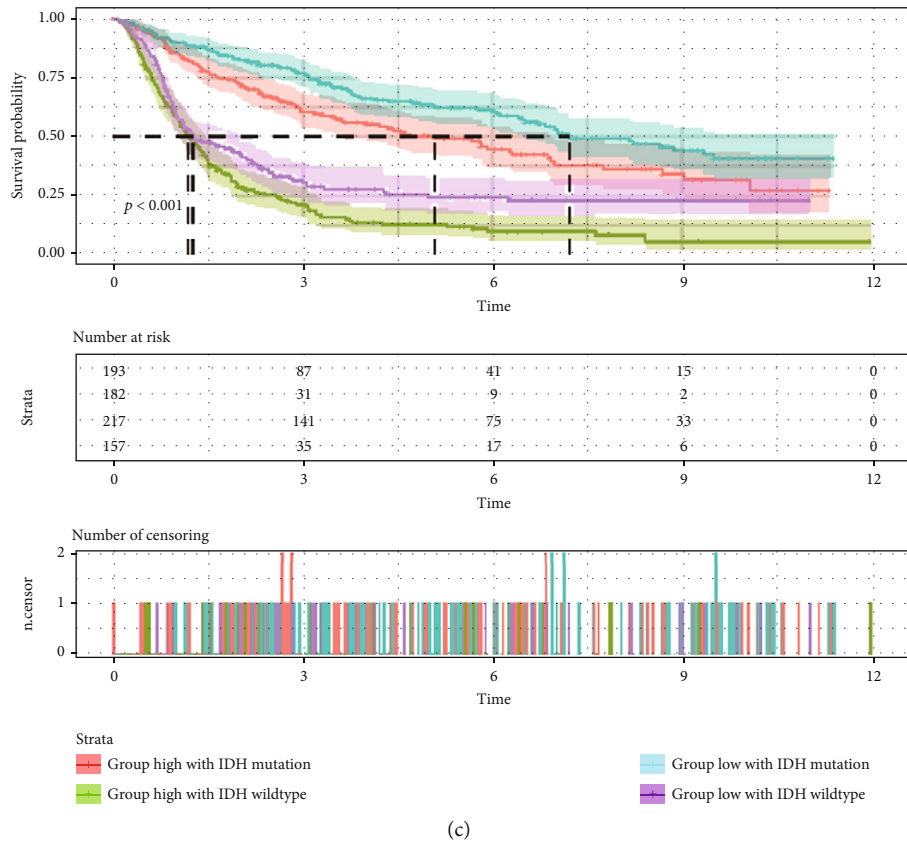
FIGURE 3: Survival analysis of CGGA patients to SPTSSA expression compared with (a) radiotherapy, (b) chemotherapy, (c) IDH mutation, and (d) 1p19q status.

(a)



(b)
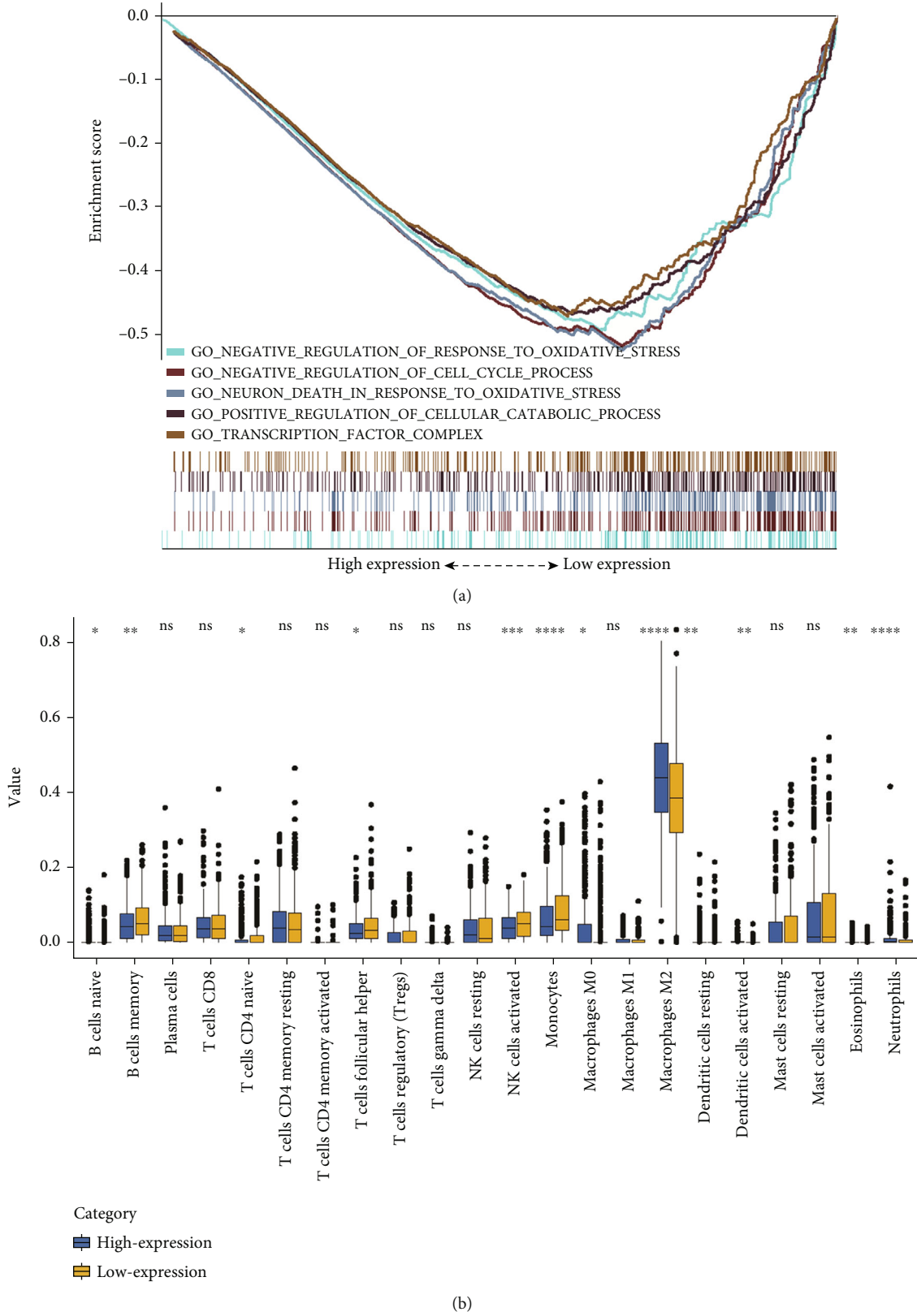
FIGURE 4: (a) Enrichment plots from the Gene Set Enrichment Analysis. (b) The proportions of 22 tumor-infiltrating immune cells in high-SPTSSA and low-SPTSSA expression group.

(a)



(b)

FIGURE 5: Continued.

(c)
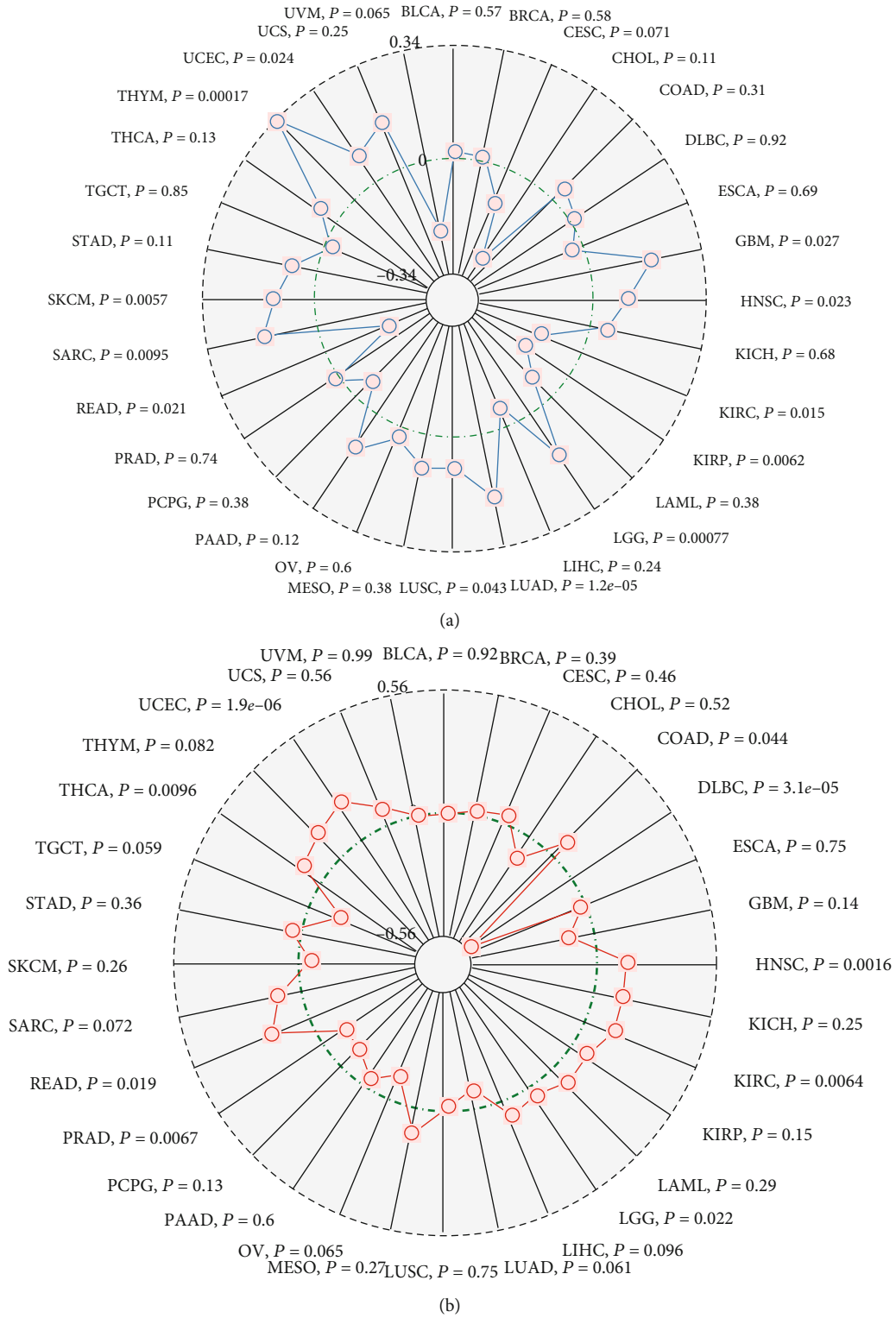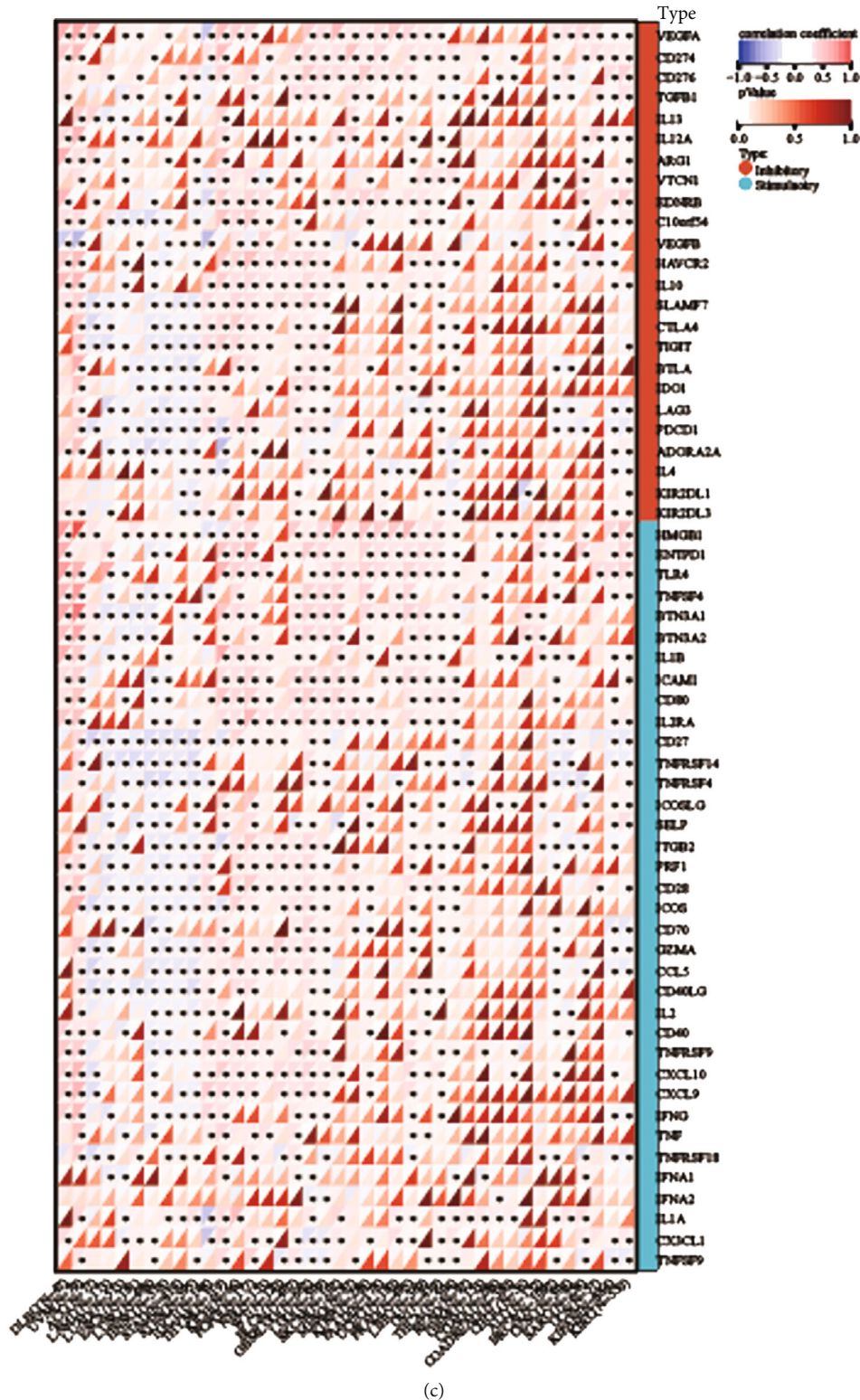
FIGURE 5: Correlations between SPTSSA and (a) TMB and (b) MSI. (c) Relationships between SPTSSA expression level and immune checkpoints.

biologists with limited computational programming skills to perform large scale gene expression analyses [36, 37]. For current study, we utilized both differential expression analysis and patient survival analysis functions of GEPIA. We found that SPTSSA was not only differentially expressed in GBM but also associated with GBM survival. We further validated our findings by searching CGGA database.
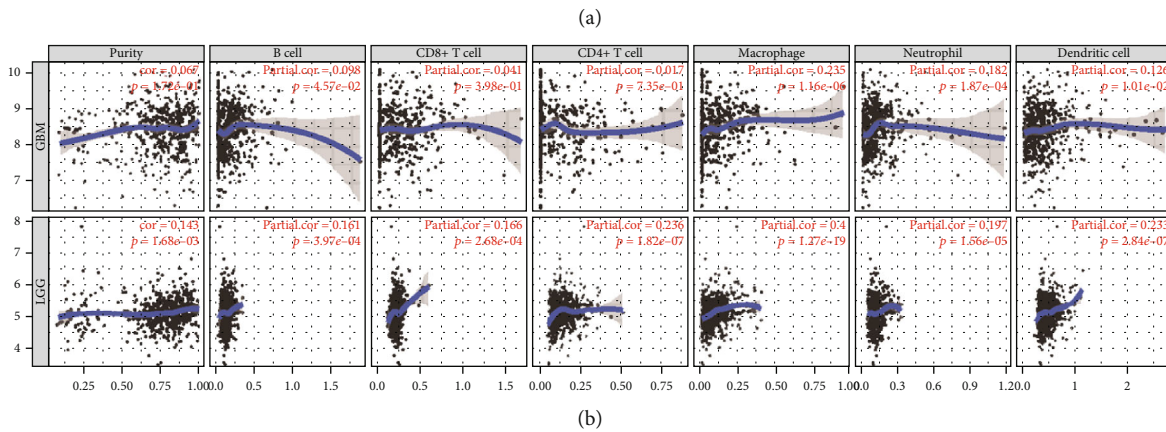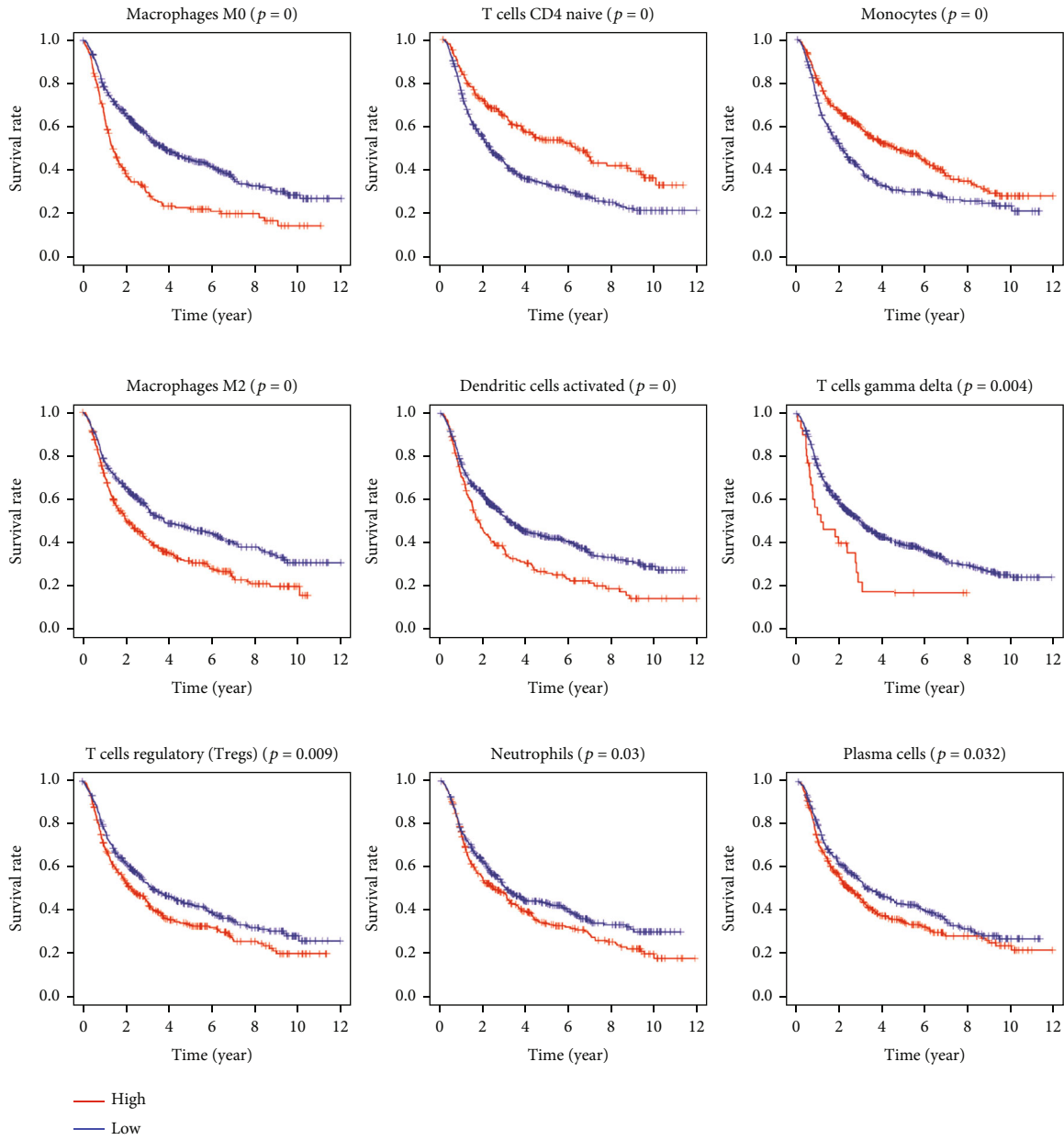
FIGURE 6: (a) Macrophage M0 cells ($p < 0.001$), T cells CD4 naive ($p < 0.001$), monocytes ($p < 0.001$), macrophages M2 ($p < 0.001$), dendritic cells activated ($p < 0.001$), T cells gamma delta ($p = 0.004$), T cells regulatory (Tregs) ($p = 0.009$), neutrophils ($p = 0.030$), and plasma cells ($p = 0.032$) were significantly associated with survival. (b) The correlation of immune cells and SPTSSA expression in the TIMER database.
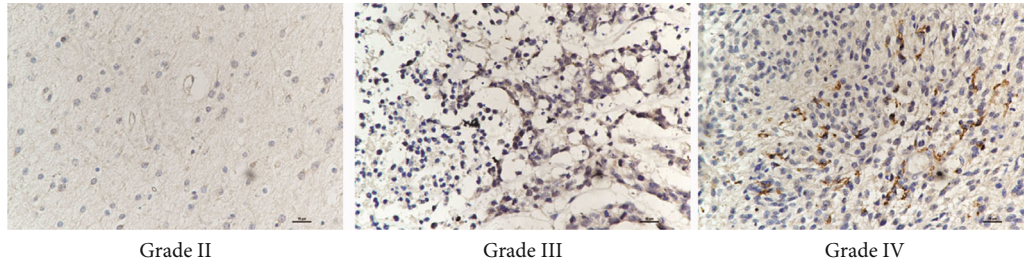
Grade II　　　　　　　　　Grade III　　　　　　　　　Grade IV

(a)
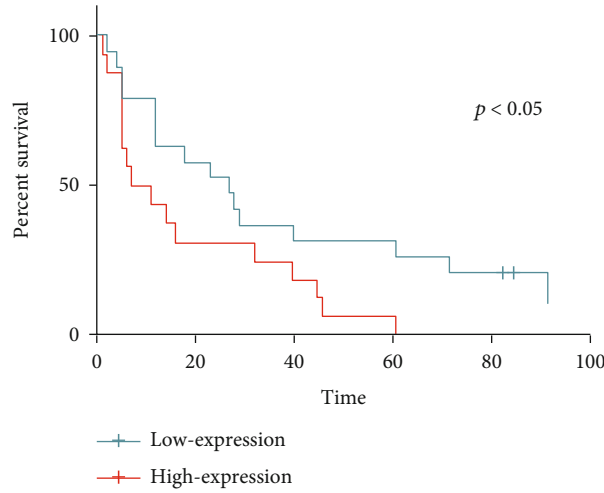


(b)
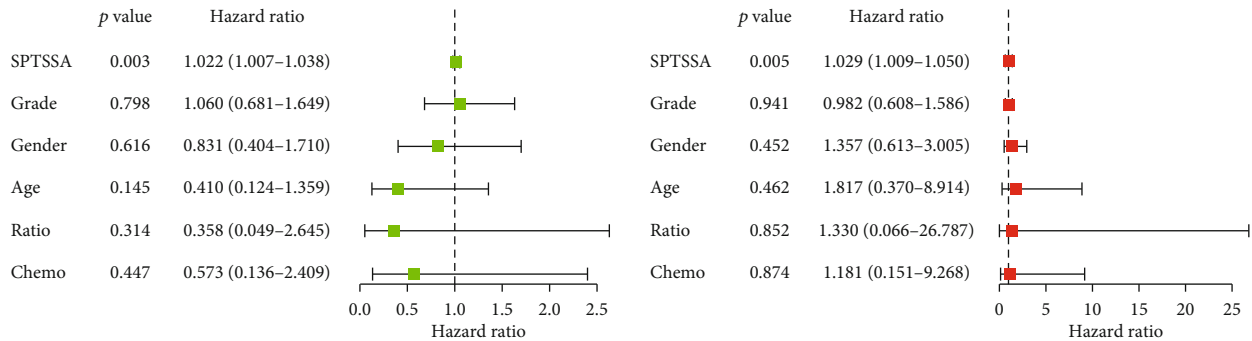


(c)

FIGURE 7: (a) IHC of grade II~IV glioma sample from the Affiliated Hospital of Nantong University. (a) Using TMA, grouped in median, high expression of SPTSSA predicts poor prognosis. (b, c) Univariate and multivariate Cox analyses indicated that SPTSSA was an independent predictor for OS.

SPTSSA is the gene encoding the small subunit A of serine palmitoyltransferase (SPT). It catalyzes the formation of sphingoid long-chain base backbone of sphingolipids [38]. Sphingolipids are structural compounds of biological membranes, and recent studies suggest sphingolipids can also serve as secondary messengers, participating in apoptosis, proliferation, senescence, angiogenesis, and vesicular trafficking [39, 40]. Because alterations in bioactive sphingolipids have been linked to cancer progression and prognosis, their key metabolic enzymes have been actively pursued as novel targets in cancer drug development [41–44]. The enrichment analysis showed that SPTSSA is related to the oxidative stress. Numerous studies have shown that the role of oxidative stress in glioma is quite important. To our best knowledge, this is the first study reporting the connection between SPTSSA expression, a key catalytic enzyme in sphingolipids synthesis, and cancer progression and prognosis.

To further shed light on the potential function of SPTSSA in GBM progression, we used both CIBERSORT and TIMER to correlate tumor-infiltrating immune cells with SPTSSA expression in GBM. Tumor-infiltrating immune cells are major member of the tumor microenvironment. They correlate with tumor prognosis and response to

therapy. Traditionally, immunohistochemistry [45–47] and flow cytometry [48] are used to analyze and enumerate different subsets of immune cells. However, these methods are limited by the availability of markers and antibodies for identification of subtypes of immune cells as well as high quality tumor samples. CIBERSORT is a type of in silico tissue dissection method for enumerating different cell fractions from undissected tissue gene expression profiles through computational deconvolution analysis [20, 49]. Using pure immune cell subtype expression profiles, CIBERSORT can accurately estimate the immune cell subtypes of a tumor biopsy and enable the discovery of biomarkers and novel immunotherapeutic targets. Using CIBERSORT, we showed that the numbers of CD4 T memory cells and macrophage cells were positively correlated with *SPTSSA* expression, while the number of activated mast cells were negatively correlated with *SPTSSA* expression.

The TIMER was used to analyze the association with the survival of the tumor-infiltrating immune cells, to confirm our observations further. Cancer biologists are enabled in quantifying the abundance of tumor-infiltrating immune cells through a flexible and comprehensive mode by the TIMER (Tumor Immune Estimation Resource) [21]. The abundance of tumor-infiltrating immune cells from the bulk gene expression profiles are identified using the TIMER computational deconvolution methods. The association with overall survival was provided together with the association of the tumor-infiltrating immune cell abundance with the gene expression by TIMER. We successfully identified that higher numbers of activated mast cells were associated with better survival in GBM, that higher numbers of macrophage M0/M1 cells were associated with poor survival in GBM, and that there was a correlation with the *SPTSSA* expression of the tumor-infiltrating dendritic cells, neutrophils, and the macrophages. The *SPTSSA* being an independent prognosis factor and dysregulated in glioma was confirmed from the IF and the IHC results.

## 5. Conclusion

In summary, using bioinformatics tools, we identified high-*SPTSSA* expression in GBM tissues, and high-*SPTSSA* expression was associated with poor survival. In silico tumor-infiltrating immune cell analysis suggests that high-*SPTSSA* expression was associated with high number of specific subtype immune cells. Future experimental studies are needed to explore the potential of *SPTSSA* as prognostic marker as well as novel immunotherapy target for GBM.

## Abbreviations

SPTSSA: The small subunit A of serine palmitoyltransferase
GEPIA: Gene Expression Profiling Interactive Analysis
CGGA: The Chinese Glioma Genome Atlas
TCGA: The Cancer Genome Atlas
GSEA: Gene Set Enrichment Analysis
TIMER: Tumor Immune Estimation Resource

IHC: Immunohistochemistry
GBM: Glioblastoma multiforme
TAMs: Tumor-associated macrophages
TINs: Tumor-infiltrating neutrophils
DEGs: Differentially expressed genes
GSEA: Gene Set Enrichment Analysis
NES: Normalized enrichment score
TMA: Tissue microarray
IF: Immunofluorescence.

## Data Availability

The datasets generated during and/or analyses during the current study are available in the Chinese Glioma Genome Atlas Network (CGGA).

## Ethical Approval

The experimental protocol was established, according to the ethical guidelines of the Helsinki Declaration. Ethical permissions were granted by the Human Ethics Committee at Affiliated Hospital of Nantong University (No. 2018-K020).

## Consent

Written informed consent was obtained from individual or guardian participants included in this study.

## Conflicts of Interest

All authors declare no competing nonfinancial/financial interests.

## Authors' Contributions

Ziheng Wang and Xinqi Ge contributed equally to this work.

## Acknowledgments

## Supplementary Materials

*Supplementary 1.* Supplementary Figure 1: (A) a nomogram showing SPTSSA and different clinical characteristics based on TCGA database. (B) Calibration curve of nomogram.

*Supplementary 2.* Supplementary Figure 2: Cox analysis of SPTSSA between (A) OS, (B) PFI, (C) DSS, and (D) DFI.

*Supplementary 3.* Supplementary Figure 3: (A) the immunotherapy response between SPTSSA expression level and anti-PD-1/CTLA4. (B) AUC curve of SPTSSA expression level in anti-PD-1/CTLA4 immunotherapy. (C) Relationships between SPTSSA and immune suppressive factors, immune promoting factors, MHC factors, chemokine, and receptors.

*Supplementary 4.* Supplementary Figure 4: connections between *SPTSSA* and genomic alteration in (A) GBM and (B) LGG.

*Supplementary 5.* Supplementary Figure 5: (A) correlations between SPTSSA and DNA methyltransferase. (B) The analysis of glioma tumor-infiltrating lymphocytes (TILs) with different SPTSSA expression.

# References

[1] A. Bohn, A. Braley, P. Rodriguez De La Vega, J. C. Zevallos, and N. C. Barengo, "The association between race and survival in glioblastoma patients in the US: a retrospective cohort study," *PLoS One*, vol. 13, no. 6, p. e0198581, 2018.

[2] J. P. Thakkar, T. A. Dolecek, C. Horbinski et al., "Epidemiologic and molecular prognostic review of glioblastoma," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 23, no. 10, pp. 1985–1996, 2014.

[3] A. M. Molinaro, J. W. Taylor, J. K. Wiencke, and M. R. Wrensch, "Genetic and molecular epidemiology of adult diffuse glioma," *Nature Reviews. Neurology*, vol. 15, no. 7, pp. 405–417, 2019.

[4] P. D. Delgado-Lopez and E. M. Corrales-Garcia, "Survival in glioblastoma: a review on the impact of treatment modalities," *Clinical & Translational Oncology*, vol. 18, no. 11, pp. 1062–1071, 2016.

[5] R. Batash, N. Asna, P. Schaffer, N. Francis, and M. Schaffer, "Glioblastoma multiforme, diagnosis and treatment; recent literature review," *Current Medicinal Chemistry*, vol. 24, no. 27, pp. 3002–3009, 2017.

[6] N. R. Smoll, K. Schaller, and O. P. Gautschi, "Long-term survival of patients with glioblastoma multiforme (GBM)," *Journal of Clinical Neuroscience*, vol. 20, no. 5, pp. 670–675, 2013.

[7] Q. T. Ostrom, H. Gittleman, P. Farah et al., "CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2006-2010," *Neuro Oncology*, vol. 15, Supplement 2, pp. ii1–ii56, 2013.

[8] N. P. Patel, K. A. Lyon, and J. H. Huang, "The effect of race on the prognosis of the glioblastoma patient: a brief review," *Neurological Research*, vol. 41, no. 11, pp. 967–971, 2019.

[9] Y. Yang, "Cancer immunotherapy: harnessing the immune system to battle cancer," *The Journal of Clinical Investigation*, vol. 125, no. 9, pp. 3335–3337, 2015.

[10] J. A. Trapani and P. K. Darcy, "Immunotherapy of cancer," *Australian Family Physician*, vol. 46, no. 4, pp. 194–199, 2017.

[11] A. Tivnan, T. Heilinger, E. C. Lavelle, and J. H. Prehn, "Advances in immunotherapy for the treatment of glioblastoma," *Journal of Neuro-Oncology*, vol. 131, no. 1, pp. 1–9, 2017.

[12] Z. Kong, Y. Wang, and W. Ma, "Vaccination in the immunotherapy of glioblastoma," *Human Vaccines & Immunotherapeutics*, vol. 14, no. 2, pp. 255–268, 2018.

[13] Y. Komohara, H. Horlad, K. Ohnishi et al., "Importance of direct macrophage-tumor cell interaction on progression of human glioma," *Cancer Science*, vol. 103, no. 12, pp. 2165–2172, 2012.

[14] A. Gieryng, D. Pszczolkowska, K. A. Walentynowicz, W. D. Rajan, and B. Kaminska, "Immune microenvironment of gliomas," *Laboratory Investigation*, vol. 97, no. 5, pp. 498–518, 2017.

[15] Q. Wang, B. Hu, X. Hu et al., "Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment," *Cancer Cell*, vol. 32, no. 1, pp. 42–56, 2017.

[16] Y. Kim, H. Jeon, and H. Othmer, "The role of the tumor microenvironment in glioblastoma: a mathematical model," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 3, pp. 519–527, 2017.

[17] 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.

[18] Z. Tang, B. Kang, C. Li, T. Chen, and Z. Zhang, "GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis," *Nucleic Acids Research*, vol. 47, no. W1, pp. W556–W560, 2019.

[19] W. Wang, Z. Zhao, F. Wu et al., "Bioinformatic analysis of gene expression and methylation regulation in glioblastoma," *Journal of Neuro-Oncology*, vol. 136, no. 3, pp. 495–503, 2018.

[20] B. Chen, M. S. Khodadoust, C. L. Liu, A. M. Newman, and A. A. Alizadeh, "Profiling tumor infiltrating immune cells with CIBERSORT," *Methods in Molecular Biology (Clifton, NJ)*, vol. 1711, pp. 243–259, 2018.

[21] Q. Jia, T. Xue, Q. Zhang et al., "CCN3 is a therapeutic target relating enhanced stemness and coagulation in hepatocellular carcinoma," *Scientific Reports*, vol. 7, no. 1, pp. 13846–13846, 2017.

[22] A. Subramanian, H. Kuehn, J. Gould, P. Tamayo, and J. P. Mesirov, "GSEA-P: a desktop application for Gene Set Enrichment Analysis," *Bioinformatics*, vol. 23, no. 23, pp. 3251–3253, 2007.

[23] Y. Sun, Y. Zhang, S. Ren et al., "Low expression of RGL4 is associated with a poor prognosis and immune infiltration in lung adenocarcinoma patients," *International Immunopharmacology*, vol. 83, p. 106454, 2020.

[24] H. Shen, Z. Wang, S. Ren et al., "Prognostic biomarker MITD1 and its correlation with immune infiltrates in hepatocellular carcinoma (HCC)," *International Immunopharmacology*, vol. 81, p. 106222, 2020.

[25] Y. Shi and T. Lammers, "Combining Nanomedicine and Immunotherapy," *Accounts of Chemical Research*, vol. 52, no. 6, pp. 1543–1554, 2019.

[26] A. Leivas, A. Perez-Martinez, M. J. Blanchard et al., "Novel treatment strategy with autologous activated and expanded natural killer cells plus anti-myeloma drugs for multiple myeloma," *Oncoimmunology*, vol. 5, no. 12, article e1250051, 2016.

[27] F. L. Boer, M. L. G. Ten Eikelder, E. H. Kapiteijn, C. L. Creutzberg, K. Galaal, and M. I. E. van Poelgeest, "Vulvar malignant melanoma: pathogenesis, clinical behaviour and management: review of the literature," *Cancer Treatment Reviews*, vol. 73, pp. 91–103, 2019.

[28] L. K. Vestergaard, D. N. P. Oliveira, C. K. Høgdall, and E. V. Høgdall, "Next generation sequencing technology in the clinic and its challenges," *Cancers (Basel)*, vol. 13, no. 8, p. 1751, 2021.

[29] Z. Huang, T. S. Johnson, Z. Han et al., "Deep learning-based cancer survival prognosis from RNA-seq data: approaches and evaluations," *BMC Medical Genomics*, vol. 13, Supplement 5, p. 41, 2020.

[30] K. Aasebø, A. Dragomir, M. Sundström et al., "Consequences of a high incidence of microsatellite instability and BRAF-

mutated tumors: a population-based cohort of metastatic colorectal cancer patients," *Cancer Medicine*, vol. 8, no. 7, pp. 3623–3635, 2019.

[31] J. Zhan, M. Niu, P. Wang et al., "Elevated HOXB9 expression promotes differentiation and predicts a favourable outcome in colon adenocarcinoma patients," *British Journal of Cancer*, vol. 111, no. 5, pp. 883–893, 2014.

[32] R. Dienstmann, G. Villacampa, A. Sveen et al., "Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer," *Annals of Oncology*, vol. 30, no. 10, pp. 1622–1629, 2019.

[33] R. Margueron and D. Reinberg, "Chromatin structure and the inheritance of epigenetic information," *Nature Reviews. Genetics*, vol. 11, no. 4, pp. 285–296, 2010.

[34] A. P. Feinberg and M. D. Fallin, "Epigenetics at the crossroads of genes and the environment," *Jama*, vol. 314, no. 11, pp. 1129-1130, 2015.

[35] J. Chen, O. Odenike, and J. D. Rowley, "Leukaemogenesis: more than mutant genes," *Nature Reviews. Cancer*, vol. 10, no. 1, pp. 23–36, 2010.

[36] L. H. Zhang, Z. Wang, L. H. Li et al., "Vestigial like family member 3 is a novel prognostic biomarker for gastric cancer," *World Journal of Clinical Cases*, vol. 7, no. 15, pp. 1954–1963, 2019.

[37] M. Tong, W. Lu, H. Liu et al., "Evaluation of MT family isoforms as potential biomarker for predicting progression and prognosis in gastric cancer," *BioMed Research International*, vol. 2019, Article ID 2957821, 15 pages, 2019.

[38] L. Zhao, S. Spassieva, K. Gable et al., "Elevation of 20-carbon long chain bases due to a mutation in serine palmitoyltransferase small subunit b results in neurodegeneration," *Proceedings of the National Academy of Sciences*, vol. 112, no. 42, pp. 12962–12967, 2015.

[39] Y. A. Hannun and L. M. Obeid, "Sphingolipids and their metabolism in physiology and disease," *Nature Reviews. Molecular Cell Biology*, vol. 19, no. 3, pp. 175–191, 2018.

[40] Y. Ilan, "Compounds of the sphingomyelin-ceramide-glycosphingolipid pathways as secondary messenger molecules: new targets for novel therapies for fatty liver disease and insulin resistance," *American Journal of Physiology-Gastrointestinal and Liver Physiology*, vol. 310, no. 11, pp. G1102–G1117, 2016.

[41] B. Ogretmen, "Sphingolipid metabolism in cancer signalling and therapy," *Nature Reviews. Cancer*, vol. 18, no. 1, pp. 33–50, 2018.

[42] B. Ogretmen, "Sphingolipids in cancer: regulation of pathogenesis and therapy," *FEBS Letters*, vol. 580, no. 23, pp. 5467–5476, 2006.

[43] N. Govindarajah, R. Clifford, D. Bowden, P. A. Sutton, J. L. Parsons, and D. Vimalachandran, "Sphingolipids and acid ceramidase as therapeutic targets in cancer therapy," *Critical Reviews in Oncology/Hematology*, vol. 138, pp. 104–111, 2019.

[44] D. Bottai, R. Adami, R. Paroni, and R. Ghidoni, "Brain cancer-activated microglia: a potential role for sphingolipids," *Current Medicinal Chemistry*, vol. 27, no. 24, pp. 4039–4061, 2019.

[45] 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.

[46] E. K. Nduom, J. Wei, N. K. Yaghi et al., "PD-L1 expression and prognostic impact in glioblastoma," *Neuro-Oncology*, vol. 18, no. 2, pp. 195–205, 2016.

[47] M. Rahman, J. Kresak, C. Yang et al., "Analysis of immunobiologic markers in primary and recurrent glioblastoma," *Journal of Neuro-Oncology*, vol. 137, no. 2, pp. 249–257, 2018.

[48] K. Gabrusiewicz, B. Rodriguez, J. Wei et al., "Glioblastoma-infiltrated innate immune cells resemble M0 macrophage phenotype," *JCI Insight*, vol. 1, no. 2, 2016.

[49] A. M. Newman and A. A. Alizadeh, "High-throughput genomic profiling of tumor-infiltrating leukocytes," *Current Opinion in Immunology*, vol. 41, pp. 77–84, 2016.

*Research Article*

# Machine Learning Assistants Construct Oxidative Stress-Related Gene Signature and Discover Potential Therapy Targets for Acute Myeloid Leukemia

**Jinhua Zhang,[1] Zhenfan Chen,[1] Fang Wang,[2,3] Yangbo Xi,[1] Yihan Hu,[1] and Jun Guo ⬝[1]**

[1]*The First Clinical Medical College of Jinan University, First Affiliated Hospital of Jinan University, Guangzhou 510630, China*
[2]*School of Life Science, Northwestern Polytechnical University, Xian, 710072, China*
[3]*Department of Hematology and Immunology, Myeloma Center Brussels, Vrije Universiteit Brussel, Brussels B-1090, Belgium*

Correspondence should be addressed to Jun Guo; guojun2009@jnu.edu.cn

*Background.* Oxidative stress (OS) is associated with the development of acute myeloid leukemia (AML). However, there is lack of relevant research to confirm that OS-related genes can guide patients in risk stratification and predict their survival probability. *Method.* First, we Data from three public databases, respectively. Then, we use batch univariate Cox regression and machine learning to select important characteristic genes; next, we build the model and use receiver operating characteristic curve (ROC) to evaluate the accuracy. Moreover, GSEAs were performed to discover the molecular mechanism and conduct nomogram visualization. In addition, the relative importance value was used to identify the hub gene, and GSE9476 was to validate hub gene difference expression. Finally, we use symptom mapping to predict the candidate herbs, targeting the hub gene, and put these candidate herbs into Traditional Chinese Medicine Systems Pharmacology (TCMSP) to identify the main small molecular ingredients and then docking hub proteins with this small molecular. *Results.* A total of 313 candidate oxidative stress-related genes could affect patients' outcomes and machine learning to select six potential genes to construct a gene signature model to predict the overall survival (OS) of AML patients. Patients in a high group will obtain a short survival time when compared with the low-risk group (HR = 3.97, 95% CI: 2.48-6.36; $p < 0.001$). ROC results demonstrate the model has better prediction efficiency with AUC 0.873. GSEA suggests that this gene is enriched in several important signaling pathways. Nomogram is constructed and is robust. PLA2G4A is a hub gene of signature and associated with prognosis, and Nobiletin could target PLA2G4A for therapy AML. *Conclusion.* We use two different machine learning methods to build six oxidative stress-related gene signatures that could assist clinical decisions and identify PLA2G4A as a potential biomarker for AML. Nobiletin, targeting PLA2G4, may provide a third pathway for therapy AML.

## 1. Introduction

Acute myeloid leukemia (AML) was derived from abnormal stem cell precursors of the myeloid lineage [1]. These proliferative clonal hematopoietic precursor cells damage the normal hematopoiesis and cause a series of clinical symptoms. Although AML is a rare disease compared with other newly diagnosed cancers, it accounts for more than 15% of acute leukemia with more than 30% mortality [2, 3]. Currently, molecular and cytogenetic features are recognised as key prognostic factors for the disease diagnosis. Hematopoietic

stem cell transplantation (HSCT) remains the only way to cure this disease, but the overall survival (OS) is still stagnant due to severe infection and acute graft-versus-host disease (aGVHD) after transplantation [3, 4]. More than 70% of patients receiving conventional chemotherapy will eventually relapse or become refractory leukemia [5]. Thus, it is of great importance to gain an understanding of the genetic variations in this disease and explore new targets for predicting prognosis and direct treatment.

Oxidative stress (OS) refers to a state of imbalance between reactive oxygen species (ROS) and antioxidant

(a)



(b)



TOP 10 important genes

| PLA2G4A | MMP7 | RCAN1 |
| MB | RYR1 | FKBP5 |
| PIK3CA | SIGMAR1 | NCF4 |
| SOCS1 | | |

(c)



(d)



(e)



Random forest results

28    6    9

Lasso regression results

Intersection of two machine learning analysis results
Venn diagram

(f)
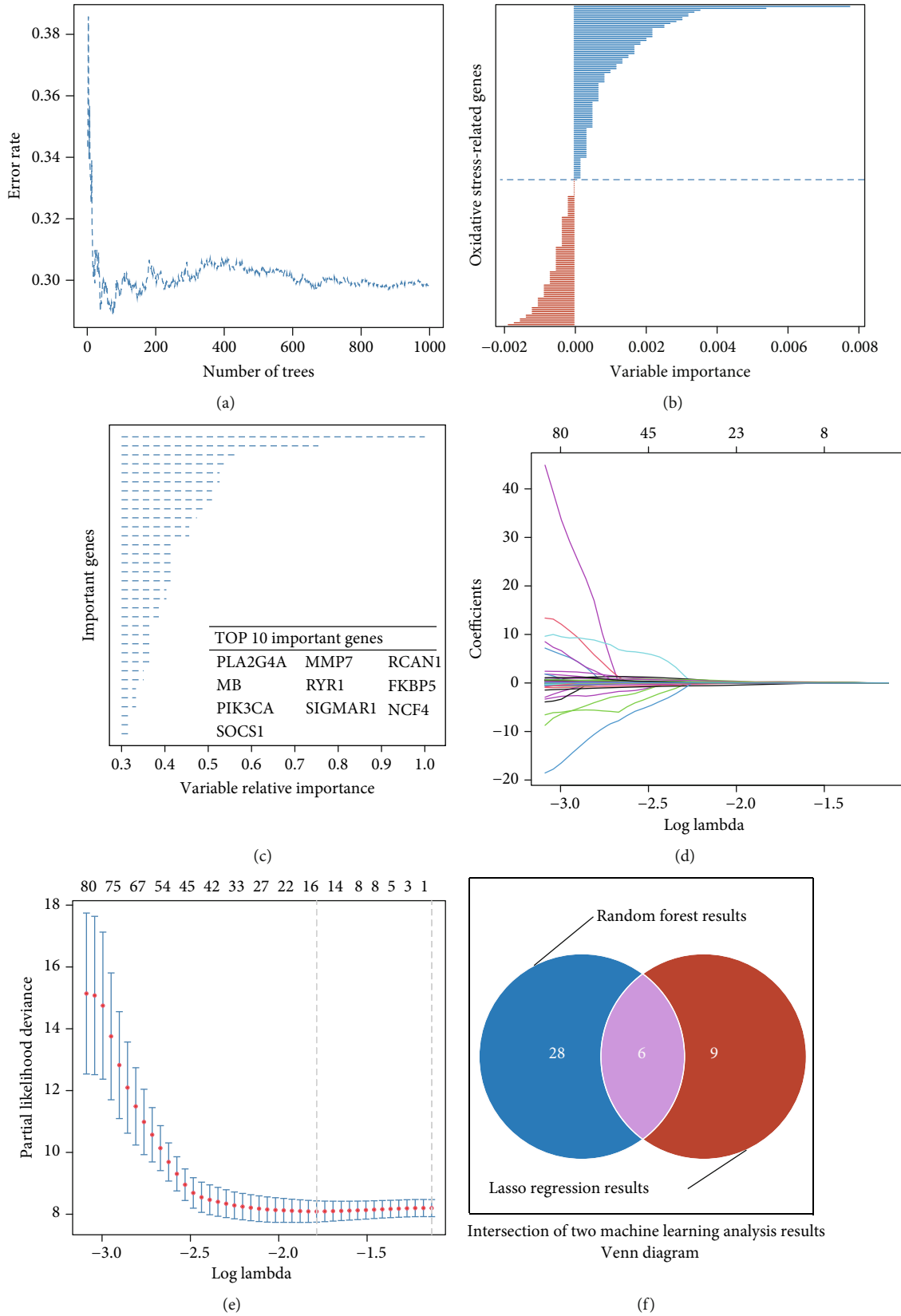
FIGURE 1: Machine learning to identify important oxidative stress-related genes for the prognosis model. (a, b) The error rate of the random forest model is 29.83%, and the better and poor prognosis genes are ordered by their importance. (c) Thirty-four prognosis-related genes are important more than 0.3. (d, e) Fifteen candidate genes are extracted from the expression profile by the lasso regression model. (f) Six potential genes were identified after the merge of the results of the above two different machine learning algorithms.

TABLE 1: Multivariate Cox regression for model genes.

| Gene symbol | Coef | HR | p value | 95% CI Lower | Upper |
|---|---|---|---|---|---|
| AGRN | -0.391 | 0.676 | 0.507 | 0.902 | 0.008 |
| ETFB | 0.827 | 2.289 | 1.051 | 4.985 | 0.037 |
| PLA2G4A | 0.236 | 1.266 | 0.965 | 1.661 | 0.088 |
| RYR1 | 0.650 | 1.917 | 1.239 | 2.966 | 0.003 |
| SIGMAR1 | 0.404 | 1.498 | 1.026 | 2.187 | 0.036 |
| SOCS1 | 0.473 | 1.605 | 1.177 | 2.189 | 0.003 |

effects in the body, which is characterized as inflammatory infiltration of neutrophils, increased secretion of proteases, and the production of a large number of oxidative compounds [6]. ROS have been confirmed to be associated with cellular signaling and gene expression in the normal cellular process [7, 8]. However, once the endogenous ROS is not adequately eliminated by the antioxidant system, the prooxidant/antioxidant balance is lost and leads to the occurrence of OS, which will damage the biological processes and the DNA repair mechanisms, leading to kinds of diseases and carcinogenesis, such as neurological disease, cardiovascular disease, breast cancer, prostate cancer, and hematologic malignancies [9–14]. On the other hand, ROS can undertake an opposite role in tumor therapy by inducing cell apoptosis [14]. Therefore, clarifying the dual role of OS in the pathogenesis and treatment of AML and selecting the potential beneficial population would lay the foundation for individualized precision treatment.

In the present work, we use two different machine learning methods to select candidate prognosis genes in oxidative stress-related gene sets and build a six-gene signature model to predict the AML patients' outcomes. In addition, we also identify PLA2G4A as a hub gene of signature and associated with prognosis and found that Nobiletin, a type of traditional Chinese medicine, targeting PLA2G4 may provide a third pathway for therapy AML.

## 2. Methods and Materials

*2.1. Data Obtaining and Prepared.* Public datasets were applied to this study: The Cancer Genome Atlas Program (TCGA), Genotype-Tissue Expression (GTEx), and Gene Expression Omnibus (GEO), respectively. We obtain AML data from TCGA database, which includes RNAseq records and clinical information of patients. Original count data are transferred to TPM style and extract 1399 oxidative stress-related genes, recorded by gene card database, from the expression profile to construct a new oxidative stress-related gene matrix for model building. Overall survival (OS) was defined as the endpoint. RNAseq data of donor bone marrow are from the GTEx database. GSE9476, including 38 donors and 26 AML samples, was used to validate hub gene differential expression between healthy hematopoietic cells and leukemic blasts. The 3D structure of hub protein and small-molecule structures are sourced from PDB and Pub Chem, respectively.

*2.2. Batch Univariate Cox Regression to Identify Prognosis-Related Genes.* Too many genes will affect patients' outcomes; here, we use batch univariate Cox regression to identify prognosis-related genes of oxidative stress. After analysis, genes with a p value less than 0.05 were identified as significant factors, and this gene will be input into the next model to perform dimensionality reduction.

*2.3. Machine Learning to Select Important Characteristic Genes.* Random forest and lasso regression are performed to identify important characteristic genes of oxidative stress. The importance of the random forest algorithm was defined as 0.3, and then, all of these requirement genes were put into the lasso regression model, which model set significant criterion was lambda is minimal.

*2.4. Predictive Model Construct and Validation.* For building the final prediction model, we use multivariate Cox regression to analyze the significant oxidative stress-related genes; as above standard, the p value is also set as less than 0.05. After selecting all of these requirement genes, we build the final model for the prediction of patient outcome, according to the regression coefficients. Each patient will obtain one risk score, patients will be divided into high- and low-risk groups, according to the median value. Survival analyses were conducted by log-rank test. Forty percent of the total data were set as a test dataset to validate robust of the above model. Area under ROC was used to evaluate the predictive accuracy.

*2.5. Difference Expression Genes between High-Risk and Low-Risk Groups.* To identify the differential expression genes between the high-risk and low-risk groups, the limma package was used to conduct this procedure, and the criteria of significant genes were set as absolute of log fold change more than 2 and p value less than 0.05.

*2.6. GSEA of Differential Expression Genes.* The molecular mechanism of differential expression genes between the high-risk and low-risk groups is unclear; here, we use the GSEA function, which provides by the cluster profile package, to do GO and KEGG pathway enrichment analysis. p value less than 0.05 was identified as a significant enrichment result.

*2.7. Nomogram Construction and Evaluation.* Nomogram is more than eyes and provides a convenient tool for clinical physicians to assist clinical decisions. Significant genes from the above multivariate Cox regression results will be considered and put into the VRPM package to conduct visualization. The area under the curve (AUC) of receiver operating characteristic (ROC) and calibration curve were used to evaluate the model's robustness.

*2.8. Hub Gene Selection and Validation.* In the gene panel, not all the genes play an important role in the model, so we select the importance value of the model gene and order it from high to low, and the biggest value of the important gene was identified as the hub gene. Donor patients' bone marrow, which obtains from the GTEx database, will be
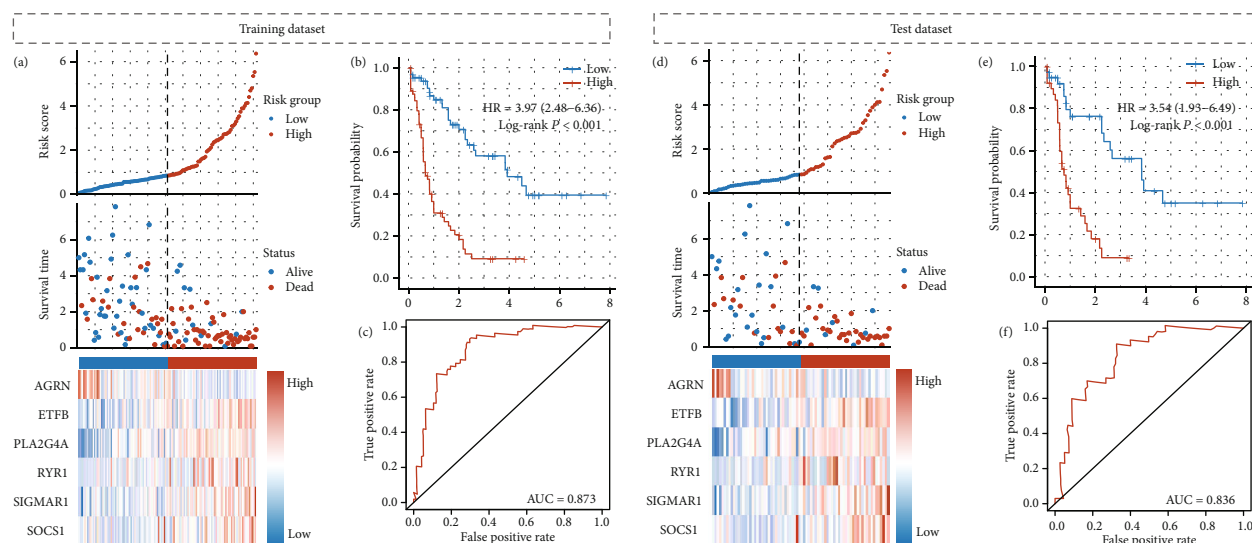
FIGURE 2: Six oxidative stress-related gene signature construction and internal validation. (a) Each patient will obtain a risk score, which is calculated by the model formula. (b) Patients in a high group will obtain a short survival time when compared with the low-risk group, and this difference is significant (HR = 3.97, 95% CI: 2.48-6.36; $p < 0.001$). (c) ROC demonstrates the model has better prediction efficiency with AUC is 0.873. (d–f) Internal validation results also support the above conclusions, and the AUC is 0.836.

compared with AML patients' bone marrow to validate hub gene difference expression. In addition, sample from GSE9476 also repeats the above operation. Moreover, survival analysis was performed to conduct by survminer package, to compare survival probability difference between high and low expression patients.

2.9. Screening Candidate Herbs Targeting Hub Protein. Traditional Chinese medicine has been confirmed that has the potential ability to target tumor markers. We use a symptom mapping (symMap Version 2.0) database to predict the candidate herbs, which will target the hub gene. Herbs with FDR less than 0.05 will be selected. The top 10 requirement herbs will be selected and also checked by the previous literature, which has been reported to have the ability of anticancer function for preparing docking with hub proteins. Candidate herbs will be input into Traditional Chinese Medicine Systems Pharmacology (TCMSP) to identify the main small molecular ingredients, according to oral bioavailability (OB) more than 30 and drug-likeliness (DL) more than 0.18.

2.10. Docking Structures between Proteins and Small Molecular Drugs. Before docking both structures, we need prepared ligand and receptor structures. So, we download the hub protein's 3D structure from the protein data bank (PDB; https://www.rcsb.org/) and obtain small molecular drug structure form Pub Chem database (https://pubchem.ncbi.nlm.nih.gov/), respectively. Then, an online tool will conduct a docking program (https://cadd.labshare.cn/cb-dock2/php/index.php).

## 3. Results

3.1. Prognosis Genes in Oxidative Stress-Related Gene Set of AML. We use a batch univariate regression model to filter no significant genes, which could not affect patients' out-comes, and some prognosis-related genes in the oxidative stress-related gene set were selected. The results of the batch univariate regression model show that a total of 313 candidates' oxidative stress-related genes could affect patients' outcomes (Supplement Table 1).

3.2. Machine Learning to Select Candidate Model Genes. Random forest and lasso regression models were used to select candidate model genes. On the one hand, we put the expression matrix to the random forest model. The results show that the error rate of the random forest model is 29.83%, and the better and poor prognosis genes are ordered by importance (Figures 1(a) and 1(b)). When we set variable relative importance to more than 0.3, 34 prognosis-related genes are selected, and the top 10 significant are shown in Figure 1(c) (Supplement Table 2). On the other hand, the above expression matrix is also put into the lasso regression model, when the model selects the minimal lambda value; 15 candidate genes are extracted from the expression profile (Figures 1(d) and 1(e)). Then, we merge the results of the above two different machine learning algorithms. Six potential genes were identified to construct a gene signature model (Figure 1(f)).

3.3. Six-Gene Signature Could Predict the OS of AML Patients. After multivariate Cox regression model analysis, we find six genes from machine learning methods are included in the model (Table 1). So, in the next step, we build a six-gene signature model to predict the OS of AML patients, according to the coefficient of multivariate Cox regression. After building the signature model, every patient will obtain a risk score, which calculated by the model formula, risk score = $(-0.391) \times$ AGRN + $(0.827) \times$ ETFB + $(0.236) \times$ PLA2G4A + $(0.650) \times$ RYR1 + $(0.404) \times$ SIGMAR1 + $(0.473) \times$ SOCS1. After the count, patients will be divided into low-risk and high-risk groups, based on median
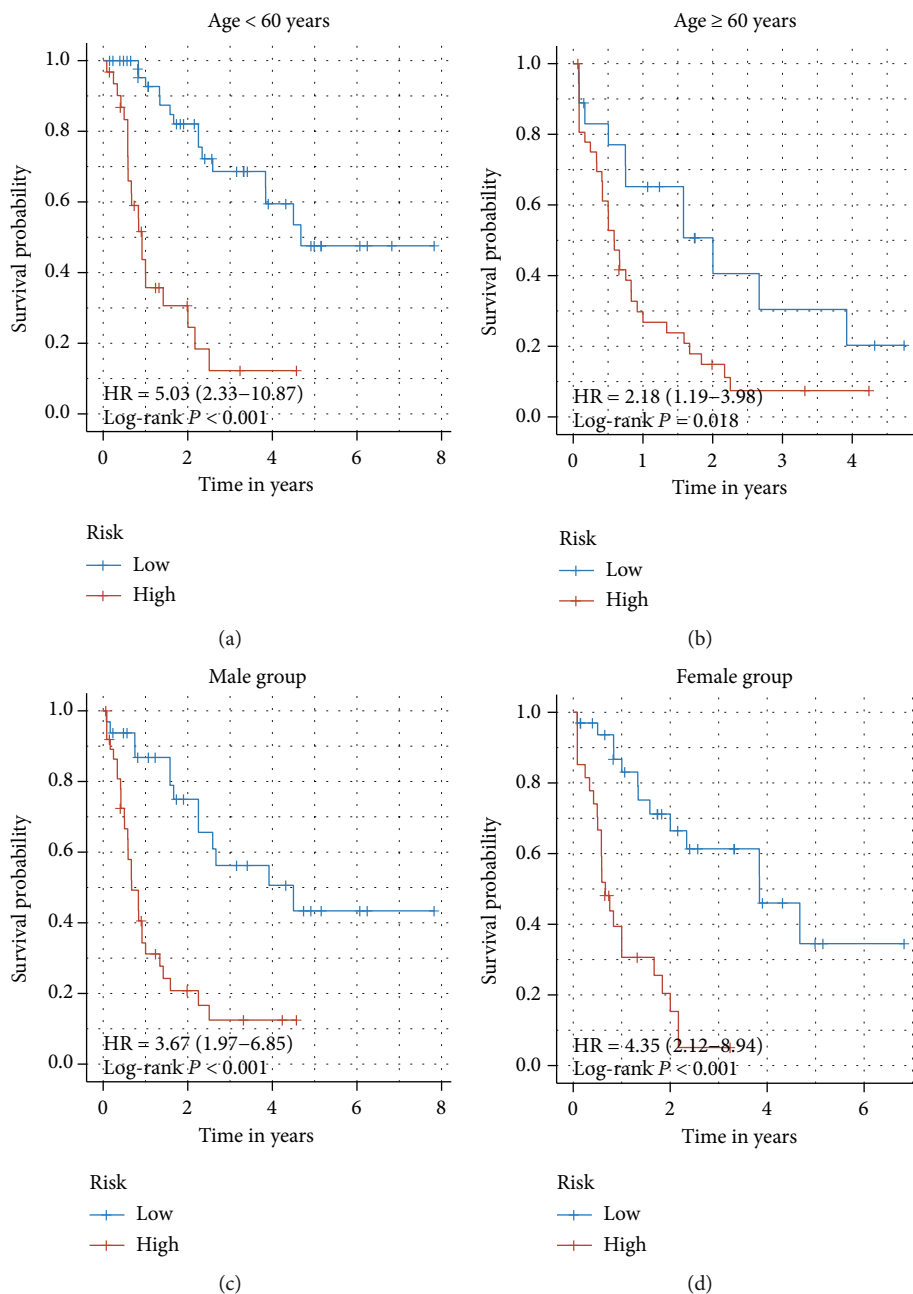
Figure 3: Signature with clinical variables. (a) In the age subgroup, patients under 60 years old, the high-risk score means shorter survival time, when compared with the low-risk group (HR = 5.03, 95% CI: 2.33-10.87; $p < 0.001$). (b) This conclusion also confirmed by patients more than 60 years old (HR = 2.18, 95% CI: 1.19-3.98; $p = 0.018$). (c, d) As for the sex subgroup, patients with low-risk scores always represent a better prognosis, when compared with a high-risk score, no matter which sex they are (HR = 3.67, 95% CI: 1.97-6.85; $p < 0.001$ vs. HR = 4.35, 95% CI: 2.12-8.94; $p < 0.001$, respectively).

value (Figure 2(a)). Figure 2(b) shows that patients in the high group will obtain a short survival time when compared with the low-risk group, and this difference is significant (HR = 3.97, 95% CI: 2.48-6.36; $p < 0.001$). We use ROC to evaluate the prediction accuracy of the model, and the results demonstrate the model has better prediction efficiency with AUC is 0.873 (Figure 2(c)). Internal validation results also support the above conclusions. Patients with low-risk scores mean longer living times when compared with high-risk score patients. In addition, the ROC of the

validation dataset also shows the model could predict patients' outcome accuracy with AUC equal to 0.836 (Figures 2(d)–2(f)).

3.4. Six-Gene Signature with Clinical Factors. Age and sex are both important clinical characteristics for AML patients. Here, we perform survival analysis to discover the difference between clinical subgroups. In the age subgroup, we find that in patients under 60 years old, the high-risk score means a shorter survival time, when compared with the low-risk
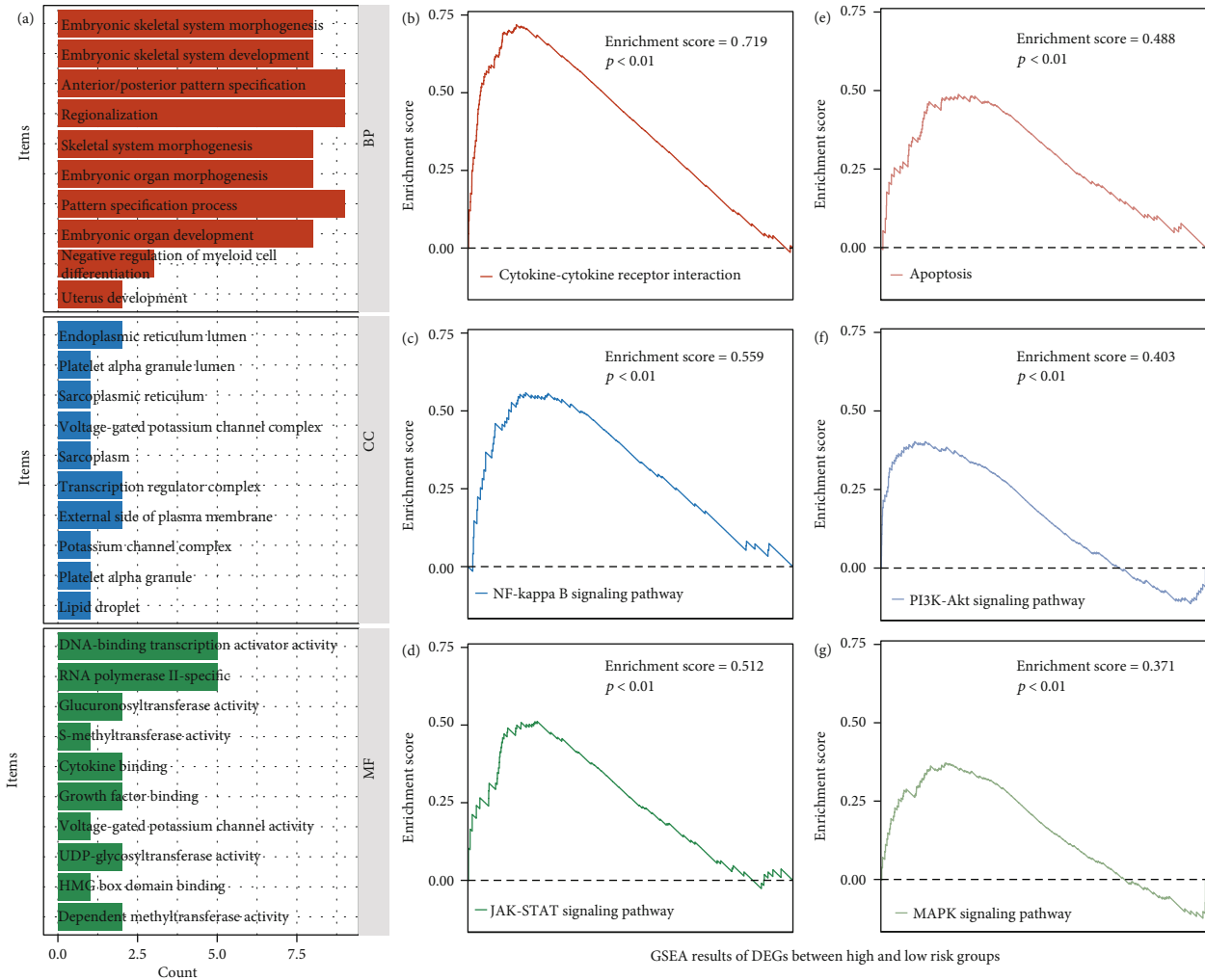
FIGURE 4: GSEA results of differential expression genes. (a) GSEA shows that this gene is enriched in embryonic skeletal system morphogenesis, endoplasmic reticulum lumen, and RNA polymerase II-specific, and the (b–g) involved pathways are cytokine-cytokine receptor interaction pathway, NF-kappa B signaling pathway, JAK-STAT signaling pathway, apoptosis, PI3K-AKT signaling pathway, and MAPK signaling pathway.

group (HR = 5.03, 95% CI: 2.33-10.87; $p < 0.001$) (Figure 3(a)). This conclusion also confirmed by patients more than 60 years old (HR = 2.18, 95% CI: 1.19-3.98; $p = 0.018$) (Figure 3(b)). As for the sex subgroup, patients with low-risk scores always represent have a better prognosis, no matter which sex they are (HR = 3.67, 95% CI: 1.97-6.85; $p < 0.001$ vs. HR = 4.35, 95% CI: 2.12-8.94; $p < 0.001$, respectively) (Figures 3(c) and 3(d)).

### 3.5. Different Expression Genes and Enrichment Analysis.
Patients with different risk scores have different prognoses. Identifying different expression genes between two groups is good to discover the molecular mechanism in the future. The different expression analyses demonstrate a total of 49 different expression genes between low- and high-risk groups (Supplement Table 3). After doing GSEA, we found that this gene is enriched in embryonic skeletal system morphogenesis, endoplasmic reticulum lumen, and RNA polymerase II-specific (Figure 4(a)), and the involved

pathways are cytokine-cytokine receptor interaction pathway, NF-kappa B, PI3K-AKT, and MAPK signaling pathway (Figures 4(b)–4(g)).

### 3.6. Nomogram Is a Useful Tool for Assistant Clinical Decision.
We have demonstrated that six gene signatures could predict patients' outcome accuracy, so we build a nomogram, based on six gene expressions, to assist clinical decisions. This model is shown in Figure 5(a), and according to six gene expressions, patients will obtain six score values and accumulate six values to become one total score and projection onto the survival axis to obtain patients' 1-year, 3-year, and 5-year survival probability, respectively. The ROC of the model is 0.761, and the C index of this nomogram is 0.774. Calibration curve analysis results suggest that survival prediction results of 1-year, 3 year, and 5-year survival probability were close to the ideal line (Figure 5(b)).

### 3.7. PLA2G4A Is a Hub Gene of Signature and Associated with Prognosis.
We extract the relative importance of six

(a)



(b)
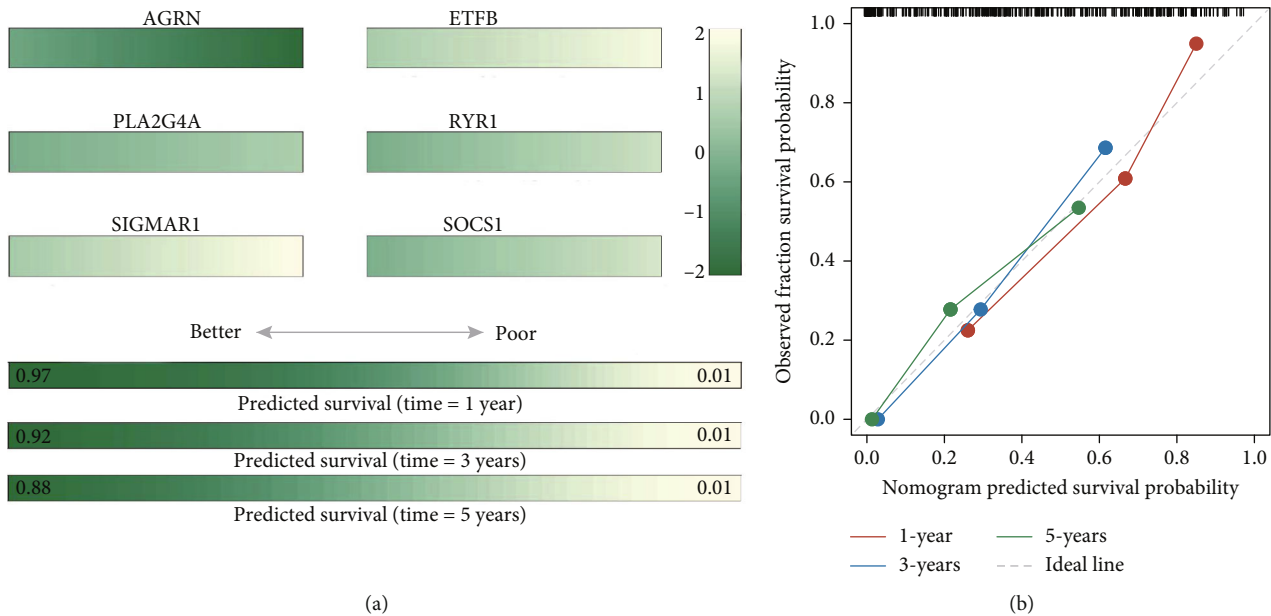
FIGURE 5: Nomogram for six model genes. (a) Nomogram, based on six-gene expression, to an assistant clinical decision, according to six-gene expression, patients will obtain six score values and accumulate six values to become one total score and projection onto the survival axis to obtain patients' 1-year, 3-year, and 5-year survival probability, respectively. (b) Calibration curve analysis results suggest that survival prediction probability was close to the ideal line.

model genes, and the PLA2G4A has the biggest value, so it was confirmed as the model hub gene (Figure 6(a)); to validate the potential value, we use external data set, GTEx, and GSE9476, to observe the expression difference between donor bone marrow and leukemic blasts from AML patients. The result demonstrates that PLA2G4 is a high expression in leukemic blasts and low expression in healthy hematopoietic cells ($p = 8.9e - 34$ vs. $p = 5.5e - 06$, respectively) (Figures 6(b) and 6(c)). Survival analysis also shows that high expression of this gene will lead to a poor outcome when compared with low expression patients (HR = 3.03, 95% CI: 1.97-4.67; $p < 0.001$) (Figure 6(d)).

*3.8. Nobiletin Targeting PLA2G4A Provides a Third Pathway for Therapy AML.* A total of 41 required herbs had been predicted, and the top 10 herbs are listed in Figure 6(e) and Supplement Table 4. Of these ten herbs, Zhiqiao has been reported to have a potential function as an anticancer. To identify which ingredients are important for these herbs, we input in into the TCMSP database. The results show that Hesperetin, Nobiletin, Naringenin, Marmin, and Beta-sitosterol are the main components of this drug (Table 2). During these components, Nobiletin has the best OB (61.67%) and highest DL (0.52); it was selected as a candidate small molecular drug to target PLA2G4A. The docking results also demonstrate the above conclusion (Figure 6(f)).

## 4. Discussion

Although advances have been achieved in the therapeutic options which greatly improve the overall response rate, the long-term prognosis of this disease remains dismal, espe-

cially among elder patients [15]. Besides, the complex molecular and cytogenetic abnormalities make AML a kind of heterogeneous disease with differential prognosis even in the same risk group by clinical practice guidelines [16]. All these revealed that insight into the genetic landscape of AML would benefit more patients. The development of various sequencing technologies has provided more information on the mechanism of pathogenesis, chemoresistance, and more refined prognostic stratification of AML in the past decades. Recently, Mer et al. [15] proposed a unique subtype of NPM1-mutated AML with different biological and therapeutical implications based on a stem cell signature. A set of mitochondrial metabolism proteins was also identified as potential targets associated with leukemia progression by multiomics [17]. Furthermore, some gene-based signatures have been constructed to predict the prognosis of AML as in other cancers [18, 19]. All these have brought new opportunities for the treatment of AML.

The maintenance of the quiescent state of hematopoietic stem cells (HSCs) depends on a condition of anaerobic glycolysis with low ROS generation, while compelling evidence has indicated that leukemia stem cells (LSCs), which are considered the main part of drug resistance, are more dependent on oxidative respiration with high ROS levels companied by an imbalance of oxidative and antioxidant, which promote the progression of leukemia by activating the pathways involved in the cell proliferation, survival, and invasion [20, 21]. Previous studies have proved that the redundant ROS could be a risk factor for tumorigenesis and the drug resistance role of ROS in varied leukemia modes [22, 23]. Interestingly, LSCs are more susceptible to external antioxidants, and ROS and lipid peroxidation by-products can trigger cell apoptosis, which also brings new chemotherapy

(a)



(b)



(c)



(d)

| | ID | Name | FDR |
|---|---|---|---|
| HERBs | SMHB00146 | Gua di | <0.001 |
| | SMHB00190 | Huo ma ren | 0.007 |
| | SMHB00102 | Di fuzi | 0.011 |
| | SMHB00485 | Zhi qiao | 0.017 |
| | SMHB00258 | Lu lu tong | 0.018 |
| | SMHB00181 | Hua ju hong | 0.021 |
| | SMHB00399 | Tian gua zi | 0.021 |
| | SMHB00323 | Qie geng | 0.025 |
| | SMHB00391 | Tian dong | 0.026 |
| | SMHB00008 | Ba qia | 0.028 |

Source from sym map database
(HerbsTarget PLA2G4A)

(e)



Source from nobiletin and PDB database
(docking of Nobiletin with PLA2G4A protein)

(f)

Figure 6: PLA2G4A is a candidate therapeutic target for AML. (a) PLA2G4A has the biggest value of relative importance in six model genes, and (b, c) it is a high expression in leukemic blasts when compared with healthy hematopoietic cells ($p = 8.9e - 34$ vs. $p = 5.5e - 06$, respectively). (d) High expression of this gene will lead to a poor outcome when compared with low expression patients (HR = 3.03, 95% CI: 1.97-4.67; $p < 0.001$). (e) The top 10 herbs target PLA2G4A; (f) docking Nobiletin, the main components of this Zhiqiao, with PLA2G4A protein.

TABLE 2: Ingredients of Chinese traditional medicine Zhiqiao.

| Mol ID | Molecule name | OB (%) | DL | MW | A log$p$ | FASA- | HL |
|---|---|---|---|---|---|---|---|
| MOL002341 | Hesperetin | 70.31 | 0.27 | 302.3 | 2.28 | 0.33 | 15.78 |
| MOL005828 | Nobiletin | 61.67 | 0.52 | 402.43 | 3.04 | 0.13 | 16.2 |
| MOL004328 | Naringenin | 59.29 | 0.21 | 272.27 | 2.3 | 0.4 | 16.98 |
| MOL013381 | Marmin | 38.23 | 0.31 | 332.43 | 3.11 | 0.31 | 4.68 |
| MOL000358 | Beta-sitosterol | 36.91 | 0.75 | 414.79 | 8.08 | 0.23 | 5.36 |

options. These all suggest a bidirectional role of ROS in leukemia [24]. Given the above complex mechanism, we aim to construct an accurate model which can provide precise predictivity and guide stratification therapy for clinical application.

As shown in our work, we used machine learning to select six hub OS-related genes, which demonstrated robust predictive ability in AML populations. Agrin (AGRN) has been described as a multifunctional heparan sulfate proteoglycan, which can regulate angiogenesis and has a board-ranging impact on the tumor microenvironment (TME) in HCC and papillary thyroid carcinoma (PTC) [25, 26]. However, nothing has been reported on the exact function of AGRN in AML. Electron transfer flavoprotein $\beta$ subunit (ETFB) can transfer the electrons to the electron transport chain (ETC) and maintain the generation of ATP. Caplan et al. [17] recently found that increased mitochondrial stress and apoptosis in AML mouse models can be induced by silencing ETFB, which suggests that ETFB could be a potential therapeutic target for AML. Moreover, the placental phospholipase A24A (PLA2G4A), sigma 1 receptor (Sig1R), and suppressors of cytokine signaling 1 (SOCS1) have been reported in varied diseases which are involved in the stress response biological procedure [27–29]. However, to the best of our knowledge, abscisic acid receptor (PYR1) has been shown as a stomatal regulation response to drought stress in plants, and no studies regarding abscisic acid receptor (PYR1) have been reported in human diseases or cancers [30]. Among these six genes, PLA2G4A has the most important in our model, which is one of the cytosolic placental phospholipases A2 (cPLA2) family and can catalyze the hydrolysis of membrane phospholipids to release arachidonic acid (AA) and lysophospholipid. It has been identified as an index in response to oxidative stress in preeclampsia and might be due to the oxidation of AA [27]. Higher expression of PLA2G4A is positively correlated with the migration and invasion of lung cancer cells and unfavorable prognosis in breast cancers [31, 32]. Previous studies also revealed that PLA2G4A expression could be an independent diagnostic and prognostic marker in patients with non-M3/NPM1 WT AML patients [33], which was also confirmed in our study. Nevertheless, whether this differential prognosis is caused by PLA2G4A through oxidative stress remains to be further investigated.

To explore the specific mechanism of OS-related genes in AML, we carried out functional annotation of DEGs between the high and low gene expression groups. The OS-related genes were enriched in the cytokine-cytokine receptor interaction pathway, NF-kappa B signaling pathway, JAK-STAT signaling pathway, apoptosis, PI3K-AKT signaling pathway, and MAPK signaling pathway as revealed by the GSEA results, which have been identified by previous studies serving as an important role in the pathogenesis and progression of AML. The PI3K/AKT pathway was proved to play important roles in regulating cell proliferation, differentiation, apoptosis, and migration in kinds of human diseases and cancers, such as diabetes, colorectal cancer, and AML [34, 35]. Some scholars also reported the PI3K/AKT pathway is associated with oxidative stress-mediated survival of melanoma and when targeting the PI3K/AKT and MAPK/ERK signaling pathway exerts an anticancer effect in leukemia cells by induction of oxidative stress and the cellular antioxidant defense mechanisms, which suggest PI3K/AKT and MAPK/ERK signaling pathway might involve in the leukemia cell apoptosis caused by oxidative stress [36, 37].

However, there are some limitations to our study. First, we did not distinguish between OS-related genes that promote leukemia proliferation and invasion with genes that mediate leukemia cell apoptosis via chemotherapy-induced OS. More datasets with pre/posttreatment information need to be included to clarify this bidirectional effect of OS. Second, our results need to be validated in a clinical trial in the further.

## 5. Conclusion

Our six oxidative stress-related gene signatures could predict AML patients' outcome accuracy, and this model is robust. It may become a useful tool to assist clinical decisions. In addition, we identify PLA2G4A as a potential biomarker for AML. Nobiletin, targeting PLA2G4, may provide a third pathway for therapy AML.

## Data Availability

All data used in the study are available in The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases.

## Conflicts of Interest

The authors declare no potential conflict of interest.

## Authors' Contributions

## Acknowledgments

## Supplementary Materials

Supplement Table 1: prognosis-related oxidative stress genes by filter batch univariate Cox regression. Supplement Table 2: genes with a relative importance of more than 0.3 in the random forest model. Supplement Table 3: different expression genes between low- and high-risk groups. Supplement Table 4: candidate herbs targeting PLA2G4A protein. *(Supplementary Materials)*

## References

[1] J. E. Rubnitz, B. Gibson, and F. O. Smith, "Acute myeloid leukemia," *Pediatric Clinics of North America*, vol. 55, no. 1, pp. 21–51, 2008.

[2] Y. Xie, S. M. Davies, Y. Xiang, L. L. Robison, and J. A. Ross, "Trends in leukemia incidence and survival in the United States (1973–1998)," *Cancer: Interdisciplinary International Journal of the American Cancer Society*, vol. 97, no. 9, pp. 2229–2235, 2003.

[3] A. Pelcovits and R. Niroula, "Acute myeloid leukemia: a review," *Rhode Island Medical Journal*, vol. 103, no. 3, pp. 38–40, 2020.

[4] A. Burnett, M. Wetzler, and B. Lowenberg, "Therapeutic advances in acute myeloid leukemia," *Journal of Clinical Oncology*, vol. 29, no. 5, pp. 487–494, 2011.

[5] X.-X. Chen, Z. P. Li, J. H. Zhu, H. T. Xia, and H. Zhou, "Systematic analysis of autophagy-related signature uncovers prognostic predictor for acute myeloid leukemia," *DNA and Cell Biology*, vol. 39, no. 9, pp. 1595–1605, 2020.

[6] H. Sies and D. P. Jones, "Reactive oxygen species (ROS) as pleiotropic physiological signalling agents," *Nature Reviews Molecular Cell Biology*, vol. 21, no. 7, pp. 363–383, 2020.

[7] R. Gerschman, D. L. Gilbert, S. W. Nye, P. Dwyer, and W. O. Fenn, "Oxygen poisoning and x-irradiation: a mechanism in common," *Science*, vol. 119, no. 3097, pp. 623–626, 1954.

[8] M. Valko, D. Leibfritz, J. Moncol, M. T. D. Cronin, M. Mazur, and J. Telser, "Free radicals and antioxidants in normal physiological functions and human disease," *The International Journal of Biochemistry & Cell Biology*, vol. 39, no. 1, pp. 44–84, 2007.

[9] J. Clerkin, R. Naughton, C. Quiney, and T. G. Cotter, "Mechanisms of ROS modulated cell survival during carcinogenesis," *Cancer Letters*, vol. 266, no. 1, pp. 30–36, 2008.

[10] S. C. Bondy, "The relation of oxidative stress and hyperexcitation to neurological disease," *Proceedings of the Society for Experimental Biology and Medicine*, vol. 208, no. 4, pp. 337–345, 1995.

[11] S. V. Lakshmi, *Oxidative Stress in Cardiovascular Disease*, 2009.

[12] N. S. Brown and R. Bicknell, "Hypoxia and oxidative stress in breast cancer oxidative stress-its effects on the growth, metastatic potential and response to therapy of breast cancer," *Breast Cancer Research*, vol. 3, no. 5, pp. 323–327, 2001.

[13] G. Gupta-Elera, A. R. Garrett, R. A. Robison, and K. L. O'Neill, "The role of oxidative stress in prostate cancer," *European Journal of Cancer Prevention*, vol. 21, no. 2, pp. 155–162, 2012.

[14] U. K. Udensi and P. B. Tchounwou, "Dual effect of oxidative stress on leukemia cancer induction and treatment," *Journal of Experimental & Clinical Cancer Research*, vol. 33, no. 1, pp. 106–115, 2014.

[15] A. S. Mer, E. M. Heath, S. A. Madani Tonekaboni et al., "Biological and therapeutic implications of a unique subtype of NPM1 mutated AML," *Nature Communications*, vol. 12, no. 1, pp. 1–13, 2021.

[16] M. Bienz, M. Ludwig, B. U. Mueller et al., "Risk assessment in patients with acute myeloid leukemia and a normal karyotype," *Clinical Cancer Research*, vol. 11, no. 4, pp. 1416–1424, 2005.

[17] M. Caplan, K. J. Wittorf, K. K. Weber et al., "Multi-omics reveals mitochondrial metabolism proteins susceptible for drug discovery in AML," *Leukemia*, vol. 36, no. 5, pp. 1296–1305, 2022.

[18] W.-L. Chen, J. H. Wang, A. H. Zhao et al., "A distinct glucose metabolism signature of acute myeloid leukemia with prognostic value," *Blood, The Journal of the American Society of Hematology*, vol. 124, no. 10, pp. 1645–1654, 2014.

[19] W. Shi, D. Hu, S. Lin, and R. Zhuo, "Five-mRNA signature for the prognosis of breast cancer based on the ceRNA network," *BioMed Research International*, vol. 2020, 17 pages, 2020.

[20] U. Testa, C. Labbaye, G. Castelli, and E. Pelosi, "Oxidative stress and hypoxia in normal and leukemic stem cells," *Experimental Hematology*, vol. 44, no. 7, pp. 540–560, 2016.

[21] C. L. Bigarella, R. Liang, and S. Ghaffari, "Stem cells and the impact of ROS signaling," *Development*, vol. 141, no. 22, pp. 4206–4218, 2014.

[22] Y.-F. Chen, H. Liu, X. J. Luo et al., "The roles of reactive oxygen species (ROS) and autophagy in the survival and death of leukemia cells," *Critical Reviews in Oncology/Hematology*, vol. 112, pp. 21–30, 2017.

[23] G. Bossis, J. E. Sarry, C. Kifagi et al., "The ROS/SUMO axis contributes to the response of acute myeloid leukemia cells to chemotherapeutic drugs," *Cell Reports*, vol. 7, no. 6, pp. 1815–1823, 2014.

[24] M. E. Irwin, N. Rivera-Del Valle, and J. Chandra, "Redox control of leukemia: from molecular mechanisms to therapeutic opportunities," *Antioxidants & Redox Signaling*, vol. 18, no. 11, pp. 1349–1383, 2013.

[25] C.-C. Wu, J. D. Lin, J. T. Chen et al., "Integrated analysis of fine-needle-aspiration cystic fluid proteome, cancer cell secretome, and public transcriptome datasets for papillary thyroid cancer biomarker discovery," *Oncotarget*, vol. 9, no. 15, pp. 12079–12100, 2018.

[26] S. Chakraborty, M. Lakshmanan, H. L. F. Swa et al., "An onco-genic role of Agrin in regulating focal adhesion integrity in hepatocellular carcinoma," *Nature Communications*, vol. 6, no. 1, pp. 1–16, 2015.

[27] M. Brien, J. Larose, K. Greffard, P. Julien, and J. F. Bilodeau, "Increased placental phospholipase $A_2$ gene expression and free $F_2$-isoprostane levels in response to oxidative stress in pre-eclampsia," *Placenta*, vol. 55, pp. 54–62, 2017.

[28] J. Wang, A. Saul, P. Roon, and S. B. Smith, "Activation of the molecular chaperone, sigma 1 receptor, preserves cone function in a murine model of inherited retinal degeneration," *Proceedings of the National Academy of Sciences*, vol. 113, no. 26, pp. E3764–E3772, 2016.

[29] X. Liu, M. G. Mameza, Y. S. Lee et al., "Suppressors of cytokine-signaling proteins induce insulin resistance in the retina and promote survival of retinal cells," *Diabetes*, vol. 57, no. 6, pp. 1651–1658, 2008.

[30] Z. Jin, S. Xue, Y. Luo et al., "Hydrogen sulfide interacting with abscisic acid in stomatal regulation responses to drought stress in Arabidopsis," *Plant Physiology and Biochemistry*, vol. 62, pp. 41–46, 2013.

[31] W. Zhang, X. Wang, L. Zhang et al., "Inhibition of PLA2G4A reduces the expression of lung cancer-related cytokines," *DNA and Cell Biology*, vol. 37, no. 12, pp. 1076–1081, 2018.

[32] S. Mishra, M. Charan, R. K. Shukla et al., "cPLA2 blockade attenuates S100A7-mediated breast tumorigenicity by inhibiting the immunosuppressive tumor microenvironment," *Journal of Experimental & Clinical Cancer Research*, vol. 41, no. 1, pp. 1–19, 2022.

[33] H. Bai, M. Zhou, M. Zeng, and L. Han, "PLA2G4A is a potential biomarker predicting shorter overall survival in patients with non-M3/NPM1 wildtype acute myeloid leukemia," *DNA and Cell Biology*, vol. 39, no. 4, pp. 700–708, 2020.

[34] M. Y. Follo, L. Manzoli, A. Poli, J. A. McCubrey, and L. Cocco, "PLC and PI3K/Akt/mTOR signalling in disease and cancer," *Advances in biological regulation*, vol. 57, pp. 10–16, 2015.

[35] S. A. Danielsen, P. W. Eide, A. Nesbakken, T. Guren, E. Leithe, and R. A. Lothe, "Portrait of the PI3K/AKT pathway in colo-rectal cancer," *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, vol. 1855, no. 1, pp. 104–121, 2015.

[36] H. G. Hambright, P. Meng, A. P. Kumar, and R. Ghosh, "Inhibition of PI3K/AKT/mTOR axis disrupts oxidative stress-mediated survival of melanoma cells," *Oncotarget*, vol. 6, no. 9, pp. 7195–7208, 2015.

[37] E. Jasek-Gajda, H. Jurkowska, M. Jasińska, and G. J. Lis, "Targeting the MAPK/ERK and PI3K/AKT signaling pathways affects NRF2, Trx and GSH antioxidant systems in leukemia cells," *Antioxidants*, vol. 9, no. 7, p. 633, 2020.

*Research Article*

# Systematic Oxidative Stress Indexes Associated with the Prognosis in Patients with T Lymphoblastic Lymphoma/Leukemia

**Liqin Ping** (ORCID)**, Yan Gao, Yanxia He, Xiaoxiao Wang, Bing Bai, Cheng Huang, and Huiqiang Huang** (ORCID)

*Department of Medical Oncology, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, Guangdong, China*

Correspondence should be addressed to Huiqiang Huang; huanghqsysucc@163.com

*Background.* T lymphoblastic lymphoma/leukemia (T-LBL/ALL) is an aggressive malignant tumor with 5-year overall survival (OS) rate reached 80% after high-dose chemotherapy. Due to the relatively low incidence of T-LBL/ALL, only a few risk factors have been identified. The occurrence and prognosis of malignant tumors are closely related to oxidative stress, but the prognostic relationship between T-LBL/ALL and systematic oxidative stress indexes has not been reported yet. *Methods.* A total of 258 T-LBL/ALL patients were retrospectively analyzed. The relationship between systematic oxidative stress indexes, such as creatinine (CRE), gamma-glutamyl transpeptidase ($\gamma$-GGT), albumin (ALB), alkaline phosphatase (ALP), fibrinogen (FBG), C-reactive protein (CRP) and total bilirubin (TBIL), and survival of T-LBL/ALL patients, was analyzed. The weight of indexes was used to calculate the patients' oxidative stress risk score. The independent prognostic value of oxidative stress risk grouping (OSRG) was analyzed. *Results.* Higher CRE, CRP, and lower ALB were associated with poorer OS in T-LBL/ALL patients. The OSRG established by CRE, ALB, and CRP was an independent prognostic factor for OS of T-LBL/ALL patients. Patients in the high-risk group were more likely to be patients older than 14 years old and patients with superior vena cava obstruction syndrome (SVCS), pleural effusion, pericardial effusion, and mediastinal mass. *Conclusion.* For OS in T-LBL/ALL patients, OSRG was observed as an independent prognostic factor, which provided a new idea for accurate prognostic prediction.

## 1. Introduction

T lymphoblastic lymphoma/leukemia (T-LBL/ALL) is a malignant tumor originating from progenitor T cells, accounting for only 2% of the lymphoma incidence. T-LBL/ALL mostly occurs in children and adolescents [1]. After intensive chemotherapy, the current 5-year overall survival (OS) rate can reach up to 80% [2]. However, the prognosis of refractory and recurrent T-LBL/ALL patients is very poor, and the 5-year OS is less than 20% [3]. Early detection of relapse may improve the outcome. Furthermore, high-dose chemotherapy can improve the survival and overall response rate (ORR), but treatment-related death (TRD) is one of the important factors affecting patients' outcomes [2]. Therefore, how to accurately predict

the prognosis by performing risk stratification according to simple and feasible indicators before treatment is of great significance to guide treatment intensity and follow-up frequency. Due to relatively low incidence rate of T-LBL/ALL, few risk factors of T-LBL/ALL have been identified by the large samples.

Oxidative stress is considered an important factor in the occurrence and development of cancer, and it also affects the patient's prognosis [4, 5]. In the evolution of tumors, oxidative stress can cause dysregulation of tumor microenvironment metabolism and signaling pathways and lead to malignant transformation of cells [6]. Reactive oxygen species (ROS) may play an important role in T cell activation and signal transduction and may be associated with the development of T cell lymphoma [7]. However, in T cell

lymphoma, tumor cells avoid ROS damage by down-regulating oxidative stress signaling pathways [8]. These evidences suggest that oxidative stress is closely related to T-LBL/ALL.

Previous studies had showed that hematologic indicators could reflect the status of systematic oxidative stress. Creatinine (CRE), C-reactive protein (CRP), and total bilirubin (TBIL) increased in oxidative stress mouse models and critically ill polytrauma patients, while decreasing after antioxidant treatment [9–12]. Albumin (ALB) is an important antioxidant in the body and is a marker commonly used to reflect oxidative stress levels [13]. Gamma-glutamyl transferase ($\gamma$-GGT), a commonly used indicator of oxidative stress, decreased after patients received antioxidant therapy [14]. Increased alkaline phosphatase (ALP) activity is associated with oxidative stress in colitis models [15]. Fibrinogen is the main plasma protein of oxidative modification and can be used as a marker to assess oxidative stress [16, 17]. Previously, we reported that systematic oxidative stress indexes are associated with the survival of breast cancer patients [18]. However, systematic oxidative stress indexes and prognosis of T-LBL/ALL have not been reported.

Therefore, the purpose of this study is to explore the relationship between systematic oxidative stress indexes and T-LBL/ALL. Afterward, it provides a new idea for predicting the prognosis of T-LBL/ALL, guiding the intensity of treatment and the frequency of follow-up.

## 2. Methods

*2.1. Patients and Study Design.* This study retrospectively enrolled patients who were pathologically diagnosed with T-LBL/ALL at Sun Yat-Sen University Cancer Center (SYSUCC) from January 2010 to December 2020. The inclusion criteria included as follows: (1) diagnosed as T-LBL/ALL by pathology in SYSUCC; (2) received standard antitumor therapy in SYSUCC; (3) patients underwent biochemical examination such as CRE, GGT, ALB, ALP, FBG, CRP, and TBIL within 3 days before chemotherapy; and (4) complete follow-up records. The exclusion criteria included as follows: (1) patients had received antitumor therapy before being transferred to our hospital; (2) absence of follow-up or biochemical information; (3) the liver and kidney dysfunction; and (4) accompanied by acute and chronic inflammatory diseases. The study is in line with the Declaration of Helsinki and the Ethics Committee of SYSUCC (identifier: B2022-155-01) as well.

*2.2. Data Collection.* We collected age, sex, Eastern Cooperative Oncology Group (ECOG) status, "B" symptoms, pathological diagnosis, clinical stage, superior vena cava obstruction syndrome (SVCS), central nervous system (CNS) involvement, mediastinal mass, pericardial effusion, pleural effusion, lactate dehydrogenase (LDH), and bone marrow (BM) biopsy examination of patients with T-LBL/ALL. The results of CRE, GGT, ALB, ALP, FBG, CRP, and TBIL were obtained within 3 days before the chemotherapy. An experienced pathologist at SYSUCC was requested to reassess the pathological diagnosis of the patients. The

patient's age referred to the age of pathological diagnosis. The "B" symptoms referred to fever >38°C for more than 3 days, severe night sweats, and/or 10% weight loss without apparent cause in the last 6 months. The clinical stage was based on the Ann Arbor stage. Furthermore, mediastinal mass, pleural effusion, and pericardial effusion were also determined based on imaging findings. SVCS referred to edema of the upper limb, neck, face, and superficial varicose veins of the upper body caused by mediastinal mass. The biochemical test was analyzed by an automatic biochemical analyzer (Hitachi 7600 series, Tokyo, Japan). As for treatment, 205 patients (79.5%) received the BFM-90/95 regimen, in which pegaspargase was produced by Jiangsu Hengrui Pharmaceutical Co., Ltd. In addition, 13 patients (5.0%) received ECOG 2003 and 12 patients (4.7%) received hyper-CVAD/MA regimen. Another 28 patients (10.9%) received other regimens, such as GD 2008ALL and SCCLG-ALL-2016. After standard treatment, patients received a regular telephone or outpatient follow-up visits after the treatment. The last follow-up of the patients included in this study was conducted in June 2021.

*2.3. Statistical Analysis.* R software (version 4.0.2) and SPSS 24.0 were used to perform analysis of the data statistics of this study. The best cut-off values of GGT, ALB, CRP, TBIL, FBG, CRE, and BUN were determined by "survminer" package in "R" software. If the value is higher than the cut-off value, the value is 2; if the value is lower than the cut-off value, the value is 1. The univariate and multivariate Cox regression analyses were used to conclude the OS-independent prognostic factors. The correlation coefficient of systematic oxidative stress indexes was determined by the lowest AIC (Akaike information criterion) value [19]. Then, the systematic oxidative stress risk score of each patient was calculated. The Kaplan-Meier survival analysis was used to analyze the patients' survival. The time-dependent receiver operating characteristic (ROC) curve was used to compare the prognostic value of oxidative stress risk grouping (OSRG) and clinical prognostic indicators. To compare the correlations between variables, the Chi-square test with $p$ value <0.05 was used to consider statistically significant in the two-tailed test.

## 3. Results

*3.1. Information on Patients and Oxidative Stress Indicators.* A total of 266 T-LBL/ALL patients at SYSUCC were retrospectively registered in our study. Out of 266, 8 patients died due to treatment-related complications. To ensure the prediction accuracy of this risk grouping, these 8 patients with TRD were excluded from the study. Of the 258 patients enrolled, 109 patients were younger, while 149 patients were older than 14 years. The patients' clinical features are presented in Table 1. Based on "Surv_cutpoint" analysis of R software, the optimal cut-offs of CRE, GGT, ALB, ALP, FBG, CRP, and TBIL were 44.6 $\mu$mol/L, 21.2 U/L, 38.7 g/L, 106.7 U/L, 3.09 g/L, 12.28 mg/L, and 11.8 $\mu$mol/L, respectively (Supplementary Figures 1(a)–1(g)). Table 2 shows

TABLE 1: Clinical characteristics of study population.

| Characteristics | Number (n) | Percentage (%) |
| --- | --- | --- |
| *Sex* | | |
| Male | 196 | 76.0 |
| Female | 62 | 24.0 |
| *Age* | | |
| ≤14 | 109 | 42.2 |
| >14 | 149 | 57.8 |
| *ECOG* | | |
| <2 | 249 | 96.5 |
| ≥2 | 9 | 3.5 |
| *B symptoms* | | |
| Yes | 56 | 21.7 |
| No | 202 | 78.3 |
| *SVCS involvement* | | |
| Yes | 103 | 39.9 |
| No | 155 | 60.1 |
| *Pleural effusion* | | |
| Yes | 96 | 37.2 |
| No | 162 | 62.8 |
| *Pericardial effusion* | | |
| Yes | 53 | 20.5 |
| No | 205 | 79.5 |
| *CNS involvement* | | |
| Yes | 10 | 3.9 |
| No | 248 | 96.1 |
| *Mediastinal involvement* | | |
| Yes | 164 | 63.6 |
| No | 94 | 36.4 |
| *BM involvement* | | |
| Yes | 102 | 39.5 |
| No | 156 | 60.5 |
| *LDH status* | | |
| Elevated | 136 | 52.7 |
| Normal | 122 | 47.3 |
| *Ann Arbor stage* | | |
| I/II | 18 | 7.0 |
| III/IV | 240 | 93.0 |
| *ASCT/HSCT at CR* | | |
| Yes | 12 | 4.7 |
| No | 246 | 95.3 |
| *Treatment response* | | |
| CR | 229 | 88.8 |
| Non-CR | 29 | 11.2 |

TABLE 2: oxidative stress indexes of study population.

| Oxidative stress indexes | Number (n) | Percentage (%) |
| --- | --- | --- |
| *CRE (μmol/L)* | | |
| ≤44.6 | 89 | 34.5 |
| >44.6 | 169 | 65.5 |
| *GGT (U/L)* | | |
| ≤21.2 | 132 | 51.2 |
| >21.2 | 126 | 48.8 |
| *ALB (g/L)* | | |
| ≤38.7 | 98 | 38.0 |
| >38.7 | 160 | 62.0 |
| *ALP (U/L)* | | |
| ≤106.7 | 154 | 59.7 |
| >106.7 | 104 | 40.3 |
| *FBG (g/L)* | | |
| ≤3.09 | 190 | 73.6 |
| >3.09 | 68 | 26.4 |
| *CRP (mg/L)* | | |
| ≤12.28 | 195 | 75.6 |
| >12.28 | 63 | 24.4 |
| *TBIL (μmol/L)* | | |
| ≤11.8 | 212 | 82.2 |
| >11.8 | 46 | 17.8 |

defined as 2 if it was higher than cut-offs, and 1 if they were lower. The univariate Cox regression analysis revealed that only ALP was not correlated with OS, while all other factors including CRE, GGT, ALB, FBG, CRP, and TBIL were correlated with OS (Figure 1(a)). In multivariate Cox regression analysis, the elevated CRP and CRE were related to worse outcomes in the T-LBL/ALL patients, while the elevated ALB was associated with better outcomes (Figure 1(b)). Then, to calculate the weight of oxidative stress indexes in the model, we calculated the correlation coefficient of indexes based on AIC. Oxidative stress risk score = -0.59 × value of ALB + 1.06 × value of CRE + 0.72 × value of CRP (Figure 1(c)). Then, OSRG of T-LBL/ALL patients was identified based on the median value of oxidative stress risk score. Patients with a risk score greater than the median value were in high-risk group, and patients with a risk score less than or equal to the median value were in low-risk group.

3.3. Relationship between OSRG and Clinical Features. The group with high-risk was correlated with worse OS in the patients with T-LBL/ALL (Figure 2(a)). Meanwhile, we calculated the relationship between OSRG and progression-free survival (PFS). The results showed that the high-risk group had shorter PFS as compared to the low-risk group (Figure 2(b)). Similarly, in subgroup analysis, we found that OS and PFS were worse in the high-risk group than that in the low-risk group in both patients younger than 14 years old (Figures 3(a) and 3(c)) and patients older than 14 years old (Figures 3(b) and 3(d)). To further explore the relationship between OSRG and clinical features, we performed the

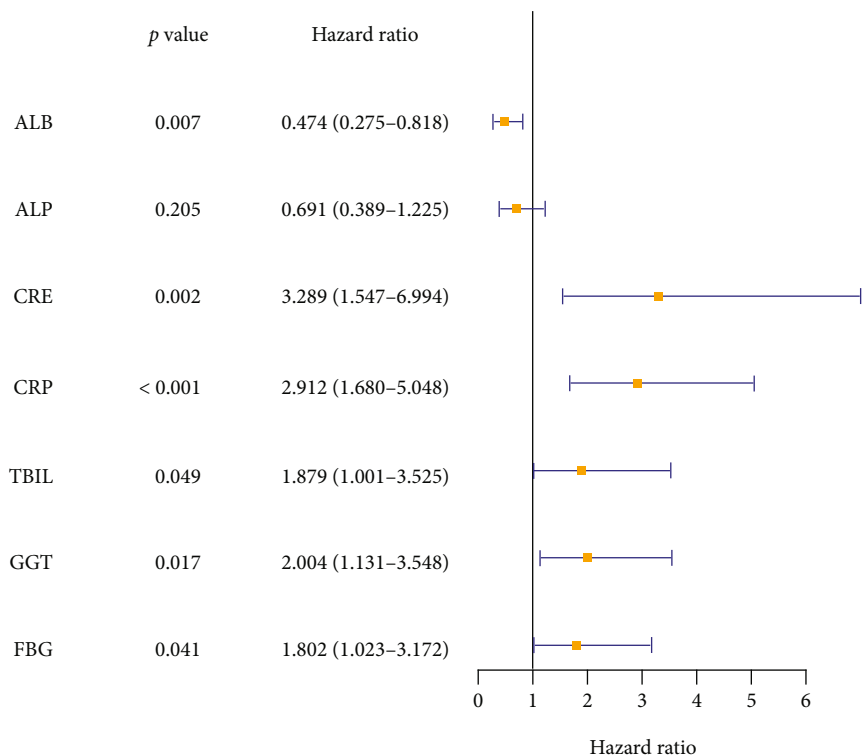the distribution of systematic oxidative stress indexes of patients.

3.2. Calculation of Oxidative Stress Risk Score. Conversion of CRE, GGT, ALB, ALP, FBG, CRP, and TBIL into dichotomy based on the cut-offs was done as described above. The values of "1" and "2" were used for scoring. The values were

(a)
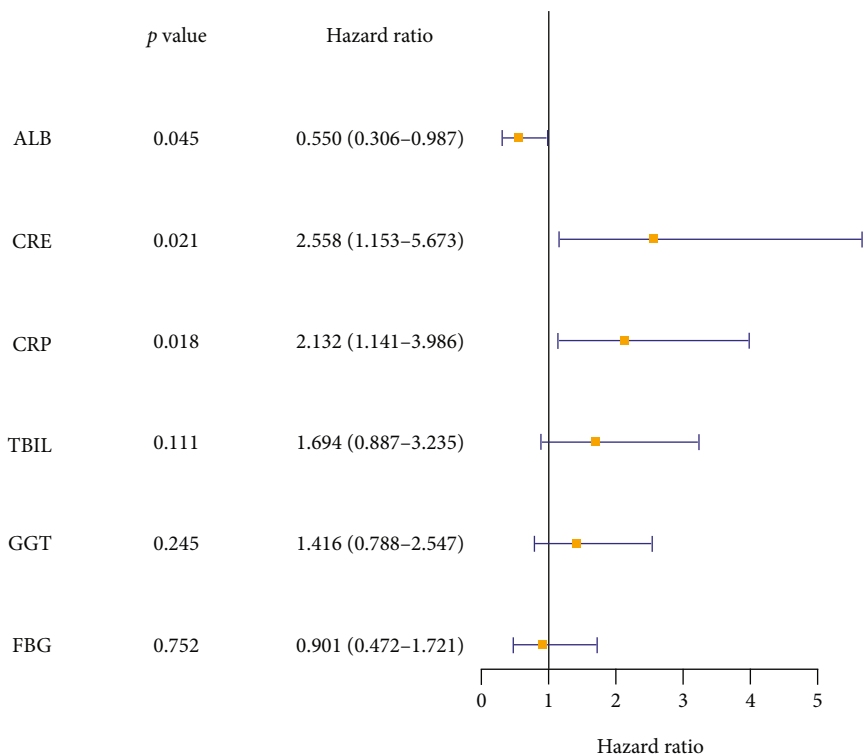


(b)

$$\text{Oxidative stress risk score} = -0.59 \times \text{value of ALB} + 1.06 \times \text{value of CRE} + 0.72 \times \text{value of CRP}$$

(c)

FIGURE 1: Establishment of oxidative stress risk score. The univariate (a) and multivariate (b) Cox regression analysis of oxidative stress indexes. (c) Calculation formula of oxidative stress risk score.
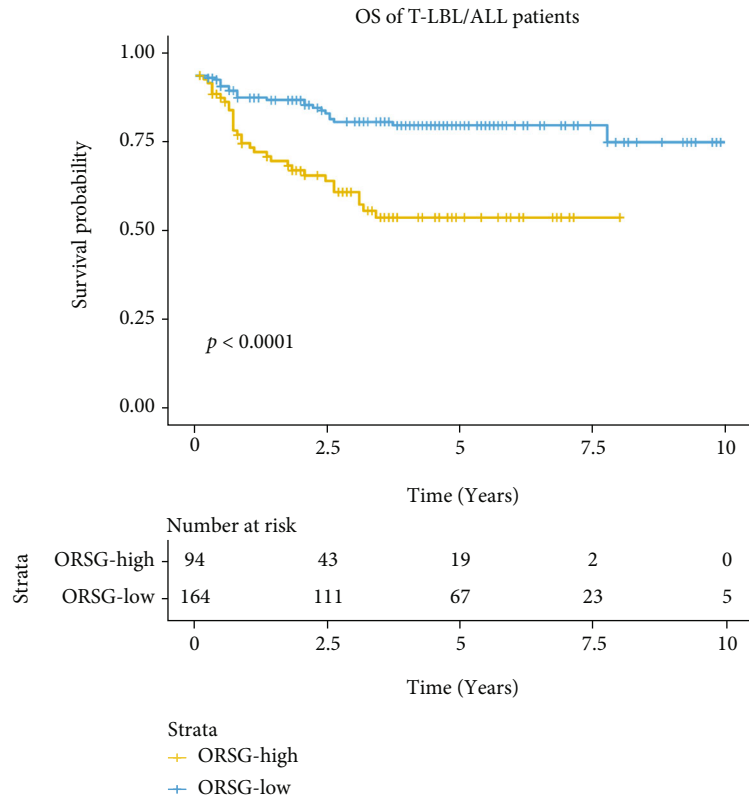
(a)
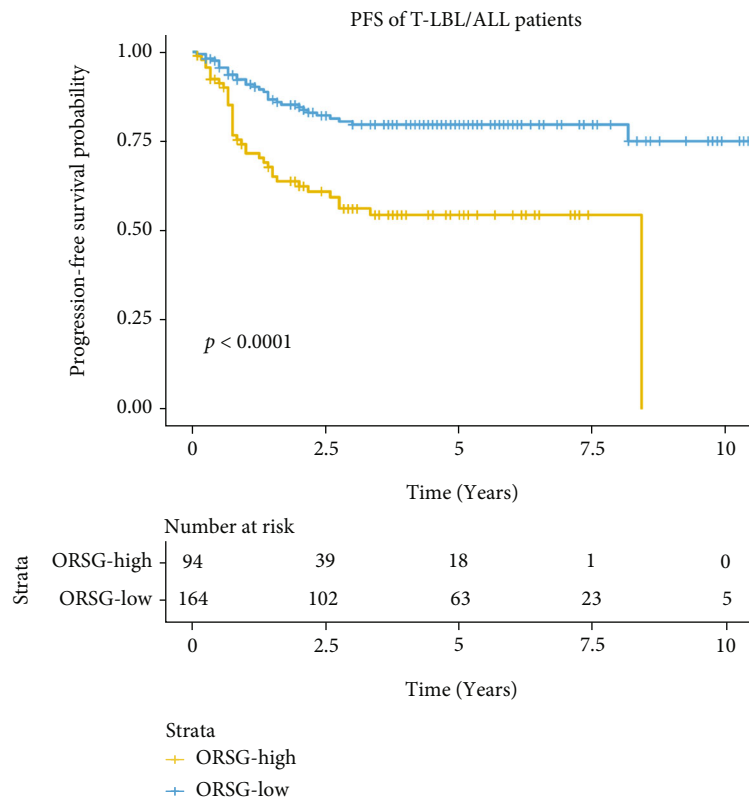


(b)

FIGURE 2: OSRG was related to prognosis in T-LBL/ALL patients. Kaplan-Meier curves revealed that the OS (a) and PFS (b) of high-OSRG patients were shorter than that of low-OSRG patients.

(a)



(b)
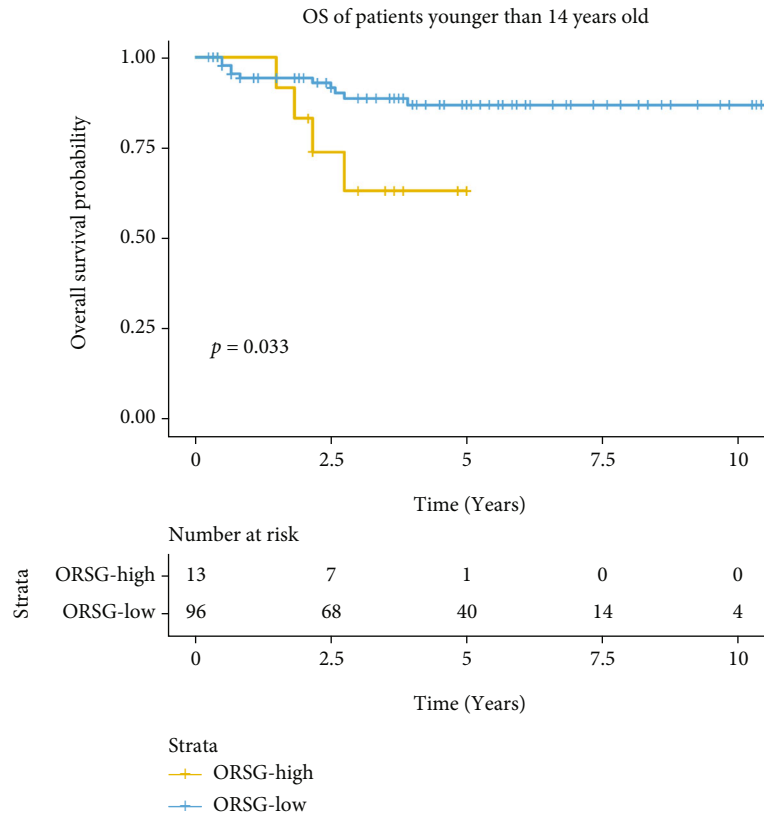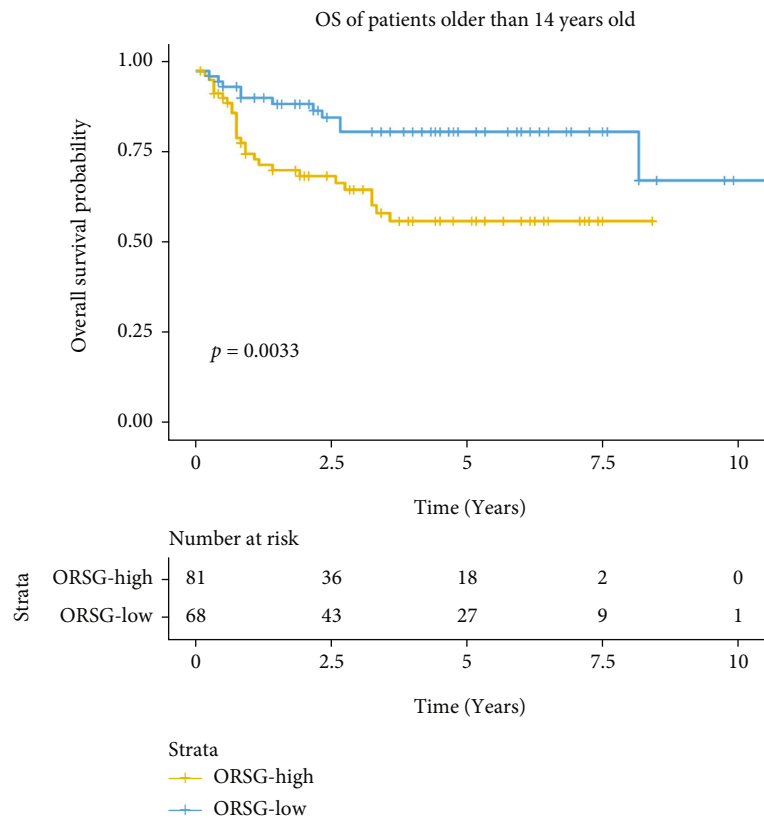
FIGURE 3: Continued.
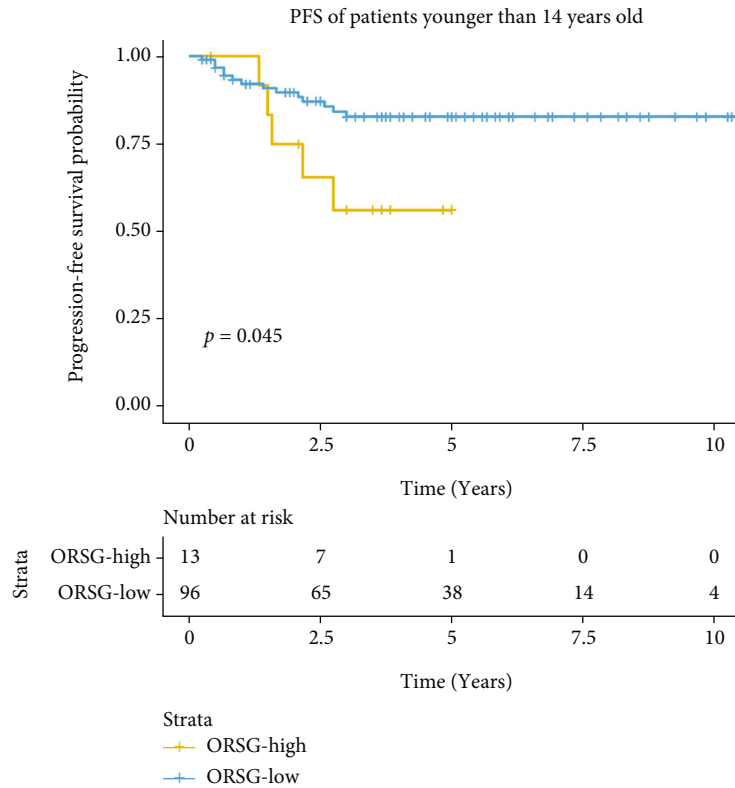
(c)
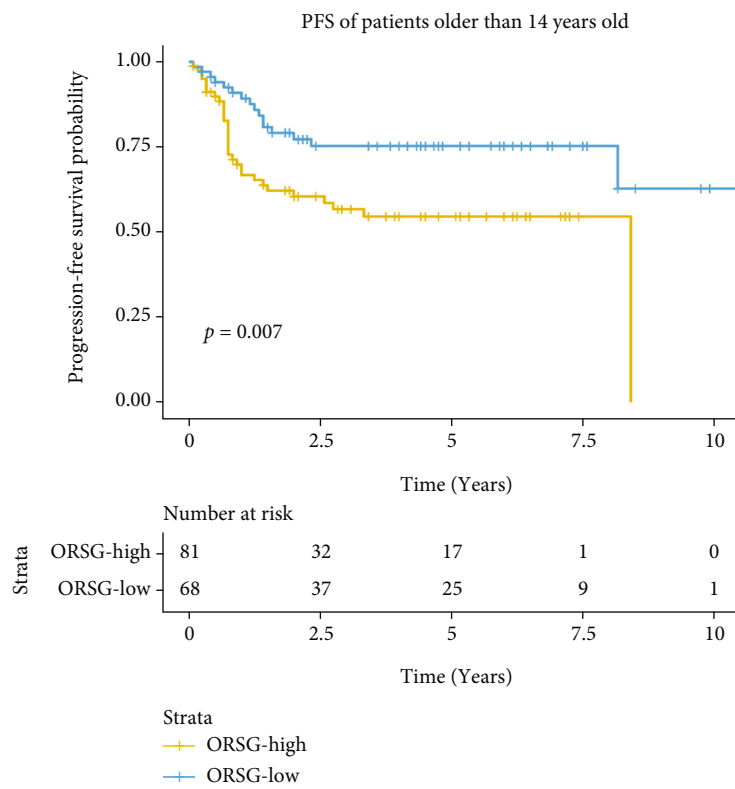


(d)

Figure 3: Subgroup survival analysis for T-LBL/ALL patients. (a) Kaplan-Meier analysis for the OS of patients younger than 14 years old. (b) Kaplan-Meier analysis for the OS of patients older than 14 years. (c) Kaplan-Meier analysis for the PFS of patients younger than 14 years old. (d) Kaplan-Meier analysis for the PFS of patients older than 14 years.

Chi-square test, and the results revealed that the patients in high-risk group were more likely to be patients older than 14 years old, and patients with SVCS, pleural effusion, pericardial effusion, and mediastinal mass. At the same time, the complete response (CR) rate was higher in the low-risk group (Table 3).

*3.4. Independent Prognostic Significance of OSRG.* In Kaplan-Meier survival analysis, the high-risk was associated with worse outcomes in T-LBL/ALL patients. To further explore its prognostic value, the univariate and multivariate Cox analysis was calculated for all the indicators that might affect the prognosis of the patients. The univariate Cox regression analysis showed that OSRG, age, CNS involvement, and BM involvement were correlated with the prognosis of the patients. The multivariate Cox regression analysis showed that OSRG, CNS involvement, and BM involvement were independent prognostic factors for T-LBL/ALL patients (Table 4). Further ROC curves revealed that the predictive accuracy of OSRG is better than factors of age, CNS involvement, and BM involvement. The area under the ROC curves was used to compare the prognostic value of OSRG and clinical prognostic indicators for predicting 2-year OS (Figure 4(a)) and 5-year OS (Figure 4(b)).

# 4. Discussion

Oxidative stress is closely related to cancer occurrence and cancer development as well as prognosis[4, 5, 20, 21]. Previously, we have reported that systematic oxidative stress indexes are related to the prognosis of breast cancer patients [18]. However, systematic oxidative stress indexes and prognosis of T-LBL/ALL have not been reported.

Albumin (ALB) is an important protein present in human plasma. In cancer patients with normal liver synthesis function, ALB is often associated with the patient's nutritional status and inflammation level [22]. It also has the effect of antioxidant stress [23]. Elevated ALB is often associated with better survival in cancer patients [24]. This study also showed that higher ALB was suggestive of longer OS in T-LBL/ALL patients. Creatinine (CRE) in the blood is derived from exogenous and endogenous. Endogenous CRE is the product of the metabolism of muscle tissue in the body. It has been shown in other studies that CRE is correlated to the level of oxidative stress in the body, which decreases after patients receive antioxidant treatment [12]. Higher CRE level is associated with a worse prognosis for patients with malignant tumor [25]. C-reactive protein (CRP) is an acute protein that reflects the level of inflammation in the body. Increased CRP can be observed in many malignant tumors and is associated with poor prognosis of tumor patients [26, 27].

The present investigation is the first to analyze the prognostic value of indicators of systematic oxidative stress on T-LBL/ALL. Among the 7 indicators of systematic oxidative stress, ALB, CRE, and CRP were independent prognostic factors for T-LBL/ALL patients. The multivariate Cox regression analysis showed that the OSRG established according to the weight of ALB, CRE, and CRP was an inde-

TABLE 3: Relationship between OSRG and clinical characteristics.

| Characteristics | OSRG-low ($n = 164$) | | OSRG-high ($n = 94$) | | $P$ value |
|---|---|---|---|---|---|
| *Sex* | | | | | |
| Male | 123 | 75.0 | 73 | 77.7 | 0.630 |
| Female | 41 | 25.0 | 21 | 22.3 | |
| *Age* | | | | | |
| ≤14 | 96 | 58.5 | 13 | 13.8 | **0.000** |
| >14 | 68 | 41.5 | 81 | 86.2 | |
| *ECOG* | | | | | |
| <2 | 156 | 95.1 | 93 | 98.9 | 0.108 |
| ≥2 | 8 | 4.9 | 1 | 1.1 | |
| *B symptoms* | | | | | |
| Yes | 32 | 19.5 | 24 | 25.5 | 0.259 |
| No | 132 | 80.5 | 70 | 74.5 | |
| *SVCS* | | | | | |
| Yes | 56 | 34.1 | 47 | 50.0 | **0.012** |
| No | 108 | 65.9 | 47 | 50.0 | |
| *Pleural effusion* | | | | | |
| Yes | 50 | 30.5 | 46 | 48.9 | **0.003** |
| No | 114 | 69.5 | 48 | 51.1 | |
| *Pericardial effusion* | | | | | |
| Yes | 17 | 10.4 | 36 | 38.3 | **0.000** |
| No | 147 | 89.6 | 58 | 61.7 | |
| *CNS involvement* | | | | | |
| Yes | 5 | 3.0 | 5 | 5.3 | 0.363 |
| No | 159 | 97.0 | 89 | 94.7 | |
| *Mediastinal involvement* | | | | | |
| Yes | 94 | 57.3 | 70 | 74.5 | **0.006** |
| No | 70 | 42.7 | 24 | 25.5 | |
| *BM involvement* | | | | | |
| Yes | 66 | 40.2 | 36 | 38.3 | 0.758 |
| No | 98 | 59.8 | 58 | 61.7 | |
| *LDH status* | | | | | |
| Elevated | 85 | 51.8 | 51 | 54.3 | 0.707 |
| Normal | 79 | 48.2 | 43 | 45.7 | |
| *Ann Arbor stage* | | | | | |
| I/II | 15 | 9.1 | 3 | 3.2 | 0.071 |
| III/IV | 149 | 90.9 | 91 | 96.8 | |
| *ASCT/HSCT at CR* | | | | | |
| Yes | 6 | 3.7 | 6 | 6.4 | 0.317 |
| No | 158 | 96.3 | 88 | 93.6 | |
| *Treatment response* | | | | | |
| CR | 152 | 92.7 | 77 | 81.9 | **0.008** |
| Non-CR | 12 | 7.3 | 17 | 18.1 | |

pendent prognostic factor for the T-LBL/ALL patients. This study also confirmed that indicators of systematic oxidative stress is also associated with the prognosis of T-LBL/ALL. At the same time, we found the association between OSRG and clinical characteristics of T-LBL/ALL patients. The high-risk group was associated with older age, presence of

TABLE 4: Results of the univariate and multivariate Cox regression analysis for OS.

| Variables | Univariate Cox analysis | | Multivariate Cox analysis | |
| --- | --- | --- | --- | --- |
| | HR (95% CI) | P value | HR (95% CI) | P value |
| *Sex* | | | | |
| Male | Reference | 0.905 | | |
| Female | 0.961 (0.504-1.833) | | | |
| *Age* | | | | |
| ≤14 | Reference | 0.008 | Reference | 0.102 |
| >14 | 2.305 (1.248-4.258) | | 1.797 (0.890-3.628) | |
| *ECOG* | | | | |
| <2 | Reference | 0.771 | | |
| ≥2 | 1.234 (0.300-5.079) | | | |
| *B symptoms* | | | | |
| No | Reference | 0.582 | | |
| Yes | 1.193(0.637-2.234) | | | |
| *SVCS involvement* | | | | |
| No | Reference | 0.224 | | |
| Yes | 1.402 (0.813-2.415) | | | |
| *Pleural effusion* | | | | |
| No | Reference | 0.492 | | |
| Yes | 1.215 (0.698-2.115) | | | |
| *Pericardial effusion* | | | | |
| No | Reference | 0.131 | | |
| Yes | 1.604 (0.868-2.961) | | | |
| *CNS involvement* | | | | |
| No | Reference | <0.001 | Reference | <0.001 |
| Yes | 4.395 (1.872-10.322) | | 4.367 (1.845-10.339) | |
| *Mediastinal involvement* | | | | |
| No | Reference | 0.212 | | |
| Yes | 1.466(0.804-2.673) | | | |
| *BM involvement* | | | | |
| No | Reference | 0.030 | Reference | 0.010 |
| Yes | 1.834 (1.060-3.173) | | 2.095 (1.198-3.666) | |
| *LDH status* | | | | |
| Elevated | Reference | 0.277 | | |
| Normal | 1.360 (0.782-2.368) | | | |
| *Ann Arbor stage* | | | | |
| I/II | Reference | 0.698 | | |
| III/IV | 0.832 (0.329-2.105) | | | |
| *ASCT/HSCT at CR* | | | | |
| No | Reference | 0.735 | | |
| Yes | 1.244 (0.381-3.930) | | | |
| *OSRG* | | | | |
| Low | Reference | <0.001 | Reference | 0.002 |
| High | 3.377 (1.929-5.911) | | 2.759 (1.470-5.180) | |

SVCS, pleural effusion, pericardial effusion, and mediastinal mass. Additionally, the CR rate was also lower in the high-risk group than in the low-risk group. Patients with relapsed/refractory T-LBL/ALL have a poor prognosis, and early detection of disease progression is of great significance to patients. Further analysis showed that OSRG could pre-

dict PFS in patients, and patients in the high-risk group had a higher recurrence rate.

At present, there are few studies on the analysis of prognostic factors, which make this study important to prognosis prediction for T-LBL/ALL patients. In this study, the data of 258 patients with T-LBL/ALL were included and analyzed,
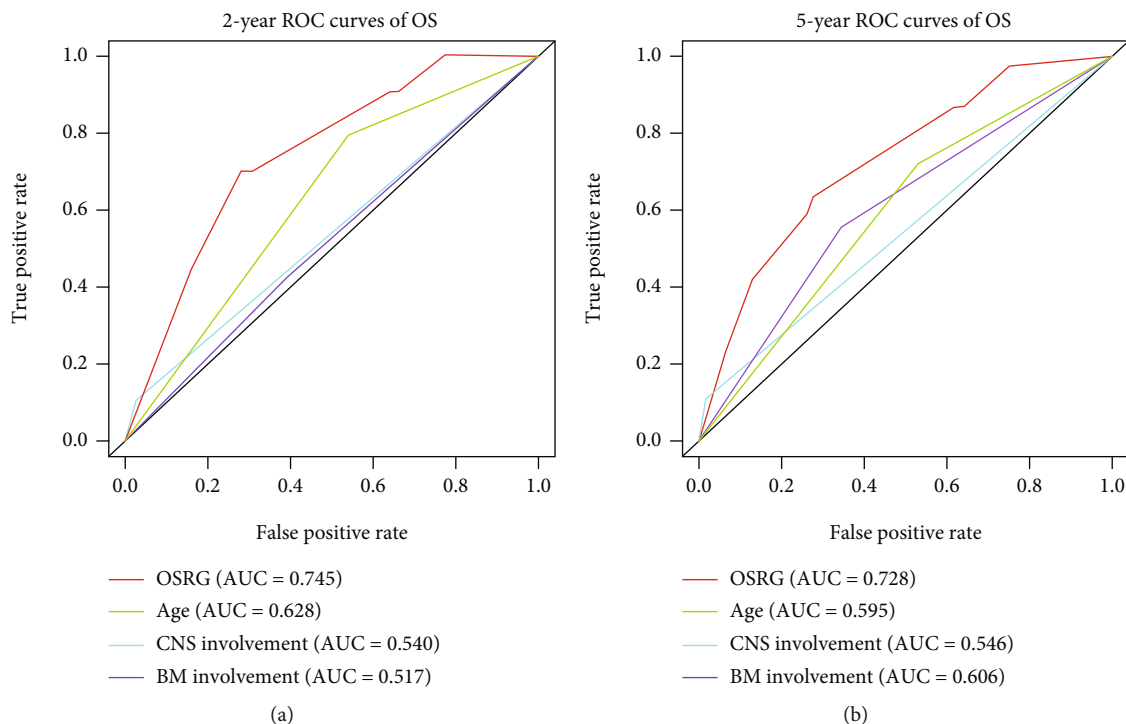
FIGURE 4: The predictive accuracy of OSRG is better than that of clinical prognostic indicators. The area under the ROC curves was used to compared the prognostic value of OSRG and clinical prognostic indicators for predicting 2-year OS (a) and 5-year OS (b).

confirming that the OSRG is an independent prognostic factor for T-LBL/ALL patients. T-LBL/ALL is one of the malignant tumors with an acceptable prognosis after intensive chemotherapy, but still, its TRD and disease progression are the factors causing the poor prognosis. This study can further provide clinicians with references for precise treatment and follow-up frequency for T-LBL/ALL patients.

However, the present study is a single-center retrospective study, and the specific mechanism of oxidative stress in tumor development and outcome is yet not clear. Due to the retrospective analysis, it failed to match the consistency of OSRG with the 2', 7'-dichlorofluorescin (DCFH) assay in detecting ROS. Furthermore, basic research and prospective study are needed in the future.

## Data Availability

These data are available by individual application to corresponding author.

## Ethical Approval

Our study was approved by the Ethics Committee of Sun-Yat-Sen University Cancer Center Health Authority (identifier: B2022-155-01).

## Conflicts of Interest

The authors declare that there is no conflict of interest.

## Authors' Contributions

All authors participated in this research, including conception and design (PLQ, HYX, and HHQ), data acquisition (PLQ, GY, HYX, WXX, BB, HC, and HHQ), data analysis and interpretation (PLQ, GY, HYX, and HHQ), material support (PLQ, HYX, and HHQ), and study supervision (PLQ, HYX, and HC), as well as drafting the article or critically revising (PLQ, WXX, and HHQ). The final version ensured and approved by all authors. Liqin Ping, Yan Gao, and Yanxia He contributed equally to this work and share first authorship.

## Acknowledgments

## Supplementary Materials

Supplementary Figure 1. Identification of optimal cut-off values of oxidative stress indexes. Supplementary Figure 1. Identification of optimal cut-off values of oxidative stress indexes. The optimal cut-off values of ALB (A), ALP (B), CRE (C), CRP (D), TBIL (E), GGT (F), and FBG (G) were identified. (Supplementary materials)

## References

[1] B. Burkhardt, S. Mueller, T. Khanam, and S. L. Perkins, "Current status and future directions of T-lymphoblastic lymphoma in children and adolescents," British Journal of Haematology, vol. 173, no. 4, pp. 545–559, 2016.

[2] E. Landmann, B. Burkhardt, M. Zimmermann et al., "Results and conclusions of the European intergroup EURO-LB02 trial in children and adolescents with lymphoblastic lymphoma," *Haematologica*, vol. 102, no. 12, pp. 2086–2096, 2017.

[3] A. Candoni, D. Lazzarotto, F. Ferrara et al., "Nelarabine as salvage therapy and bridge to allogeneic stem cell transplant in 118 adult patients with relapsed/refractory T-cell acute lymphoblastic leukemia/lymphoma. A CAMPUS ALL study," *American journal of hematology*, vol. 95, no. 12, pp. 1466–1472, 2020.

[4] C. Gorrini, I. S. Harris, and T. W. Mak, "Modulation of oxidative stress as an anticancer strategy," *Drug discovery*, vol. 12, no. 12, pp. 931–947, 2013.

[5] B. Yang and Q. Chen, "Cross-Talk between Oxidative Stress and m6A RNA Methylation in Cancer," *Oxidative Medicine and Cellular Longevity*, vol. 2021, Article ID 6545728, 26 pages, 2021.

[6] A. K. Maurya and M. Vinayak, "PI-103 and quercetin attenuate PI3K-AKT signaling pathway in T-cell lymphoma exposed to hydrogen peroxide," *PloS one*, vol. 11, no. 8, article e0160686, 2016.

[7] A. Larbi, J. Kempf, and G. Pawelec, "Oxidative stress modulation and T cell activation," *Experimental Gerontology*, vol. 42, no. 9, pp. 852–858, 2007.

[8] S. Kumar, B. Dhamija, D. Attrish et al., "Genetic alterations and oxidative stress in T cell lymphomas," *Pharmacology & Therapeutics*, vol. 236, p. 108109, 2022.

[9] S. Periasamy, D. Z. Hsu, Y. H. Fu, and M. Y. Liu, "Sleep deprivation-induced multi-organ injury: role of oxidative stress and inflammation," *EXCLI Journal*, vol. 14, pp. 672–683, 2015.

[10] A. Teixeira-Gomes, B. Laffon, V. Valdiglesias et al., "Exploring early detection of frailty syndrome in older adults: evaluation of oxi-immune markers, clinical parameters and modifiable risk factors," *Antioxidants (Basel, Switzerland)*, vol. 10, no. 12, 2021.

[11] W. Chen, S. Tumanov, D. J. Fazakerley et al., "Bilirubin deficiency renders mice susceptible to hepatic steatosis in the absence of insulin resistance," *Redox Biology*, vol. 47, p. 102152, 2021.

[12] M. Sandesc, A. F. Rogobete, O. H. Bedreag et al., "Analysis of oxidative stress-related markers in critically ill polytrauma patients: an observational prospective single-center study," *Bosnian Journal of Basic Medical Sciences*, vol. 18, no. 2, pp. 191–197, 2018.

[13] L. Liu, K. Xie, M. Yin et al., "Serum potassium, albumin and vitamin B12 as potential oxidative stress markers of fungal peritonitis," *Annals of Medicine*, vol. 53, no. 1, pp. 2132–2141, 2021.

[14] S. Ahmed, A. Zahoor, M. Ibrahim et al., "Enhanced efficacy of direct-acting antivirals in hepatitis C patients by coadministration of black cumin and ascorbate as antioxidant adjuvants," *Oxidative Medicine and Cellular Longevity*, vol. 2020, Article ID 7087921, 10 pages, 2020.

[15] F. Sánchez de Medina, O. Martínez-Augustin, R. González et al., "Induction of alkaline phosphatase in the inflamed intestine: a novel pharmacological target for inflammatory bowel disease," *Biochemical pharmacology*, vol. 68, no. 12, pp. 2317–2326, 2004.

[16] M. Becatti, R. Marcucci, G. Bruschi et al., "Oxidative modification of fibrinogen is associated with altered function and structure in the subacute phase of myocardial infarction," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 34, no. 7, pp. 1355–1361, 2014.

[17] J. G. McLarnon, "A leaky blood-brain barrier to fibrinogen contributes to oxidative damage in Alzheimer's disease," *Antioxidants (Basel, Switzerland)*, vol. 11, no. 1, 2022.

[18] K. Zhang, L. Ping, T. Du et al., "A novel systematic oxidative stress score predicts the prognosis of patients with operable breast cancer," *Oxidative Medicine and Cellular Longevity*, vol. 2021, Article ID 9441896, 14 pages, 2021.

[19] S. I. Vrieze, "Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC)," *Psychological Methods*, vol. 17, no. 2, pp. 228–243, 2012.

[20] M. Morotti, C. E. Zois, R. El-Ansari et al., "Increased expression of glutamine transporter SNAT2/SLC38A2 promotes glutamine dependence and oxidative stress resistance, and is associated with worse prognosis in triple-negative breast cancer," *British Journal of Cancer*, vol. 124, no. 2, pp. 494–505, 2021.

[21] A. Leone, M. S. Roca, C. Ciardiello, S. Costantini, and A. Budillon, "Oxidative stress gene expression profile correlates with cancer patient poor prognosis: identification of crucial pathways might select novel therapeutic approaches," *Oxidative Medicine and Cellular Longevity*, vol. 2017, Article ID 2597581, 18 pages, 2017.

[22] A. Eckart, T. Struja, A. Kutz et al., "Relationship of nutritional status, inflammation, and serum albumin levels during acute illness: a prospective study," *The American journal of medicine*, vol. 133, no. 6, pp. 713–722.e7, 2020.

[23] D. A. Belinskaia, P. A. Voronina, V. I. Shmurak, R. O. Jenkins, and N. V. Goncharov, "Serum albumin in health and disease: esterase, antioxidant, transporting and signaling properties," *International journal of molecular sciences*, vol. 22, no. 19, p. 10318, 2021.

[24] T. Yamamoto, K. Kawada, and K. Obama, "Inflammation-related biomarkers for the prediction of prognosis in colorectal cancer patients," *International journal of molecular sciences*, vol. 22, no. 15, p. 8002, 2021.

[25] J. Panotopoulos, F. Posch, P. T. Funovics et al., "Elevated serum creatinine and low albumin are associated with poor outcomes in patients with liposarcoma," *Journal of Orthopaedic Research*, vol. 34, no. 3, pp. 533–538, 2016.

[26] M. W. Socha, B. Malinowski, O. Puk et al., "C-reactive protein as a diagnostic and prognostic factor of endometrial cancer," *Critical Reviews in Oncology/Hematology*, vol. 164, p. 103419, 2021.

[27] E. M. Thurner, S. Krenn-Pilko, U. Langsenlehner et al., "The elevated C-reactive protein level is associated with poor prognosis in prostate cancer patients treated with radiotherapy," *European journal of cancer (Oxford, England: 1990)*, vol. 51, no. 5, pp. 610–619, 2015.