

## Research Article

# Identification of Prognostic Immune-Related Genes in Pancreatic Adenocarcinoma and Establishment of a Prognostic Nomogram: A Bioinformatic Study

Guolin Wu,<sup>1</sup> Zhenfeng Deng ,<sup>1</sup> Zongrui Jin,<sup>1</sup> Jilong Wang ,<sup>1</sup> Banghao Xu ,<sup>1</sup> Jingjing Zeng,<sup>2</sup> Minhao Peng ,<sup>1</sup> Zhang Wen ,<sup>1</sup> and Ya Guo <sup>1</sup>

<sup>1</sup>Department of Hepatobiliary Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China

<sup>2</sup>Department of Pathology, The First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China

Correspondence should be addressed to Ya Guo; guoygxmu@163.com

Received 28 February 2020; Accepted 13 May 2020; Published 9 June 2020

Academic Editor: Yujiang Fang

Copyright © 2020 Guolin Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Background.** The prognosis of pancreatic adenocarcinoma (PAAD) is extremely poor and has not been improved. Thus, an effective method to assess the prognosis of patients must be established to improve their survival rate. **Method.** This study investigated immune-related genes that could be used as potential therapeutic targets for PAAD. Level 3 gene expression data from the PAAD cohort and the relevant clinical information were obtained from The Cancer Genome Atlas (TCGA) database. For validation, other PAAD datasets (DSE62452) were downloaded from the Gene Expression Omnibus (GEO) database. The PAAD datasets from TCGA and GEO were used to screen immune-related genes through the Molecular Signatures Database using gene set enrichment analysis. Then, the overlapping immune-related genes of the two datasets were identified. Coexpression networks of the immune-related genes were constructed. **Results.** A signature of three immune-related genes (CKLF, ERAP2, and EREG) was identified in patients with PAAD. The signature could be used to divide the patients with PAAD into high- and low-risk groups based on their median risk score. Multivariate Cox regression analysis was performed to determine the independent prognostic factors of PAAD. Time-dependent receiver operating characteristic (ROC) curve analysis was conducted to assess the prediction accuracy of the prognostic signature. Last, a nomogram was established to assess the individualized prognosis prediction model based on the clinical characteristics and risk score of the TCGA PAAD dataset. The accuracy of the prognostic signature was further evaluated through functional evaluation and principal component analysis. **Conclusions.** The results indicated that the signature of three immune-related genes had excellent predictive value for PAAD. These findings might help improve personalized treatment and medical decisions.

## 1. Introduction

Pancreatic cancer is a leading cause of death in developed countries, and it is a common malignant tumor worldwide [1]. The main tumor type of pancreatic cancer is pancreatic adenocarcinoma (PAAD), which accounts for approximately 85% of cases [2]. The prognosis of pancreatic cancer is extremely poor, and the 5-year survival rate is estimated to be less than 5% [1]. Pancreatic cancer is expected to surpass breast cancer and become the third leading cause of cancer death [3]. Studies have predicted that the mortality rate of pancreatic cancer in malignant tumors worldwide will have

ranked second by 2030 [4]. An important reason for the low survival rate of pancreatic cancer is that most patients are diagnosed at the end stage of the disease [5]. Smoking, high-fat diet, obesity, alcoholism, diabetes, and chronic pancreatitis are risk factors for pancreatic cancer [6–8]. The pathogenesis of familial pancreatic cancer is closely related to *CDKN2A*, *BRCA1*, *BRCA2*, and *PALB2* [9]. The only possible treatment for pancreatic cancer is surgical resection, but only a few patients with early-stage pancreatic cancer are eligible to undergo resection [10]. Chemotherapy, targeted therapy, and immunotherapy for pancreatic cancer treatment improve a patient's survival time [11]. Immunotherapy

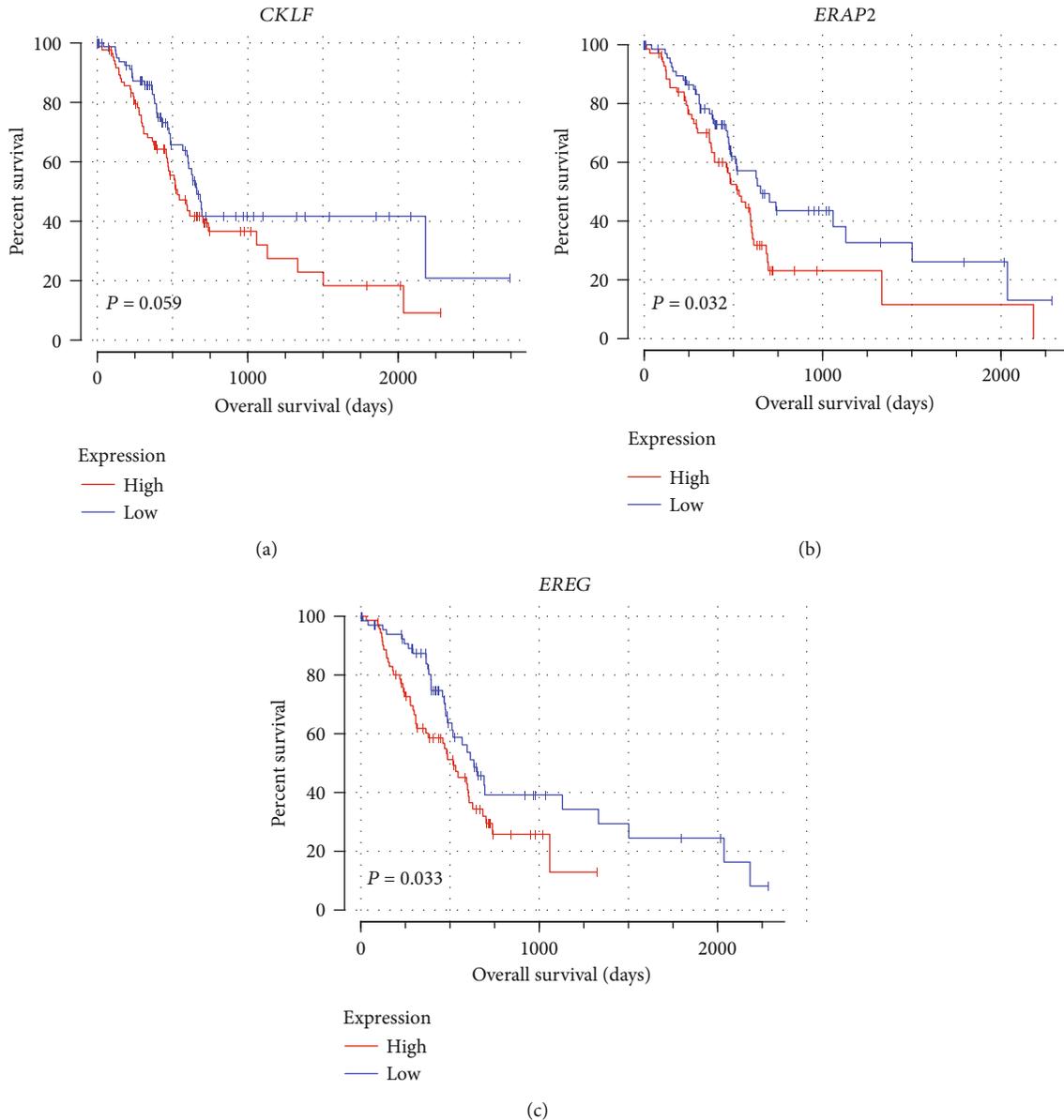


FIGURE 1: Kaplan-Meier curves of three prognostic immune-related genes in PAAD. The order is as follows: (a) *CKLF*,  $P = 0.059$ ; (b) *ERAP2*,  $P = 0.032$ ; and (c) *EREG*,  $P = 0.033$ .

has become a new pillar of cancer treatment for more than a decade, and it has offered new hope for reducing the morbidity and mortality of this refractory disease [12]. The development of immunotherapy for PAAD treatment faces challenges because of the poor immunogenic nature of PAAD [13]. Nevertheless, a large percentage of patients with PAAD may benefit from immunotherapy in the future [14]. With remarkable progress in bioinformatics, prognostic gene expression characteristics have been extensively developed for PAAD [15]. The development of tumor molecular biology has further promoted tumor therapy based on immune-related genes. Therefore, the abnormal expression of immune-related genes may have prognostic value for patients with PAAD and provide a new basis for administering tumor immunotherapy for PAAD. In this study, immune-related genes associated with the prognosis of

PAAD were identified on the basis of RNA-seq data from TCGA through the Molecular Signatures Database (MSigDB) [16], and a risk score model for PAAD prognosis was constructed. A prognostic nomogram that combined prognostic gene trait risk models and clinical prognostic factors was established to predict overall survival (OS). The reliability of this method was verified through the GEO database.

## 2. Materials and Methods

**2.1. Datasets Source.** The level 3 gene expression data from the PAAD cohort and corresponding clinical information were obtained from the data portal of TCGA (<https://portal.gdc.cancer.gov/>, accessed October 11, 2019). Relevant clinical information, such as gender, age, radical resection, grade, alcohol history, survival, and outcome, was also obtained

TABLE 1: Clinical and pathologic characteristics of PAAD patients and prognostic analysis.

Variables	Events	Total (n = 142)	MST (days)	HR (95% CI)	P value
Age (years)					0.344
≤60	22	45	593	1	
>60	56	97	568	1.270 (0.774-2.086)	
Gender					0.441
Female	40	67	532	1	
Male	38	75	614	0.838 (0.536-1.312)	
Alcohol history <sup>a</sup>					0.742
No	27	49	532	1	
Yes	43	81	598	1.084 (0.670-1.756)	
Tumor stage <sup>b</sup>					0.660
I	5	13	598	1	
II	70	122	568	1.292 (0.519-3.221)	
III + IV	3	6	545	1.086 (0.252-4.667)	
Histologic grade <sup>c</sup>					0.140
G1	8	19	627	1	
G2	44	82	603	1.425 (0.648-3.135)	
G3 + G4	26	40	473	1.632 (0.734-3.628)	
Radical resection <sup>d</sup>					0.014
R0	43	81	627	1	
R1 + R2 + RX	32	53	394	1.812 (1.126-2.916)	
Risk score					0.009
Low	33	71	691	1	
High	45	71	486	1.852 (1.165-2.944)	

Abbreviations: PAAD: pancreatic adenocarcinoma; MST: media survival time; HR: hazard ratio; CI: confidence interval.

Notes: <sup>a</sup>Information of alcohol history were not acquired in 12 patients; <sup>b</sup>Information of tumor stage were not acquired in 1 patient; <sup>c</sup>Information of histologic grade were not acquired in 1 patient; <sup>d</sup>Information of radical resection were not acquired in 8 patients.

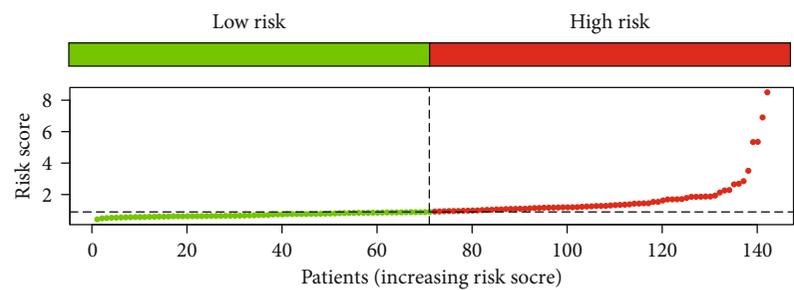
from the data portal of TCGA. For validation, gene expression microarray datasets (DSE62452) were downloaded from the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). The data used in this study were downloaded from GEO and TCGA, so data were acquired and applied in accordance with GEO and TCGA publishing guidelines and data access policies. Therefore, no additional approval from an ethics committee was required.

**2.2. Immune-Related Gene Screening.** A list of immune-related genes was extracted from the datasets of immune system process (M13664) and immune response (M19817) from MSigDB (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>) [16]. The expression data of these genes were screened from the PAAD cases of TCGA and GEO. Immune-related prognostic genes were further evaluated with univariate Cox proportional hazard regression by using a “survival” package (version 2.44-1.1) on the R platform (version 3.6.1). Genes with  $P < 0.05$  and  $|\text{hazard ratio (HR)}| > 1.00$  were considered to be prognostic risk genes, and their expression levels were significantly associated with OS in PAAD.

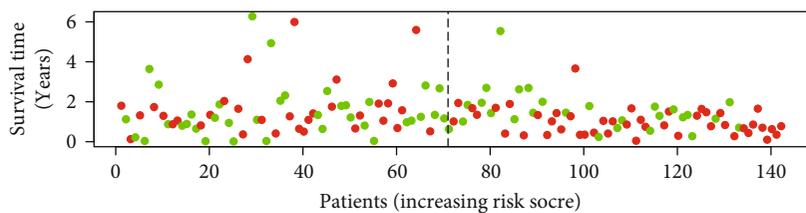
**2.3. Signature Development.** The prognostic immune-related genes were analyzed using multivariate Cox regression anal-

ysis with OS as the dependent variable to evaluate their roles in predicting PAAD survival. A prognostic risk score model was prepared via the linear combination of the expression levels of immune-related genes with the multivariate Cox regression coefficient ( $\beta$ ) as the weight [17]. The risk scores were calculated using the prognostic gene signatures. The risk score formula was as follows: risk score = expression of gene1  $\times \beta_1$  + expression of gene2  $\times \beta_2$  + ... expression of gene n  $\times \beta_n$  [18, 19]. A total of 142 cases were divided into high- and low-risk groups based on the median risk score.  $|\text{HR}| > 1.0$  and  $P < 0.05$  were selected among the TCGA and GEO datasets as a cut-off. Then, three genes were chosen for signature development. A receiver operating characteristic (ROC) curve was established over time on the R platform to assess the accuracy of the risk score model for predicting the prognosis of PAAD [20].

**2.4. Functional and Pathway Enrichment Analyses.** The functional enrichment analyses of the immune-related genes mainly involving Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were carried out using the “cluster Profiler” R package [21]. GO analysis revealed the functions of the immune-related genes in biological processes (BP), cellular components (CC), and molecular functions (MF), and the KEGG

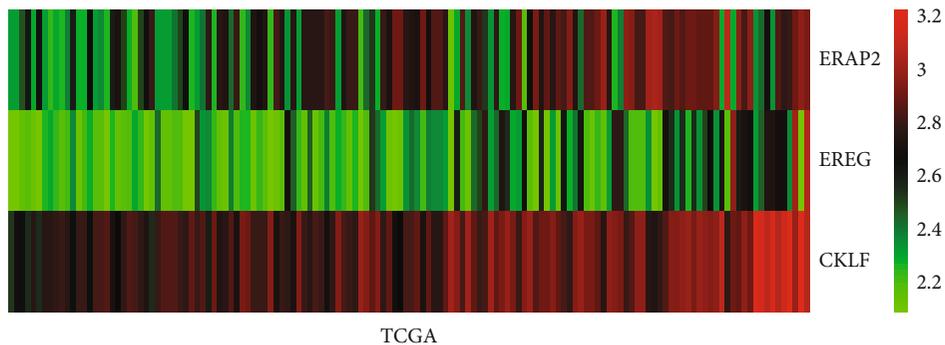


(a)

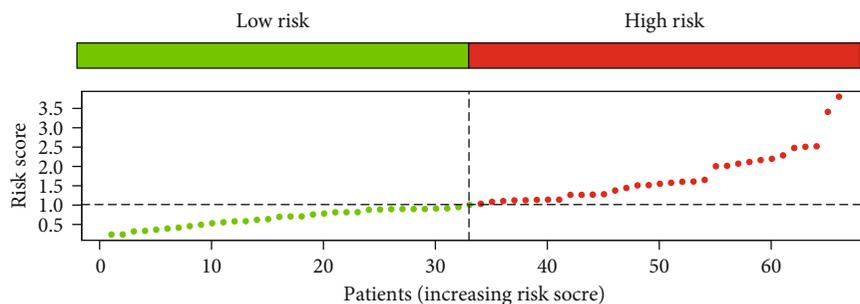


● Dead  
● Alive

(b)



(c)



(d)

FIGURE 2: Continued.

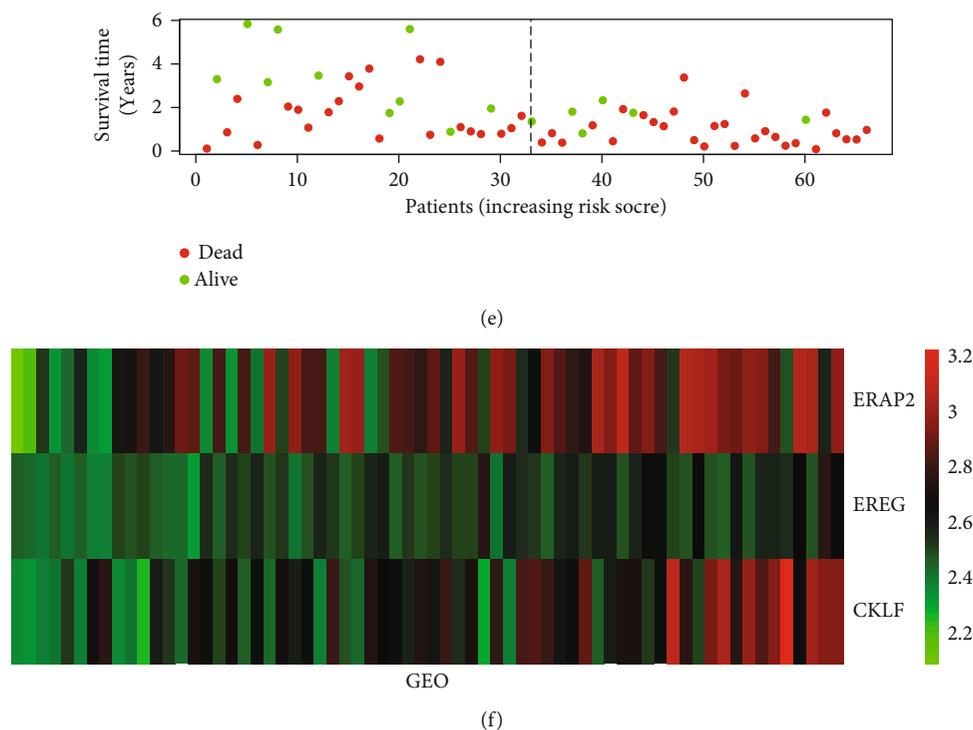


FIGURE 2: Prognostic risk score model analysis of three prognostic immune genes in PAAD patients. In the database from TCGA, from top to bottom are (a) the risk score, (b) patient survival status distribution, and (c) three hub gene expression heat maps for the low- and high-risk groups. In the GEO database, from top to bottom are (d) the risk score, (e) patient survival status distribution, and (f) three hub gene expression heat maps for the low- and high-risk groups.

analysis showed the pathway enrichment of the immune-related genes.  $P < 0.05$  was considered statistically significant.

**2.5. Predictive Nomogram Construction and Validation.** A prognostic signature based on the expression of immune-related genes was comprehensively analyzed to further assess the prognostic model. After collinearity was tested, a Nomogram was predicted using a stepwise Cox regression model to predict the 1-, 3-, and 5-year OS rates of the patients with PAAD the datasets from TCGA and GEO. Kaplan-Meier analysis and area under the curve (AUC) comparison of the ROC curve were applied to predict and observe the OS rate for assessing the performance of the prognostic nomogram. We not only compared the clinical outcomes of the low-risk and high-risk groups but also evaluated the prognostic value of PAAD with a risk score through a nomogram. The potential application of risk scores in the prediction of clinical status was also explored.

**2.6. Statistical Analysis (OS Curve).** Kaplan-Meier survival analysis by log-rank test was conducted to identify the immune-related genes associated with the prognosis of PAAD. Univariate, multivariate, and Cox regression analyses and principal component analysis (PCA) were performed in R and SPSS version 22.0 (Chicago, IL, USA). Univariate and multivariate Cox regression analyses as well as performed to assess survival. HRs and 95% confidence intervals (CIs) were

calculated to identify OS-associated genes. Statistical significance was set at  $P < 0.05$ .

### 3. Results

**3.1. Identification of Immune-Related Genes with Prognostic Value.** The clinical information and gene expression profiles of 142 PAAD cases were downloaded from the database from TCGA for further analysis. A total of 332 immune-related genes were selected from MSigDB v4.0 [16] (immune system process, immune response); (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>). For validation, gene expression microarray datasets (DSE62452; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62452>) were downloaded from the GEO dataset. Immune-related genes were screened in the same way. Immune-related prognostic genes were further evaluated via univariate Cox proportional hazard regression by using the survival software package. Coexpression networks of the identified immune-related genes were constructed. Then, three genes (chemokine-like factor [CKLF], endoplasmic reticulum aminopeptidase 2 [ERAP2], and epi-regulin [EREG]) that indicated a high risk in the TCGA, and GEO databases were highlighted (Table S1). The Kaplan-Meier analysis results of CKLF, ERAP2, and EREG are shown in Figure 1.

**3.2. Prognostic Model Construction and ROC Curve Analysis.** A signature of three immune-related genes was developed

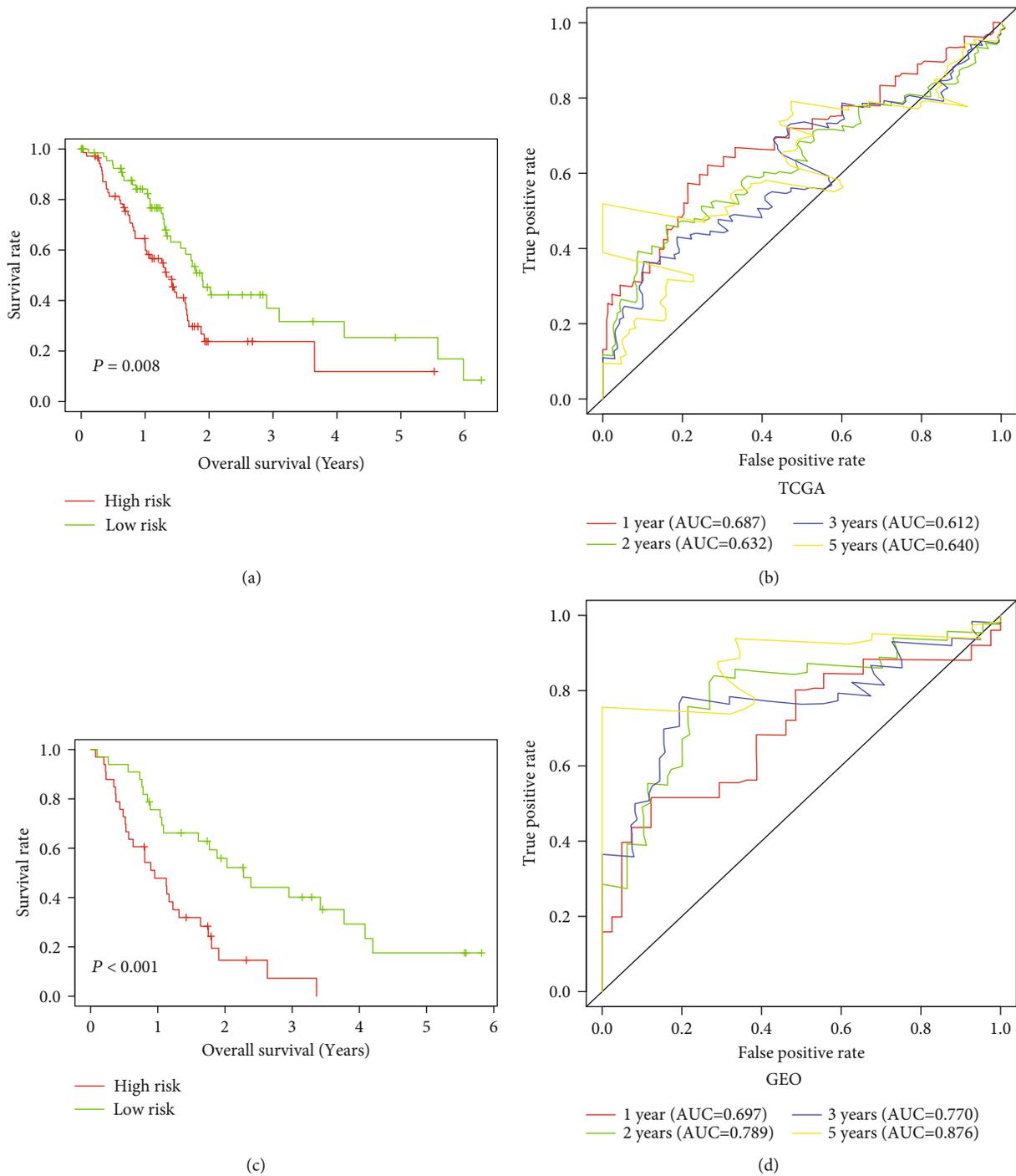


FIGURE 3: Kaplan-Meier curves for low-risk and high-risk populations in different databases. ROC curve for predicting survival in PAAD patients by the risk score (a) Kaplan-Meier curves for low- and high-risk groups using the database from TCGA ( $P = 0.008$ ). (b) The ROC curve for predicting the survival rate of PAAD based on the risk score of the TCGA database. (c) Kaplan-Meier curves for low- and high-risk groups in GEO ( $P < 0.001$ ). (d) The ROC curve for predicting the survival rate of PAAD based on the risk score of the GEO database.

using a risk score method [18, 19]. The risk score formula was as follows: risk score = expression of *CKLF*  $\times$  (0.9452) + expression of *ERAP2*  $\times$  (0.2968) + expression of *EREG*  $\times$  (0.3896). The 142 patients in the database from TCGA were divided into high- and low-risk groups based on their median risk score. Survival analysis showed that the OS of the patients in the high-risk group was shorter (high risk and

low risk: 486 vs. 691 days) than it was in the low-risk group. The risk of death significantly increased in the patients with high-risk scores ( $P$  value = 0.009; HR = 1.852; 95%CI = 1.165 – 2.944; Table 1, Figures 2(a) and 2(b)). In addition, the risk heat maps of the gene expression profiles of *CKLF*, *ERAP2*, and *EREG* indicated that the expression levels of these genes were higher in the high-risk group than in the

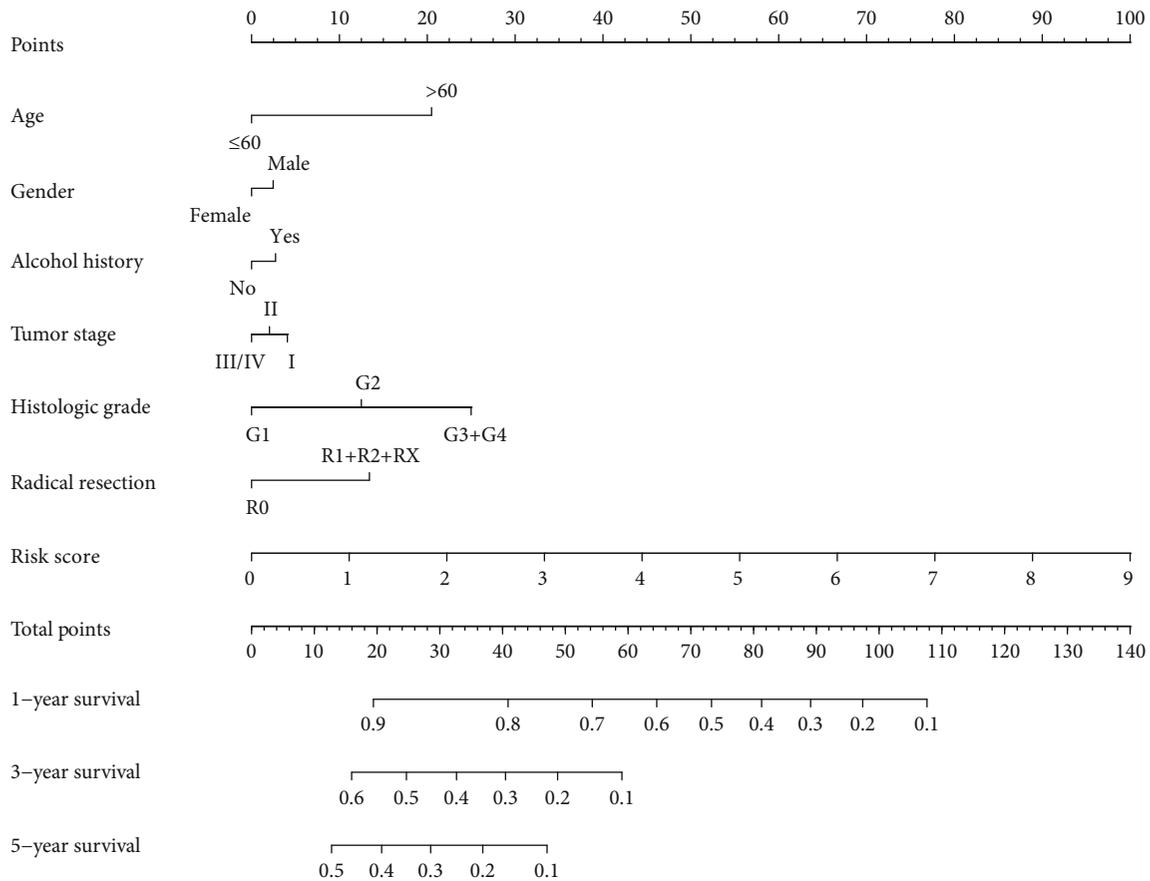


FIGURE 4: The relationship between risk score and clinical information. A prognostic nomogram predicting 1-, 3-, and 5-year OS of PAAD in the database from TCGA.

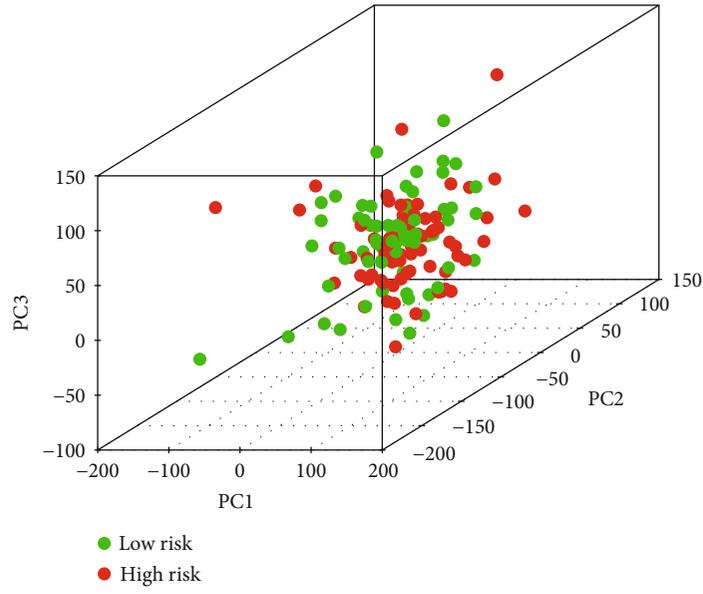
low-risk group (Figure 2(c)). Validation by the GEO database further confirmed that the survival rate of the high-risk score group was lower than that of the low-risk group (Figures 2(d)–2(f)). The Kaplan-Meier curves of the OS rates of the patients with PAAD from the different groups were stratified in terms of the signature from the TCGA and GEO datasets. In the database from TCGA, the OS of the low-risk group was longer than that of the high-risk group ( $P = 0.008$ , Figure 3(a)). The same result was obtained from the GEO database ( $P < 0.001$ , Figure 3(c)). Time-dependent ROC curve analysis was carried out to assess the prediction accuracy of the prognostic signature. Our results showed that the prognostic signature from the database from TCGA in the current study performed well in predicting 1-, 2-, and 3-year survival rates. The area under the curve values for 1-, 2-, and 3-year survival were 0.687, 0.632, and 0.612, respectively (Figure 3(b)). The prognostic signatures of the GEO database also performed well in predicting 1-, 2-, and 3-year survival rates (Figure 3(d)).

**3.3. Predictive Nomogram Construction and Validation.** A comprehensive nomogram survival analysis was conducted to investigate the relationship between the risk scores and clinical characteristics of OS for patients with PAAD. A nomogram was drawn with RMS and its auxiliary packages based on the clinical information of PAAD and the risk score.

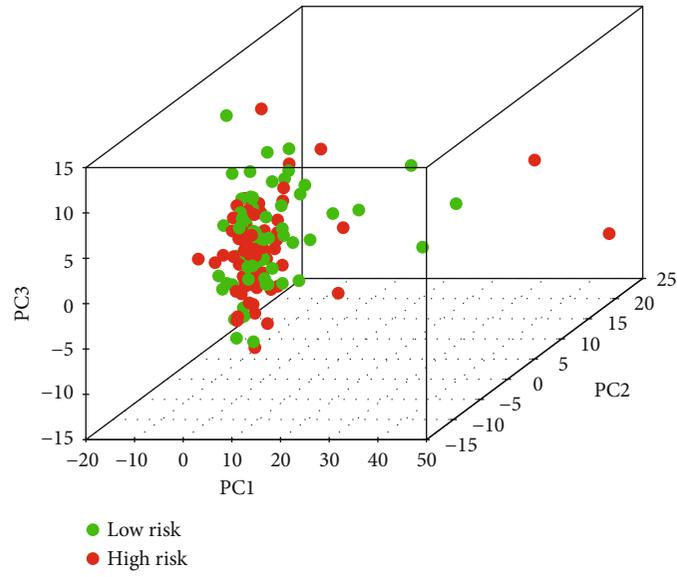
The results confirmed that the prognostic markers of the risk score significantly influenced the risk points, whereas other clinical features had a lower effect on the risk points (Figure 4). In our nomogram, the shortcoming was that the prognostic signature of tumor stage could not perform well in PAAD.

**3.4. Low- and High-Risk Groups Displayed Different Immune Status.** PCA was performed to study the differences between low- and high-risk populations based on the expression profiles of all genes, immune-related genes, and risk-related genes (Figures 5(a)–5(c)). Our results indicated that low- and high-risk groups were usually distributed in different directions. According to the prognostic signature of immune-related genes, patients in the high-risk group could be clearly distinguished from patients in the low-risk group. Therefore, the immune status of PAAD with a specific gene signature was different from other genes.

**3.5. Functional Enrichment Analysis of Genes.** Immune-related genes were subjected to functional enrichment analysis by applying the cluster Profiler R package. The immune-related genes from the database from TCGA in the BP group were mainly enriched in T cell activation and regulation of leukocyte activation. The genes in the CC group were significantly enriched on the side of the membrane and the



(a)



(b)

FIGURE 5: Continued.

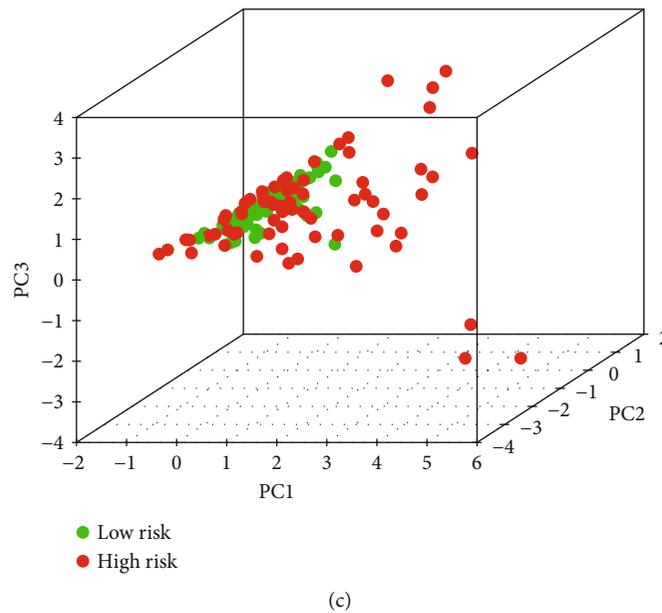


FIGURE 5: Principal component analysis between low-risk and high-risk groups based on different classification methods. (a) All genes. (b) Immune genes. (c) Risk genes.

external side of the plasma membrane. The genes in the MF group were mainly enriched in cytokine receptor activity and cytokine receptor binding ( $P < 0.05$ ; Figure 6(a)). KEGG analysis suggested that most of the immune-related gene pathways were significantly linked to cytokine-cytokine receptor interaction, Th17 cell differentiation, and hematopoietic cell lineage ( $P < 0.05$ ; Figure 6(b)). The workflow of this study is shown in Figure 7.

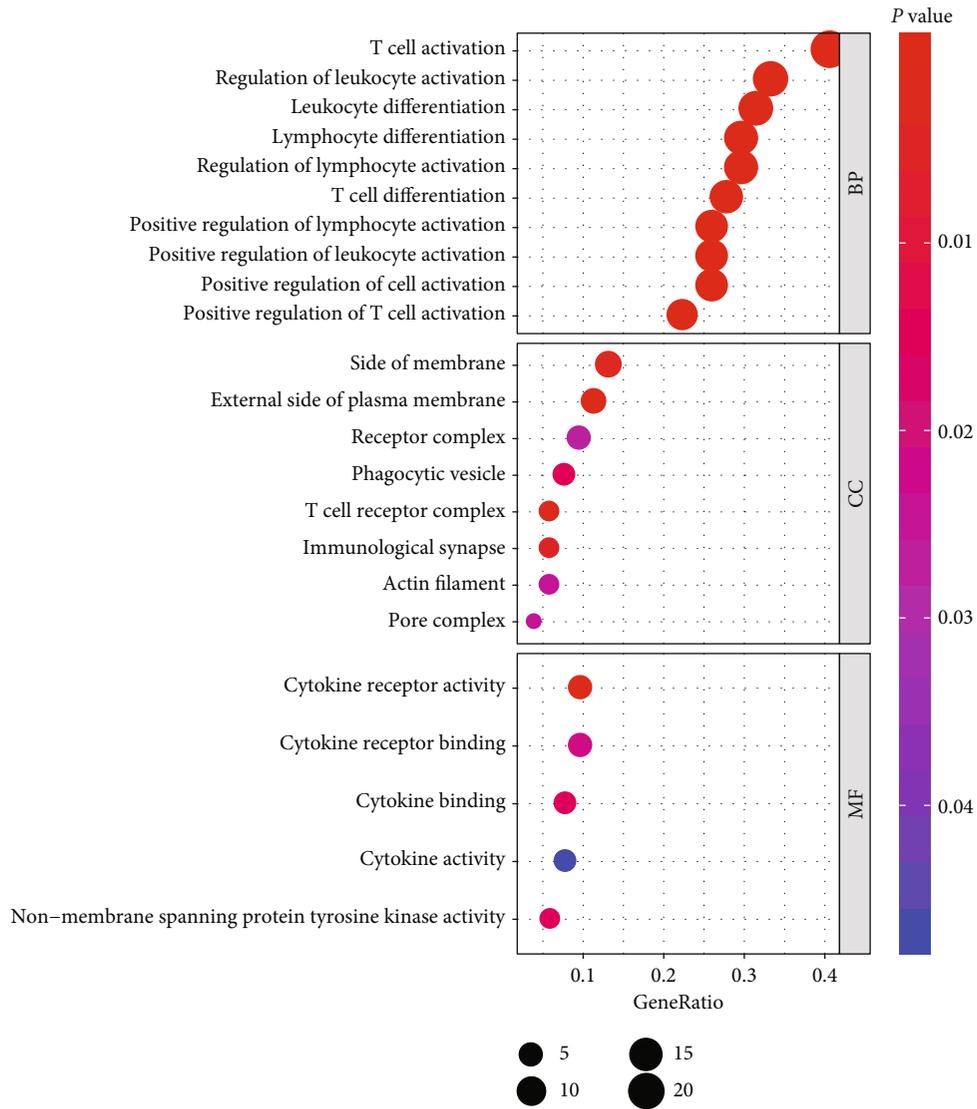
#### 4. Discussion

Pancreatic cancer is a common cause of death and has poor prognosis. The 5-year survival rate of patients with this disease is approximated to be less than 5% [1]. PAAD is treated with numerous strategies, including surgery, neoadjuvant therapy, chemotherapy, targeted molecular therapy, radiation therapy, and immunotherapy. However, the effects of treatment methods are limited, and novel methods for PAAD treatment should be further explored to provide patients with personalized treatment and improve their survival. With remarkable progress in bioinformatics, the mining of TCGA databases has been increasingly applied to predict cancer prognosis in many studies [15, 22, 23]. In the current study, we attempted to identify immune-related genes that contributed to the OS of patients with PAAD using a database from TCGA.

First, two datasets (TCGA and DSE62452) were collected to study the prognosis of immune-related genes in patients with PAAD. A total of 332 immune-related genes were extracted from a database from TCGA. Immune-related genes were also extracted from DSE62452 and identified by constructing a coexpression network of immune-related genes. Genes with  $P < 0.05$  and  $|HR| > 1.00$  were considered to be prognostic risk genes. Then, three identified genes, *CKLF*, *ERAP2*, and *EREG*, were found to be associated with

a high risk in the databases from TCGA and GEO. The signature of three immune-related genes was developed using a risk score method, and the patients with PAAD were divided into low- and high-risk groups based on their median risk score. The results showed that the prognosis of the patients in the high-risk group was worse than that in the low-risk group. Multivariate Cox regression analysis was conducted to determine the independent prognostic factors of PAAD. Then, the prognostic signature was comprehensively analyzed on the basis of immune-related gene expression. Kaplan-Meier analysis and the AUC comparison of the ROC curve confirmed that the three immune-related gene signatures were reliable for OS prediction. Next, a nomogram was established and integrated with a signature of three immune-related genes and clinical data, and OS was accurately predicted. The PCA results revealed that the prognostic signature of immune-related genes could clearly distinguish patients in the high-risk group from those in the low-risk group. The KEGG pathway analysis indicated that most of the immune-related genes were significantly associated with cytokine-cytokine receptor interaction, Th17 cell differentiation, and hematopoietic cell lineage. Therefore, the nomogram could be used as a progression indicator and predictor of the OS of patients with PAAD.

In our current study, three genes associated with poor prognosis of PAAD were identified: *CKLF*, *ERAP2*, and *EREG*. *CKLF* is a protein-encoding gene whose product is a cytokine. Cytokines are small proteins that play important roles in immune and inflammatory responses. The protein encoded by *CKLF* is an effective chemoattractant of neutrophils, monocytes, and lymphocytes [24, 25]. Some studies have shown that high neutrophil levels are associated with the prognosis of patients with PAAD [26, 27]. *CKLF* has four isoforms, designated *CKLF1–4*; among them, *CKLF1* has the highest expression level [24]. Previous reports suggested that



(a)

FIGURE 6: Continued.

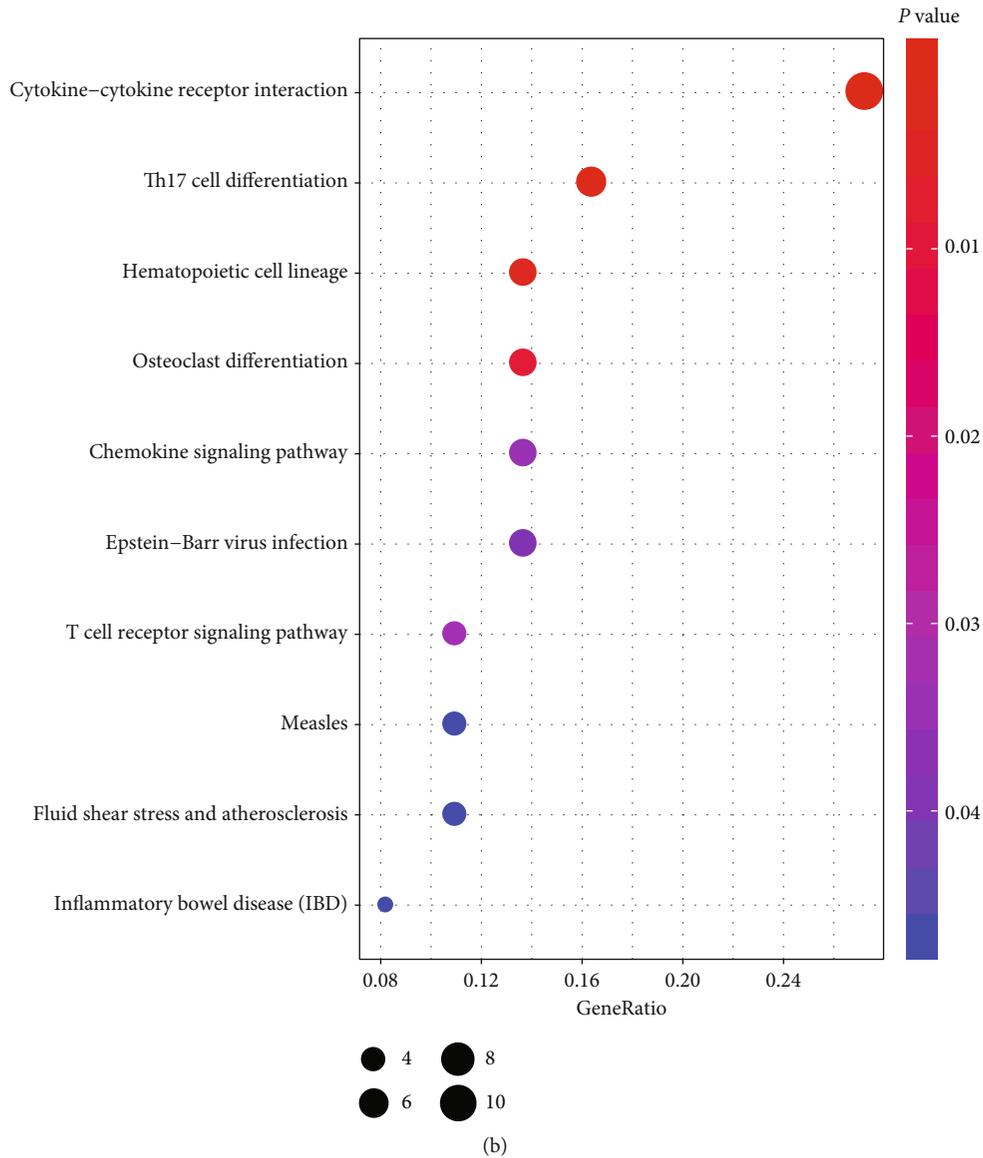


FIGURE 6: Functional enrichment analysis results of immune genes. (a) GO term enrichment results. (b) KEGG pathway analysis.

CKLF1 expression may play an essential role in the development of atopic dermatitis [28] and psoriasis [29]. The GO annotation related to this gene included chemokine activity. Chemokines not only participate in cancer-associated inflammation but also promote tumor development and progression [30]. CKLF1 is highly expressed in malignant ovarian cancer, providing a new basis for the clinical diagnosis and treatment of tumors [31]. Therefore, the high *CKLF* expression in PAAD might be the cause of PAAD pathogenesis and progression.

The GO annotations related to *ERAP2* included metallopeptidase and aminopeptidase activities. Human *ERAP2* was initially identified as a homolog of human placental leucine aminopeptidase or insulin-regulated aminopeptidase [32]. *ERAP2* increases susceptibility to autoimmune diseases, infectious diseases, and cancer because of its genetic variability [33]. *ERAP2* is associated with several immune-mediated diseases, including ankylosing spondylitis, psoriasis, and

Crohn’s disease [32, 34]. *ERAP2*-related pathways include class I MHC-mediated antigen processing and presentation and the innate immune system. Data from the cBioPortal website (<http://www.cbioportal.org>) showed that *ERAP2* is highly expressed in pancreatic cancer. *ERAP1* and *ERAP2* may be important targets that enhance T and NK cell-mediated immune responses against established cancers [35].

*EREG* is closely related to pancreatic cancer development [36]. *EREG* is a member of the epidermal growth factor (EGF) family of peptide growth factors [37]. The stimulation of the EGFR pathway also promotes tumor cell migration, adhesion, and metastasis [38]. *AREG* and *EREG* are required for autocrine EGFR signaling, indicating that *EREG* plays an important role in tumor progression [39]. *EREG* is expressed in a variety of adult tissues, and its increased expression or activity appears to promote the progression of several different human malignancies [40]. Studies have shown that *EREG* enhances the migration and chemotaxis ability of adipose-

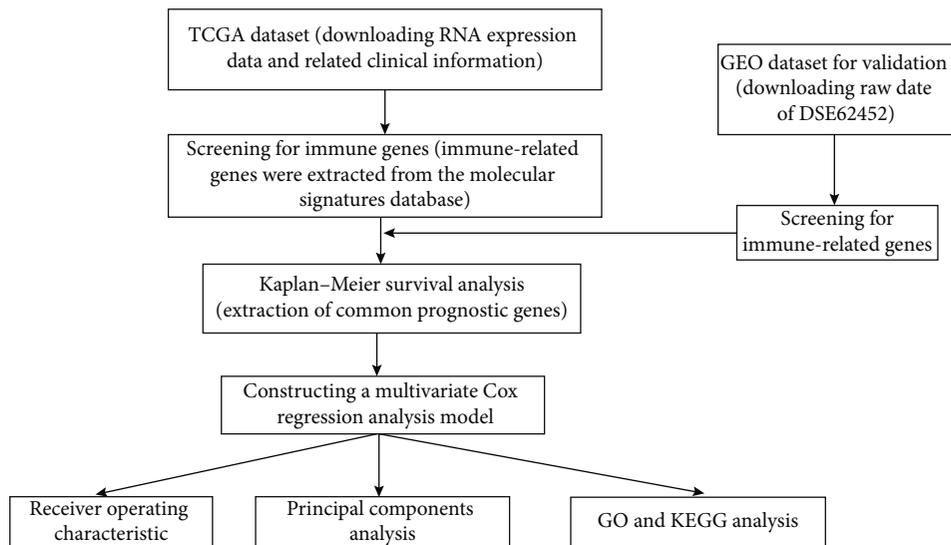


FIGURE 7: Flow chart of data preparation processing analysis and validation.

derived stem cells [41]. In colorectal cancer, EREG serves as a biomarker of anti-EGFR therapy [42]. The inhibition of EGFR signaling in pancreatic cancer may lead to a decrease in the growth and invasion of pancreatic tumors [43]. Previous studies revealed that EREG is upregulated in pancreatic cancer and stimulates the growth of pancreatic cancer cells [36]. Thus, our research method could be reliably used to identify prognosis-related genes for PAAD.

These risk genes could be potential molecular targets for PAAD treatment. The results of gene screening demonstrated that *EREG* was associated with the prognosis of pancreatic cancer, so our approach could be used to accurately screen prognostic genes. For the newly discovered *CKLF* and *ERAP2*, further experiments are needed to determine whether they are related to the prognosis of PAAD and verify our results. Cancer immunotherapy, which relies on the immune system to eliminate primary tumors, has shown unique advantages for cancer treatment [44]. The 2018 Nobel Prize in physiology or medicine was awarded to pioneers in the field of cancer immunotherapy, which has been a tremendously successful area of work [45]. Tumor immunotherapy has been widely explored in various fields, such as nanotechnology-enhanced immunotherapy [46, 47].

With the continuous improvement of immunotherapy technology, the indication of PAAD immunotherapy needs to be further studied. Immunotherapy is recommended for patients with pancreatic cancer with MSI/MMR molecular characteristics and distant metastasis [48]. Most patients diagnosed with PAAD for the first time are in advanced stages, missed the best time for surgical treatment, or the patient cannot tolerate surgery. PAAD is composed of dense connective tissue and highly infiltrating immune cells, which is very easy to induce chemotherapy resistance [49]. After treatment with conventional chemotherapy drugs, the patient's survival status and quality of life did not improve significantly. Therefore, in order to obtain a better prognosis for patients, immunotherapy can be selected according to the immune microenvironment of PAAD. Some studies have

shown that the establishment of an “immune score” system for expressing pancreatic cancer could be used to assess the degree of immune cell infiltration in the tumor immune microenvironment [50]. Improving our understanding of how PAAD immune and stromal components interact and the tumor microenvironment can help improve our immunotherapy [51, 52]. Future strategies using immunotherapy to treat pancreatic cancer include changing immune checkpoint inhibitors from monotherapy to combination therapy and combining immunotherapy with chemotherapy, radiation therapy, and targeted therapy [53]. Due to the obvious heterogeneity among individuals with PAAD, the uses of immunotherapy will be based on the results of genetic testing, so that a personalized treatment plan can be implemented to improve the efficacy of the treatment [54]. In the future, gene sequencing technology is expected to detect and identify high-risk immune genes of PAAD and provide new directions for precise immunotherapy of PAAD.

Immune-related genes are associated with survival and may be used as biomarkers to assess the suitability of various immunotherapies. Our immune prognosis gene signature provided a novel idea and methods related to the molecular mechanism and prognosis prediction of PAAD. This signature might help solve the problem of poor immunogenicity in PAAD and improve the effectiveness and safety of cancer immunotherapy. Fortunately, because of remarkable advancements in whole-genome sequencing technology and bioinformatics, some high-throughput tumor databases have been generated and can be used for public academic research. The pathways involved in the prognosis of PAAD can also be further studied. The risk gene signaling pathway of PAAD can be inhibited to achieve the purpose of immunotherapy. The nomogram also indicated that the risk score of our prognosis genes could reliably predict the OS of patients with PAAD. However, our study had certain limitations compared with previous studies. First, our clinical information was mainly obtained from databases from TCGA and GEO. Some patients' clinical information was incomplete, and

detailed data on patient prognosis were unavailable. Second, a signature of three immune-related genes was generated, and a normal tissue control was not prepared. As such, our data were not convincing enough to establish a prognostic model. The expression level of prognosis-related genes and their molecular mechanisms in the pathogenesis and progression of PAAD should be further explored experimentally. The screened genes could be verified through real-time PCR and Western blot.

## 5. Conclusions

In summary, we constructed immune-related gene coexpression networks and identified a signature of three immune-related genes that had prognostic value for patients with PAAD. The prognosis for patients in the high-risk group was worse than that in the low-risk group. Further research on these immune-related genes would help fully understand the potential links of the immune system and responses to the prognosis of PAAD. The prognostic nomogram could reliably predict the OS of patients with PAAD and might be used as a guide for the diagnosis and immunotherapy of PAAD. However, our results were not further verified experimentally because of the limitations of this investigation, so more in-depth studies should be conducted to reveal the relationship between the prognosis of PAAD and immune-related genes.

## Abbreviations

PAAD:	Pancreatic adenocarcinoma
TCGA:	The Cancer genome atlas
GEO:	Gene expression omnibus
MSigDB:	Molecular signatures database
GSEA:	Gene set enrichment analysis
HR:	Hazard ratio
CI:	Confidence intervals
OS:	Overall survival
ROC:	Receiver operating characteristic
AUC:	The area under curve
PCA:	Principal components analysis
GO:	Gene ontology
KEGG:	Kyoto encyclopedia of genes and genomes
BP:	Biological process
CC:	Cellular component
MF:	Molecular function.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest, and all authors should confirm its accuracy.

## Authors' Contributions

Guolin Wu and Zhenfeng Deng contributed equally to this work. Guolin Wu and Zhenfeng Deng designed the study protocol. Jilong Wang, Banghao Xu, Zongrui Jin, Jingjing Zeng, and Minhao Peng conducted the study and analyzed the data. Guolin Wu and Zhenfeng Deng wrote the manuscript. Zhang Wen and Ya Guo revised the manuscript. Zhang Wen, Ya Guo, and Jilong Wang strived for financial support. Guolin Wu and Zhenfeng Deng contributed equally to this work.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 81560387, 81902983), the Guangxi Natural Science Foundation of China (No. 2018JJB140382), "Medical Excellence Award" funded by the creative research development grant from the first affiliated hospital of Guangxi Medical University (No. 180327), the Guangxi medical and health technology development and application project (No. S2019097, S2018100) and the Emergency and Medical Rescue Talent Small Highland in Guangxi Zhuang Autonomous Region•Open Topics of Key Laboratory for Emergency Medicine in Guangxi Universities (No. GXJZ201501).

## Supplementary Materials

Table S1: the univariate survival analysis results of the TCGA and GEO cohort. (*Supplementary Materials*)

## References

- [1] M. Ilic and I. Ilic, "Epidemiology of pancreatic cancer," *World Journal of Gastroenterology*, vol. 22, no. 44, pp. 9694–9705, 2016.
- [2] L. Y. Chen, C. S. Cheng, C. Qu et al., "Overexpression of CBX3 in pancreatic adenocarcinoma promotes cell cycle transition-associated tumor progression," *International Journal of Molecular Sciences*, vol. 19, no. 6, pp. 1768–1782, 2018.
- [3] 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.
- [4] L. Rahib, B. D. Smith, R. Aizenberg, A. B. Rosenzweig, J. M. Fleshman, and L. M. Matrisian, "Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States," *Cancer Research*, vol. 74, no. 11, pp. 2913–2921, 2014.
- [5] T. Kamisawa, L. D. Wood, T. Itoi, and K. Takaori, "Pancreatic cancer," *The Lancet*, vol. 388, no. 10039, pp. 73–85, 2016.
- [6] C. Yuan, V. Morales-Oyarvide, A. Babic et al., "Cigarette smoking and pancreatic cancer survival," *Journal of Clinical Oncology*, vol. 35, no. 16, pp. 1822–1828, 2017.
- [7] V. Bagnardi, M. Rota, E. Botteri et al., "Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis," *British Journal of Cancer*, vol. 112, no. 3, pp. 580–593, 2015.

- [8] S. N. Camara, T. Yin, M. Yang et al., "High risk factors of pancreatic carcinoma," *Journal of Huazhong University of Science and Technology-Medical Sciences*, vol. 36, no. 3, pp. 295–304, 2016.
- [9] D. B. Zhen, K. G. Rabe, S. Gallinger et al., "BRCA1, BRCA2, PALB2, and CDKN2A mutations in familial pancreatic cancer: a PACGENE study," *Genetics in Medicine*, vol. 17, no. 7, pp. 569–577, 2015.
- [10] T. E. Clancy, "Surgery for pancreatic cancer," *Hematology-Oncology Clinics of North America*, vol. 29, no. 4, pp. 701–716, 2015.
- [11] V. P. Balachandran, G. L. Beatty, and S. K. Dougan, "Broadening the impact of immunotherapy to pancreatic cancer: challenges and opportunities," *Gastroenterology*, vol. 156, no. 7, pp. 2056–2072, 2019.
- [12] J. M. Fritz and M. J. Lenardo, "Development of immune checkpoint therapy for cancer," *Journal of Experimental Medicine*, vol. 216, no. 6, pp. 1244–1254, 2019.
- [13] I. H. Sahin, G. Askan, Z. I. Hu, and E. M. O'Reilly, "Immunotherapy in pancreatic ductal adenocarcinoma: an emerging entity?," *Annals of Oncology*, vol. 28, no. 12, pp. 2950–2961, 2017.
- [14] P. Bailey, D. K. Chang, M.-A. Forget et al., "Exploiting the neoantigen landscape for immunotherapy of pancreatic ductal adenocarcinoma," *Scientific Reports*, vol. 6, no. 1, p. 35848, 2016.
- [15] Kim, Jo, Jang et al., "PAC-5 Gene Expression Signature for Predicting Prognosis of Patients with Pancreatic Adenocarcinoma," *Cancers*, vol. 11, no. 11, p. 1749, 2019.
- [16] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, pp. 15545–15550, 2005.
- [17] I. S. Lossos, D. K. Czerwinski, A. A. Alizadeh et al., "Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes," *The New England Journal of Medicine*, vol. 350, no. 18, pp. 1828–1837, 2004.
- [18] W. Wang, Z. Zhao, F. Yang et al., "An immune-related lncRNA signature for patients with anaplastic gliomas," *Journal of Neuro-Oncology*, vol. 136, no. 2, pp. 263–271, 2018.
- [19] X. Liao, G. Zhu, R. Huang et al., "Identification of potential prognostic microRNA biomarkers for predicting survival in patients with hepatocellular carcinoma," *Cancer Management and Research*, vol. 10, pp. 787–803, 2018.
- [20] C. Combescurie, T. V. Perneger, D. C. Weber, J. P. Daurès, and Y. Foucher, "Prognostic ROC curves a method for representing the overall discriminative capacity of binary markers with right-censored time-to-event endpoints," *Epidemiology*, vol. 25, no. 1, pp. 103–109, 2014.
- [21] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *Omics-A Journal of Integrative Biology*, vol. 16, no. 5, pp. 284–287, 2012.
- [22] B. Li, K. Pu, and X. Wu, "Identifying novel biomarkers in hepatocellular carcinoma by weighted gene co-expression network analysis," *Journal of Cellular Biochemistry*, vol. 120, no. 7, pp. 11418–11431, 2019.
- [23] J. X. Gu, X. Zhang, R. C. Miao et al., "Six-long non-coding RNA signature predicts recurrence-free survival in hepatocellular carcinoma," *World Journal of Gastroenterology*, vol. 25, no. 2, pp. 220–232, 2019.
- [24] W. Han, Y. Lou, J. Tang et al., "Molecular cloning and characterization of chemokine-like factor 1 (CKLF1), a novel human cytokine with unique structure and potential chemotactic activity," *Biochemical Journal*, vol. 357, no. 1, pp. 127–135, 2001.
- [25] Y. Lou, D. Xia, W. Han, Y. Wang, and D. J. G. Ma, "Molecular cloning and characterization of rat chemokine-like factor 1 and 2," *Gene*, vol. 307, no. 1, pp. 125–132, 2003.
- [26] Q. Qi, Y. Geng, M. Sun, P. Wang, and Z. Chen, "Clinical implications of systemic inflammatory response markers as independent prognostic factors for advanced pancreatic cancer," *Pancreatology*, vol. 15, no. 2, pp. 145–150, 2015.
- [27] T. Chao, E. E. Furth, and R. H. Vonderheide, "CXCR2-dependent accumulation of tumor-associated neutrophils regulates T-cell immunity in pancreatic ductal adenocarcinoma," *Cancer Immunology Research*, vol. 4, no. 11, pp. 968–982, 2016.
- [28] G. Y. Yang, X. Chen, Y. C. Sun, C. L. Ma, and G. Qian, "Chemokine-like factor 1 (CLFK1) is over-expressed in patients with atopic dermatitis," *International Journal of Biological Sciences*, vol. 9, no. 8, pp. 759–765, 2013.
- [29] Y. Tan, Y. Wang, L. Li, J. Xia, S. Peng, and Y. He, "Chemokine-like factor 1-derived C-terminal peptides induce the proliferation of dermal microvascular endothelial cells in psoriasis," *PLoS One*, vol. 10, no. 4, article e0125073, 2015.
- [30] N. Mukaida and T. Baba, "Chemokines in tumor development and progression," *Experimental Cell Research*, vol. 318, no. 2, pp. 95–102, 2012.
- [31] D.-D. Liu, X.-Y. Song, P.-F. Yang et al., "Progress in pharmacological research of chemokine like factor 1 (CKLF1)," *Cytokine*, vol. 102, pp. 41–50, 2018.
- [32] A. Hattori and M. Tsujimoto, "Endoplasmic reticulum aminopeptidases: biochemistry, physiology and pathology," *Journal of Biochemistry*, vol. 154, no. 3, pp. 219–228, 2013.
- [33] Y. Yao, N. Liu, Z. Zhou, and L. Shi, "Influence of ERAP1 and ERAP2 gene polymorphisms on disease susceptibility in different populations," *Human Immunology*, vol. 80, no. 5, pp. 325–334, 2019.
- [34] P. C. Robinson, E. Lau, P. Keith et al., "ERAP2 functional knockout in humans does not alter surface heavy chains or HLA-B27, inflammatory cytokines or endoplasmic reticulum stress markers," *Annals of the Rheumatic Diseases*, vol. 74, no. 11, pp. 2092–2095, 2015.
- [35] M. Compagnone, L. Cifaldi, and D. Fruci, "Regulation of ERAP1 and ERAP2 genes and their dysfunction in human cancer," *Human Immunology*, vol. 80, no. 5, pp. 318–324, 2019.
- [36] Z. Zhu, J. Kleeff, H. Friess et al., "Epiregulin is up-regulated in pancreatic cancer and stimulates pancreatic cancer cell growth," *Biochemical and Biophysical Research Communications*, vol. 273, no. 3, pp. 1019–1024, 2000.
- [37] M. Massip-Copiz, M. Clauzure, A. G. Valdivieso, and T. A. Santa-Coloma, "Epiregulin (EREG) is upregulated through an IL-1 $\beta$  autocrine loop in Caco-2 epithelial cells with reduced CFTR function," *Journal of Cellular Biochemistry*, vol. 119, no. 3, pp. 2911–2922, 2018.
- [38] R. S. Herbst, "Review of epidermal growth factor receptor biology," *International Journal of Radiation Oncology • Biology • Physics*, vol. 59, no. 2, pp. S21–S26, 2004.
- [39] B. L. Carpenter, M. Chen, T. Knifley et al., "Integrin  $\alpha\beta 4$  promotes autocrine epidermal growth factor receptor (EGFR) signaling to stimulate migration and invasion toward hepatocyte growth factor (HGF)," *Journal of Biological Chemistry*, vol. 290, no. 45, pp. 27228–27238, 2015.

- [40] D. J. Riese and R. L. Cullum, "Epiregulin: roles in normal physiology and cancer," *Seminars in Cell & Developmental Biology*, vol. 28, pp. 49–56, 2014.
- [41] Y. Cao, L. Wang, H. Yang et al., "Epiregulin promotes the migration and chemotaxis ability of adipose-derived mesenchymal stem cells via mitogen-activated protein kinase signaling pathways," *Journal of Cellular Biochemistry*, vol. 119, no. 10, pp. 8450–8459, 2018.
- [42] F. Bormann, S. Stinzing, S. Tierling et al., "Epigenetic regulation of Amphiregulin and Epiregulin in colorectal cancer," *International Journal of Cancer*, vol. 144, no. 3, pp. 569–581, 2018.
- [43] C. M. Sloss, F. Wang, M. A. Palladino, and J. C. Cusack Jr., "Activation of EGFR by proteasome inhibition requires HB-EGF in pancreatic cancer cells," *Oncogene*, vol. 29, no. 21, pp. 3146–3152, 2010.
- [44] W. Sang, Z. Zhang, Y. Dai, and X. Chen, "Recent advances in nanomaterial-based synergistic combination cancer immunotherapy," *Chemical Society Reviews*, vol. 48, no. 14, pp. 3771–3810, 2019.
- [45] X. Zang, "2018 Nobel Prize in medicine awarded to cancer immunotherapy: immune checkpoint blockade - a personal account," *Genes & Diseases*, vol. 5, no. 4, pp. 302–303, 2018.
- [46] S. Lim, J. Park, M. K. Shim et al., "Recent advances and challenges of repurposing nanoparticle-based drug delivery systems to enhance cancer immunotherapy," *Theranostics*, vol. 9, no. 25, pp. 7906–7923, 2019.
- [47] M. S. Goldberg, "Improving cancer immunotherapy through nanotechnology," *Nature Reviews Cancer*, vol. 19, no. 10, pp. 587–602, 2019.
- [48] S. Macherla, S. Laks, A. Naqash, A. Bulumulle, E. Zervos, and M. Muzaffar, "Emerging Role of Immune Checkpoint Blockade in Pancreatic Cancer," *International Journal of Molecular Sciences*, vol. 19, no. 11, p. 3505, 2018.
- [49] I. M. Stromnes, K. E. DelGiorno, P. D. Greenberg, and S. R. Hingorani, "Stromal reengineering to treat pancreas cancer," *Carcinogenesis*, vol. 35, no. 7, pp. 1451–1460, 2014.
- [50] J. Galon, B. Mlecnik, G. Bindea et al., "Towards the introduction of the 'Immunoscore' in the classification of malignant tumours," *The Journal of Pathology*, vol. 232, no. 2, pp. 199–209, 2014.
- [51] R. J. Torphy, R. D. Schulick, and Y. Zhu, "Understanding the immune landscape and tumor microenvironment of pancreatic cancer to improve immunotherapy," *Molecular Carcinogenesis*, 2020.
- [52] Y. Sunami and J. Kleeff, "Immunotherapy of pancreatic cancer," *Progress in Molecular Biology and Translational Science*, vol. 164, pp. 189–216, 2019.
- [53] J. Wu and J. Cai, "Dilemma and Challenge of Immunotherapy for Pancreatic Cancer," *Digestive diseases and sciences*, 2020.
- [54] J. S. Bowers, S. R. Bailey, M. P. Rubinstein, C. M. Paulos, and E. R. Camp, "Genomics meets immunity in pancreatic cancer: Current research and future directions for pancreatic adenocarcinoma immunotherapy," *Oncology Reviews*, vol. 13, no. 2, p. 430, 2019.