

Research Article

Development of a Gene Risk Signature for Patients of Pancreatic Cancer

Tao Liu ^{1,2}, Long Chen,³ Guili Gao,⁴ Xing Liang,¹ Junfeng Peng,¹ Minghui Zheng,¹ Judong Li,¹ Yongqiang Ye ², and Chenghao Shao ¹

¹Department of Pancreatic-biliary Surgery, Changzheng Hospital, Navy Medical University, Shanghai, China

²Department of Hepatobiliary Surgery, Heze Municipal Hospital, No. 2888, Caozhou Road, Mudan District, Heze 274000, Shandong, China

³Department of Gastrointestinal Surgery, Heze Municipal Hospital, No. 2888, Caozhou Road, Mudan District, Heze 274000, Shandong, China

⁴Department of Cardiology, Heze Municipal Hospital, No. 2888, Caozhou Road, Mudan District, Heze 274000, Shandong, China

Correspondence should be addressed to Yongqiang Ye; qe2420404@163.com and Chenghao Shao; gq0042402@163.com

Received 16 September 2021; Accepted 25 October 2021; Published 7 January 2022

Academic Editor: Kalidoss Rajakani

Copyright © 2022 Tao Liu 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. Pancreatic cancer is a highly malignant solid tumor with a high lethality rate, but there is a lack of clinical biomarkers that can assess patient prognosis to optimize treatment. **Methods.** Gene-expression datasets of pancreatic cancer tissues and normal pancreatic tissues were obtained from the GEO database, and differentially expressed genes analysis and WGCNA analysis were performed after merging and normalizing the datasets. Univariate Cox regression analysis and Lasso Cox regression analysis were used to screen the prognosis-related genes in the modules with the strongest association with pancreatic cancer and construct risk signatures. The performance of the risk signature was subsequently validated by Kaplan–Meier curves, receiver operating characteristic (ROC), and univariate and multivariate Cox analyses. **Result.** A three-gene risk signature containing CDKN2A, BRCA1, and UBL3 was established. Based on KM curves, ROC curves, and univariate and multivariate Cox regression analyses in the TRAIN cohort and TEST cohort, it was suggested that the three-gene risk signature had better performance in predicting overall survival. **Conclusion.** This study identifies a three-gene risk signature, constructs a nomogram that can be used to predict pancreatic cancer prognosis, and identifies pathways that may be associated with pancreatic cancer prognosis.

1. Introduction

Pancreatic cancer is a highly malignant solid tumor [1], and its incidence and mortality rates continue to increase [2]. The most common symptoms in patients with pancreatic cancer are abdominal pain, anorexia, fatigue, and weight loss [3]; pancreatic cancer lacks specific biomarkers [4], and the main serum markers commonly used today are carcinoembryonic antigen and carbohydrate antigen 19-9; however, their sensitivity is not ideal [3]. Surgery is the most important approach in the treatment of pancreatic cancer. Due to atypical symptoms and the lack of effective screening tools, many patients have progressed to an unresectable state at the time of diagnosis. With the development of research,

radiotherapy, chemotherapy, targeted therapy, and immunotherapy have been applied in the clinical treatment of pancreatic cancer with some success. However, for individual patients, a model that can effectively predict prognosis is still needed to guide clinical selection of treatment. The development of high-throughput sequencing technology has made it possible to discover prognosis-related biomarkers.

Weighted gene co-expression network analysis (WGCNA) has been used to detect correlations between gene modules consisting of highly correlated gene clusters and specific clinical features [5] and has been widely used to identify gene modules associated with clinical features of various cancers. In the present study, we identified gene

modules highly correlated with pancreatic cancer tissue by WGCNA. In addition to this, we further identified genes associated with prognosis by univariate Cox regression analysis and Lasso Cox regression analysis.

2. Materials and Methods

2.1. Gene Expression Dataset Collection and Processing. Download datasets related to pancreatic cancer gene expression from GEO [6] (<https://www.ncbi.nlm.nih.gov/geo/>). The selection criteria in this article are (1) pancreatic cancer samples and normal samples were obtained from human samples; (2) the training and validation datasets needed to contain survival data; (3) using microarray gene-expression technology or RNA-Seq technology. Datasets GSE15471 [7], GSE16515 [8], GSE28735 [9], and GSE57495 [10] were selected for differential analysis and weighted gene co-expression analysis datasets, in which the GSE28735 dataset containing survival data was used to construct the prognostic gene signature; GSE78229 [11] was selected as the external validation dataset. Using the “sva” package [12] of R package, the GSE15471, GSE16515, GSE28735, and GSE57495 datasets are merged and normalized.

2.2. Differentially Expressed Gene Analysis and Weighted Gene Co-Expression Analysis (WGCNA). Identification of differentially expressed genes (DEGs) by “limma” [13] package in R, setting $|\log \text{Fold change} (\log \text{FC})| \geq 1$ and adjusted $p < 0.05$ as standard. And we used “ggplot2” package [14] and “pheatmap” package [15] to plot heatmap and volcano map of DEGs. GO and KEGG enrichment analysis of the differential genes is carried out by R package “clusterProfiler” [16] and “GOplot” [17].

GSE15471, GSE16515, GSE28735, and GSE57495 data were merged, and weighted gene co-expression analysis was used to identify co-expressed gene modules using the “WGCNA” package [5] of R. GO and KEGG analysis was then applied to the genes within the module with the highest correlation to tumorigenesis.

2.3. Construction of Risk Signature. The univariate Cox regression and Lasso regression analyses of genes within the module were performed by the “survival” package [18] and “glmnet” package [19] to screen for prognosis-related genes within the module and construct a risk signature. Kaplan–Meier analysis was used to examine the survival outcomes of the high-risk and low-risk groups, and the predictive power of the risk signature was assessed using the area under the curve (AUC) of the controlled operating characteristic (ROC) curve. Prognosis-related genes were subsequently calculated in relation to risk score.

2.4. Construction and Valuation of Nomogram. Evaluation of prognostic factors are important for stage, grade, and risk score in the GSE78229 dataset by univariate and multifactor cox regression analysis using the “forestplot” package [20] in R. Nomogram which was

drawn through “rms” package [21] and “regplot” package [22] to examine the accuracy of the nomogram by measuring the performance of the nomogram by the C-index. Calibration curves at 1, 3, and 5 years survival. Diagonal lines are used as a reference for best prediction. The R package “timeROC” was used by graph receiver operating characteristic curves (ROC) to determine the prognostic performance of the gene signature and nomogram.

3. Results

3.1. Differentially Expressed Genes’ (DEGs) Identification. The GSE15471, GSE16515, GSE28735, and GSE57495 datasets were merged and normalized by the R package “sva” [12]. Subsequent differentially expressed gene analysis using the “limma” package [13] identified 77 DEGs containing 52 upregulated and 25 downregulated genes. We next performed GO and KEGG enrichment analysis of differential genes and plotted the circles. The GO analysis (Figure 1(c)) of their biological process (BP) was mainly enriched in extracellular structure organization and extracellular matrix organization; the cellular component (CC) was mainly enriched in proteinaceous extracellular matrix and extracellular matrix; the molecular function (MF) was mainly involved in extracellular matrix structural constituent and platelet-derived growth factor binding. However, no pathways were enriched in KEGG analysis.

3.2. Weighted Gene Co-Expression Network Construction and Key Module Identification. Weighted gene co-expression networks of GSE15471, GSE16515, GSE28735, and GSE57495 were constructed by the “WGCNA” package in R (version 4.0). The samples were clustered, and the sample clustering tree was drawn after removing the outliers (Figure 2(a)). We chose $\beta = 7$ ($R^2 = 0.9$) to construct the scale-free network (Figure 2(b)). Eight co-expression modules were finally identified (which contained a grey module composed of genes that could not be categorized) (Figures 2(c) and 2(d)). Next, module-clinical feature correlation heat maps were drawn to assess the correlation between modules and clinical features (tumor vs. normal). The brown module had the strongest correlation with tumor tissue ($r = 0.53$ and $p = 7e - 22$). Therefore, the brown module was selected as the key module for further analysis.

3.3. Construction of a Multigene Signature. Univariate Cox regression analysis was performed on 166 genes within the brown module to screen 30 genes associated with survival at $p < 0.05$, followed by Lasso Cox regression analysis in GSE28735 to calculate risk scores for pancreatic cancer patients. Risk score = $(\text{CDKN2A} \times 0.672) + (\text{BRCA1} \times -0.142) + (\text{UBL3} \times -0.185)$.

Patients were divided into a high-risk group ($n = 21$) and a low-risk group ($n = 21$) according to the median risk score. There was a significant difference in overall survival (OS) between the high- and low-risk groups ($p = 1.385e - 02$) (Figure 3(c)). The areas under the curve at 1, 2, and 3 years

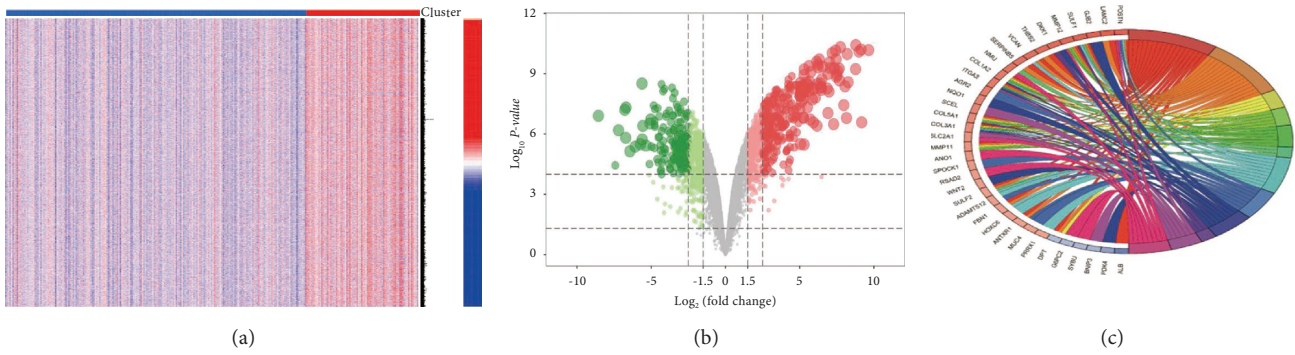


FIGURE 1: Differentially expressed genes' (DEGs) identification. Heat map (a) and volcano map (b) of gene-expression profiles of pancreatic cancer tissues and normal tissues after merge of four datasets, GSE15471, GSE16515, GSE28735, and GSE57495. Differentially expressed genes were screened using $|\log_{2}FC| \geq 1$ and adjusted using $p < 0.05$, with red representing upregulated genes and blue representing downregulated genes. (c) GO enrichment analysis of differentially expressed genes.

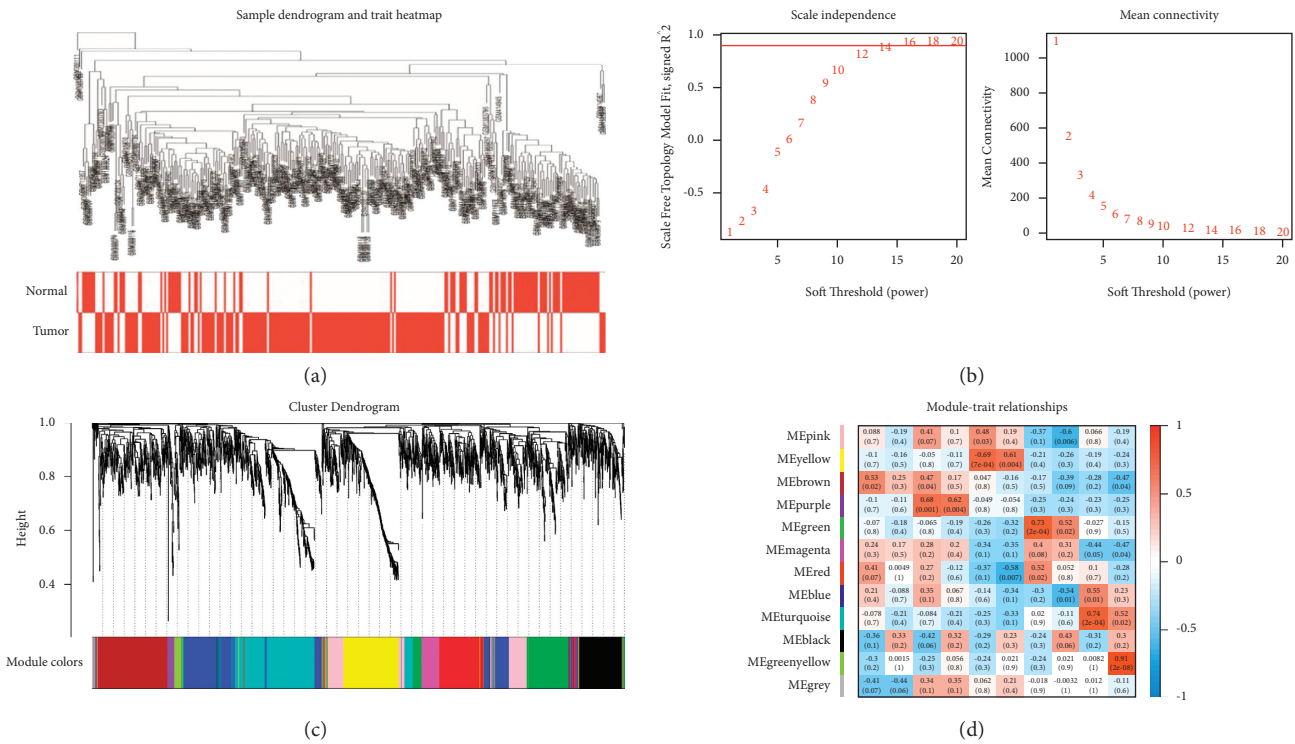


FIGURE 2: Weighted gene co-expression network construction and key module identification. (a) Cluster dendrogram of pancreatic cancer samples and normal samples. (b) According to the scale-free index and the mean connectivity to screen soft threshold. (c) The cluster dendrogram of co-expression network modules. (d) Relationships between module and trait.

was 0.75, 0.893, and 0.733, respectively (Figure 3(d)). The prognostic and prognostic accuracy of the three-gene signature was subsequently validated using GSE78229 as a test cohort. There was also a significant difference in OS between the high-risk and low-risk groups in the test cohort ($p = 5.418e - 03$) (Figure 3(e)), with areas under the curve at 1, 2, and 3 years of 0.773, 0.731, and 0.741, respectively (Figure 3(f)).

The performance of the risk signature was further evaluated by univariate (Figure 4(a)) and multivariate (Figure 4(b)) Cox regression analysis in the train and test

cohort, respectively. The results showed that the risk score was significantly associated with OS. Multivariate cox regression analysis revealed that three-gene signatures were independent predictors of outcome in pancreatic cancer patients.

In addition, we created a prognostic nomogram to help physicians predict overall patient survival in the clinic (Figure 4(c)). The calibration curve (Figure 4(d)) of the nomogram and the area under the curve of ROC (Figure 4(e)) showed a good concordance between prediction and observation.

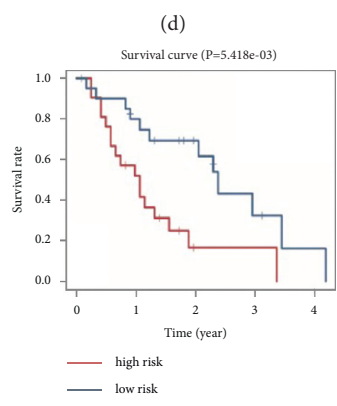
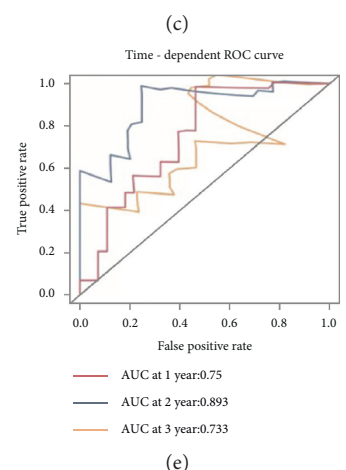
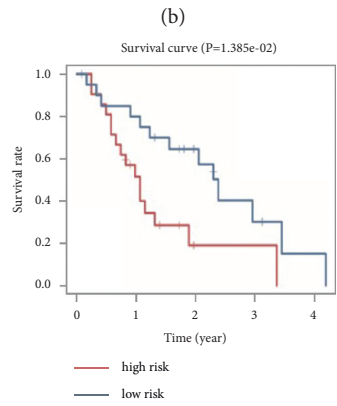
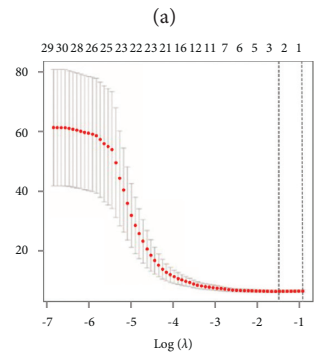
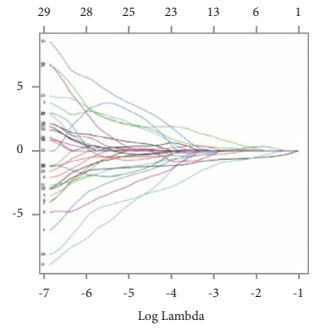
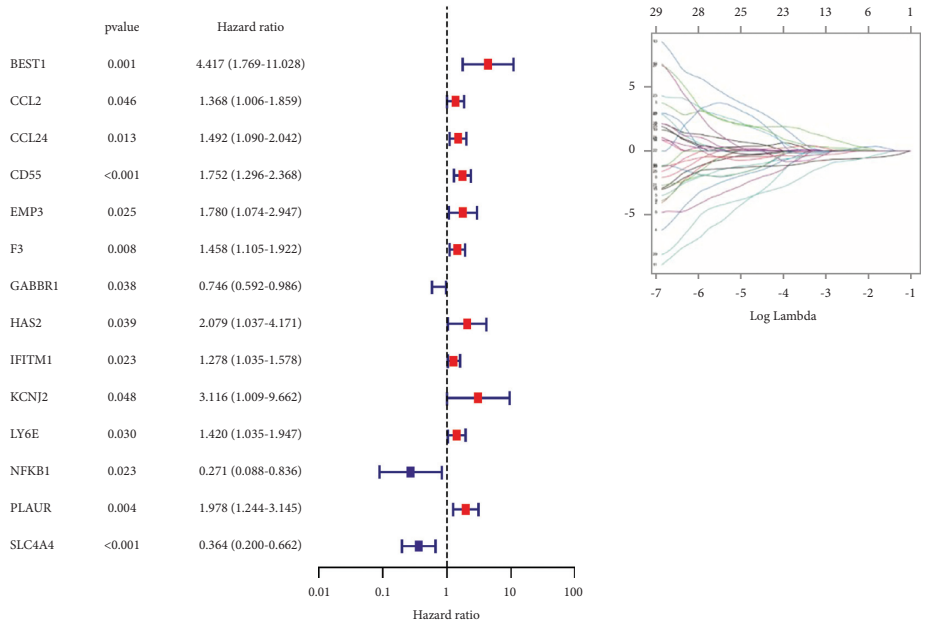


FIGURE 3: Continued.

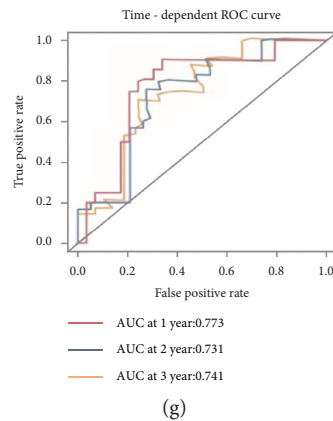


FIGURE 3: Construction of a multigene signature. (a) Univariate Cox analysis of genes within the brown module, screening of prognosis-related genes, and forest mapping. (b) The Lasso coefficient profiles of prognosis-related genes. (c) The partial likelihood deviance is plotted against $\log(\lambda)$. Kaplan–Meier plot of overall survival of patients is in the high-risk and low-risk groups in the train cohort (d) and test cohort (f). ROC curves for three-gene signatures in train cohort (e) and test cohort (g).

4. Discussion

Pancreatic cancer is highly malignant, lacks reliable early screening methods, and has a poor prognosis, with an expected 5-year survival rate of approximately 9% [1]. Therefore, there is an urgent need to find biomarkers that affect the prognosis of pancreatic cancer in clinical treatment, which will facilitate the assessment of patient prognosis and will help to improve the prognosis by tailoring the treatment to the individual patient.

Tumor development is the result of multigene interactions, and therefore, an increasing number of risk signatures are used to predict prognosis [23–26]. In this study, we proposed a three-gene (CDKN2A, BRCA1, and UBL3) risk signature by WGCNA and Lasso Cox regression analysis for predicting overall survival in pancreatic cancer patients, with statistically significant differences in overall survival between high- and low-risk groups in the train cohort and test cohort. We then evaluated the prognostic performance of risk signature with the AUC of ROC, and the results showed that the risk signature could predict overall survival of pancreatic cancer patients accurately. Subsequent univariate and multivariate Cox analyses showed that the risk score could predict prognosis as an independent prognostic factor. In addition, we combined clinical characteristics to construct nomogram that can be used in the clinic to guide personalized treatment.

Among the risk genes we identified, CDKN2A (cytoskeleton-associated protein 2-like) was significantly highly expressed in the high-risk group and positively correlated with risk score; BRCA1 (glutathione S-transferase Mu 5) and UBL3 (Ubiquitin-like 3) were significantly down regulated in the high-risk group and negatively correlated with the risk score. CDKN2A has been reported to promote lung adenocarcinoma invasion and is correlated with poor prognosis [27]. Monteverde et al. found that CDKN2A could promote non-small cell lung cancer (NSCLC) progression by regulating transcriptional elongation, and targeting CDKN2A could enhance therapeutic response in patients with NSCLC

[28]. Li et al. also found that CDKN2A knockdown inhibited proliferation, migration, invasion, and epithelial mesenchymal transition in glioblastoma cells [29]. Bioinformatics studies have found that CDKN2A is also associated with breast cancer [30], prostate cancer [31], and colorectal cancer [32]. BRCA1 has been reported to play an oncogenic role in bladder cancer, with significantly lower expression levels in cancer tissues than in normal tissues, and overexpression of BRCA1 reduced cell proliferation, migration, and colony-forming ability [33]. In contrast, in bladder cancer, upregulation of BRCA1 was able to resist oxidative stress, thereby promoting bladder cancer cell growth [34]. Pitt et al. also found mutations in BRCA1 in thyroid cancer. In addition to this, bioinformatics studies have found that BRCA1 is also associated with ovarian [35], colorectal [36], and gastric [37] cancers. Consistent with our speculation, Zhao et al. found that, in NSCLC, UBL3 acts as a tumor suppressor gene to inhibit cancer cell proliferation [38].

GSEA analysis revealed differences in 2 key signaling pathways between high- and low-risk groups. Base excision repair (BER) removes endogenous DNA damages that occur at all times in human cells, and its defects are associated with tumorigenesis [39], but cancer cells are also able to tolerate oxidative stress through increased BER activity, and targeting BER can improve the efficacy of radio/chemotherapy [40]. Our results show that the BER pathway is enriched in the high-risk group, suggesting that the BER pathway is active in high-risk patients, possibly leading to shorter survival by affecting their sensitivity to clinical treatment. An increasing number of studies have found that abnormal metabolism affects patient prognosis [41, 42], and the enrichment of propanoate metabolism pathway in the low-risk group suggests that the risk signature may affect patient prognosis through tumor metabolism.

In summary, our study identified a 3-gene risk signature for predicting prognosis, and the value of this risk signature was validated in an external test cohort. By combining this risk signature with clinical tumor pathology staging, a visual nomogram was created to facilitate the prediction of survival

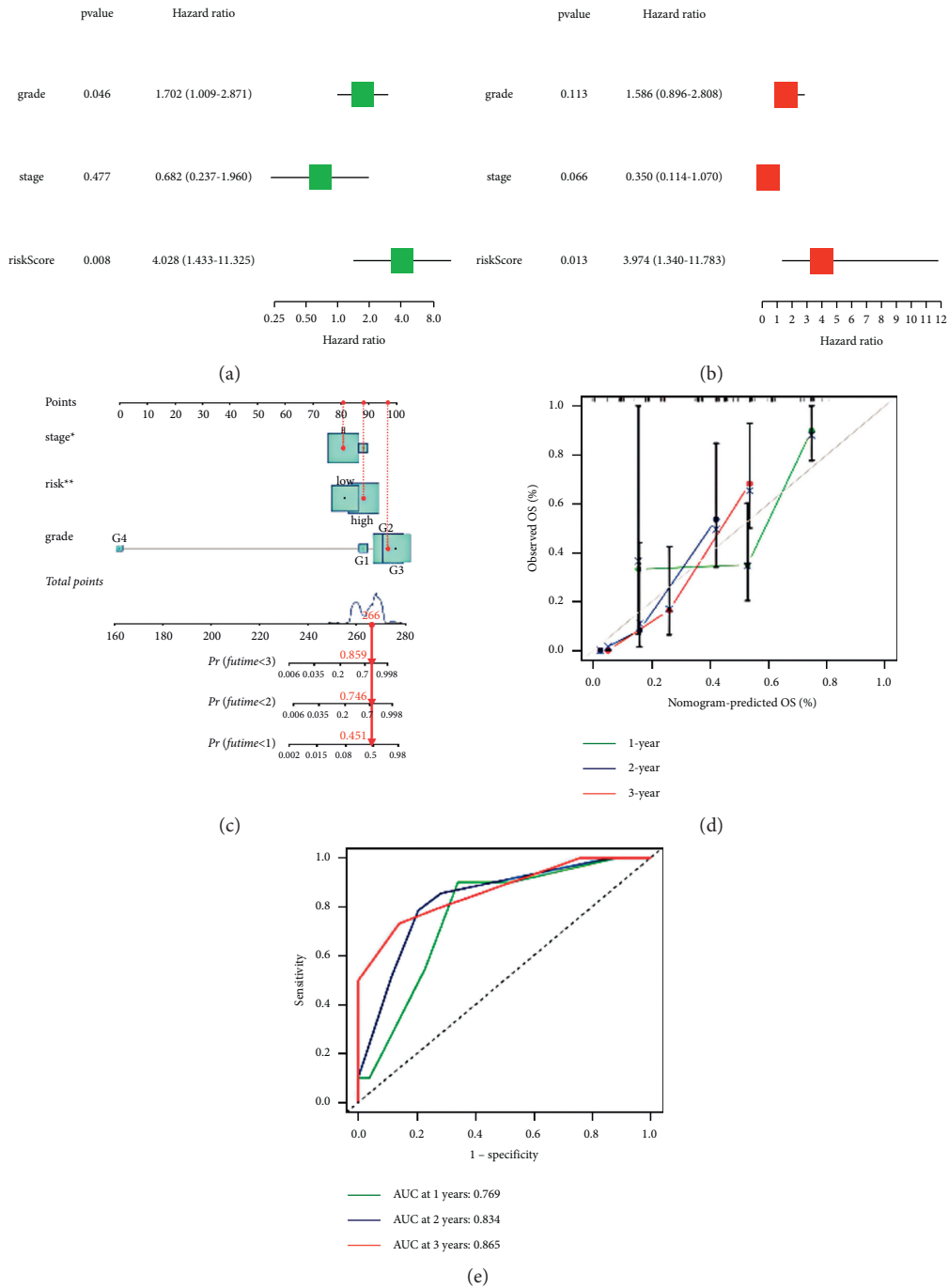


FIGURE 4: Evaluation of the predictive value of three-gene signatures and the creation of nomogram. (a) Test cohort risk signature univariate. (b) Multivariate Cox regression analysis forest plot. (c) Nomogram. (d) Calibration plot. (e) ROC curve for test cohort.

outcomes. We further analyzed risk signature-related signaling pathways and risk genes' expression and survival analysis in a variety of cancers, providing a research direction for exploring prognosis-related genes in in vivo and in vitro experiments. We further analyzed risk signature-related signaling pathways and risk gene expression and survival analysis in a variety of cancers, providing a research direction for exploring prognosis-related genes in cytological and animal experiments.

Data Availability

The analyzed datasets generated during the present study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest in the present study.

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] 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.
- [3] H. Zhu, T. Li, Y. Du, and M. Li, "Pancreatic cancer: challenges and opportunities," *BMC Medicine*, vol. 16, no. 1, p. 214, 2018.
- [4] L. Zhang, S. Sanagapalli, and A. Stoita, "Challenges in diagnosis of pancreatic cancer," *World Journal of Gastroenterology*, vol. 24, no. 19, pp. 2047–2060, 2018.
- [5] P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," *BMC Bioinformatics*, vol. 9, no. 1, p. 559, 2008.
- [6] E. Clough and T. Barrett, "The gene expression omnibus database," *Methods in Molecular Biology*, vol. 1418, pp. 93–110, 2016.
- [7] L. Badea, V. Herlea, S. O. Dima, T. Dumitrascu, and I. Popescu, "Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia," *Hepato-gastroenterology*, vol. 55, no. 88, pp. 2016–2027, 2008.
- [8] H. Pei, L. Li, B. L. Fridley et al., "FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt," *Cancer Cell*, vol. 16, no. 3, pp. 259–266, 2009.
- [9] G. Zhang, A. Schetter, P. He et al., "DPEP1 inhibits tumor cell invasiveness, enhances chemosensitivity and predicts clinical outcome in pancreatic ductal adenocarcinoma," *PLoS One*, vol. 7, no. 2, Article ID e31507, 2012.
- [10] D.-T. Chen, A. H. Davis-Yadley, P.-Y. Huang et al., "Prognostic fifteen-gene signature for early stage pancreatic ductal adenocarcinoma," *PLoS One*, vol. 10, no. 8, Article ID e0133562, 2015.
- [11] J. Wang, S. Yang, P. He et al., "Endothelial nitric oxide synthase traffic inducer (NOSTRIN) is a negative regulator of disease aggressiveness in pancreatic cancer," *Clinical Cancer Research*, vol. 22, no. 24, pp. 5992–6001, 2016.
- [12] L. C. Chu, M. G. Goggins, and E. K. Fishman, "Diagnosis and detection of pancreatic cancer," *The Cancer Journal*, vol. 23, no. 6, pp. 333–342, 2017.
- [13] I. Diboun, L. Wernisch, C. A. Orengo, and M. Koltzenburg, "Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma," *BMC Genomics*, vol. 7, no. 1, p. 252, 2006.
- [14] L. Wilkinson, "ggplot2: elegant graphics for data analysis by WICKHAM, H.," *Biometrics*, vol. 67, no. 2, pp. 678–679, 2011.
- [15] H. C. Ho, A. Knudby, and W. A. Huang, "Spatial framework to map heat health risks at multiple scales," *International Journal of Environmental Research and Public Health*, vol. 12, no. 12, pp. 16110–23, 2015.
- [16] 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.
- [17] W. Walter, F. Sánchez-Cabo, and M. Ricote, "GOplot: an R package for visually combining expression data with functional analysis: Figure 1," *Bioinformatics*, vol. 31, no. 17, pp. 2912–2914, 2015.
- [18] C. S. Groeneveld, V. S. Chagas, S. J. M. Jones et al., "RTNsurvival: an R/Bioconductor package for regulatory network survival analysis," *Bioinformatics*, vol. 35, no. 21, pp. 4488–4489, 2019.
- [19] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for cox's proportional hazards model via coordinate descent," *Journal of Statistical Software*, vol. 39, no. 5, pp. 1–13, 2011.
- [20] J. Kim, K. S. Kaufman, and H. Bang, "Graphing ratio measures on forest plot," *Journal of the American College of Cardiology*, vol. 71, no. 5, pp. 585–586, 2018.
- [21] B. J. Wells, K. M. Lenoir, J. F. Diaz-Garelli et al., "Predicting current glycated hemoglobin values in adults: development of an algorithm from the electronic health record," *JMIR Medical Informatics*, vol. 6, no. 4, p. e10780, 2018.
- [22] R. Marshall, *Regplot: Enhanced Regression Nomogram Plot*, 2020.
- [23] Z. Meng, D. Ren, K. Zhang, J. Zhao, X. Jin, and H. Wu, "Using ESTIMATE algorithm to establish an 8-mRNA signature prognosis prediction system and identify immunocyte infiltration-related genes in Pancreatic adenocarcinoma," *Aging*, vol. 12, no. 6, pp. 5048–5070, 2020.
- [24] X.-H. Shi, X. Li, H. Zhang et al., "A five-microRNA signature for survival prognosis in pancreatic adenocarcinoma based on TCGA data," *Scientific Reports*, vol. 8, no. 1, p. 7638, 2018.
- [25] G.-W. Huang, Y.-J. Xue, Z.-Y. Wu et al., "A three-lncRNA signature predicts overall survival and disease-free survival in patients with esophageal squamous cell carcinoma," *BMC Cancer*, vol. 18, no. 1, p. 147, 2018.
- [26] Z. Meng, Q. Yuan, J. Zhao et al., "The m6A-related mRNA signature predicts the prognosis of pancreatic cancer patients," *Molecular Therapy-Oncolytics*, vol. 17, pp. 460–470, 2020.
- [27] G. Xiong, L. Li, X. Chen et al., "Up-regulation of CKAP2L expression promotes lung adenocarcinoma invasion and is associated with poor prognosis," *OncoTargets and Therapy*, vol. 12, pp. 1171–1180, 2019.
- [28] T. Monteverde, S. Sahoo, M. La Montagna et al., "CKAP2L promotes non-small cell lung cancer progression through regulation of transcription elongation," *Cancer Research*, vol. 81, no. 7, pp. 1719–1731, 2021.
- [29] Y. F. Li, W. C. Tsai, C. H. Chou et al., "CDKN2A knockdown exerts antitumor effects by increasing miR-4496 in

- glioblastoma cell lines,” *International Journal of Molecular Sciences*, vol. 22, no. 1, 2020.
- [30] L. Fan, J. Hou, and G. Qin, “Prediction of disease genes based on stage-specific gene regulatory networks in breast cancer,” *Frontiers in Genetics*, vol. 12, Article ID 717557, 2021.
- [31] T. Feng, D. Wei, Q. Li et al., “Four novel prognostic genes related to prostate cancer identified using Co-expression structure network analysis,” *Frontiers in Genetics*, vol. 12, Article ID 584164, 2021.
- [32] Y. Zhang, J. Luo, Z. Liu et al., “Identification of hub genes in colorectal cancer based on weighted gene co-expression network analysis,” *Bioscience Reports*, vol. 41, no. 7, Article ID BSR20211280, 2021.
- [33] Y. C. Jou, S. C. Wang, Y. C. Dia et al., “Anti-cancer effects and tumor marker role of glutathione S-transferase Mu 5 in human bladder cancer,” *International Journal of Molecular Sciences*, vol. 22, no. 6, Article ID 3056, 2021.
- [34] T. Li, D. Jiang, and K. Wu, “p62 promotes bladder cancer cell growth by activating KEAP1/NRF2-dependent antioxidative response,” *Cancer Science*, vol. 111, no. 4, pp. 1156–1164, 2020.
- [35] Y. Wang, L. Lei, Y. G. Chi, L. B. Liu, and B. P. Yang, “A comprehensive understanding of ovarian carcinoma survival prognosis by novel biomarkers,” *European Review for Medical and Pharmacological Sciences*, vol. 23, no. 19, pp. 8257–8264, 2019.
- [36] Y. L. Sun, Y. Zhang, Y. C. Guo, Z. H. Yang, and Y. C. Xu, “A prognostic model based on six metabolism-related genes in colorectal cancer,” *BioMed Research International*, vol. 2020, Article ID 5974350, 2020.
- [37] Y. Chen, B. Li, J. Wang et al., “Identification and verification of the prognostic value of the glutathione S-transferase Mu genes in gastric cancer,” *Oncology Letters*, vol. 20, no. 4, p. 100, 2020.
- [38] X. Zhao, Y. Zhou, Q. Hu et al., “Identification of a potential tumor suppressor gene, UBL3, in non-small cell lung cancer,” *Cancer Biology and Medicine*, vol. 17, no. 1, pp. 76–87, 2020.
- [39] S. S. Wallace, D. L. Murphy, and J. B. Sweasy, “Base excision repair and cancer,” *Cancer Letters*, vol. 327, no. 1-2, pp. 73–89, 2012.
- [40] G. J. Grundy and J. L. Parsons, “Base excision repair and its implications to cancer therapy,” *Essays in Biochemistry*, vol. 64, no. 5, pp. 831–843, 2020.
- [41] T. E. Oni, G. Biffi, L. A. Baker et al., “SOAT1 promotes mevalonate pathway dependency in pancreatic cancer,” *The Journal of Experimental Medicine*, vol. 217, no. 9, 2020.
- [42] C. Li, X. Li, G. Li et al., “Identification of a prognosis-associated signature associated with energy metabolism in triple-negative breast cancer,” *Oncology Reports*, vol. 44, no. 3, pp. 819–837, 2020.