Establishment of a Prognostic Prediction System Based on Tumor Microenvironment of Pancreatic Cancer

Bo Liu

Pidu district people's hospital of Chengdu

Tingting Fu

The third affiliated hospital of Chengdu medical college

Ping He

Pidu district people's hospital of Chengdu

Ke Xu (✉ xuke@cmc.edu.cn)

The First Affiliated Hospital of Chengdu Medical College: Chengdu Medical College The First Affiliated Hospital <https://orcid.org/0000-0002-5520-5754>

Research

Keywords: Pancreatic cancer, prognostic system, risk score, tumor microenvironment, overall survival

Posted Date: October 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-955286/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: Pancreatic cancer (PC) is an inflammatory tumor. Tumor microenvironment (TME) plays an important role in the development of PC. This study aims to explore hub genes of TME and establish a prognostic prediction system for PC.

Methods: High throughput RNA-sequencing and clinical data of PC were downloaded from TCGA and ICGC database, respectively. PC Patients were divided into High- and low-score group by using stromal, immune scores system based on ESTIMATE. Differentially expressed genes (DEGs) between High- and low-score patients were screened and survival related DEGs were identified as candidate genes by univariate COX regression analysis. Final variables for establishment of the prognostic prediction system were determined by LASSO analysis and multivariate COX regression analysis. The predictive power of the prognostic system was evaluated by internal and external validation.

Results: A total of 210 candidate genes were identified by stromal, immune scores system and survival analyses. Finally, the prognostic risk score system was constructed by the following genes: FAM57B, HTRA3, CXCL10, GABRP, SPRR1B, FAM83A, LY6D. In process of internal validation, Harrell's C-index of this prognostic risk score system was 0.73, and the area under the receiver operating characteristic curve (AUC) value of 1-year, 2-year and 3-year OS period was 0.67, 0.76 and 0.86, respectively. In the external validation set, the survival prediction C-index was 0.71, and the AUC was 0.81, 0.72, 0.78 at 1-year, 2-year and 3-year, respectively.

Conclusion: This prognostic risk score system based on TME demonstrated a good predictive capacity to the prognosis of PC.

Background

Pancreatic cancer (PC) is the seventh leading cause of death and may even rise to third in the future¹. In contrast to the overall declining trends for other malignancies, the stable mortality rate of PC is considered to be caused by the slow progress of diagnosis and treatment in recent years². Currently, surgery is still the only radical treatment for PC, and TNM stage is the most widely used prognosis indicator for PC patients. Many factors, such as positive margin, adjuvant chemotherapy, postoperative complications and back pain were also found to be significantly related to the prognosis of PC³⁻⁵. It is crucial to make an accurately prognosis prediction for making individual treatment plan to PC patients. Therefore, it is necessary to use different principles to explore rational and optimized prognosis models.

Tumor microenvironment (TME) is composed of tumor cells, inflammatory cells, endothelial cells, hematopoietic cells, fibroblasts, extracellular matrix and various bioactive secretions⁶. In the field of PC, it has been found that the surrounding microenvironment of PC plays an important role in tumor proliferation, immune escape, drug resistance and metastasis, thus affecting the prognosis of the patients⁷⁻¹⁴. Different tumor samples were given immune score and stromal score based on the

expression of relevant genes involved in TME to help us explore potential information about TME by ESTIMATE¹⁵. In this study, we used ESTIMATE to mine the genes related to TME, and construct a prognostic risk score system based on these genes for PC.

Discussion

The prognosis of pancreatic cancer was extremely poor. The overall 5-year survival rate was only 9%, and merely 37% in stage I or IIA postoperative PC patients². There were many prognostic factors for PC patients, and it was difficult to predict the prognosis accurately by relying on solo prognostic indicator. Some prognostic prediction systems were constructed with clinical parameters to improve the predictive capability, classical ones such as Botsis score, Heidelberg score, MSKCC score, etc.^{3,16,17}. However, their performances were still not entirely unsatisfactory. More accurate indicators and prediction systems should be used as the reference for individualized clinical decision-making in the future. As the mechanism of TME was gradually revealed, it was a reasonable and feasible way to explain clinical characteristics and improve clinical practice guide by using the information of TME. ESTIMATE score system was developed in 2013. The principle was to use the gene expression data of immune and stromal cells to infer the contents of them in tumor samples¹⁵. Currently, RNA sequencing data of 25 cancer species from TCGA had been analyzed by ESTIMATE. In this study, a total of 210 candidate genes related to TME and prognosis were identified by using ESTIMATE and survival analysis. These genes were significantly enriched in the construction of extracellular matrix and Cytokine-cytokine receptor interaction by GO and KEGG enrichment analysis. Finally, seven genes were screened for the construction of the prognostic risk score system which were rare reported. Now, these genes would be briefly described below. The mechanism of FAM57B and LY6D in malignant tumors has not been reported. HTRA3, which expresses a pro-apoptotic protease, promotes drug-induced cytotoxic effects and was thought to have the function of anti-tumor¹⁸. It showed inhibition of tumor proliferation and metastasis in lung cancer cells and endometrial cancer cells, and high expression of HTRA3 was associated with longer disease-free survival in non-small cell lung cancer patients^{19–21}. SPRR1B is overexpressed in oral tumor stem cells and has been found to induce tumor proliferation by activating MAPK signaling, but its specific functional mechanism in PC has not been investigated²². CXCL10, also known as IP-10, high expression in the TME of PC could influence lymphocytes recruitment and was correlate with poor survival²³. GABRP, significantly higher expression in PC tissue than normal tissues, could promote tumor growth and metastasis. Higher expression of GABRP was also associated with poor prognosis of PC. In mechanism, GABRP interacts with KCNN4 to induce calcium influx, activate the nuclear factor κ B signaling, and ultimately promote macrophage infiltration by inducing the expression of CXCL5 and CCL20²⁴. Overexpression of FAM83A activated TGF- β signaling pathway and the Wnt/ β -catenin signaling. In vitro and vivo, FAM83A expression was associated with the characteristics of PC tumor stem cell and the generation of chemotherapy resistance²⁵. The role of these novel genes in tumorigenesis and tumor development should be further investigated.

The predictive capabilities of the prognostic risk score systems were different by construction using different principles and angles. In the past decade, with the rapid development of bioinformatics, tumor molecular biology and tumor immunology, a large number of molecular markers have been discovered and were considered to have potential diagnostic and therapeutic value. The number of the prognostic models constructed by tumor molecular markers were gradually increasing, and most of them showed acceptable performance²⁶⁻³⁰. However, there was still no perfect prognostic model due to the heterogeneous tumors and unknown pathogenesis. Current research on the pathogenesis of malignant tumors, a large number of TME related molecular markers have been revealed had the relationship with tumor proliferation, immune escape, drug resistance, metastasis and so on. They also play as mediators for tumor - environment interaction. Many researchers have transferred their focuses from tumor cells to the surrounding environment of tumor cells. PC, as an inflammatory tumor, has changes in the TME during its development of inflammatory - cold process^{14,31,32}. In clinical research on TME and PC, phase Ib study of PEGylated human recombinant hyaluronidase (PEGPH20) in combination with gemcitabine in advanced PC patients has shown promising efficacy³³. Phase I study of CCR2 inhibition in combination with FOLFIRINOX in borderline resectable and locally advanced PC patients also has yielded a good clinical outcome³⁴. It is very important to study TME related prognostic genes of PC, which can further reveal the tumor development and progression. Establishment of a prognostic system based on TME related prognostic genes can more accurately and conveniently evaluate the prognosis of PC patients. Pu et al. analyzed the TME related prognostic genes through the similar way, and identified 53 genes related to immune score and 17 genes related to stromal score, which were related to recurrence free survival. Unfortunately, prognostic model had not further constructed³⁵. To our knowledge, this prognostic prediction system was the first model constructed by ESTIMATE for PC. According to this prognostic risk score system, Harrell's C-index was 0.73, and AUC value of 1-year, 2-year and 3-year OS period was 0.67, 0.76 and 0.86 in the internal validation set. Harrell's C-index was 0.71, and AUC value of 1-year, 2-year and 3-year OS period was 0.81, 0.72 and 0.78 in the external validation set. By reviewing other reported prognostic prediction systems of PC, this prognostic system contains fewer variables and has a good predictive capacity²⁶⁻²⁹. Furthermore, it was found this prognostic risk score system was significantly associated with the expression of some ICGs, suggesting that it may have predictive implications for immunotherapy response and efficacy.

Since the data were based on TCGA and ICGC, the limited sample size and the quality of data set were the biggest bottleneck of this study. In addition, the functions and mechanisms of the novel genes should be further explored by basic experiments. The effectiveness of this prognostic prediction system needs to be confirmed by multi-center and prospective research.

Conclusion

This prognostic risk score system based on TME demonstrated a good predictive capacity to the prognosis of PC.

Methods

1. Data acquisition

RNA sequencing and clinical data (age, gender, differentiation, T staging, N staging, survival data) of PC patients were downloaded from The Cancer Genome Atlas (TCGA). A total of 170 patients were included in this study after exclusion of the patients with overall survival (OS) < 30 days. Subsequently, Stromal and immune scores of these 170 PC patients were obtained from ESTIMATE (<https://bioinformatics.mdanderson.org/estimate/>)¹⁵. And then, RNA sequencing and survival data of PC patients were also downloaded from International Cancer Genome Consortium (ICGC) (<https://dcc.icgc.org/>). A total of 63 patients were constituted the external validation data set for the study after exclusion of the patients with OS < 30 days.

2. Identification of candidate genes

Kaplan-Meier method was used to assess OS of patients with different stromal and immune scores. X-Tile software was used to determine the best thresholds for grouping respectively³⁶. The limma package was used to analyze the differentially expressed genes (DEGs) between the groups with high and low stromal scores, and high and low immune scores³⁷. $|\log_2FC| \geq 1$ and adjusted $p < 0.05$ was considered had statistically significant difference. Co-DEGs were obtained by taking intersection of these two DEGs sets. Finally, survival related DEGs were screened by univariate cox analysis and identified as candidate genes.

3. Gene Ontology (GO) term enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis

DAVID online tool (<https://david.ncifcrf.gov/>), ClusterProfiler package were used to carry out GO and KEGG enrichment analysis by inputting candidate genes^{38,39}. GO analysis includes three parts: cellular component (CC), molecular function (MF) and biological process (BP). Visualize the analysis results through the GOplot package⁴⁰. The two terms with adjusted $p < 0.05$ were considered to be statistically significant.

4. Establishment and validation of prognostic risk score system

170 PC patients from TCGA were divided into two groups in a ratio of 3:2. The process of grouping was in accordance with the principle of randomization. One group with 102 patients was set as the training set, and another group with 68 patients was set as the internal validation set. LASSO analysis was performed by using glmnet package to narrow down candidate genes⁴¹. Stepwise method was adopted in multivariate Cox regression analysis to determine the final genes for establishment of the prognostic risk score system. The risk score calculation formula was as follows:

$$\text{Risk score} = \beta_1 * \text{Exp}\beta_1 + \beta_2 * \text{Exp}\beta_2 + \beta_3 * \text{Exp}\beta_3 + \beta_i * \text{Exp}\beta_i$$

The prognostic risk score system was validated in training set, internal validation set and external validation set. The performance of the risk score system was investigated by calculation of the Harrell's concordance index (C-index) ⁴². TimeROC package was used for construction of the time dependent receiver operating characteristic (ROC) and calculation of the area under the ROC curve (AUC) at 1-year, 2-year and 3-year OS ⁴³. Univariate and multivariate COX regression survival analysis were performed by adding the factor of risk score in the training set.

5. Relationship between risk score system and clinical characteristics, immune cell infiltration and immune checkpoint coding genes (ICGs)

Wilcoxon test was used to compare the difference of risk score between different characteristics of clinical parameters. 170 PC patients from TCGA were divided into high- and low-risk group based on the median prognostic score. The difference of 22 immune cell infiltration between the two groups was analyzed by CIBERSORT algorithm ⁴⁴. Furthermore, the expression of ICGs between high- and low-risk group were also investigated.

6. Statistical analysis

R software v3.6.3 (www.r-project.org) was used for statistical analyses in this study. The numerical variables were described as means $\pm$ SD. Kaplan-Meier method was used in survival analysis, and log-rank test was used to compare the differences between groups. Wilcoxon test was used for nonparametric comparisons. All statistical tests were two-tailed, $p < 0.05$ were considered to be significantly different.

Results

1. Basic information of the PC patients

170 PC patients with completed information were extracted from TCGA. The age was 64.45 ± 10.83 years old, the proportion of males was 54.1%, median OS was 15.35 months. After random grouping according to 3:2, the age of the training set was 63.08 ± 10.23 years old, the proportion of males was 54.9%, and median OS was 15.27 months. The age of the internal validation set was 66.51 ± 11.44 years, the proportion of males was 52.9% males, and median OS was 15.68 months. 63 PC patients from ICGC had a median OS of 14.10 months. The detail of each clinical parameter was shown in Table 1.

2. Grouping and survival analysis based on immune and stromal scores

-180.59 was calculated as the best cut-off value for grouping based on immune scores, and there was a significant difference in OS between high and low immune score group ($p=0.013$). In the same way, -542.33 was set as the best cut-off value for grouping based on stromal scores, and there was also a

significant difference in OS between the high and low stromal scores group ($p=0.038$). As shown in Figure 1A.

3. Identification of candidate genes

By analyzing the DEGs between high and low immune score group, a total of 486 DEGs were identified, including 423 up-regulated genes and 63 down-regulated genes. 531 up-regulated genes and 146 down-regulated genes were screened when high and low stromal scores group was compared. A total of 351 DEGs were obtained after taking intersection of the two DEGs sets, including 308 up-regulated genes and 43 down-regulated genes. Finally, 210 candidate genes were identified by univariate COX regression analysis. As shown in Figure 1B.

4. GO enrichment analysis, KEGG pathway enrichment analysis

When GO enrichment analysis was carried out on 210 candidate genes of PC, extracellular matrix organization, extracellular matrix and extracellular matrix structural constituent was the most significant functional enrichment of BP, CC and MF, respectively. Cytokine-cytokine receptor interaction was the most relevant pathway by KEGG enrichment analysis. See Figure 2 and Table 2.

5. Establishment of the prognostic risk score system

A total of 12 genes were screened from the 210 candidate genes by using LASSO analysis, and final 7 variables were determined for construction the prognostic risk score system by multivariate Cox regression analysis, as shown in Figure 3. The calculation formula was as follows:

$$\text{Risk score} = (0.331 * \text{Exp FAM57B}) + (-0.375 * \text{Exp HTRA3}) + (0.320 * \text{Exp CXCL10}) + (0.205 * \text{Exp GABRP}) + (0.139 * \text{Exp SPRR1B}) + (0.217 * \text{Exp FAM83A}) + (0.092 * \text{Exp LY6D})$$

6. Validation of the prognostic risk score system

Patients were divided into high- and low-risk group by setting the median as the cutoff value in each data set. It was found that the high- and low-risk group all had significant survival differences in training set, internal validation and external validation ($p=0.000$, $p=0.021$, $p=0.001$). In the training set, Harrell's C-index of this prognostic risk score system was 0.78, and the AUC was 0.87, 0.91, 0.87 at 1-year, 2-year and 3-year, respectively. In process of internal validation, the survival prediction C-index was 0.73, and AUC value of 1-year, 2-year and 3-year OS period was 0.67, 0.76 and 0.86, respectively. In the external validation set, the survival prediction C-index was 0.71, and the AUC was 0.81, 0.72, 0.78 at 1-year, 2-year and 3-year, respectively. As shown in Figure 4. Univariate and multivariate regression analyses were carried out by adding the factor of risk score in the training set. Only risk score was independent factors for OS ($p < 0.001$). As shown in Figure 5.

7. Relationship between risk score system and clinical characteristics, immune cell infiltration and immune checkpoint coding genes (ICGs)

PC Patients with the differentiation of G3/G4 had a higher prognostic risk score than patients with G1/G2 ($p=0.003$), and no significant difference was found among patients with different age, gender and TNM staging. The proportions of M0 and M1 macrophage were higher in high risk score patients, while the proportion of monocytes was less than the patients with low risk score ($p=0.008$, $p=0.001$, $p=0.008$). The prognostic risk score was significantly correlated with the expression of the following ICGs: BTLA, CD80, LGALS9, NT5C, TNFSF9 ($P=0.040$, $P=0.033$, $P=0.000$, $P=0.035$, $P=0.000$). As shown in Figures 6A, B, C.

Abbreviations

PC, Pancreatic cancer; TME, Tumor microenvironment; DEGs, Differentially expressed genes; AUC, area under the curve; C-index, Harrell's concordance index; TCGA, The Cancer Genome Atlas; ICGC, International Cancer Genome Consortium; FC, fold changes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; CC, cellular component (CC); MF, molecular function, BP, biological process; ROC, receiver operating characteristic; ICGs, immune checkpoint coding genes; PEGPH20, PEGylated human recombinant hyaluronidase.

Declarations

Acknowledgements

None.

Funding

The project was funded by the Project of Chengdu Medical Research (Grant No. 2021015).

Competing interests

The authors declare no potential conflicts of interest.

Ethics approval and consent to participate

The ethical approval did not refer to this study and this study did not need the informed consent.

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and analyzed during the current study are available in TCGA official website repository (<https://portal.gdc.cancer.gov/>) and International Cancer Genome Consortium (ICGC) (<https://dcc.icgc.org/>).

Code availability

Not applicable

Authors' contributions

Bo Liu: Conceptualization, Methodology. Ke Xu: Writing - Original draft preparation. Tingting Fu and Ping He: Software, Data curation, visualization and Investigation.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7-30.
3. Ferrone CR, Kattan MW, Tomlinson JS, Thayer SP, Brennan MF, Warshaw AL. Validation of a postresection pancreatic adenocarcinoma nomogram for disease-specific survival. *J Clin Oncol*. 2005;23(30):7529-7535.
4. Watanabe Y, Nishihara K, Matsumoto S, Okayama T, Abe Y, Nakano T. Effect of postoperative major complications on prognosis after pancreatectomy for pancreatic cancer: a retrospective review. *Surg Today*. 2017;47(5):555-567.
5. Adamu M, Nitschke P, Petrov P, et al. Validation of prognostic risk scores for patients undergoing resection for pancreatic cancer. *Pancreatology*. 2018.
6. Dougan SK. The Pancreatic Cancer Microenvironment. *Cancer J*. 2017;23(6):321-325.
7. Zhu Y, Herndon JM, Sojka DK, et al. Tissue-Resident Macrophages in Pancreatic Ductal Adenocarcinoma Originate from Embryonic Hematopoiesis and Promote Tumor Progression. *Immunity*. 2017;47(3):597.
8. Porembka MR, Mitchem JB, Belt BA, et al. Pancreatic adenocarcinoma induces bone marrow mobilization of myeloid-derived suppressor cells which promote primary tumor growth. *Cancer Immunol Immunother*. 2012;61(9):1373-1385.
9. Neesse A, Algul H, Tuveson DA, Gress TM. Stromal biology and therapy in pancreatic cancer: a changing paradigm. *Gut*. 2015;64(9):1476-1484.
10. Condeelis J, Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell*. 2006;124(2):263-266.
11. Coffelt SB, Hughes R, Lewis CE. Tumor-associated macrophages: effectors of angiogenesis and tumor progression. *Biochim Biophys Acta*. 2009;1796(1):11-18.

12. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. *Nat Rev Cancer*. 2009;9(4):239-252.
13. Ozdemir BC, Pentcheva-Hoang T, Carstens JL, et al. Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell*. 2014;25(6):719-734.
14. Liu Q, Liao Q, Zhao Y. Chemotherapy and tumor microenvironment of pancreatic cancer. *Cancer Cell Int*. 2017;17:68.
15. Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
16. Botsis T, Anagnostou VK, Hartvigsen G, Hripcsak G, Weng C. Modeling prognostic factors in resectable pancreatic adenocarcinomas. *Cancer Inform*. 2010;7:281-291.
17. Hartwig W, Hackert T, Hinz U, et al. Pancreatic cancer surgery in the new millennium: better prediction of outcome. *Ann Surg*. 2011;254(2):311-319.
18. Wenta T, Rychlowski M, Jurewicz E, et al. The HtrA3 protease promotes drug-induced death of lung cancer cells by cleavage of the X-linked inhibitor of apoptosis protein (XIAP). *Febs J*. 2019;286(22):4579-4596.
19. Zhao J, Feng M, Liu D, et al. Antagonism between HTRA3 and TGFbeta1 Contributes to Metastasis in Non-Small Cell Lung Cancer. *Cancer Res*. 2019;79(11):2853-2864.
20. Lv Q, Yang B, Ning C, et al. Hypoxia is involved in the reduction of HtrA3 in patients with endometrial hyperplasia and cancer. *Biochem Biophys Res Commun*. 2018;503(4):2918-2923.
21. Zhao J, Zhang J, Zhang X, Feng M, Qu J. High temperature requirement A3 (HTRA3) expression predicts postoperative recurrence and survival in patients with non-small-cell lung cancer. *Oncotarget*. 2016;7(26):40725-40734.
22. Michifuri Y, Hirohashi Y, Torigoe T, et al. Small proline-rich protein-1B is overexpressed in human oral squamous cell cancer stem-like cells and is related to their growth through activation of MAP kinase signal. *Biochem Biophys Res Commun*. 2013;439(1):96-102.
23. Lunardi S, Jamieson NB, Lim SY, et al. IP-10/CXCL10 induction in human pancreatic cancer stroma influences lymphocytes recruitment and correlates with poor survival. *Oncotarget*. 2014;5(22):11064-11080.
24. Jiang SH, Zhu LL, Zhang M, et al. GABRP regulates chemokine signalling, macrophage recruitment and tumour progression in pancreatic cancer through tuning KCNN4-mediated Ca(2+) signalling in a GABA-independent manner. *Gut*. 2019;68(11):1994-2006.
25. Chen S, Huang J, Liu Z, Liang Q, Zhang N, Jin Y. FAM83A is amplified and promotes cancer stem cell-like traits and chemoresistance in pancreatic cancer. *Oncogenesis*. 2017;6(3):e300.
26. Wu M, Li X, Zhang T, Liu Z, Zhao Y. Identification of a Nine-Gene Signature and Establishment of a Prognostic Nomogram Predicting Overall Survival of Pancreatic Cancer. *Front Oncol*. 2019;9:996.

27. Yan X, Wan H, Hao X, et al. Importance of gene expression signatures in pancreatic cancer prognosis and the establishment of a prediction model. *Cancer Manag Res.* 2019;11:273-283.
28. Song W, Miao DL, Chen L. Nomogram for predicting survival in patients with pancreatic cancer. *Onco Targets Ther.* 2018;11:539-545.
29. Huang L, Balavarca Y, van der Geest L, et al. Development and validation of a prognostic model to predict the prognosis of patients who underwent chemotherapy and resection of pancreatic adenocarcinoma: a large international population-based cohort study. *BMC Med.* 2019;17(1):66.
30. Liu Y, Zhu D, Xing H, Hou Y, Sun Y. A 6gene risk score system constructed for predicting the clinical prognosis of pancreatic adenocarcinoma patients. *Oncol Rep.* 2019;41(3):1521-1530.
31. Clark CE, Hingorani SR, Mick R, Combs C, Tuveson DA, Vonderheide RH. Dynamics of the immune reaction to pancreatic cancer from inception to invasion. *Cancer Res.* 2007;67(19):9518-9527.
32. Koorstra JB, Feldmann G, Habbe N, Maitra A. Morphogenesis of pancreatic cancer: role of pancreatic intraepithelial neoplasia (PanINs). *Langenbecks Arch Surg.* 2008;393(4):561-570.
33. Hingorani SR, Harris WP, Beck JT, et al. Phase Ib Study of PEGylated Recombinant Human Hyaluronidase and Gemcitabine in Patients with Advanced Pancreatic Cancer. *Clin Cancer Res.* 2016;22(12):2848-2854.
34. Nywening TM, Wang-Gillam A, Sanford DE, et al. Targeting tumour-associated macrophages with CCR2 inhibition in combination with FOLFIRINOX in patients with borderline resectable and locally advanced pancreatic cancer: a single-centre, open-label, dose-finding, non-randomised, phase 1b trial. *Lancet Oncol.* 2016;17(5):651-662.
35. Pu N, Chen Q, Gao S, et al. Genetic landscape of prognostic value in pancreatic ductal adenocarcinoma microenvironment. *Ann Transl Med.* 2019;7(22):645.
36. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res.* 2004;10(21):7252-7259.
37. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
38. Dennis G, Jr., Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 2003;4(5):P3.
39. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* 2012;16(5):284-287.
40. Walter W, Sanchez-Cabo F, Ricote M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics.* 2015;31(17):2912-2914.
41. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010;33(1):1-22.
42. Long J, Zhang L, Wan X, et al. A four-gene-based prognostic model predicts overall survival in patients with hepatocellular carcinoma. *J Cell Mol Med.* 2018;22(12):5928-5938.

43. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. 2013;32(30):5381-5397.
44. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-457.

Tables

Table 1

The clinical characteristics of PC patients

Clinical characteristic	TCGA training set		TCGA internal validation set		ICGC external validation set	
	Total (n=102)	Percent (%)	Total (n=68)	Percent (%)	Total (n=63)	Percent (%)
Age(mean±SD)						
	63.08±10.23		66.51±11.44		65.89±11.24	
Gender						
Male	56	54.9	36	52.9	31	49.2
Female	46	45.1	32	47.1	32	50.8
Grade						
G1+G2	74	72.5	45	66.2	40	63.5
G3+G4	28	27.5	21	30.9	22	34.9
T stage						
T1+T2	13	12.7	14	20.6	4	6.3
T3+T4	89	87.3	52	76.5	57	90.5
N status						
N0	24	23.5	23	33.8	16	25.4
N+	73	71.6	41	60.3	45	71.4
NA	5	4.9	4	5.9	2	3.2
TNM stage						
I+II	95	93.1	65	95.6	0	0
III+IV	7	6.9	1	1.5	0	0
NA	0	0	2	2.9	63	100
Survival status						
Alive	48	47.1	32	47.1	26	41.3
Dead	54	52.9	36	52.9	37	58.7

NA not available.

Table 2.

GO and KEGG enrichment analysis of candidate TME related genes in PC ranked by p value. (TOP 5)

Category	GO or KEGG ID	GO or KEGG term	p.adjust	Count
BP	GO:0030198	extracellular matrix organization	5.60E-21	35
BP	GO:0043062	extracellular structure organization	3.92E-19	35
BP	GO:0030199	collagen fibril organization	3.72E-11	12
BP	GO:0043588	skin development	1.60E-08	24
BP	GO:0035987	endodermal cell differentiation	2.27E-08	10
CC	GO:0031012	extracellular matrix	1.49E-24	42
CC	GO:0062023	collagen-containing extracellular matrix	4.45E-20	35
CC	GO:0044420	extracellular matrix component	4.03E-12	12
CC	GO:0005581	collagen trimer	1.87E-11	14
CC	GO:0005583	fibrillar collagen trimer	1.52E-10	7
MF	GO:0005201	extracellular matrix structural constituent	9.34E-20	25
MF	GO:0030020	extracellular matrix structural constituent conferring tensile strength	5.99E-11	11
MF	GO:0005539	glycosaminoglycan binding	5.27E-09	17
MF	GO:0008201	heparin binding	5.47E-08	14
MF	GO:0005125	cytokine activity	7.40E-08	14
KEGG	hsa04512	ECM-receptor interaction	8.21E-06	10
KEGG	hsa04657	IL-17 signaling pathway	8.21E-06	10
KEGG	hsa04974	Protein digestion and absorption	8.21E-06	10
KEGG	hsa04061	Viral protein interaction with cytokine and cytokine	9.50E-	9

		receptor	05	
KEGG	hsa05146	Amoebiasis	9.50E-05	9

Figures

A

B

Figure 1

Survival analysis based on immune and stromal scores and identification of co-DEGs. A. Kaplan-Meier survival curves of high and low immune score group, and high and low stromal scores group. B. The venn diagram of significantly up-regulated and down-regulated genes.

Figure 3

Risk score distribution, survival status, and expression patterns of 7 included genes in risk score system for high- and low-risk patients from training set, internal validation set and external validation set.

TCGA training set

TCGA internal validation set

ICGC external validation set

Figure 4

Kaplan-Meier survival curves of high- and low-risk patients and the time dependent ROC at 1-year, 2-year, and 3-year OS from verification set, internal validation set and external validation set.

Figure 5

Univariate and multivariate Cox regression analysis of OS.

Figure 6

Relationship between risk score system and clinical characteristics (A), immune cell infiltration (B) and the expression of ICGs (C).