

# COL5A3 is a Prognostic Biomarker and Correlated with Immune Infiltrates in Pancreatic Cancer

Yongjie Li (✉ [11753599@qq.com](mailto:11753599@qq.com))

School of Pharmacy, Shaoyang University

Min Zeng

Department of Respiratory Medicine, Shaoyang Central Hospital

Ting Wang

The Second Affiliated Hospital, Hengyang Medical School, University of South China

Feng Jiang

Department of Nutrition, Taizhou Central Hospital

---

## Research Article

**Keywords:** Pancreatic cancer, COL5A3, tumor, immune infiltration, ClusterProfiler

**Posted Date:** October 14th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-889396/v2>

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

[Read Full License](#)

---

# Abstract

Pancreatic cancer is a malignant tumor of digestive system with high fatality rate, and its prognosis is very poor. Type I collagen  $\alpha 3$  (COL5A3) is highly expressed in a variety of tumor tissues, but its prognostic value and immune infiltration in pancreatic cancer are still unclear. Therefore, we evaluated the prognostic role of COL5A3 in pancreatic cancer and its correlation with immune infiltration. The transcriptional expression profiles of COL5A3 in pancreatic cancer and normal tissues were downloaded from the Cancer Genome Atlas (TCGA). In the GEO (Gene expression omnibus) dataset (GSE16515), we compared the expression of COL5A3 in normal and tumor tissues. The expression of COL5A3 protein was evaluated by the human protein atlas (THPA). The effect of COL5A3 on survival was evaluated by Kaplan-Meier method. Receiver operating characteristic (ROC) curve was used to distinguish pancreatic cancer from paracancerous normal tissues. Protein-protein interaction (PPI) network was constructed by the STRING. Use the "ClusterProfiler" package for functional enrichment analysis. Tumor immune estimation resource (TIMER) and tumor-immune system interaction database (TISIDB) were used to determine the relationship between COL5A3 mRNA expression and immune infiltration. Compared with normal tissues, COL5A3 is highly expressed in pancreatic cancer tissues. The high expression of COL5A3 is related to the poor clinicopathological features and poor prognosis of pancreatic cancer. Kaplan-Meier survival analysis showed that the prognosis of pancreatic cancer patients with high expression of COL5A3 was worse than that of patients with low expression of COL5A3. ROC curve shows that COL5A3 has high sensitivity and specificity in the diagnosis of pancreatic cancer. Correlation analysis showed that the expression of COL5A3 mRNA was related to immune cell infiltration. This study reveals for the first time that COL5A3 may be a new prognostic biomarker related to immune infiltration and provide a new target for the diagnosis and treatment of pancreatic cancer.

## Background

Pancreatic cancer is a highly malignant tumor disease of the digestive system, and its incidence is on the rise in recent years. In the United States, the number of deaths from pancreatic cancer is increasing. Because of its poor prognosis, its 5-year survival rate is about 10%, which is increasingly becoming a common cause of cancer death<sup>1</sup>. According to the 2018 global cancer morbidity and mortality estimates compiled by the International Agency for Research on Cancer, there are 458918 new cases of pancreatic cancer and 432242 deaths worldwide, accounting for 2.5 per cent of newly diagnosed cancers and 4.5 per cent of all cancer deaths that year<sup>2</sup>. Due to the clinical characteristics of patients with pancreatic cancer, such as hidden onset, difficult early diagnosis, rapid progress, low resection rate, easy recurrence and metastasis after operation, clinical diagnosis and treatment is very challenging. With the introduction of new surgical techniques and medical methods, such as laparoscopy and neoadjuvant radiotherapy and chemotherapy, the treatment of pancreatic cancer is gradually developing. However, the improvement of therapeutic effect on patients with pancreatic cancer is very limited<sup>3</sup>. In recent years, immunotherapy has achieved significant efficacy in the treatment of a variety of malignant tumors<sup>4-7</sup>, in which the treatment of pancreatic cancer is beginning to dawn, but immunotherapy is not suitable for all cancer

patients<sup>8</sup>. In addition, targeted anticancer therapy for patients with pancreatic cancer has also made significant progress, which is expected to become the future trend of precision oncology<sup>9</sup>. In view of the limitations of the current treatment of pancreatic cancer and the ineffectiveness of treatment results, and little is known about the potential pathogenesis of pancreatic cancer, the detection of biomarkers of early pancreatic cancer is very important in the study of pancreatic cancer.

The collagen gene family is the main component of the extracellular matrix, accounting for 30% of the total protein mass of mammals, while type V collagen alpha3 (COL5A3) is a member of the collagen family and plays a key role in regulating fiber formation in the extracellular matrix<sup>10</sup>. At present, the related research on COL5A3 is rarely reported in the literature. It has been reported that the expression of collagen gene COL5A3 plays an important role in bone formation<sup>11</sup>. Some studies have shown that COL5A3 is involved in the occurrence and development of a variety of tumors, including breast cancer<sup>12</sup>, uveal melanoma<sup>13</sup>, prostate cancer<sup>14</sup>, renal cell carcinoma<sup>15</sup> and so on. Although it has been reported that COL5A3 is highly expressed in many tumors by high-throughput sequencing, the correlation between the expression of COL5A3 in pancreatic cancer and its prognostic value has not been reported.

Therefore, in this study, we used several online databases, such as TCGA (The cancer genome atlas) and GEO (Gene expression omnibus), and used their clinical indicators and survival data to evaluate the significance of COL5A3 expression in patients with pancreatic cancer. In addition, we also studied the relationship between COL5A3 mRNA levels and tumor infiltrating immune cells. In summary, our study closely linked the overexpression of COL5A3 to the low survival rate of pancreatic cancer.

## Materials And Methods

### TCGA datasets

TCGA (The Cancer Genome Atlas) is a landmark cancer genomics project, including 33 cancer types, a total of more than 20,000 samples collected. The transcriptional expression data and clinical information of COL5A3 we studied were downloaded from the official website of TCGA. In addition, we also studied the expression level of COL5A3 gene in cancer and paracancerous tissues combined with the integrated data of TCGA and GTEx database. The RNAseq data in TPM (transcripts per million reads) format were analyzed and compared after log<sub>2</sub> transformation. Finally, in the GEO dataset (GSE16515), we compared the expression of COL5A3 in normal and tumor tissues.

### RNA-Sequencing data of COL5A3 in pancreatic cancer

Download the RNA-Seq expression data of COL5A3 in pancreatic cancer from TCGA website. Because there are few normal tissue samples of pancreatic cancer in TCGA database, we integrated the normal tissue samples of pancreatic cancer in GTEx database, and finally retained the data of 171 cases of adjacent normal tissue and 179 cases of tumor tissue. The selected clinical samples contain COL5A3

gene expression data and related clinical information, such as patient sex, age, smoking and drinking history, tumor TNM stage and initial treatment outcome.

### **The human protein atlas (THPA)**

THPA (<https://www.proteinatlas.org/>) is the human protein map, which aims to provide information on the tissue and cell distribution of all 24000 human proteins. In this study, we used THPA to observe the expression of COL5A3 in normal tissues and pancreatic cancer by immunohistochemical (IHC) images.

### **Analysis of prognostic indexes**

In order to explore the clinical value of COL5A3 gene in the prognosis of patients with pancreatic cancer, we analyzed the survival indexes such as OS (overall survival) and DSS (disease specific survival). The RNA sequencing data and corresponding clinical information of pancreatic cancer patients were downloaded from TCGA database and visualized by Kaplan-Meier curve. The expression of COL5A3 was divided into two groups: low expression group and high expression group, and the P value was obtained by Log-rank test and Cox regression analysis.

### **Protein-Protein Interaction (PPI) networks and functional enrichment analysis**

STRING (<https://www.string-db.org/>) is a commonly used database for searching known protein-protein interactions and predicting protein-protein interactions. By studying protein-protein interaction networks, it is helpful to mine core regulatory genes. In this study, we obtained the top ten COL5A3 related genes through STRING database and constructed PPI network map. We use GO and KEGG enrichment analysis, mainly ggplot2 package (version 3.3.3) for visualization, and clusterProfiler package (version 3.14.3) to analyze the selected data.

### **Tumor immune estimation resource database (TIMER)**

TIMER (<https://cistrome.shinyapps.io/timer/>) uses RNA-Seq expression profile data to detect the infiltration of immune cells in tumor tissues. At present, it mainly provides the infiltration of six kinds of immune cells. In this study, we used TIMER database to determine the expression of COL5A3 and tumor purity and the correlation between six kinds of immune infiltrating cells (B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells).

### **Tumor immune system interaction database (TISIDB)**

TISIDB (<http://cis.hku.hk/TISIDB>) is a website for studying the interaction between tumors and the immune system. It covers a number of tumor-related databases. In this study, we determined the expression of COL5A3 and 28 kinds of tumor infiltrating lymphocytes (TILs) in different human cancers, and used TISIDB to make the relationship between the expression of COL5A3 and a variety of TILs abundance.

### **Statistical analysis**

The statistical analysis and visualization of all our data are carried out in R (version 3.6.3), and the main R package involved: ggplot2 package (version 3.3.3) is used for visual analysis. Mann-Whitney U test was used to determine the difference between pancreatic cancer tissue and adjacent normal tissue. In order to evaluate the effect of COL5A3 expression on survival, we used Kaplan-Meier and log-rank tests for statistical analysis of survival data, mainly survival package (version 3.2-10), and survminer package (version 0.4.9) for visualization. The ROC curve was analyzed by pROC package (version 1.17.0.1).

## Results

### Expression pattern of COL5A3 in pan-cancer perspective

To assess the mRNA expression of COL5A3 in different cancer types, we analyzed datasets from 33 cancer types. As shown in Figure 1A, compared with normal tissues, COL5A3 was significantly upregulated in 28 of all 33 cancer types. This data indicated the mRNA expression of COL5A3 was differentially expressed in multiple tumor types.

### COL5A3 mRNA and protein expression was upregulated in patients with pancreatic cancer

In order to determine the expression level of COL5A3 mRNA and protein in pancreatic cancer, we analyzed the expression data of COL5A3 from TCGA data and HPA immunohistochemistry, and further analyzed and verified the expression of COL5A3 using GEO database. As shown in figure 1B, the expression level of COL5A3 in pancreatic cancer is significantly higher than that in normal tissues ( $P < 0.001$ ). Consistent with figure 1C, we can see that the expression of COL5A3 is up-regulated in the GSE16515 data set compared with normal tissues ( $P < 0.05$ ). Finally, as shown in figure 1D, the immunohistochemical staining of HPA also showed that the expression of COL5A3 protein was up-regulated in pancreatic cancer tissues. The above results showed that the mRNA and protein expression levels of COL5A3 in pancreatic cancer tissues were up-regulated.

### The relationship between COL5A3 expression and clinical pathological characteristics in pancreatic cancer patients

We studied the clinicopathological features of pancreatic cancer patients with differential expression of COL5A3. As shown in Table 1, compared with the group with low expression of COL5A3, the group with high expression of COL5A3 had a significantly worse initial therapeutic effect ( $P = 0.004$ ), and the group with high expression of COL5A3 had a more significant history of alcohol ( $P = 0.036$ ). However, there was no statistical difference in other clinicopathological features, such as age, sex, TNM stage, pathological stage and so on. To sum up, the high expression of COL5A3 is closely related to some poor clinicopathological features, and also provides a new idea for the study of single gene molecules related to tumor drug resistance.

**Table 1.** Clinical characteristics of the pancreatic cancer patients (TCGA).

Characteristic	Low expression of COL5A3	High expression of COL5A3	p
n	89	89	
T stage, n (%)			0.250
T1	6 (3.4%)	1 (0.6%)	
T2	12 (6.8%)	12 (6.8%)	
T3	68 (38.6%)	74 (42%)	
T4	2 (1.1%)	1 (0.6%)	
N stage, n (%)			0.754
N0	26 (15%)	24 (13.9%)	
N1	59 (34.1%)	64 (37%)	
M stage, n (%)			0.390
M0	31 (36.9%)	48 (57.1%)	
M1	3 (3.6%)	2 (2.4%)	
Pathologic stage, n (%)			0.758
Stage I	12 (6.9%)	9 (5.1%)	
Stage II	71 (40.6%)	75 (42.9%)	
Stage III	2 (1.1%)	1 (0.6%)	
Stage IV	3 (1.7%)	2 (1.1%)	
Radiation therapy, n (%)			0.058
No	66 (40.5%)	52 (31.9%)	
Yes	17 (10.4%)	28 (17.2%)	
Primary therapy outcome, n (%)			0.004**
PD	16 (11.5%)	33 (23.7%)	
SD	8 (5.8%)	1 (0.7%)	
PR	6 (4.3%)	4 (2.9%)	
CR	40 (28.8%)	31 (22.3%)	
Gender, n (%)			1.000
Female	40 (22.5%)	40 (22.5%)	

Characteristic	Low expression of COL5A3	High expression of COL5A3	p
Male	49 (27.5%)	49 (27.5%)	
Race, n (%)			0.741
Asian	6 (3.4%)	5 (2.9%)	
Black or African American	4 (2.3%)	2 (1.1%)	
White	77 (44.3%)	80 (46%)	
Age, n (%)			0.368
<=65	50 (28.1%)	43 (24.2%)	
>65	39 (21.9%)	46 (25.8%)	
Residual tumor, n (%)			0.720
R0	55 (33.5%)	52 (31.7%)	
R1	24 (14.6%)	28 (17.1%)	
R2	2 (1.2%)	3 (1.8%)	
Histologic grade, n (%)			0.124
G1	21 (11.9%)	10 (5.7%)	
G2	43 (24.4%)	52 (29.5%)	
G3	22 (12.5%)	26 (14.8%)	
G4	1 (0.6%)	1 (0.6%)	
Anatomic neoplasm subdivision, n (%)			0.369
Head of Pancreas	66 (37.1%)	72 (40.4%)	
Other	23 (12.9%)	17 (9.6%)	
Smoker, n (%)			0.604
No	35 (24.3%)	30 (20.8%)	
Yes	38 (26.4%)	41 (28.5%)	
Alcohol history, n (%)			0.036*
No	25 (15.1%)	40 (24.1%)	
Yes	57 (34.3%)	44 (26.5%)	
History of diabetes, n (%)			0.774

Characteristic	Low expression of COL5A3	High expression of COL5A3	p
No	52 (35.6%)	56 (38.4%)	
Yes	20 (13.7%)	18 (12.3%)	
History of chronic pancreatitis, n (%)			0.507
No	61 (43.3%)	67 (47.5%)	
Yes	8 (5.7%)	5 (3.5%)	
Family history of cancer, n (%)			0.271
No	26 (23.6%)	21 (19.1%)	
Yes	27 (24.5%)	36 (32.7%)	
Age, mean $\pm$ SD	63.6 $\pm$ 11.37	65.9 $\pm$ 10.13	0.155

\*p<0.05, \*\*p<0.01

### Predictive value of COL5A3 for pancreatic cancer diagnosis and prognosis

In order to evaluate the effect of COL5A3 on the clinical prognosis of pancreatic cancer, we used K-M curve analysis to verify the prediction of clinical outcome of COL5A3. As shown in figure 2A-2B, the survival rate of OS in the high expression group was significantly lower than that in the low expression group of COL5A3 ( $p = 0.019$ ), consistent with that in the DSS group, the survival rate in the high expression group was also significantly lower than that in the low expression group ( $p = 0.026$ ). Next, we studied the clinical benefits of COL5A3, and we used ROC curves to demonstrate its value in the differential diagnosis of pancreatic cancer. As shown in figure 2C, the area under the curve (AUC) is 0.843, suggesting that COL5A3 has high sensitivity and specificity in the diagnosis of pancreatic cancer. The above results show that the up-regulation of COL5A3 indicates a worse prognosis and has high diagnostic value for clinical diagnosis.

### COL5A3-associated PPI network and functional enrichment

In order to construct COL5A3-associated PPI networks and functional annotations, we used STRING database, GO and KEGG analysis. As shown in figure 3A, the PPI network shows that some genes are closely related to COL5A3, such as ADAMTS14, ADAMTS2, ADAMTS3, BMP1, COL11A1, COL1A2, LUM, P4HA3, PCOLCE and PCOLCE2. As shown in Figure 3B, COL5A3-related genes were involved in many biological processes (BP), cellular compositions (CC), and molecular functions (MF). The results of GO and KEGG enrichment analysis of COL5A3-related genes were mainly involved in extracellular structure organization, extracellular matrix organization, collagen fibril organization, protein digestion and absorption. The GO and KEGG analysis demonstrated that COL5A3 might promote tumor proliferation,

migration and invasion by affecting extracellular matrix. Analysis of the correlation between COL5A3 expression and co-expression genes in pancreatic cancer patients based on TCGA database, as shown in figure 4A-J.

### **Analysis of the relationship between the expression of COL5A3 and the level of immune cell infiltration in pancreatic carcinoma**

We used TIMER database to analyze the expression of COL5A3 and tumor purity and the correlation between six kinds of infiltrating immune cells (B cells, CD8 T cells, CD4 T cells, macrophages, neutrophils and dendritic cells). As shown in figure 5A, the results displayed that the expression level of COL5A3 had obviously positive correlation with infiltrating levels of CD4+T cells ( $r=0.213$ ,  $P=5.40e-03$ ), macrophage cells ( $r=0.414$ ,  $P=1.82e-08$ ), neutrophils ( $r=0.399$ ,  $P=6.52e-08$ ), dendritic cells ( $r=0.375$ ,  $P=4.45e-07$ ) in pancreatic cancer, but no association with tumor purity and CD8+ T cells.  $P < 0.05$  was defined as statistically significant. Figure 5B shows the relationship between COL5A3 and the expression of 28 kinds of tumor infiltrating lymphocytes in human cancer. As shown in figures 5C-H, the expression of COL5A3 is related to the abundance of Treg ( $r=0.37$ ,  $P=4.43e-07$ ), Tgd ( $r=0.36$ ,  $P=8.8e-07$ ), Tcm CD4 ( $r=0.35$ ,  $P=1.83e-06$ ), NK ( $r=0.348$ ,  $P=2.15e-06$ ), pDC ( $r=0.339$ ,  $P=4.03e-06$ ) and Tcm CD8 ( $r=0.257$ ,  $P=0.000526$ ). These results suggest that the expression of COL5A3 may play an important role in the immune infiltration of pancreatic cancer.

## **Discussion**

Although with the continuous development of medicine and the continuous emergence of new technologies and new treatments, the therapeutic effect on patients with pancreatic cancer is very limited. Pancreatic cancer is still a malignant tumor with a very high fatality rate<sup>16</sup>. Therefore, the discovery of new biomarkers of pancreatic cancer plays an important role in its early diagnosis, treatment and prognosis.

As there are few studies on COL5A3 in cancer, we use bioinformatics analysis to find out its biological function in pancreatic cancer. In this study, we found that COL5A3 is highly expressed in a variety of tumor tissues, including pancreatic cancer. High expression of COL5A3 is associated with poor clinicopathological features, such as initial treatment efficacy and alcohol history. ROC curve analysis shows that COL5A3 has high predictive efficiency and can be used as a promising biomarker for early diagnosis of pancreatic cancer. According to the Kaplan-Meier curve analysis, we confirmed that the high expression of COL5A3 was associated with poor OS and DSS. Therefore, COL5A3 can be used as a potential biomarker for poor prognosis of pancreatic cancer.

In order to further study the potential biological function of COL5A3, we carried out enrichment analysis through GO function and KEGG pathway. COL5A3-related genes are mainly related to extracellular structure tissue, extracellular matrix tissue, collagen fiber tissue, protein digestion and absorption. Some

studies have shown that type V collagen, including COL5A3, is highly expressed in the extracellular matrix of pancreatic cancer cells and promotes the proliferation, migration and metastasis of pancreatic cancer by binding to  $\alpha$ -2- $\beta$ -1 integrin receptor<sup>17</sup>. Some studies have also shown that the up-regulation of COL5A3 in pancreatic cancer is a significant feature of fibrosis and malignant tumor stroma<sup>18</sup>. These results suggest that COL5A3 may play a role in affecting tumor microenvironment and regulating cell proliferation, migration and metastasis.

Tumor microenvironment (TME) contains a complex cellular environment, including lymphatic endothelial cells, vascular endothelial cells, mesenchymal cells and immune cells, as well as extracellular matrix and inflammatory matrix<sup>19,20</sup>. The dynamic and symbiotic relationship between tumor cells and their microenvironment forms in the early stage of malignant tumor growth, and affects the occurrence and development of tumor<sup>21,22</sup>. Different types of tumor patients have obvious clinical effects after immunotherapy, but if all patients are given the same immunotherapy, the effect is often not ideal<sup>23,24</sup>. Because of the complexity of tumor immune environment, it is very important to predict and guide immunotherapy by analyzing the special immune microenvironment in the tumor. However, the relationship between the expression of COL5A3 and immune cell infiltration in pancreatic cancer has not been reported. Our results show that the expression of COL5A3 in TME is positively correlated with several tumor infiltrating immune cells (Treg, Tgd, Tcm CD4, NK, pDC and Tcm CD8) through Timer and TISIDB database. There is growing evidence that innate immune cells (macrophages, neutrophils, dendritic cells, innate lymphoid cells, myeloid suppressor cells and natural killer cells) and acquired immune cells (T cells and B cells) can promote tumorigenesis and development in tumor microenvironment (TME). The interaction between cancer cells and proximal immune cells eventually leads to an environment that promotes tumor growth and metastasis<sup>25-27</sup>. These studies suggest that there is a potential correlation between COL5A3 and immune infiltration of pancreatic cancer.

In this study, there are still the following shortcomings: we are based on the mining of public database samples for statistical analysis, lack of our own clinical samples for further verification. In addition, the detailed mechanism of the effect of COL5A3 on pancreatic cancer can be further studied by designing in vitro and in vivo experiments.

## Conclusions

In summary, this study found for the first time that the expression of COL5A3 was up-regulated in pancreatic cancer and was related to initial therapeutic efficacy and alcohol history. In addition, our study also found that COL5A3 has a high predictive effect on the diagnosis of pancreatic cancer and is related to the level of immune infiltration of pancreatic cancer. Therefore, COL5A3 may be used as a potential new prognostic biomarker to provide new ideas for further experiments and clinical trials.

## Declarations

### Ethics approval

All the data in this article are collected in accordance with the guidelines approved by the Medical Ethics Review Committee of Shaoyang University and are in line with the current laws of China.

### **Consent for publication.**

All authors have reviewed the manuscript and consented for publication.

### **Data availability**

The raw data of this study are derived from the TCGA database (<https://portal.gdc.cancer.gov/>) and GEO data portal (<https://www.ncbi.nlm.nih.gov/geo/>), which are publicly available databases.

### **Acknowledgements**

We acknowledge the TCGA, GEO, GTEx, THPA, TIMER, TISIDB and STRING databases for free use.

### **Author contributions**

Y.L. and M.Z. contributed to the research conception, design and visualization. T.W. and F.J. carried out data analysis and graphic generation. Y.L. wrote a draft and T.W. revised the manuscript.

### **Funding**

None.

### **Competing interests**

The authors declare that they have no competing interests.

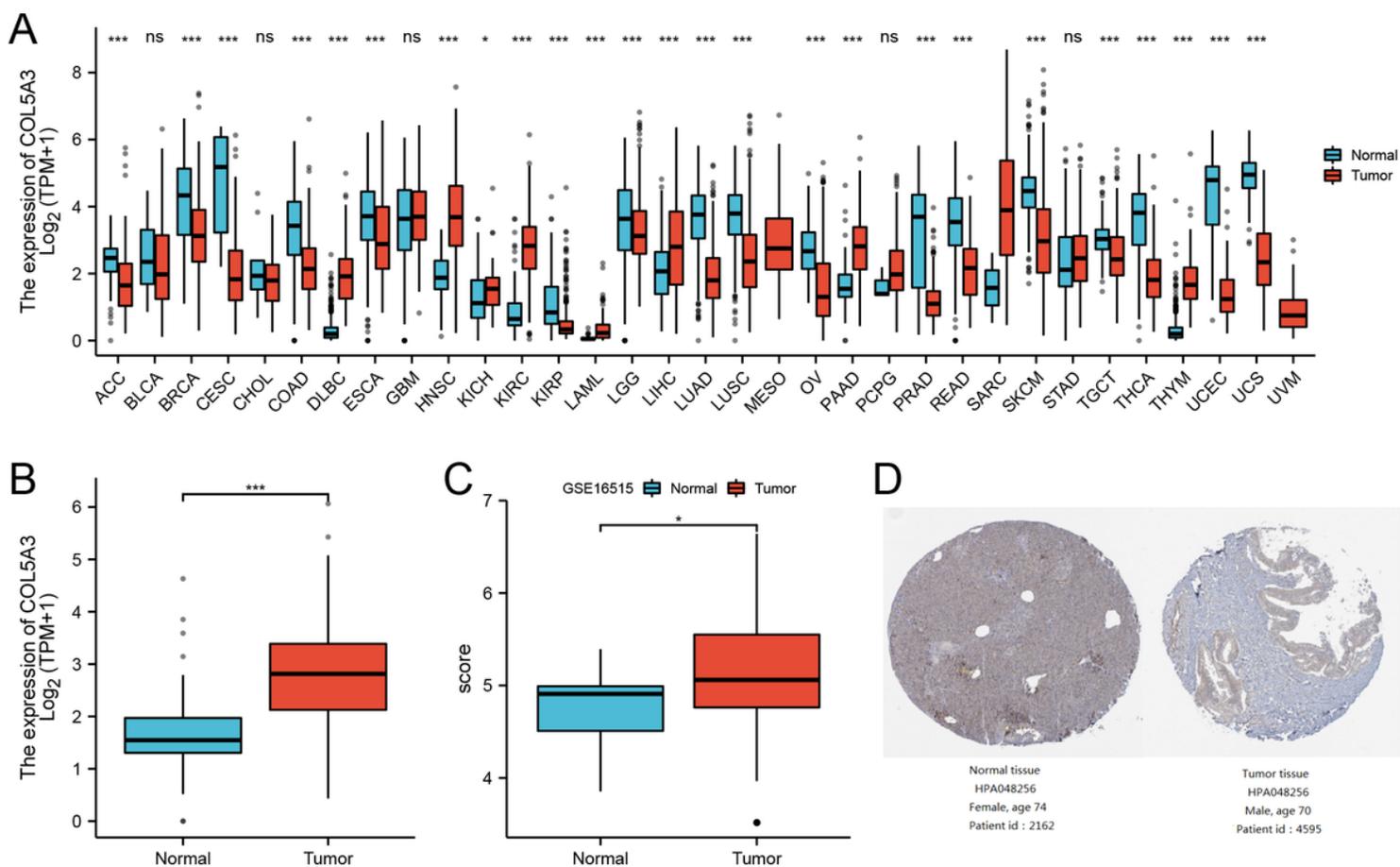
## **References**

1. Mizrahi, J. D., Surana, R., Valle, J. W., & Shroff, R. T. Pancreatic cancer. *Lancet (London, England)*, **395**(10242), 2008–2020 (2020).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, **68**(6), 394–424 (2018).
3. McGuigan, A. *et al.* Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. *World journal of gastroenterology*, **24**(43), 4846–4861 (2018).
4. Brahmer, J. R. *et al.* Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *The New England journal of medicine*, **366**(26), 2455–2465 (2012).
5. Cohen, E. *et al.* The Society for Immunotherapy of Cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC). *Journal for immunotherapy of cancer*, **7**(1), 184 (2019).

6. Rosenberg, A. R. *et al.* Skin cancer precursor immunotherapy for squamous cell carcinoma prevention. *JCI insight*, **4**(6), e125476 (2019).
7. Terranova-Barberio, M. *et al.* Exhausted T cell signature predicts immunotherapy response in ER-positive breast cancer. *Nature communications*, **11**(1), 3584 (2020).
8. Feng, M. *et al.* PD-1/PD-L1 and immunotherapy for pancreatic cancer. *Cancer letters*, **407**, 57–65 (2017).
9. Qian, Y. *et al.* Molecular alterations and targeted therapy in pancreatic ductal adenocarcinoma. *Journal of hematology & oncology*, **13**(1), 130 (2020).
10. Haq, F., Ahmed, N., & Qasim, M. Comparative genomic analysis of collagen gene diversity. *3 Biotech*, **9**(3), 83 (2019).
11. Yun-Feng, W., Matsuo, N., Sumiyoshi, H., & Yoshioka, H. Sp7/Osterix up-regulates the mouse pro-alpha3(V) collagen gene (Col5a3) during the osteoblast differentiation. *Biochemical and biophysical research communications*, **394**(3), 503–508 (2010).
12. Huang, W. *et al.* Bisphenol S induced epigenetic and transcriptional changes in human breast cancer cell line MCF-7. *Environmental pollution (Barking, Essex : 1987)*, **246**, 697–703 (2019).
13. Li, W. *et al.* Bioinformatic Analysis of Differentially Expressed Genes and Screening of Hub Genes in Uveal Melanoma Cells with BRCA1-Associated Protein 1 Related Protein 1 Depletion. *Journal of biomedical nanotechnology*, **16**(8), 1205–1218 (2020).
14. Ruiz-Deya, G. *et al.* Differential DNA Methylation in Prostate Tumors from Puerto Rican Men. *International journal of molecular sciences*, **22**(2), 733 (2021).
15. Chen, Y. *et al.* Identification of biological targets of therapeutic intervention for clear cell renal cell carcinoma based on bioinformatics approach. *Cancer cell international*, **16**, 16 (2016).
16. Neoptolemos, J. P. *et al.* Therapeutic developments in pancreatic cancer: current and future perspectives. *Nature reviews. Gastroenterology & hepatology*, **15**(6), 333–348 (2018).
17. Weniger, M., Honselmann, K. C., & Liss, A. S. The Extracellular Matrix and Pancreatic Cancer: A Complex Relationship. *Cancers*, **10**(9), 316 (2018).
18. Dey, S. *et al.* Global targetome analysis reveals critical role of miR-29a in pancreatic stellate cell mediated regulation of PDAC tumor microenvironment. *BMC cancer*, **20**(1), 651 (2020).
19. Arneth B. Tumor Microenvironment. *Medicina (Kaunas, Lithuania)*, **56**(1), 15 (2019).
20. Jarosz-Biej, M., Smolarczyk, R., Cichoń, T., & Kułach, N. Tumor Microenvironment as A "Game Changer" in Cancer Radiotherapy. *International journal of molecular sciences*, **20**(13), 3212 (2019).
21. Deepak, K. *et al.* Tumor microenvironment: Challenges and opportunities in targeting metastasis of triple negative breast cancer. *Pharmacological research*, **153**, 104683 (2020).
22. Hegde, P. S., & Chen, D. S. Top 10 Challenges in Cancer Immunotherapy. *Immunity*, **52**(1), 17–35 (2020).
23. DeBerardinis R. J. Tumor Microenvironment, Metabolism, and Immunotherapy. *The New England journal of medicine*, **382**(9), 869–871 (2020).

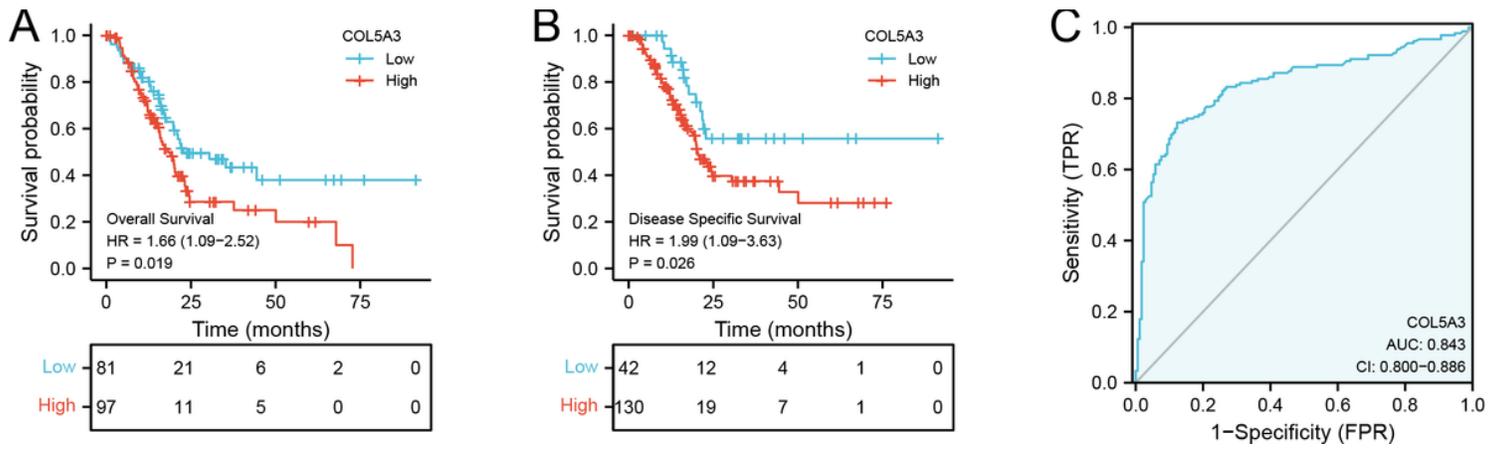
24. Frankel, T., Lanfranca, M. P., & Zou, W. The Role of Tumor Microenvironment in Cancer Immunotherapy. *Advances in experimental medicine and biology*, **1036**, 51–64 (2017).
25. Wang, J. J., Lei, K. F., & Han, F. Tumor microenvironment: recent advances in various cancer treatments. *European review for medical and pharmacological sciences*, **22**(12), 3855–3864 (2018).
26. Hinshaw, D. C., & Shevde, L. A. The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer research*, **79**(18), 4557–4566 (2019).
27. Wu, T., & Dai, Y. Tumor microenvironment and therapeutic response. *Cancer letters*, **387**, 61–68 (2017).

## Figures



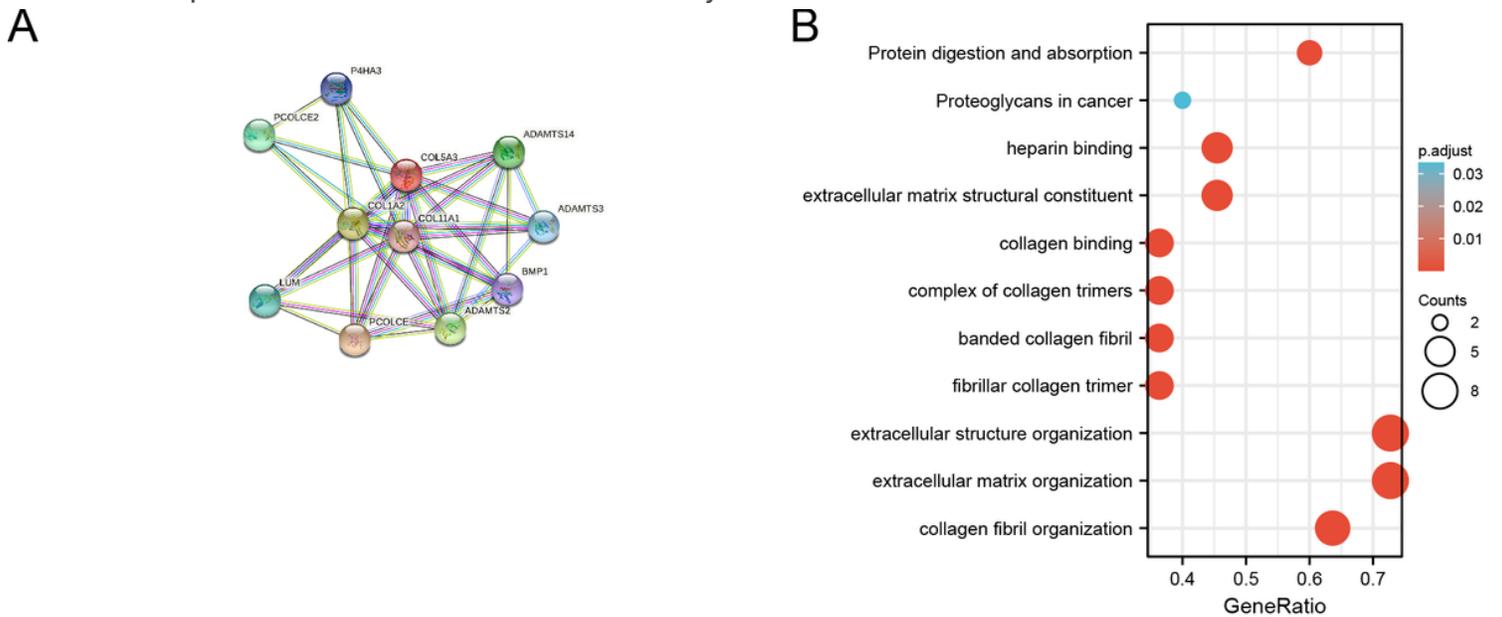
**Figure 1**

(A) Expression pattern of COL5A3 in Pan-cancer perspective. (B) The mRNA expression levels of COL5A3 in 179 pancreatic cancer samples and 171 normal samples. (C) The aberrant expression of COL5A3 based on the GEO database. COL5A3 mRNA levels in pancreatic cancer tumor tissues and normal tissues in the GSE16515 dataset. (D) The protein levels of COL5A3 based on Human Protein Atlas. (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ )



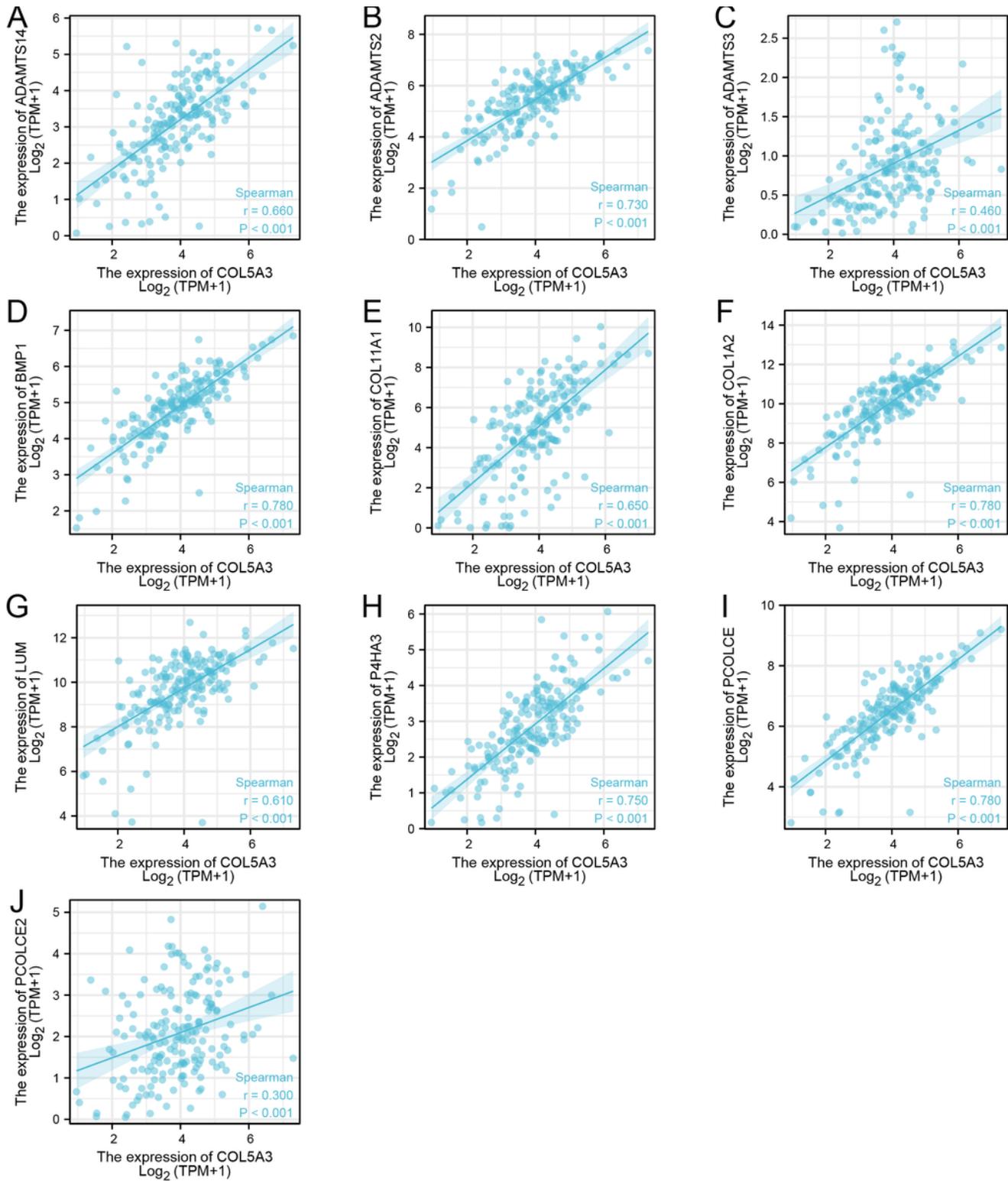
**Figure 2**

Kaplan-Meier and ROC curves for COL5A3. (A,B) Kaplan-Meier survival curves indicated that pancreatic cancer patients with high COL5A3 mRNA expression had a shorter OS ( $P = 0.019$ ) and DSS ( $P = 0.026$ ) than those with low-level of COL5A3. (C) ROC curve showed that COL5A3 had an AUC value of 0.843 to discriminate pancreatic cancer tissues from healthy controls.



**Figure 3**

PPI networks and functional enrichment analysis. (A) A network of COL5A3 and its co-expression genes. (B) Functional enrichment analysis of 11 involved genes.



**Figure 4**

The correlation analysis between the expression of COL5A3 and co-expressed genes in pancreatic cancer. (A–J)

