Novel targets identified by integrated cancer-stromal interactome analysis of pancreatic adenocarcinoma

Yukihiko Hiroshima, Rika Kasajima, Yayoi Kimura, Daisuke Komura, Shumpei Ishikawa, Yasushi Ichikawa, Michael Bouvet, Naoto Yamamoto, Takashi Oshima, Soichiro Morinaga, Shree Ram Singh, Robert M. Hoffman, Itaru Endo, Yohei Miyagi

PII: S0304-3835(19)30534-8

DOI: https://doi.org/10.1016/j.canlet.2019.10.031

Reference: CAN 114538

To appear in: *Cancer Letters*

Received Date: 7 July 2019

Revised Date: 3 October 2019

Accepted Date: 18 October 2019

Please cite this article as: Y. Hiroshima, R. Kasajima, Y. Kimura, D. Komura, S. Ishikawa, Y. Ichikawa, M. Bouvet, N. Yamamoto, T. Oshima, S. Morinaga, S.R. Singh, R.M. Hoffman, I. Endo, Y. Miyagi, Novel targets identified by integrated cancer-stromal interactome analysis of pancreatic adenocarcinoma, *Cancer Letters*, https://doi.org/10.1016/j.canlet.2019.10.031.

Abstract

The pancreatic cancer microenvironment is crucial in cancer development, progression and drug resistance. Cancer-stromal interactions have been recognized as important targets for cancer therapy. However, identifying relevant and druggable cancer-stromal interactions is challenging due to the lack of quantitative methods to analyze the whole cancer-stromal interactome. Here we studied 14 resected pancreatic cancer specimens (8 pancreatic adenocarcinoma (PDAC) patients as a cancer group and 6 intraductal papillary-mucinous adenoma (IPMA) patients as a control). Shotgun proteomics of the stromal lesion dissected with laser captured microdissection was performed, and identified 102 differentially expressed proteins in pancreatic cancer stroma. Next, we obtained gene expression data in human pancreatic cancer and normal pancreatic tissue from The Cancer Genome Atlas database (n=169) and The Genotype-Tissue Expression database (n=197), and identified 1435 genes, which were differentially expressed in pancreatic cancer cells. To identify relevant and druggable cancer-stromal-interaction targets, we applied these datasets to our in-house ligand-receptor database. Finally, we identified 9 key genes and 8 key cancer-stromal-interaction targets for PDAC patients. Furthermore, we examined FN1 and ITGA3 protein expression in pancreatic cancer tissues using the TMAs of 271 PDAC cases, and demonstrated that FN1-ITGA3 had unfavorable prognostic impact for PDAC patients.

Novel targets identified by integrated cancer-stromal interactome analysis of pancreatic adenocarcinoma

Yukihiko Hiroshima[1,*], Rika Kasajima[1], Yayoi Kimura[2], Daisuke Komura[3], Shumpei Ishikawa[3], Yasushi Ichikawa[4], Michael Bouvet[5], Naoto Yamamoto[6], Takashi Oshima[6], Soichiro Morinaga[6], Shree Ram Singh[7,*], Robert M. Hoffman[5,8,*], Itaru Endo[9], and Yohei Miyagi[1]

[1] Kanagawa Cancer Center Research Institute, Yokohama, Japan

[2] Advanced Medical Research Center, Yokohama City University, Yokohama, Japan

[3] Department of Prevention Medicine, Tokyo University, Tokyo, Japan

[4] Department of Clinical Oncology, Yokohama City University, Yokohama, Japan

[5]Department of Surgery, University of California, San Diego, CA, USA

[6]Department of Gastrointestinal Surgery, Kanagawa Cancer Center, Yokohama, Japan

[7]Basic Research Laboratory, National Cancer Institute, Frederick, MD, USA

[8]AntiCancer, Inc., San Diego, CA, USA

[9]Department of Gastroenterological Surgery, Yokohama City University, Yokohama, Japan

[*]Corresponding authors:

Yukihiko Hiroshima, M.D., Ph.D., Kanagawa Cancer Center Research Institute, Yokohama, Japan, E-mail: yhiroshiy@gancen.asahi.yokohama.jp

Robert M. Hoffman, Ph. D., AntiCancer, Inc., San Diego, CA, USA, email: all@anticancer.com

Shree Ram Singh, Ph. D., Basic Research Laboratory, National Cancer Institute, Frederick, MD, USA, email: singhshr@mail.nih.gov

Running title: Cancer-stromal interactome analysis of pancreatic cancer

Abstract

The pancreatic cancer microenvironment is crucial in cancer development, progression and drug resistance. Cancer-stromal interactions have been recognized as important targets for cancer therapy. However, identifying relevant and druggable cancer-stromal interactions is challenging due to the lack of quantitative methods to analyze the whole cancer-stromal interactome. Here we studied 14 resected pancreatic cancer specimens (8 pancreatic adenocarcinoma (PDAC) patients as a cancer group and 6 intraductal papillary-mucinous adenoma (IPMA) patients as a control). Shotgun proteomics of the stromal lesion dissected with laser captured microdissection was performed, and identified 102 differentially expressed proteins in pancreatic cancer stroma. Next, we obtained gene expression data in human pancreatic cancer and normal pancreatic tissue from The Cancer Genome Atlas database (n=169) and The Genotype-Tissue Expression database (n=197), and identified 1435 genes, which were differentially expressed in pancreatic cancer cells. To identify relevant and druggable cancer-stromal-interaction targets, we applied these datasets to our in-house ligand-receptor database. Finally, we identified 9 key genes and 8 key cancer-stromal-interaction targets for PDAC patients. Furthermore, we examined FN1 and ITGA3 protein expression in pancreatic cancer tissues using the TMAs of 271 PDAC cases, and demonstrated that FN1-ITGA3 had unfavorable prognostic impact for PDAC patients.

Keywords: Pancreatic cancer; cancer-stromal interaction; proteomics; bioinformatics

1. Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the most malignant solid tumors arising within the ducts of the pancreas. The lack of early diagnosis and its rapid progression results in advanced stages of PDAC patients when diagnosed. In the past several decades, there have been many studies of the molecular pathogenesis of PDAC, however, limited advances have been made to prolong the survival and to reduce the mortality [1, 2].

Pancreatic cancer has expensive desmoplasia, a process in which fibrous tissue infiltrates

and envelops the tumor [2, 3]. A remarkable increase in interstitial connective tissue (collagen type I and fibronectin) was observed (the mean collagen content in pancreatic cancer tissue and tumor-associated chronic pancreatitis was 3-fold higher than in normal pancreas). There were no differences in the proportion of collagen types I, III, and V among chronic pancreatitis, tumor-associated chronic pancreatitis, and pancreatic cancer tissue [4, 5]. The tumor microenvironment, including desmoplasia, plays a vital role in cancer development and progression, and cancer-stromal interactions are important targets for cancer therapy [2, 6, 7]. However, identifying relevant and druggable cancer-stromal interactions is challenging due to the lack of quantitative methods to analyze the entire cancer-stromal interactome [8].

Although several computational methods have been developed to analyze cancer-stromal interactions using microarrays or RNA-seq data from human cancer tissues [9] or cancer xenograft mouse models [10, 11], they could not evaluate individual interactions and prioritize cancer-stromal interactions as targets for cancer treatment because the expression profiles of cancer cells and stromal cells were independent. To overcome such limitations, we have previously developed CASTIN (CAncer-STromal INteractome analysis) for quantitative profiling of cancer-stromal interactome from RNA-seq data using human cancer xenograft mouse models [8]. CASTIN determines direction and strength of individual transmitting signals between two interacting cells based on the expression levels of cancer and stroma. However, CASTIN cannot fully evaluate secreted proteins and genes whose sequence is very similar between human and mouse.

In the present study, we performed a shotgun proteomic analysis using dissected stromal areas of PDAC tissues to identify differentially expressed (DE) proteins in PDAC stroma. Differentially expressed genes (DEGs) in PDAC cells were identified using the public

databases. Subsequently the DEG lists of PDAC cells and stroma were integrated to our in-house ligand-receptor interaction database to identify crucial and potentially druggable cancer-stromal interactions in PDAC.

2. Materials and methods

2.1. Patient selection for shotgun proteomic analysis and immunohistochemistry

This study population consisted of 14 patients with pancreatic tumor (8 PDAC and 6 intraductal papillary-mucinous adenoma (IPMA) patients, Table1) who underwent primary pancreatic lesion resection at the Department of Gastroenterological Surgery, Yokohama City University for shotgun proteomic analysis, and 271 PDAC patients who underwent primary pancreatic lesion resection at the Department of Gastrointestinal Surgery, Kanagawa Cancer Center for immunohistochemistry. All patients were thoroughly informed about the study and provided their written consent, in accordance with the ethical guidelines of the Institutional Review Board (IRB), Yokohama City University, and Kanagawa Cancer Center, Japan (notice of approval of IRB protocol number A180125007 and 2019EKI-38) and the Declaration of Helsinki.

2.2. Preparation of frozen tissue sections

Tissue blocks obtained from pancreatic tumor patients were embedded in Tissue-Tek O.C.T. compound (Sakura Finetek Japan Co. Ltd., Tokyo, Japan) and immediately frozen at −80 °C in dry ice. The blocks were cut with a Leica cryostat into 10 μm sections, in an environment controlled to -20 °C. The inner chamber and the stage of the cryostat were

wiped with 100 % ethanol and the blade was changed after each sample in order to avoid cross contamination. Three sections were transferred onto each of 10 Leica polyethylene naphthalate (PEN) membrane slides (Thermo Fisher Scientific, Waltham, MA, USA) designed for laser-captured microdissection (LCM). After the placement of three sections on PEN membrane slides, one extra reference section was prepared on a regular glass slide for hematoxylin and eosin (H & E) staining. Sections were maintained frozen and stored at -80 °C.

2.3. Staining and dehydration for LCM

Hematoxylin staining was used to guide the LCM in ethanol dehydrated sections. One PEN membrane slide at a time was stained and dehydrated, in preparation for each LCM session. The staining and dehydration protocol was performed with 5 clean Coplin jars, prefilled with 50 mL ethanol [70 % (jar # 1-3), 95 % (jar #3), and 100 % (jar #4)]. The protocol was performed by dipping the slide for 30 seconds in each jar of the ethanol series, following the numerical order, to obtain dehydration. After jar 2, the slide was drained, placed horizontally, washed for 10 seconds with 200 uL phosphate buffered saline (PBS) followed by hematoxylin staining for 120 seconds, then washed with PBS and transferred to jar 3, to continue dehydration.

2.4. LCM of stromal tissue

LCM samples were collected with a Leica LMD system. Each LCM session lasted a total of 60 minutes, to avoid tissue rehydration and degradation. Pancreatic cancer, adenoma

and normal ducts were identified in the tissue by visualizing in bright field and in phase-contrast with a 4x magnification (Fig. 1A, the area surrounded by yellow line), and pancreatic acinar cells and islets were identified in the tissue by visualizing in bright field, too (Fig. 1A, the area surrounded by blue line). Scans of the H&E-stained reference sections were used to map the pancreatic stromal area (Fig. 1B). Samples of pancreatic cancer, adenoma, normal duct, acinar cells and islets were micro-dissected from pancreatic tissue sections, and the remaining tissue was defined as the pancreatic stroma. Each sample was collected from 6-10 pancreatic sections. The LCM tissues were collected into 500 μL sterile RNase/DNase/Protease-free Eppendorf tubes and were resuspended with 50 μL of 50 mM NH_4HCO_3. The tubes were closed, frozen in dry ice, and subsequently maintained at -80 °C.

2.5. Protein extraction

The LCM samples were washed three times with PBS buffer containing a protease inhibitor cocktail (Roche, Switzerland), and homogenized in 50 mM NH_4HCO_3, 4 M urea, 2 M thiourea and 4 % (w/v) sodium deoxycholate with a Protease Inhibitor Mix (GE Healthcare, Piscataway, NJ, USA) and Phosphatase Inhibitor Cocktail (Sigma, Madison, WI, USA), using a Sample Grinding Kit (GE Healthcare). The homogenate was then sonicated 8 times (30 s intervals) using a UCD250 (Cosmo BIO, Tokyo, Japan).

The protein extractions were reduced with 10mM dithiothreitol at room temperature for 30 min and alkylated using 10 mM iodoacetamide at room temperature for a further 30 min. Samples (30 μg) were digested 3 hr at 37 ˚C with 0.3 μg of lysylendopeptidase (Wako, Tokyo, Japan), and subsequently 18 hr at 37 ˚C with 3 μg of TrypsinGold (Promega, Madison, WI,

USA) after diluted 4-fold with 500 mM tetraethylammonium bromide. Sodium deoxycholate, including the peptide solution, was removed using the phase-transfer surfactant method [12], then peptides were desalted using C18 Stage Tips [13].

2.6. Shotgun liquid-chromatography tandem-mass spectrometry

LC–MS/MS analysis was performed on a LTQ Orbitrap Velos (Thermo Fisher Scientific) using Xcalibur version 2.0.7 coupled to an UltiMate 3000 LC system (Dionex, LC Packings, Sunnyvale, CA, USA).

Prior to injection into the mass spectrometer, digested samples (2 μg each) were loaded online in a reverse-phase precolumn (C18 Pepmap column, LC Packings) and resolved on a nanoscale C18 Pepmap capillary column (75 μm i.d. × 15 cm) (LC Packings) with a gradient of acetonitrile / 0.1 % (v/v) formic acid at a flow rate of 300 nL/min. Peptides were separated using a 145 min gradient of 5 – 100 % solvent B (0.1 % [v/v] formic acid / 95 % [v/v] acetonitrile); solvent A was 0.1 % (v/v) formic acid / 2 % (v/v) acetonitrile. Precursor ions were subject to dynamic exclusion for 180 s using a 5 ppm window and resolution of 60 000. Full-scan mass spectra were measured from 350 – 1200 m/z on an LTQ Orbitrap Velos mass spectrometer operated in data-dependent mode using the TOP10 strategy. MS/MS scanning conditions were as follows: normalized collision energy, 35.0 %; isolation width, 2 m/z; activation time, 10 ms; activation Q, 0.25. The general mass spectrometric conditions were as follows: spray voltage, 2.1 kV; capillary temperature, 250 °C.

2.7. Label-free protein relative quantitation analysis

Label-free relative protein quantitation analysis was performed using Progenesis QI for proteomics, version 2.0 (Nonlinear Dynamics, Durham, NC, USA). Database searches to identify proteins were performed using MASCOT (version 2.5.1; Matrix Science, London, UK) against *Homo sapiens* protein sequences in the UniProt / SWISS-PROT database. The search parameters were as follows: (1) a peptide mass tolerance of ± 5 ppm; (2) a fragment mass tolerance of ± 0.5 Da; (3) + 2, + 3, or + 4 charges; (4) variable modifications (protein N-terminus acetylation/carbamylation, methionine oxidation, and cysteine carbamidomethylation); (5) a false discovery rate of 1 %; and (6) a peptide ion score of > 30. Proteins identified as being upregulated in pancreatic cancer were extracted using the following parameters in the Progenesis QI for proteomics software: (1) > 2 peptide counts used for further statistical analyses.

2.8. Gene expression data

Gene expression datasets of pancreatic ductal adenocarcinoma (PDAC, accession: TCGA, n = 169) and normal pancreatic tissue (accession: SRP012682, n = 197) were obtained from the Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov) and the Genotype-Tissue Expression (GTEx; https://gtexportal.org/home/), respectively, referring to the ReCount2 database (https://jhubiostatistics.shinyapps.io/recount/). Gene expression dataset SRP038143 containing whole transcriptome sequencing for three common pancreatic cell lines (MIA PaCa2, PANC1, HPAC) was also obtained from the ReCount2 database.

2.9. Data processing and screening of differentially expressed (DE) proteins and genes

The Empirical Bayes method was used to identify significant DE proteins (or genes) between PDAC and IPMA (or normal pancreatic) samples basing on the edgeR package in R version 3.5.2 (https ://www.r-proje ct.org/). P values were adjusted for multiple testing depending on the Benjamini-Hochberg False Discovery Rate (FDR) method. The strict thresholds for identifying DE proteins were set as FDR < 0.05, and those for DE genes were set as FDR < 0.05 and |fold change (FC)| > 16.

2.10. Statistical analysis

Unsupervised hierarchical clustering, MA plot and heatmap generation were performed using the TCC package and ggplot2 package in R version 3.5.2. Principal component analysis (PCA) was also performed using in R version 3.5.2.

2.11. Pathway and process enrichment analysis of DE proteins and genes

The following ontology sources: Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway, Gene Ontology Biological Processes (GO-BP), Reactome Gene Sets, Canonical Pathways and CORUM were applied for the functional annotation and pathway enrichment analysis of DEGs through using Metascape (http://metascape.org/gp/index.html#/main/). All statistically enriched terms were identified, and accumulative hypergeometric p-values and enrichment factors were calculated and used for filtering. Remaining significant terms were then hierarchically clustered into a tree based on Kappa-statistical similarities among their gene memberships. Then, 0.3 kappa score was applied as the threshold to cast the tree into term clusters. The terms within each cluster were shown in supplementary Table S4 and S6

("Count": the number of genes in the user-provided lists with membership in the given ontology term, "%": the percentage of all of the user-provided genes that are found in the given ontology term, "Log10(P)": the p-value in log base 10, "Log10(q)": the multi-test adjusted p-value in log base 10). Next, a subset of representative terms from this cluster were selected and converted into a network layout (Fig. 2C and 3D). Each term is represented by a circle node, where its size is proportional to the number of input gene that fall into that term, and its color represents its cluster identity. Terms with a similarity score > 0.3 are linked by an edge (the thickness of the edge represents the similarity score). The network is visualized with Cytoscape (v3.1.2) [14] with “force-directed” layout and with edge bundled for clarity. One term from each cluster is selected to have its term description shown as label.

2.12. In-house ligand-receptor database construction

We have constructed an in-house ligand-receptor database [8]. The database construction consisted of three main steps (i) extraction of localization information from Human Protein Reference Database (HPRD) [15] (ii) extraction of ligand-receptor interaction from KEGG data [16] (iii) curation by reviewing original literature. First, proteins localized primarily to extracellular spaces and the plasma membrane were selected as ligand and receptor candidates, respectively. Information of primary localization was downloaded from the Human Protein Reference Database (HPRD, release 8) [15] on 9 September 2009. Among all the pairs of ligand and receptor candidates, only those that appeared in the protein-protein interaction in KEGG pathway database [16] (release 55.0, downloaded on 7 August 2010) proceeded to the next curation step. The direction of interaction was determined according to

relations (activation, inhibition, binding/association, or indirect effect) in KEGG database. Finally, researchers in the field of biology curated each interaction by carefully reviewing the original literature attached in the KEGG database.

2.13. Kaplan-Meier survival analysis of key genes in PDAC

To evaluate the prognostic value of candidate genes, the patient samples were split into two cohorts according to the median expression value of candidate genes. Clinical information was obtained from TCGA, and overall survival (OS) was plotted via Kaplan-Meier method using the survival package in R version 3.5.2. Log-rank test was used to evaluate significance.

2.14. Immunohistochemistry

Tissue microarrays (TMAs) were constructed by sampling two 1.5-mm-diameter cores from the same tumor area. Four-micron sections were cut from each TMA block and manual immunohistochemistry was carried out according to standard protocols using anti-FN1 rabbit polyclonal antibody (1:100, ab2413, abcam, Cambridge, UK), anti-ITGA3 rabbit polyclonal antibody (1:200, ab131055, abcam) and anti-ITGA5 rabbit monoclonal antibody (1:100, ab150361, abcam). A streptavidin biotin detection system (Dako REALTM Kit, Dako, Denmark) and 3,3-Diaminobenzidine (DAB) were used according to the manufacturer's

instructions. Sections were counterstained with Mayer's hematoxylin, and coverslipped. Sections from a multi-tissue block served as positive and negative control.

2.15. Scoring procedure

Slides were scanned at ×20 magnification with an automated scanning system (Aperio CS2, Leica Biosystems, Wetzlar, Germany). The intensity of FN1 staining in stroma and ITGA3 staining in cancer cells were scored as follows: grade 0, not stained; grade 1, faintly stained; grade 2, weakly stained; and grade 3, strongly stained. The immunohistochemical evaluation was independently confirmed by two observers (Y.H. and Y.M.), and we defined grade 0 and 1 as negative, whereas grade 2 and 3 as positive for statistical analyses.

2.16. Statistical analysis

Unsupervised hierarchical clustering, MA plot and heatmap generation were performed using the TCC package and the ggplot2 package in R version 3.5.2. Principal component analysis (PCA) was also performed using in R version 3.5.2.

FN1/ITGA3 staining and clinicopathological variables were performed using the Mann-Whitney U test or Fisher's exact test as appropriate. OS time was defined as the time

from date of surgery to date of death or date of last follow-up. The relationship between OS and variable of interest was evaluated by uni- and multivariate analyses. OS curves were calculated using the Kaplan-Meier method and compared by the log-rank test. Cox's proportional hazard model was used to perform uni- and multivariate survival analyses. A P-value of less than 0.05 was defined as statistically significant. The RcmdrPlugin.EZR package in R version 3.5.2. was used for all statistical analyses.

3. Results

3.1. DE proteins in pancreatic stroma of PDAC patients

We measured peptides by shotgun LC–MS/MS, followed by relative quantitative analysis using Progenesis QI for proteomics software to determine the difference in protein composition between PDAC and IPMA groups. Consequently, a total of 5237 peptides derived from 1017 proteins were detected and identified in these samples. The complete lists of identified peptides and proteins are shown in Supplementary Table S1and S2 (1% FDR, peptide ion score of >30). Dendrogram of protein expression levels in PDAC and IPMA stromal tissues showed PDAC and IPMA samples were clearly grouped into 2 clusters with no PDAC sample found gathered in the IPMA sample cluster (Fig. 2A). DE proteins (FDR < 0.05) of the PDAC stroma were screened out based on R analysis. Relative to the IPMA stroma, the total DE protein number for PDAC stroma was 138 (Fig. 2B, Supplementary Table S3). The DE proteins were converted into 95 gene symbols (Fig. 1C).

3.2. Pathway and process enrichment analysis of DE proteins in PDAC stroma

A total of 14 overrepresent GO-BP terms related to “Extracellular matrix organization”, “Response to wounding” and “Actin cytoskeleton organization”, and only 2 signaling pathways including “Pancreatic secretion” and “Malaria”, were significantly enriched among the DE proteins (Supplementary Table S4). To further capture the relationships between the terms, a subset of enriched terms has been selected and rendered as a network plot, where terms with a similarity > 0.3 are connected by edges. The top 20 clusters with their representative enriched terms were selected according to p-value (Fig. 2C).

3.3. Identification of DEGs in PDAC cells using public databases

To explore the differences of gene expression between PDAC tissue and normal pancreatic tissue, the gene expression datasets of PDAC and normal pancreatic tissue were obtained from the public databases (the TCGA and the GTEx) as described before. The results of principal component analysis could differentiate the PDAC tissues from normal tissues directly (Fig. 3A). The first constituted principal component explained 47.4% of the variance of the variables, the second principal component explained 5.1% of the variance, and the cumulative variance that explained is 52.5%. DEGs (FDR < 0.05, $|\log2\ FC| \geq 16$) were screened out based on R analysis. A total of 4131 DEGs were identified in PDAC tissues compared to normal tissues. After that, 1434 genes, including 1240 up-regulated genes and 194 down-regulated ones, were identified as the DEGs in PDAC cells by filtering

out the genes which were not detected in the gene expression dataset of three common pancreatic cancer cell lines (SRP038143) (Fig. 1C, 3B).

3.4. Hierarchical clustering diagram and pathway and process enrichment analysis of DEGs after filtration

Hierarchical clustering diagram of the 1434 DEGs between normal pancreatic tissue (n = 197) and PDAC (n = 169) are shown in Fig. 3C. These 1434 DEGs could clearly distinguish PDAC from normal pancreatic tissues. For these 1434 DEGs, pathway and process enrichment analysis has been carried out. Top 20 clusters with their representative enriched terms, including 12 overrepresented GO-BP terms related to “Cell division”, “Regulation of mitotic cell cycle”, “Extracellular matrix organization”, “Response to wounding” and “Actin cytoskeleton organization”, and only 1 signaling pathway of “Pathways in cancer”, are shown in Supplementary Table S6 and Fig. 3D. Combined with the enriched GO-BP terms and pathways of DE proteins in PDAC stroma, only 3 GO-BP terms including “Extracellular matrix organization”, “Response to wounding” and “Actin cytoskeleton organization” were overlapped between the analyses. No overlapped pathway was detected between them. These results suggest that some genes included in these 3 GO-BP terms may be key genes of cell-stroma interaction in pancreatic cancer.

3.5. Integrated interactome analysis using multiple platforms

The CASTIN [8] evaluates and summarizes gene expression profiles of cancer and stroma from RNA-seq data using cancer xenograft models. In this study, we identified 138 DE proteins in PDAC stroma using dissected stromal areas of PDAC tissues, whereas 1434 DEGs in PDAC cells were identified using the public databases filtering out the genes which

were not detected in the gene expression dataset of three common pancreatic cell lines. To integrate these data across multiple platforms, 138 DE proteins were converted to 95 DE genes (DEGs). After that, the DEG lists of PDAC stroma and PDAC cells were integrated to our in-house ligand-receptor interaction database [8]. The ligand-receptor interactions were searched for the two directions of signal transduction, from cancer ligand to stromal receptor and from stromal ligand to cancer receptor (hereafter referred to as C-S direction and S-C direction, respectively). As shown in Table 2, 8 key cancer-stromal interactions were identified. All interactions were S-C direction, where extra-cellular matrix (ECM) related genes including tenascin (TNC), thrombospondin 1 (THBS1) and fibronectin (FN1) were identified as ligand and various subunits of integrin (ITG) were identified as receptor. No interaction of C-S direction was identified.

3.6. Kaplan–Meier survival analysis of key interactions and signature genes

To analyze the prognostic relevance of 8 key interactions, Kaplan-Meier analysis was performed with the survival package in R, stratified by median expression value of interaction associated genes. As shown in Fig. 4, FN1 ($p = 0.041$) and integrin subunit alfa 3 (ITGA3, $p = 0.012$) were negatively associated with OS. Moreover, Kaplan-Meier analysis, stratified by co-expression value of the genes associated with the 8 key interactions, suggested that FN1-ITGA3 interaction and FN1- integrin subunit alfa 5 (ITGA5) interaction were unfavorable factors of prognosis of PDAC patients (Fig. 5).

3.7. Immunohistochemistry

Figure 6 shows representative images of pancreatic cancer tissues stained for FN1, ITGA3 and ITGA5. We detected increased FN1 protein expression in stromal cells (especially in fibroblasts) and ITGA3 protein expression in cancer cells, whereas no increased ITGA5 protein expression was detected in cancer cells but in stromal cells (Fig. 6). This result suggests that ITGA5 may not play an important role as a receptor of FN1 in pancreatic cancer cells. Next, we examined FN1 and ITGA3 protein expression in pancreatic cancer tissues using the TMAs of 271 pancreatic cancer patients. We detected increased FN1 protein expression in stromal cells of 178 cases (65.7%), and ITGA3 in cancer cells of 90 cases (33.2%) out of 271 pancreatic cancer cases. Table 3 shows the association of FN1/ITGA3 co-expression with other clinicopathological parameters. FN1/ITGA3 co-expression was significantly correlated with only recurrence rate ($p = 0.04$) but not with other parameters such as tumor size, histological type, lymphatic invasion, venous invasion, intrapancreatic neural invasion, stage or curability. Lymph node metastasis was slightly correlated with FN1/ITGA3 co-expression ($p = 0.09$).

OS was estimated in all 271 patients. As shown in Figure 7, FN1-positive group, ITGA3-positive group and FN1/ITGA3-positive group showed significantly poor survival ($p = 0.019$, $p = 0.0007$ and $p > 0.0001$, respectively). Table 4 showed the prognostic factors for

OS. In the univariate analysis, age (≥ 70 years), tumor size (≥ 35 mm), lymph node metastasis, lymphatic invasion, intrapancreatic neural invasion, stage, curability, presence of adjuvant therapy and FN1/ITGA3 co-expression were significant prognostic factors for OS. In multivariate analysis, tumor size (≥ 35 mm), lymph node metastasis, stage, presence of adjuvant therapy and FN1/ITGA3 co-expression were independent prognostic factors for OS.

4. Discussion

In the present study, we identified 95 DEGs in PDAC stroma using dissected stromal areas of PDAC tissues and 1434 DEGs in PDAC cells using public databases, respectively. Subsequently, the DEGs of PDAC stroma and PDAC cells were integrated to our in-house ligand-receptor interaction database, and 9 key genes and 8 key cancer-stromal interactions for PDAC patients were identified.

Recently, studies have investigated molecular subtypes of PDAC based on integrative transcriptional profiling analysis. In the pivotal study, Collisson et al. [17] defined three PDAC subtypes: classical, quasi-mesenchymal and exocrine-like, which show differences in outcome and therapeutic responses [17]. In another study by the Australian Pancreatic Cancer Genome Initiative as part of the International Cancer Genome Consortium (ICGC), 4 subtypes: squamous, pancreatic progenitor, aberrantly differentiated endocrine exocrine (ADEX), and immunogenic have unraveled [18]. Squamous, pancreatic progenitor, and ADEX reproduce the quasi-mesenchymal, classical, and exocrine-like subtypes, respectively, from Collisson et al. [17]. However, a more recent study showed that Collisson et al.'s

"exocrine-like" or "quasi-mesenchymal" subtypes, and the ICGC's "ADEX" or "immunogenic" subtypes were associated with low-purity samples [19], reflecting the contamination of surrounding pancreatic tissue than a genuine PDAC subtype. As just described, gene expression analyses of bulk PDAC tumors are hampered by limited tumor cellularity and the presence of abundant stroma intermixed with normal endocrine and exocrine cells. To overcome these limitations, Moffitt et al. has taken into account the role of the stroma in PDAC subtyping by computationally micro-dissecting normal, tumoral, and microenvironment transcriptomic signals composing PDAC tissue [20]. They not only identified 2 tumor-specific subtypes: a basal-like subtype and a classical subtype, but also defined normal and activated stromal subtypes, which were independently associated with prognosis. "Activated" stroma was characterized by genes associated with tumor promotion, including the ECM proteins SPARC (secreted protein acidic and rich in cysteine), POSTIN (lys-pro-pro-arg), THBS (thrombospondin-1), and FN1 (fibronectin 1) which we precisely identified as the differentially expressed proteins in PDAC stroma. In this study, we used real microdissection for the proteomic analysis of PDAC (or IPMA) stroma, eliminating the contamination of pancreatic cancer, adenoma, normal duct, acinar cells and islets. For the identification of DEGs of PDAC cells, we filtered out the genes which were not detected in the gene expression dataset of common pancreatic cancer cell lines which were assumed to be purely neoplastic [21], resulting in no detection of endocrine- or exocrine-related gene in the DEGs for PDAC cells (Supplementary Table S5).

Gene Set Enrichment Analysis (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically-significant, concordant differences between two biological states [22]. In the previous study, genome-wide transcriptome

analysis by Jones et al. showed more than 21,000 genetic alteration in PDAC, which mainly affected 12 core signaling pathways including “cell adhesion”, “invasion”, “TGF-beta signaling” and “integrin signaling” [23]. Additionally, Pan et al. [9] identified the DEGs between normal tissue and PDAC of different stages using the public microarray dataset (GSE62165), and showed that the up-regulated DEGs were commonly enriched in five fundamental pathways throughout stages, including pathways in cancer, ECM-receptor interaction and focal adhesion. They also showed that LAMA3 (Laminin Subunit Alpha 3), LAMB3 (Laminin Subunit Beta 3), LAMC2 (Laminin Subunit Gamma 2), COL4A1 (Collagen type IV alpha 1) and FN1 were commonly shared by these pathways and were unfavorable factors for prognosis. Bedsides GO and KEGG pathway, enrichment analyses by Wang et al. [24] showed that 20 DEGs of PDAC tissues were enriched in ECM-receptor interaction and focal adhesion pathways, and FN1 as well as genes of collagen family was significantly enriched in these pathways, suggesting that FN1 and genes of collagen family may play an important role in PDAC progression. In our study, no pathway related to ECM-receptor interaction and adhesion was detected as in the previous studies. Meanwhile, 3 GO-BP terms including “Extracellular matrix organization”, “Response to wounding” and “Actin cytoskeleton organization” were overlapped between the DEGs of PDAC cells and stroma, where genes related to ECM-receptor interaction and adhesion pathways categorized. This was because we separately identified the DEGs of PDAC cells and stroma for interactome analysis, and these DEGs included only ligands or receptors of interactions.

FN1 is a major constituent of the ECM within the TME (tumor microenvironment) and is not only produced mainly by fibroblasts, but also by cancer cells [25]. Normally, FN1 supports cell-ECM interactions and is essential for wound healing, development, and tissue

homeostasis [26], whereas, increased cell proliferation and enhanced chemoresistance was found when FN1 adhered to pancreatic cancer cells [27]. The prognostic impact of FN1 expression in PDAC is still controversial. Javle et al. reported that high expression of FN1 correlated with p-ERK and a worsened survival [28]. On the other hand, Hu et al. reported that stromal FN1 expression was not associated with long-term survival by immunohistochemical analysis of 138 PDAC patients [29]. In the present study, we evaluated the prognostic value of FN1 by using gene expression and clinical information data obtained from TCGA, and found that the high expression value of FN1 was negatively associated with OS ($p = 0.041$). Furthermore, we examined FN1 protein expression in pancreatic cancer stroma using the TMAs of 271 pancreatic cancer patients, and found that FN1-positive group showed significantly poor survival ($p = 0.019$, Fig.7A). ITGα5β1 is the primary receptor for FN1. Abrogating FN1-ITGα5β1 interaction in various animal models of cancer inhibited both angiogenesis and tumor growth [30, 31]. However, drugs which target this interaction including PF-04605412 (specific α5 subunit neutralizer [30]), have failed in clinical trials [32]. In our immunohistochemical analysis, no increased ITGA5 protein expression was detected in cancer cells but in stromal cells (Fig. 6), suggesting that ITGA5 may not play an important role as a receptor of FN1 in pancreatic cancer cells. On the other hand, increased ITGA3 protein expression was detected in cancer cells, and FN1/ITGA3 co-expression was an independent prognostic factor for OS. In addition, the expression of ITGA3 was reported as a diagnostic and prognostic biomarker in pancreatic cancer [33]. Together with these things, an anti-ITGα3β1 antibody which specifically blocks the α3 subunit may inhibit tumor progression in PDAC patients. Further understanding of FN1 expression and function in the

context of PDAC may potentially help to improve the effectiveness of FN1 inhibition in the clinical setting.

Recent advances in mass spectrometry have allowed for proteomic profiling of various types of cancer, and integration of these data with other biological data developed a more complete understanding of specific cancers and their genetic drivers [34, 35]. Since proteins are ultimately the functional effectors of biological activity in cancer cells, we hypothesized that interactome analysis using proteome data may be an especially sensitive method for identifying potential therapeutic targets in PDAC. We succeeded in identifying the DE proteins in PDAC stroma from LMD samples. However, the proteins in cancer cells were not able to be analyzed due to low tumor cellularity of PDAC tissues. To overcome this limitation, further experimental studies with high-sensitivity mass spectrometry for very low-abundance proteins, careful microdissection, and single-cell profiling technologies are encouraged.

In conclusion, we demonstrated the integrated genomic and proteomic interactome analysis of PDAC, and identified 8 key cancer-stromal interactions. Besides, we demonstrated that FN1-ITGA3 had unfavorable prognostic impact for PDAC patients using the TMAs of 271 PDAC cases. Although our study has limitations as described above, our integrative exploration helps advance strategies and leads to discovery of clinically relevant cancer-stromal interactions in PDAC that can be targeted with drugs.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

This study was supported in part by a Japan Society for the Promotion of Science KAKENHI grant no. 26830081 to Y.H. and 19K09181 to Y.H. The authors wish to thank Mr. Yoshiyasu Nakamura and Ms. Mitsuyo Yoshihara for their technical support.

References

1. Ilic M, Ilic I. Epidemiology of pancreatic cancer. *World J Gastroenterol*. 2016;22(44):9694-9705. doi:10.3748/wjg.v22.i44.9694
2. Neoptolemos JP, Kleeff J, Michl P, Costello E, Greenhalf W, Palmer DH. Therapeutic developments in pancreatic cancer: current and future perspectives. *Nat Rev Gastroenterol Hepatol*. 2018;15(6):333-348. doi:10.1038/s41575-018-0005-x
3. Bachem MG, Schünemann M, Ramadani M, et al. Pancreatic carcinoma cells induce fibrosis by stimulating proliferation and matrix synthesis of stellate cells. *Gastroenterology*. 2005;128(4):907-921. doi:10.1053/j.gastro.2004.12.036
4. Imamura T, Iguchi H, Manabe T, et al. Quantitative analysis of collagen and collagen subtypes I, III, and V in human pancreatic cancer, tumor-associated chronic pancreatitis, and alcoholic chronic pancreatitis. *Pancreas*. 1995;11(4):357-364. doi:10.1097/00006676-199511000-00007
5. Mollenhauer J, Roether I, Kern HF. Distribution of extracellular matrix proteins in pancreatic ductal adenocarcinoma and its influence on tumor cell proliferation in vitro. *Pancreas*. 1987;2(1):14-24. doi:10.1097/00006676-198701000-00003
6. Joyce J, Quail D. Microenvironmental regulation of tumor progression and metastasis. *Nat Med*. 2013;19(11):1423-1437. doi:10.1038/nm.3394.Microenvironmental
7. Jhoti H, Williams G, Rees DC, Murray CW. The “rule of three” for fragment-based drug discovery: where are we now? *Nat Rev Drug Discov*. 2013;12(8):644-644. doi:10.1038/nrd3926-c1
8. Komura D, Isagawa T, Kishi K, et al. CASTIN: a system for comprehensive analysis of cancer-stromal interactome. *BMC Genomics*. 2016;17(1):899. doi:10.1186/s12864-016-3207-z
9. Pan Z, Li L, Fang Q, et al. Analysis of dynamic molecular networks for pancreatic ductal adenocarcinoma progression. *Cancer Cell Int*. 2018;18(1):1-18. doi:10.1186/s12935-018-0718-5
10. Creighton CJ. Analysis of Tumor-Host Interactions by Gene Expression Profiling of

Lung Adenocarcinoma Xenografts Identifies Genes Involved in Tumor Formation. *Mol Cancer Res*. 2005;3(3):119-129. doi:10.1158/1541-7786.mcr-04-0189

11. Henare K, Wang L, Wang LC, et al. Dissection of stromal and cancer cell-derived signals in melanoma xenografts before and after treatment with DMXAA. *Br J Cancer*. 2012;106(6):1134-1147. doi:10.1038/bjc.2012.63
12. Kimura Y, Yanagimachi M, Ino Y, et al. Identification of candidate diagnostic serum biomarkers for Kawasaki disease using proteomic analysis. *Sci Rep*. 2017;7(March):1-12. doi:10.1038/srep43732
13. Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc 2007 28*. 2007;2(8):1896. doi:10.1038/nprot.2007.261
14. Paul Shannon 1, Andrew Markiel 1, Owen Ozier, 2 Nitin S. Baliga, 1 Jonathan T. Wang, 2 Daniel Ramage 2, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res*. 2003;13(22):6. doi:10.1101/gr.1239303.metabolite
15. Keshava Prasad TS, Goel R, Kandasamy K, et al. Human Protein Reference Database - 2009 update. *Nucleic Acids Res*. 2009;37(SUPPL. 1):767-772. doi:10.1093/nar/gkn892
16. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28(1):27. doi:10.1093/NAR/28.1.27
17. Collisson EA, Sadanandam A, Olson P, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med*. 2011;17(4):500-503. doi:10.1038/nm.2344
18. Bailey P, Chang DK, Nones K, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*. 2016;531(7592):47-52. doi:10.1038/nature16965
19. Raphael BJ, Hruban RH, Aguirre AJ, et al. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell*. 2017;32(2):185-203.e13. doi:10.1016/j.ccell.2017.07.007
20. Moffitt RA, Marayati R, Flate EL, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet*. 2015;47(10):1168-1178. doi:10.1038/ng.3398
21. Iacobuzio-Donahue CA, Maitra A, Olsen M, et al. Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays. *Am J Pathol*. 2003;162(4):1151-1162. doi:10.1016/S0002-9440(10)63911-9
22. Subramanian A, Subramanian A, Tamayo P, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc*

Natl Acad Sci U S A. 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102

23. Jones S, Zhang X, Parsons DW, et al. Core Signaling Pathways in Human Pancreatic Cancers Revealed by Global Genomic Analyses. *Science (80-).* 2008;321(5897):1801-1806. doi:10.1126/science.1164368
24. Wang Y, Li Y. Analysis of molecular pathways in pancreatic ductal adenocarcinomas with a bioinformatics approach. *Asian Pacific J Cancer Prev*. 2015;16(6):2561-2567. doi:10.7314/APJCP.2015.16.6.2561
25. Topalovski M, Brekken RA. Matrix control of pancreatic cancer: New insights into fibronectin signaling. *Cancer Lett*. 2016;381(1):252-258. doi:10.1016/j.canlet.2015.12.027
26. Pankov R, Yamada KM. Fibronectin at a glance. *J Cell Sci*. 2002;115(Pt 20):3861-3863. doi:10.1242/jcs.00059
27. Miyamoto H, Murakami T, Tsuchida K, Sugino H, Miyake H, Tashiro S. Tumor-stroma interaction of human pancreatic cancer: acquired resistance to anticancer drugs and proliferation regulation is dependent on extracellular matrix proteins. *Pancreas*. 2004;28(1):38-44. http://www.ncbi.nlm.nih.gov/pubmed/14707728.
28. Javle MM, Gibbs JF, Iwata KK, et al. Epithelial-mesenchymal transition (EMT) and activated extracellular signal-regulated kinase (p-Erk) in surgically resected pancreatic cancer. *Ann Surg Oncol*. 2007;14(12):3527-3533. doi:10.1245/s10434-007-9540-3
29. Hu D, Ansari D, Zhou Q, Sasor A, Said Hilmersson K, Andersson R. Stromal fibronectin expression in patients with resected pancreatic ductal adenocarcinoma. *World J Surg Oncol*. 2019;17(1):1-8. doi:10.1186/s12957-019-1574-z
30. Li G, Zhang L, Chen E, et al. Dual functional monoclonal antibody PF-04605412 targets integrin α5β1 and elicits potent antibody-dependent cellular cytotoxicity. *Cancer Res*. 2010;70(24):10243-10254. doi:10.1158/0008-5472.CAN-10-1996
31. Bhaskar V, Zhang D, Fox M, et al. A function blocking anti-mouse integrin α5β1 antibody inhibits angiogenesis and impedes tumor growth in vivo. *J Transl Med*. 2007;5:1-11. doi:10.1186/1479-5876-5-61
32. Mateo J, Berlin J, De Bono JS, et al. A first-in-human study of the anti-α5β1 integrin monoclonal antibody PF-04605412 administered intravenously to patients with advanced solid tumors. *Cancer Chemother Pharmacol*. 2014;74(5):1039-1046. doi:10.1007/s00280-014-2576-8
33. Jiao Y, Li Y, Liu S, Chen Q, Liu Y. ITGA3 serves as a diagnostic and prognostic biomarker for pancreatic cancer. *Onco Targets Ther*. 2019;Volume 12:4141-4152.

doi:10.2147/OTT.S201675

34. Edwards NJ, Oberti M, Thangudu RR, et al. The CPTAC Data Portal: A Resource for Cancer Proteomics Research. *J Proteome Res*. 2015;14(6):2707-2713. doi:10.1021/pr501254j
35. Huang KL, Li S, Mertins P, et al. Corrigendum: Proteogenomic integration reveals therapeutic targets in breast cancer xenografts. *Nat Commun*. 2017;8:15479. doi:10.1038/ncomms15479
36. Jones PL, Crack J, Rabinovitch M. Regulation of tenascin-C, a vascular smooth muscle cell survival factor that interacts with the alpha v beta 3 integrin to promote epidermal growth factor receptor phosphorylation and growth. *J Cell Biol*. 1997;139(1):279-293. doi:10.1083/jcb.139.1.279
37. Huang W, Chiquet-Ehrismann R, Orend G, Moyano J V., Garcia-Pardo A. Interference of tenascin-C with syndecan-4 binding to fibronectin blocks cell adhesion and stimulates tumor cell proliferation. *Cancer Res*. 2001;61(23):8586-8594.
38. Lawler J, Hynes RO. An integrin receptor on normal and thrombasthenic platelets that binds thrombospondin. *Blood*. 1989;74(6):2022-2027. http://www.ncbi.nlm.nih.gov/pubmed/2478219.
39. Sipes JM, Krutzsch HC, Lawler J, Roberts DD. Cooperation between thrombospondin-1 type 1 repeat peptides and α(v)β3 integrin ligands to promote melanoma cell spreading and focal adhesion kinase phosphorylation. *J Biol Chem*. 1999;274(32):22755-22762. doi:10.1074/jbc.274.32.22755
40. Ding Y, Pan Y, Liu S, Jiang F, Jiao J. Elevation of MiR-9–3p suppresses the epithelial-mesenchymal transition of nasopharyngeal carcinoma cells via down-regulating FN1, ITGB1 and ITGAV. *Cancer Biol Ther*. 2017;18(6):414-424. doi:10.1080/15384047.2017.1323585
41. Gao W, Liu Y, Qin R, Liu D, Feng Q. Silence of fibronectin 1 increases cisplatin sensitivity of non-small cell lung cancer cell line. *Biochem Biophys Res Commun*. 2016;476(1):35-41. doi:10.1016/j.bbrc.2016.05.081
42. Eriksson J, Le Joncour V, Nummela P, et al. Gene expression analyses of primary melanomas reveal CTHRC1 as an important player in melanoma progression. *Oncotarget*. 2016;7(12). doi:10.18632/oncotarget.7604
43. Li W, Liu Z, Zhao C, Zhai L. Binding of MMP-9-degraded fibronectin to β6 integrin promotes invasion via the FAK-Src-related Erk1/2 and PI3K/Akt/Smad-1/5/8 pathways in breast cancer. *Oncol Rep*. 2015;34(3):1345-1352. doi:10.3892/or.2015.4103

44. Miroshnikova YA, Rozenberg GI, Cassereau L, et al. α5β1-Integrin promotes tension-dependent mammary epithelial cell invasion by engaging the fibronectin synergy site. *Mol Biol Cell*. 2017;28(22):2958-2977. doi:10.1091/mbc.e17-02-0126

FIGURE LEGENDS

Figure 1. Representative images of pancreatic cancer sections stained with hematoxylin (A) and hematoxylin and eosin (B) for LCM. Pancreatic cancer, adenoma and normal ducts were identified in both hematoxylin and H&E staining with a 4x magnification (the area surrounded by a yellow line). Pancreatic acinar cells and islets were identified in the tissue as well (the area surrounded by a blue line). Scans of the H&E stained reference sections were used to map the pancreatic stromal area (B). Areas of pancreatic cancer, adenoma, normal duct, acinar cells and islets were microdissected from pancreatic tissue sections, and the remaining area was defined as the pancreatic stroma. (C) Scheme for integration of proteomic and transcriptomic data based on differential expression clustering and cancer-stromal interaction.

Figure 2. Identification and pathway enrichment analysis of differentially-expressed proteins in PDAC stroma. (A) Dendrogram of protein expression levels in PDAC and IPMA stromal tissues showing PDAC and IPMA samples were clearly grouped into 2 clusters. (B) MA plot of differential gene expression levels in the two groups, where expression intensity is on the x-axis and differences in gene expression levels are on the y-axis, each dot represents one gene, blue dots represent genes whose abundance is differentially expressed (FDR<0.05) in PDAC stroma, and black dots non-significantly changed regulation. (C) Networks of

enriched terms. A subset of enriched terms has been selected and rendered as a network plot, where terms with a similarity > 0.3 are connected by edges. Each node is colored by cluster ID, and nodes which share the same cluster ID are typically close to each other.

Figure 3. Identification and pathway enrichment analysis of differentially expressed genes in PDAC cells. (A) Principal Component Analysis (PCA) mapping for gene expression of patients with normal pancreatic tissue or PDAC. The horizontal axis represents the score of the first principal component of each sample; the ordinate axis represents the score of the second principal component. N represents the normal pancreatic tissue samples (total of 197), T represents the PDAC samples (total of 169). (B) MA plot of differential gene expression levels in the two groups, where expression intensity is on the x-axis and differences in the gene expression levels are on the y-axis, each dot represents one gene, blue dots represent genes whose abundance is differentially expressed (FC > 16 and FDR < 0.05) in PDAC cells. (C) Hierarchical clustering diagram of differences between normal pancreatic tissue (n=197) and PDAC (n=169). The horizontal axis represents the sample name (a blue line indicates normal tissues; red line indicates the PDAC). The right ordinate axis represents the clustering condition of 1434 DEGs. Red indicates the up regulation of the gene, blue indicates the down regulation. (D) Networks of enriched terms by the 1434 DEGs. A subset of enriched terms has been selected and rendered as a network plot, where terms with a similarity >0.3 are connected by edges. Each node is colored by cluster ID, and nodes which share the same cluster ID are typically close to each other.

Figure 4. Overall survival analysis of nine candidate genes by Kaplan–Meier plotter. Log-rank test was used to evaluate significance.

Figure 5. Overall survival analysis of eight candidate interactions by Kaplan–Meier plotter. Log-rank test was used to evaluate significance.

Figure 6. Representative images of pancreatic cancer tissues stained for FN1, ITGA3 and ITGA5. Increased FN1 protein expression were detected in stromal cells and ITGA3 protein expression in cancer cells. No increased ITGA5 protein expression was detected in cancer cells but in stromal cells (Scale bar: 600 μm).

Figure 7. Overall survival for pancreatic cancer patients with tumors positive for FN1 (A),

ITGA3 (B) expression or FN1/ITGA3 (C) co-expression by Kaplan–Meier plotter. Log-rank test was used to evaluate significance.

Table 1. Sample summary for proteomic analysis

Sample	Age	Gender	Pathology	TNM
PDAC270	74	M	mod	T3N0M0
PDAC278	67	F	wel	T3N1M0
PDAC284	53	F	mod	T3N1M0
PDAC290	71	M	mod	T3N0M0
PDAC291	77	M	mod	T3N0M0
PDAC294	67	M	mod	T3N0M0
PDAC349	65	F	wel	T3N1M0
PDAC397	78	M	por	T3N1M0
IPMA281	65	F	IPMA with mild atypia	NA
IPMA285	36	M	IPMA with mild atypia	NA
IPMA296	62	F	IPMA with mild atypia	NA
IPMA297	57	F	IPMA	NA
IPMA352	75	M	IPMA	NA
IPMA392	73	M	IPMN: adenocarcinoma in adenoma	TisN0M0

mod: moderately differentiated adenocarcinoma, wel: well differentiated adenocarcinoma, por: poorly differentiated adenocarcinoma, IPMN: Intraductal papillary-mucinous neoplasm

Table 2. Mutually dependent interactions in PDAC dataset

Ligand	Receptor	Direction	Possible relevance for cancer-stroma interactions
TNC	ITGA5	S-C	TNC interacts with ITGA5 to promote epidermal growth factor receptor phosphorylation and growth [36]. Interference of TNC with syndecan-4 binding to fibronectin blocks cell adhesion and stimulates tumor cell proliferation [37].
THBS1	ITGB3	S-C	THBS1 is secreted by malignant glioma cells and interacts with ITGAVB3 and ITGA3B1 to promote migration [38]. Cooperation between THBS1 and ITGAVB3 to promote melanoma cell spreading [39].
FN1	ITGB1	S-C	Elevation of miR-9-3p suppresses the proliferation and metastases of nasopharyngeal carcinoma via downregulating FN1, ITGB1 and inhibiting the EMT process [40]. FN1 plays a role in the development of cisplatin resistance in non-small cell lung cancer, possibly by modulation of β-catenin signaling through interaction with ITGB1 [41].
FN1	ITGB3	S-C	FN1 and ITGB3 are coordinated with CTHRC1 and promote migration of melanoma cells [42].
FN1	ITGB6	S-C	Promotes breast cancer invasion [43].
FN1	ITGB7	S-C	Not reported
FN1	ITGA3	S-C	Expression of ITGA3 can be used as a diagnostic and prognostic biomarker in pancreatic cancer [33].
FN1	ITGA5	S-C	Rigid collagen fibrils potentiate PI3K activation to promote malignancy and consistent up-regulation of α5β1 integrin and FN in many tumors and their correlation with cancer aggression [44].

S-C: signal transduction from stromal ligand to cancer cell-receptor

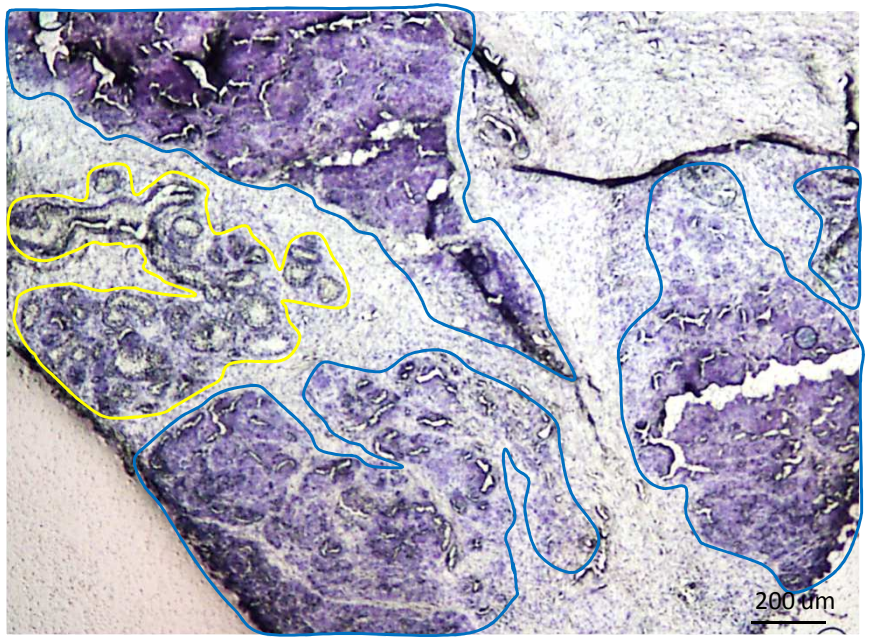
Table 3. Clinicopathological parameters and FN1/ITGA3 co-expression

	FN1/ITGA3 co-expression		
	Positive (*n* = 68)	**Negative (*n* = 203)**	***P***
Age (years) *	68.5 (44 - 86)	68 (40 - 86)	0.865†
Gender			
Male / Female	37 / 31	108 / 95	1.0‡
Tumor size (mm) *	35 (10 - 90)	35 (5 - 105)	0.247†
Histological type			
Well, Moderately / Poorly	59 / 9	178 / 25	0.834‡
Depth of invasion (UICC)			
T1, 2 / T3, 4	6 / 62	21 / 182	0.819‡
Lymph node metastasis (UICC)			
Negative / Positive	14 / 54	64 / 139	0.09‡
Lymphatic invasion			
Negative / Positive	25 / 43	78 / 125	0.886‡
Venous invasion			
Negative / Positive	14 / 54	50 / 153	0.621‡
Intrapancreatic neural invasion			
Negative / Positive	12 / 56	47 / 156	0.398‡
Stage (UICC)			
I, II / III, IV	59 / 9	189 / 14	0.13‡
Curability			
R0, 1 / R2	67 / 1	203 / 0	0.251‡
Adjuvant therapy (R0, 1)			
Present / Absent	49 / 18	152 / 51	0.635‡
Recurrence (R0, 1)			
Negative / Positive	6 / 61	40 / 163	0.04‡
***Values are mean (range).**			
UICC, International Union against Cancer Classification.			
†Mann–Whitney U test; ‡Fisher's exact test.			

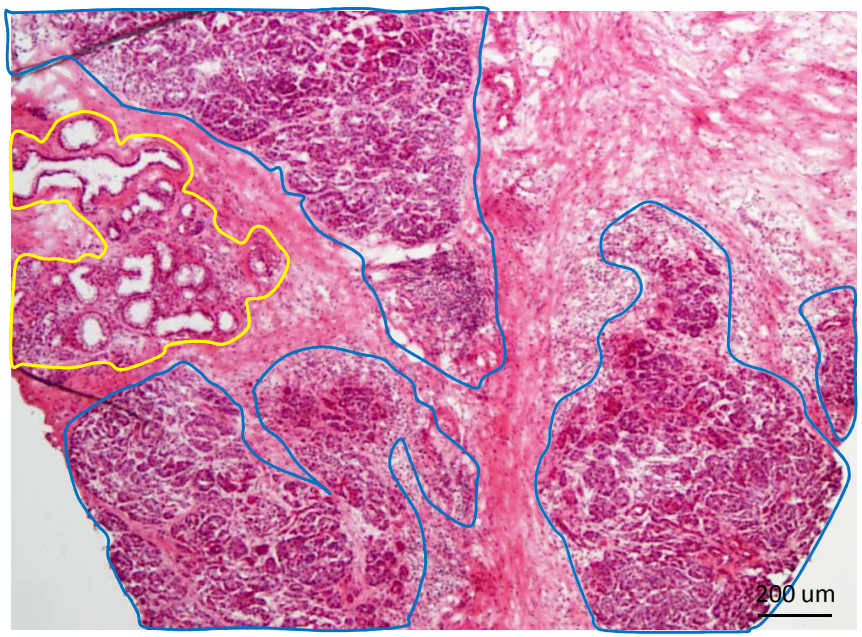
Table 4. Prognostic factors for overall survival (Cox proportional hazard regression model)

Variables	*n*	Univariate	Multivariate	
		p-value	Hazard ratio (95% CI)	*p*-value
Tumor size (≧ 35 mm)				
Negative vs Positive	144 / 127	0.03		
Gender				
Female vs Male	122 / 149	0.71		
Tumor size (≧ 35 mm)				
Negative vs Positive	72 / 199	< 0.001	1.815 (1.201 - 2.743)	0.005
Histological type				
Well, Moderately vs Poorly	237 / 34	0.05		
Depth of invasion (UICC)				
T1, 2 vs T3, 4	27 / 244	0.26		
Lymph node metastasis				
Negative vs Positive	78 / 193	< 0.001	1.811 (1.192 - 2.750)	0.005
Lymphatic invasion				
Negative vs Positive	103 / 168	0.01		
Venous invasion				
Negative vs Positive	64 / 207	0.24		
Intrapancreatic neural invasion				
Negative vs Positive	59 / 212	0.02		
Stage (UICC)				
I, II vs III, IV	248 / 23	< 0.001	2.121 (1.241 - 3.625)	0.006
Curability				
R0, R1 vs R2	270 / 1	0.006		
Adjuvant therapy (R0,R1)				
Absent vs Present	201 / 70	< 0.001	0.343 (0.241 - 0.489)	< 0.001
FN1/ITGA3 co-expression				
Negative vs Positive	68 / 203	< 0.001	1.784 (1.259 - 2.530)	< 0.001

A


200 um

B


200 um

C

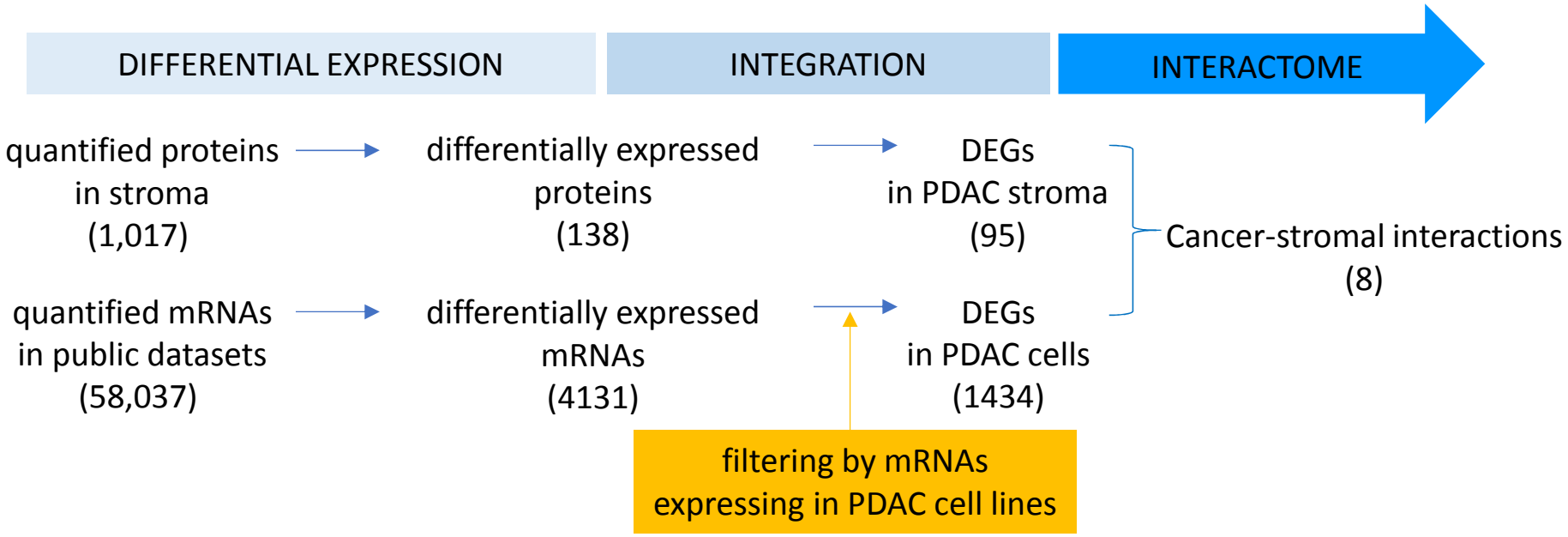
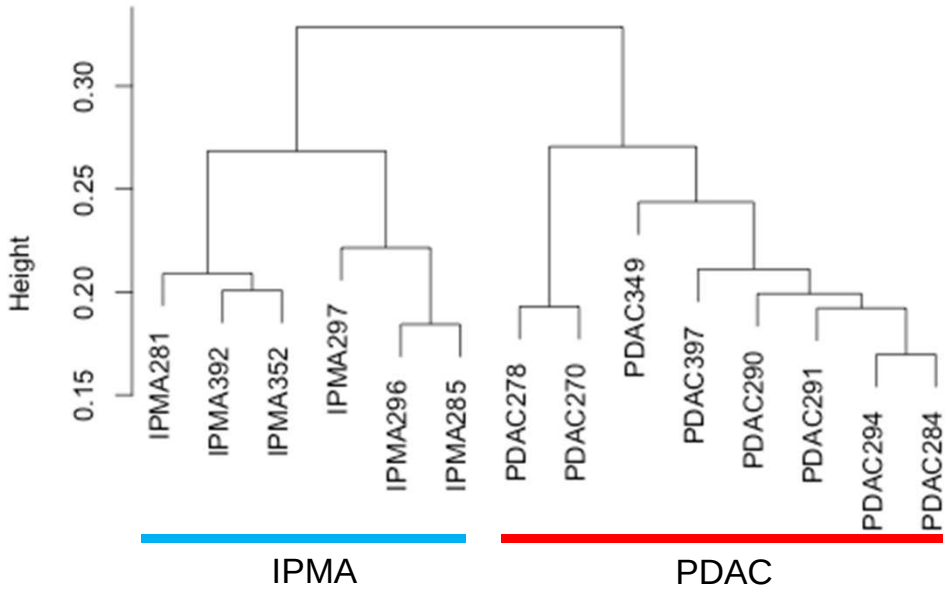
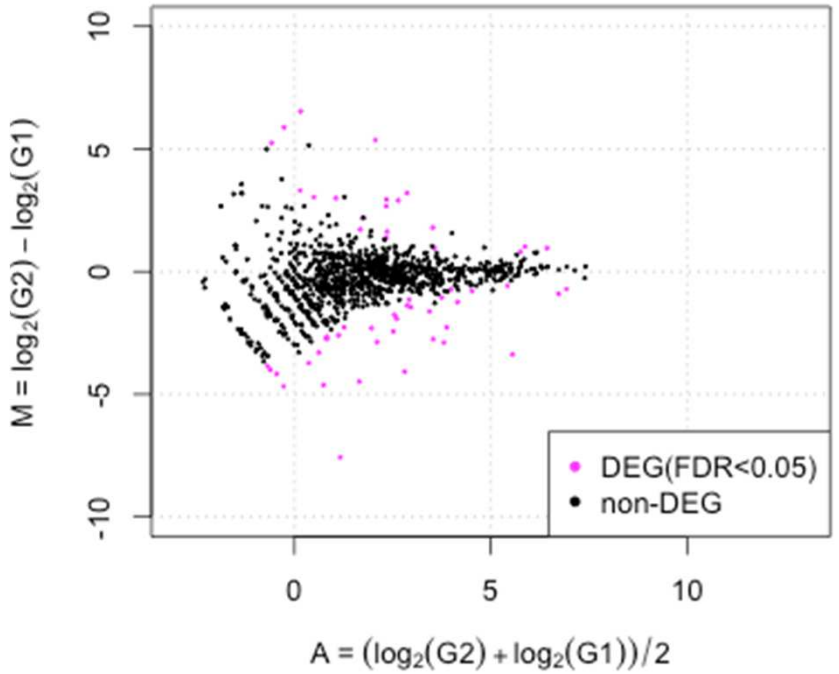

DIFFERENTIAL EXPRESSION
INTEGRATION
INTERACTOME
quantified proteins in stroma (1,017)
differentially expressed proteins (138)
DEGs in PDAC stroma (95)
quantified mRNAs in public datasets (58,037)
differentially expressed mRNAs (4131)
DEGs in PDAC cells (1434)
Cancer-stromal interactions (8)
filtering by mRNAs expressing in PDAC cell lines

Figure 1

A


Height
0.15
0.20
0.25
0.30
IPMA281
IPMA392
IPMA352
IPMA297
IPMA296
IPMA285
PDAC278
PDAC270
PDAC349
PDAC397
PDAC290
PDAC291
PDAC294
PDAC284
IPMA
PDAC

B


M = log2(G2) – log2(G1)
A = (log2(G2) + log2(G1))/2
10
5
0
-5
-10
0
5
10
DEG(FDR<0.05)
non-DEG

C

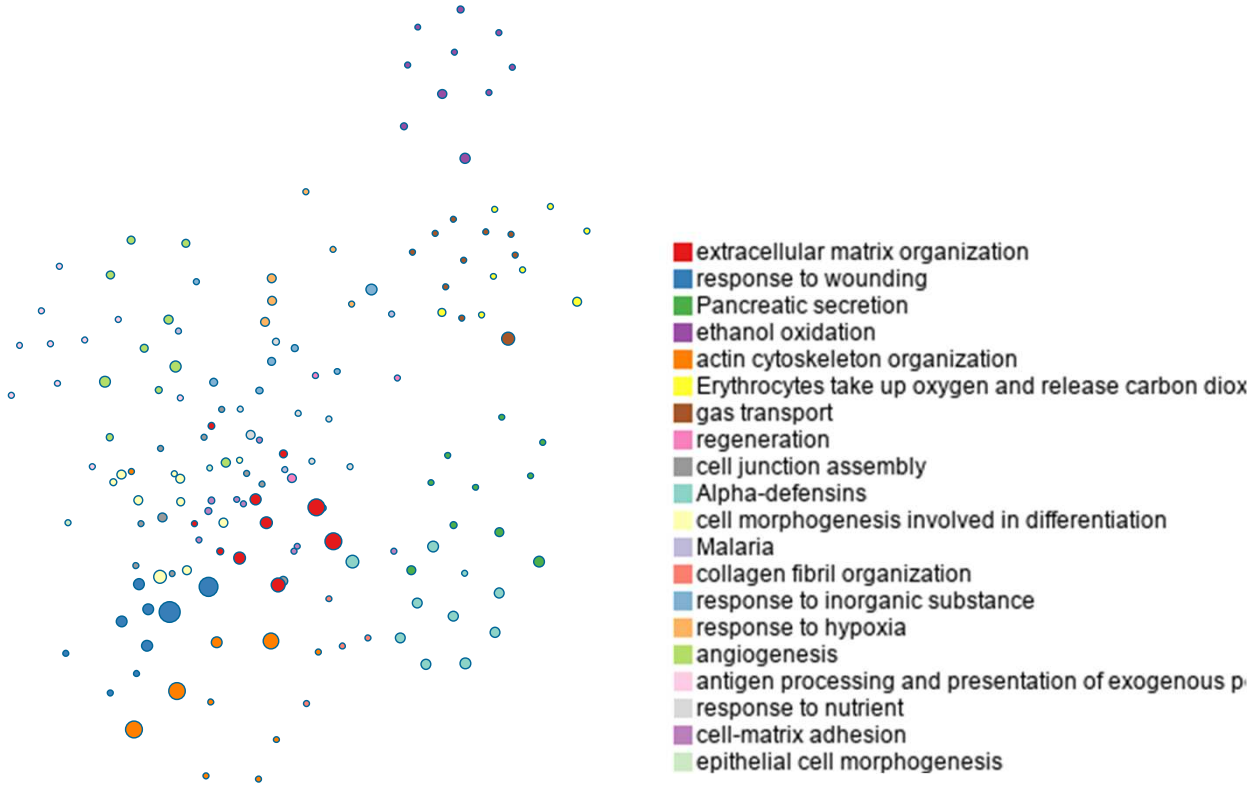

extracellular matrix organization
response to wounding
Pancreatic secretion
ethanol oxidation
actin cytoskeleton organization
Erythrocytes take up oxygen and release carbon diox
gas transport
regeneration
cell junction assembly
Alpha-defensins
cell morphogenesis involved in differentiation
Malaria
collagen fibril organization
response to inorganic substance
response to hypoxia
angiogenesis
antigen processing and presentation of exogenous p
response to nutrient
cell-matrix adhesion
epithelial cell morphogenesis

Figure 2

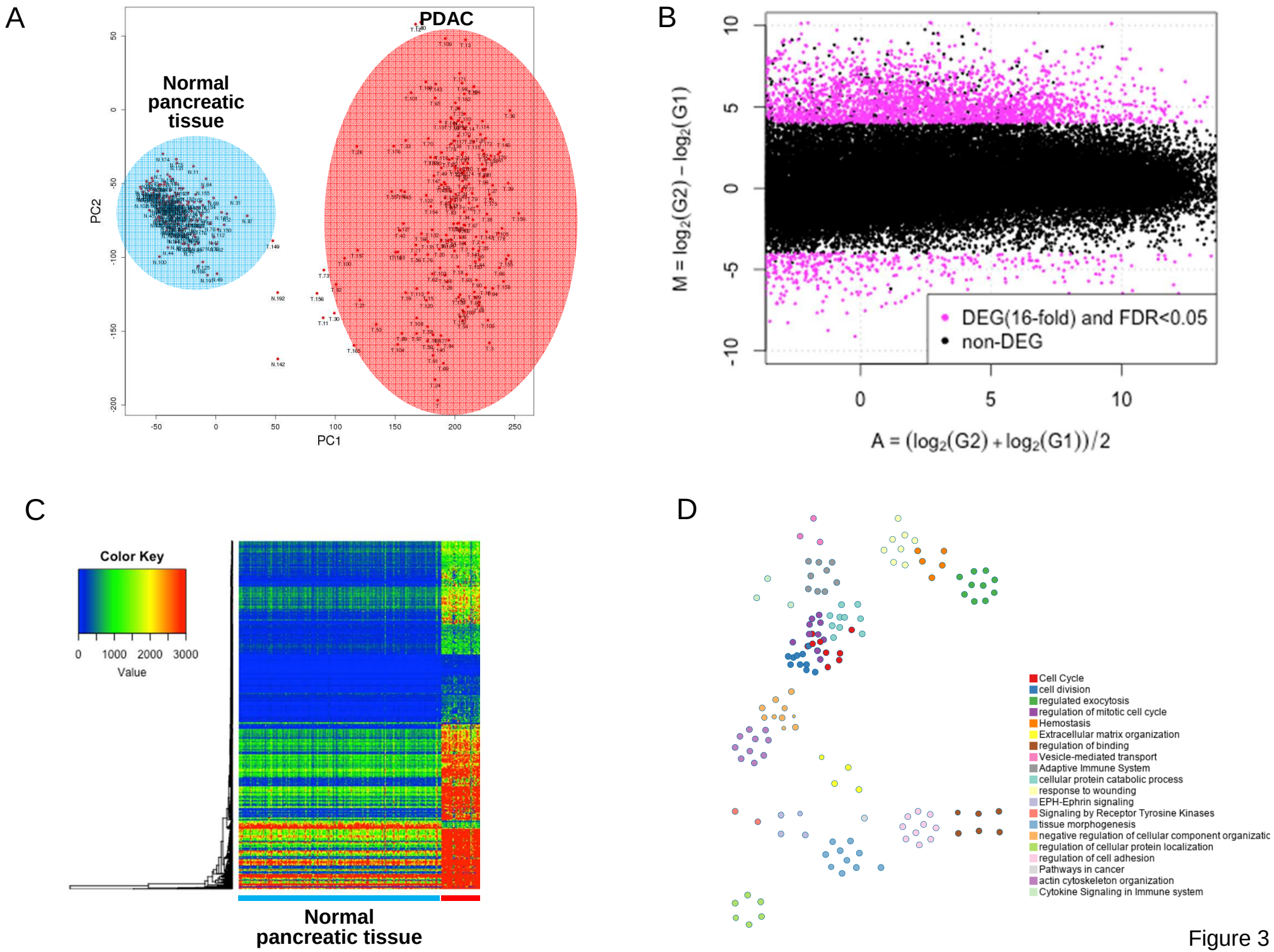

A
PDAC
Normal pancreatic tissue
PC2
PC1
B
M = log2(G2) – log2(G1)
A = (log2(G2) + log2(G1))/2
DEG(16-fold) and FDR<0.05
non-DEG
C
Color Key
0 1000 2000 3000
Value
Normal pancreatic tissue
D
Cell Cycle
cell division
regulated exocytosis
regulation of mitotic cell cycle
Hemostasis
Extracellular matrix organization
regulation of binding
Vesicle-mediated transport
Adaptive Immune System
cellular protein catabolic process
response to wounding
EPH-Ephrin signaling
Signaling by Receptor Tyrosine Kinases
tissue morphogenesis
negative regulation of cellular component organizatic
regulation of cellular protein localization
regulation of cell adhesion
Pathways in cancer
actin cytoskeleton organization
Cytokine Signaling in Immune system

Figure 3

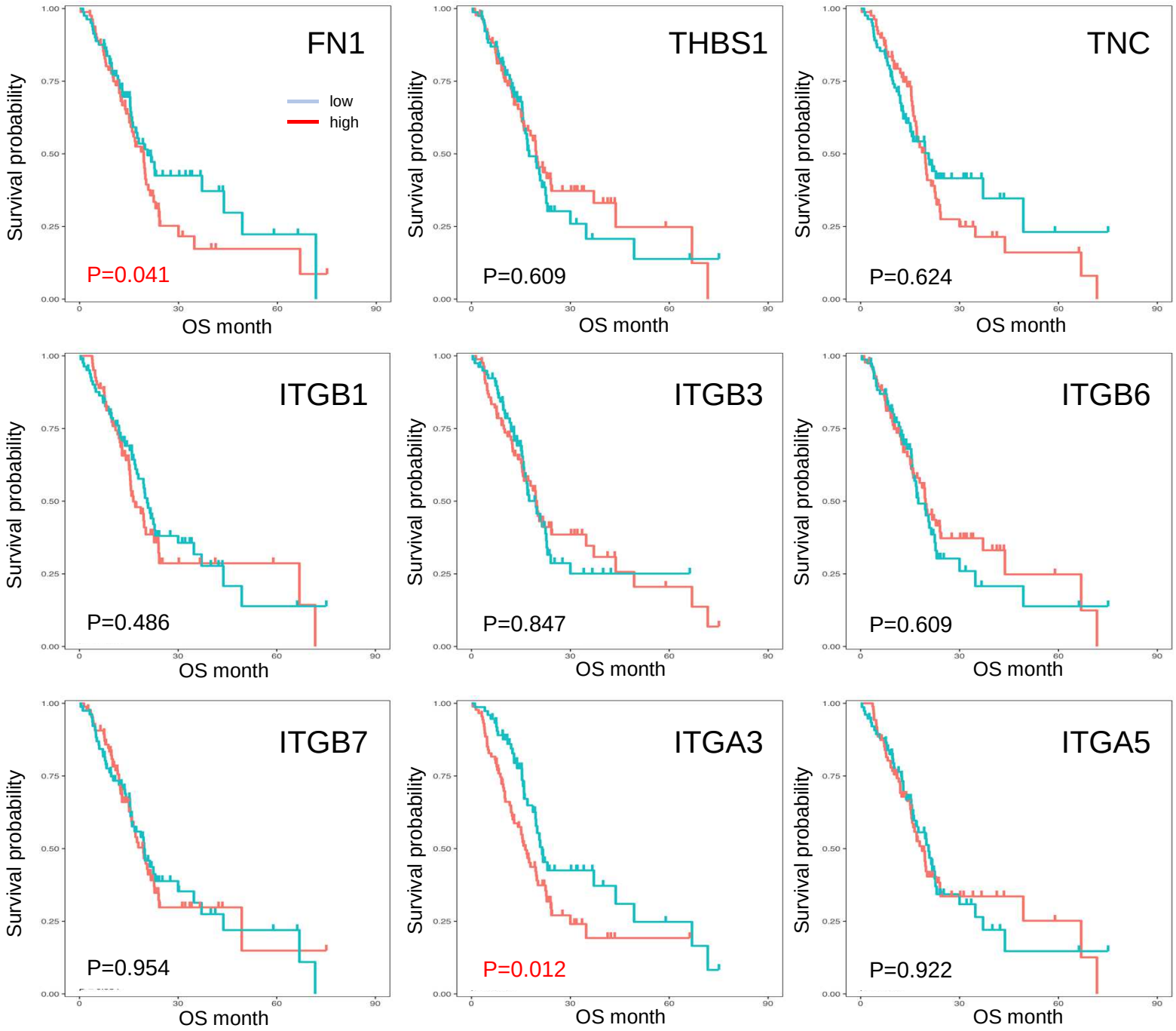

FN1
low
high
P=0.041
THBS1
P=0.609
TNC
P=0.624
ITGB1
P=0.486
ITGB3
P=0.847
ITGB6
P=0.609
ITGB7
P=0.954
ITGA3
P=0.012
ITGA5
P=0.922
Survival probability
OS month

Figure 4

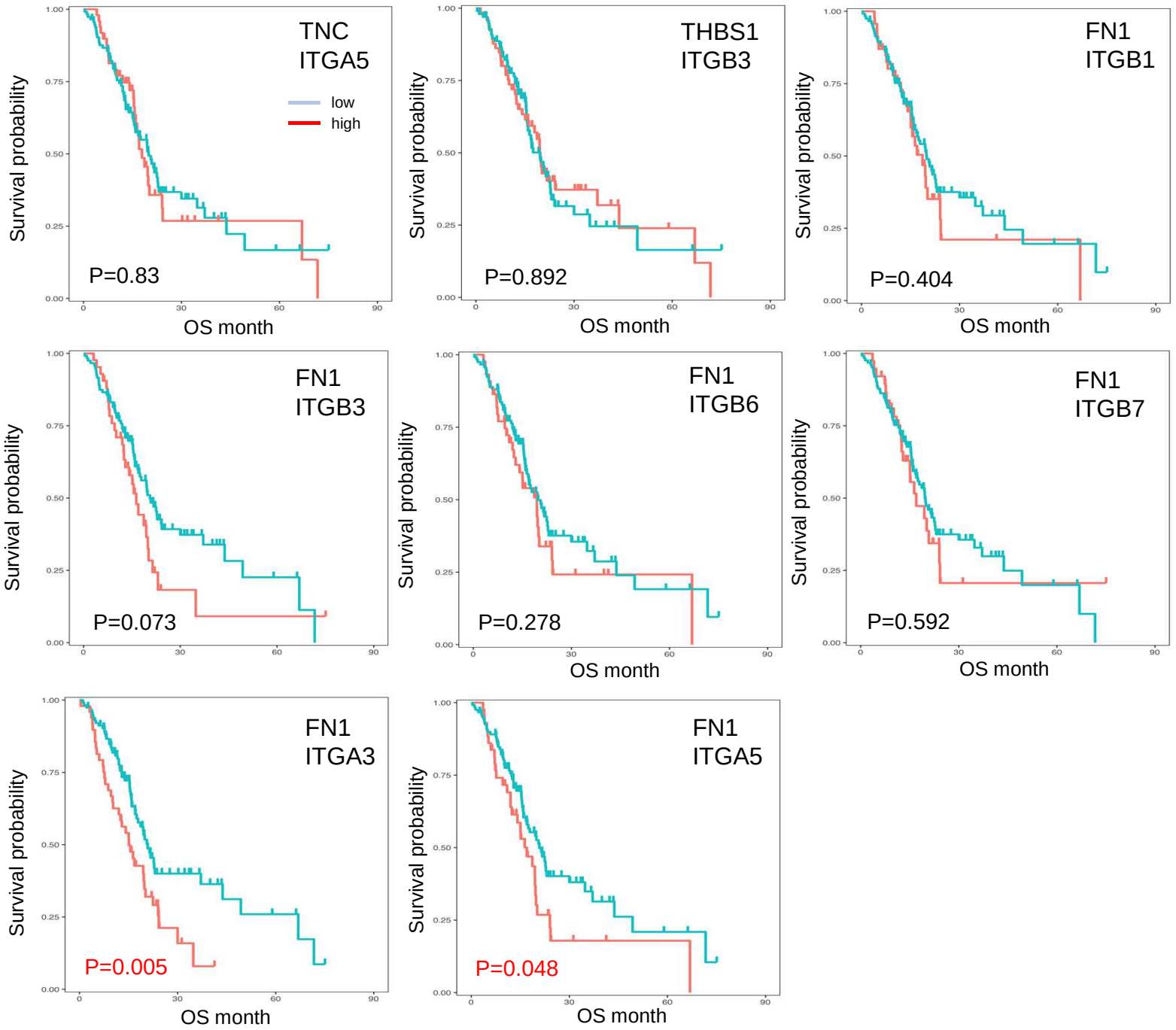

TNC
ITGA5
low
high
P=0.83
THBS1
ITGB3
P=0.892
FN1
ITGB1
P=0.404
FN1
ITGB3
P=0.073
FN1
ITGB6
P=0.278
FN1
ITGB7
P=0.592
FN1
ITGA3
P=0.005
FN1
ITGA5
P=0.048
Survival probability
OS month

Figure 5

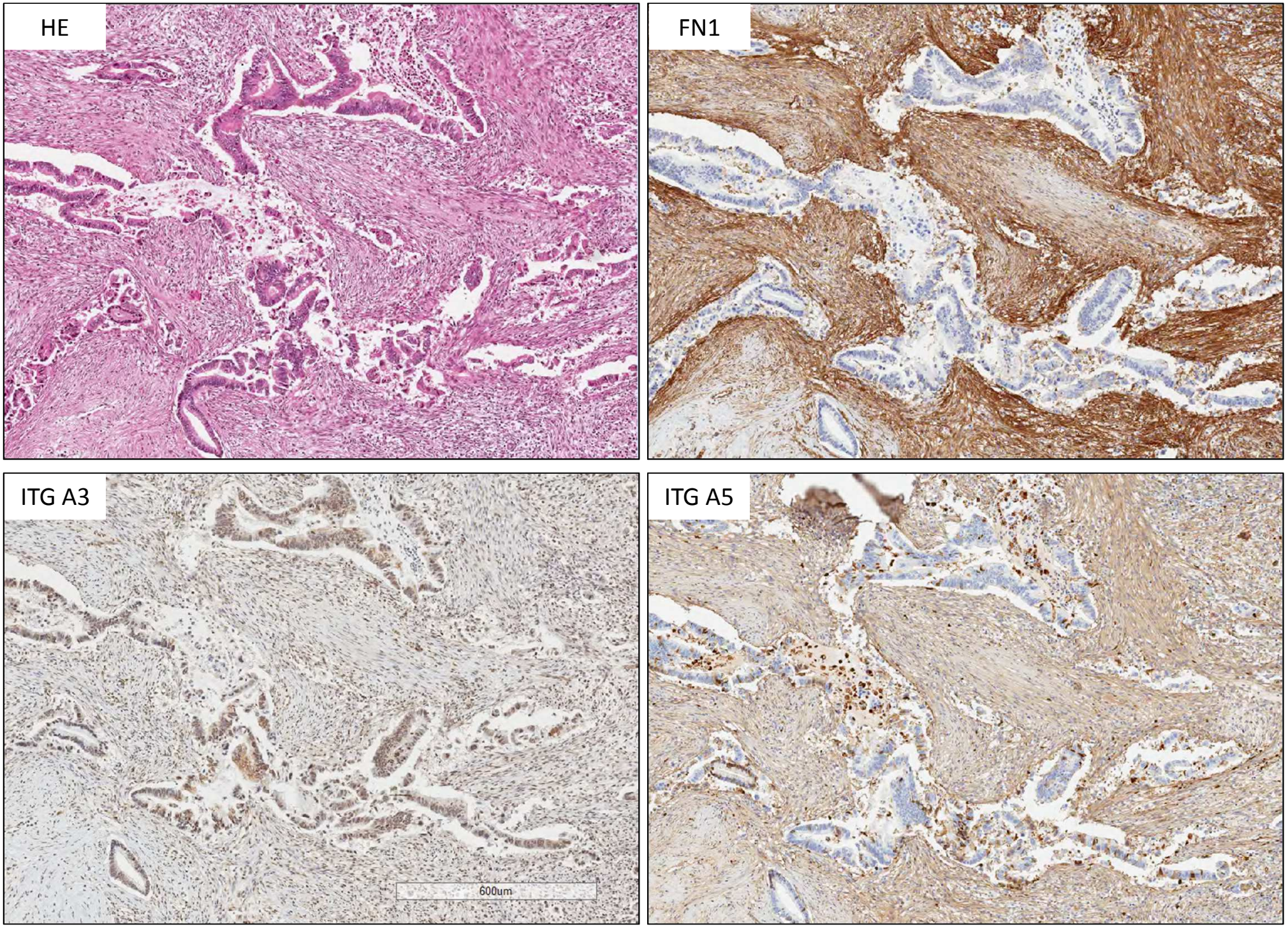


Figure 6

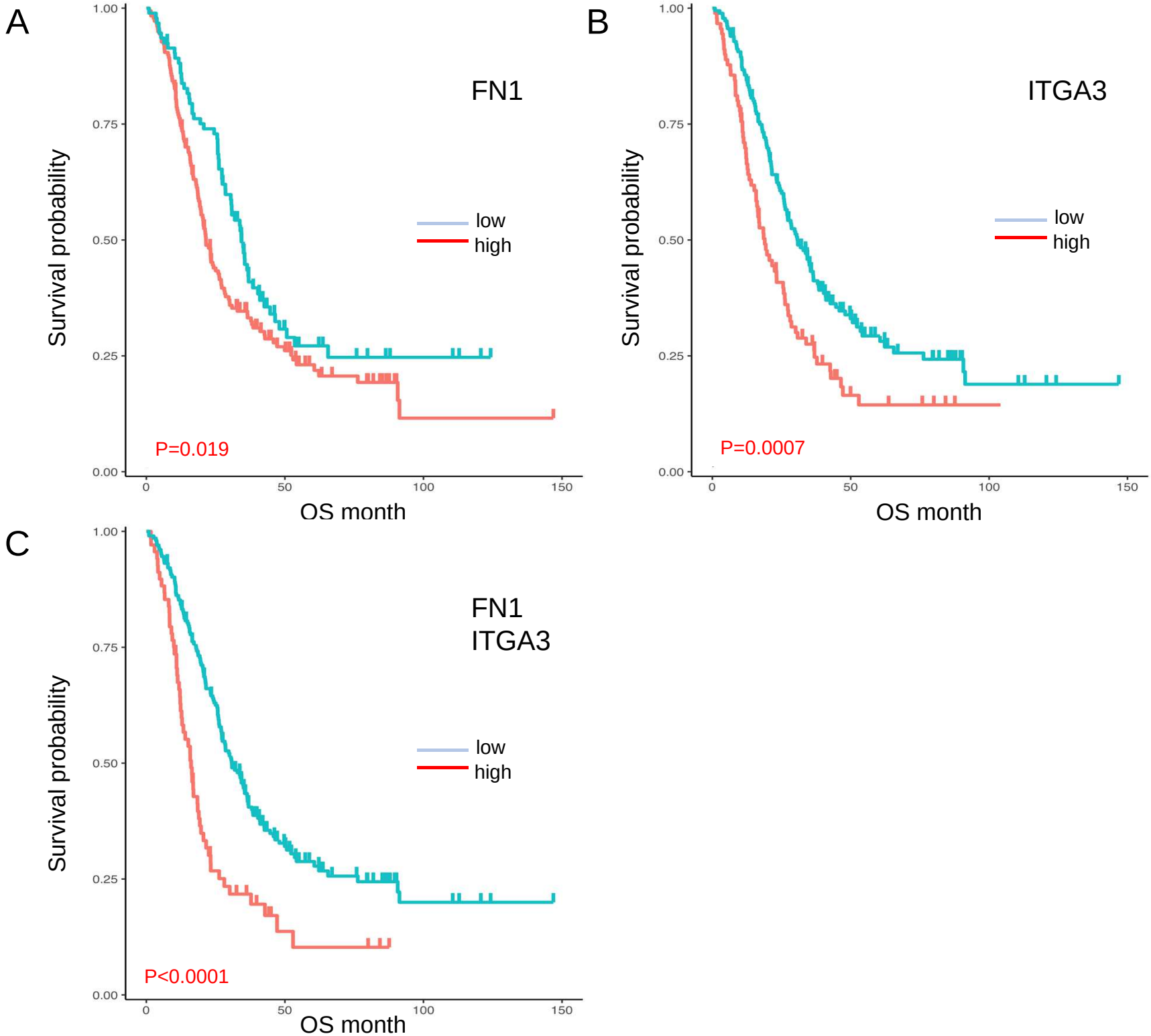

A
FN1
low
high
Survival probability
1.00
0.75
0.50
0.25
0.00
0
50
100
150
OS month
P=0.019
B
ITGA3
low
high
Survival probability
1.00
0.75
0.50
0.25
0.00
0
50
100
150
OS month
P=0.0007
C
FN1
ITGA3
low
high
Survival probability
1.00
0.75
0.50
0.25
0.00
0
50
100
150
OS month
P<0.0001

Figure 7

Highlights

- Performed a shotgun proteomic analysis using dissected stromal areas of PDAC tissues.
- Identified differentially expressed (DE) proteins in PDAC stroma.
- Identified crucial and potentially druggable cancer-stromal interactions in PDAC.
- Identified 9 key genes and 8 key cancer-stromal-interaction targets for PDAC patients.
- FN1-ITGA3 and FN1-ITGA5 have unfavorable prognostic impact for PDAC patients.

Conflict of interest statement

The authors declare that they have no competing interests.