Computational and Structural Biotechnology Journal 19 (2021) 6375–6385

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

The unique pancreatic stellate cell gene expression signatures are associated with the progression from acute to chronic pancreatitis

Cheng Hu [a,1], Liyuan Yin [b,1], Zhiyao Chen [a], Richard T. Waldron [c,d], Aurelia Lugea [c,d], Yiyun Lin [e], Xiaoqian Zhai [b], Li Wen [f], Yuan-Ping Han [g], Stephen J. Pandol [c,d], Lihui Deng [a,*], Qing Xia [a,*]

[a] Department and Laboratory of Integrated Traditional Chinese and Western Medicine, Sichuan Provincial Pancreatitis Center and West China-Liverpool Biomedical Research Center, West China Hospital, Sichuan University, Chengdu, China
[b] Lung Cancer Center, West China Hospital, Sichuan University, Chengdu, China
[c] Department of Medicine and Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, United States
[d] Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States
[e] Graduate School of Biomedical Sciences, UT MD Anderson Cancer Center, University of Texas, Houston, TX, United States
[f] Department of Gastroenterology and Shanghai Key Laboratory of Pancreatic Disease, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[g] Center for Growth, Metabolism and Aging, College of Life Sciences, Sichuan University, Chengdu, China

ARTICLE INFO

Article history:
Received 15 July 2021
Received in revised form 16 November 2021
Accepted 20 November 2021
Available online 26 November 2021

Keywords:
Pancreatic stellate cell
Acute pancreatitis
Recurrent acute pancreatitis
Chronic pancreatitis
SPARC

ABSTRACT

Chronic pancreatitis (CP) is characterized by irreversible fibro-inflammatory changes induced by pancreatic stellate cell (PSC). Unresolved or recurrent injury causes dysregulation of biological process following AP, which would cause CP. Here, we systematically identify genes whose expressions are unique to PSC by comparing transcriptome profiles among total pancreas, pancreatic stellate, acinar, islet and immune cells. We then identified candidate genes and correlated them with the pancreatic disease continuum by performing intersection analysis among total PSC and activated PSC genes, and genes persistently differentially expressed during acute pancreatitis (AP) recovery. Last, we examined the association between candidate genes and AP, and substantiated their potential as biomarkers in experimental AP and recurrent AP (RAP) models. A total of 68 genes were identified as highly and uniquely expressed in PSC. The PSC signatures were highly enriched with extracellular matrix remodeling genes and were significantly enriched in AP pancreas compared to healthy control tissues. Among PSC signature genes that comprised a fibrotic phenotype, 10 were persistently differentially expressed during AP recovery. SPARC was determined as a candidate marker for the pancreatic disease continuum, which was not only persistently differentially expressed even five days after AP injury, but also highly expressed in two clinical datasets of CP. Sparc was also validated as highly elevated in RAP compared to AP mice. This work highlights the unique transcriptional profiles of PSC. These PSC signatures' expression may help to identify patients with high risk of AP progression to CP.

1. Introduction

Chronic pancreatitis (CP) is characterized by irreversible damage of pancreatic parenchyma and ductal structures with extensive replacement with fibrosis. Currently, there is no curative therapy for CP [1,2]. Once CP is established, the pooled relative risk estimate for pancreatic cancer among patients is 13.3 [3]. Accurate and early diagnosis of CP are urgently needed. Increasingly, it is being established that a continuum exists among pancreatic disorders, including acute pancreatitis (AP), recurrent acute pancreatitis (RAP), CP and pancreatic adenocarcinoma (PDAC) [4]. Patients with an episode of AP have a 20–30% likelihood of one or more recurrent episodes, with progression to CP in an estimated 10% of the recurrent cases [5–8]. Theoretically, revealing the mechanism by which AP-RAP progresses to CP may enable us to alter the disease course or prevent disease progression.

In both CP and PDAC, pancreatic stellate cell (PSC), the major producer of extracellular matrix, drives the extensive stromal reaction and irreversible pathophysiologic transformations. Whereas PSC has been proven to be activated during AP [9], little is known about the termination of PSC activation following AP, and its role in

* Corresponding authors.
E-mail addresses: denglihui@scu.edu.cn (L. Deng), xiaqing@medmail.com.cn (Q. Xia).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.csbj.2021.11.031

the progression of the continuum of pancreatic disorders. Based on the "Necrosis-Fibrosis Sequence" theory [10,11], which indicating that repeated attacks of AP with necroinflammation result in irreversible acinar atrophy and fibrosis, we hypothesize that PSC is primed to be more sensitive and persistently activated in response to recurrent stimuli, which might be the major connecting events during pancreatic disorders. The aim of this study is to identify the genes unique to PSC, investigate their correlation to pancreatic disorders, and substantiate their role in the progression of the disease continuum.

2. Materials and methods

2.1. Identification of the PSC gene signatures

To identify the PSC gene expression signatures, 7 platform-matched transcriptome datasets covering 12 pancreatic cell and tissue types were obtained from National Center for Biotechnology Information Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo) (see Table 1). The screening conditions were as follows: 1) cells were from pancreas; 2) the data were from same platform; 3) the tissue source should be in the same. Raw scan data were obtained and converted to normalized data using background corrected and median scale normalized. Probes' annotations were downloaded from relative platforms in GEO. Multiple probes corresponding to a gene were collapsed into a median value and subsequently labelled with an official gene symbol provided by the NCBI Entrez Gene database. The PSC gene signatures were identified as the genes with median expression in either quiescent or activated stellate cells that were at least 1.5 times as high as the highest expression in all other cell and tissue types, $P < 0.05$ and adjust $P < 0.05$. Mouse genes were mapped to orthologous human genes based on the NCBI HomoloGene database, and mouse genes with no known human homologous relationships were excluded from subsequent analysis. The R package "ComplexHeatmap" [12] package was an efficient method to visualize the association between different aspects of datasets. It was used to visualize the expression level between different types of cells from the 7 datasets. PCA (principal component analysis) was conducted to evaluate the inter-sample relationships.

Table 1
Datasets Used to Derive the Pancreatic Stellate Cell Signatures.

Category	Description	Sample accession
Stellate cell	Primary pancreatic fibroblasts	GSM1511497, GSM1511498, GSM1511499
Acinar cell	Freshly isolated pancreatic exocrine acini from C57BL/6 mice	GSM1717936, GSM1717937, GSM1717938, GSM1717939
Ductal cell	Pancreatic ductal organoids from C57BL/6 mice	GSM1717940, GSM1717941, GSM1717942, GSM1717943
B cell	B220(+) B lymphocytes from C57BL/6 mice	GSM686644, GSM686645
T cell	T lymphocytes from C57BL/6 mice	GSM686646, GSM686647, GSM686648, GSM686649
NK	NK1.1 + NK cells from C57BL/6 mice	GSM686650, GSM686651
NKT	CD3ε(+)NK1.1(+) cells from C57BL/6 mice	GSM686652, GSM686653
Erythroblasts	Ter119(+) erythroblasts from C57BL/6 mice	GSM686654, GSM686655
Neutrophil	Gr-1(+) neutrophils from C57BL/6 mice	GSM686656, GSM686657
Macrophage	Mac-1(+) monocytes/macrophages from C57BL/6 mice	GSM686658, GSM686659
Islet	Freshly isolated islets from C57BL/6 mice	GSM2233120, GSM2233121, GSM2233122, GSM2233123, GSM2233124, GSM1067240, GSM1067241
Whole pancreas	Pancreas from C57BL/6 mice	GSM2064123, GSM2064124, GSM2064125, GSM2064126, GSM2064127, GSM2064128, GSM1298160, GSM1298161, GSM1298162, GSM1298163, GSM1298164

2.2. Molecular pathway analysis

Gene ontology terms of the PSC signatures were performed by GO (geneontology.org) analysis. The R package "ClusterProfile" was used to analyze the significant GO terms. The enriched items were ranked according to *P* values. Top 20 terms were visualized by bar plot.

Gene Set Enrichment Analysis (GSEA) [13] was used to define whether a defined set of genes showed statistically significant between two phenotypes. Firstly the PSC gene signatures were defined as the gene list. And then "ClusterProfile" package was used to calculate whether the disease group and normal group could be defined by this gene list. Normalized enrichment score and *P* value were calculated. At the same time, Gene Set Variation Analysis (GSVA) [14] was used to evaluate the GSVA score of each PSC signature gene in each sample. GSVA package was used.

2.3. Identification of activated PSC gene signatures

Focused on the activated PSC function, we selected the activated PSC expressed genes as a subset of PSC gene signatures. The pancreatic un-activated and activated mouse PSC gene sequence data were obtained through GEO database (GSE43770) [15]. The counts values were then transformed to FPKM values for next analysis. Gene annotation was performed with the Ensembl database. The gene expression values were log2-transformed for subsequent analysis. Genes upregulated in activated PSC were selected by the criterial fold change >=1.5. The intersection between PSC gene signature and upregulated differentially expressed genes were conducted and considered as activated PSC gene signature.

2.4. Identification of the genes associated to pancreatic disease progression

GSE99774 data [16] was downloaded from GEO database. After normalization, FPKM value transformed and gene annotation, we used the correlation.R package and Pearson's test to perform a correlation analysis between gene expression and tissue extraction time (days). *P* values, the false discovery rate (FDR) and the correlation index were obtained. Genes with correlation index ≥ 0.4 and *P* value < 0.05 were considered significantly AP recovery related genes. The intersection between PSC gene signatures, activated PSC signatures and recovery related genes were identified as pancreatic disease progression related genes.

2.5. Analysis of association between PSC signatures and overall survival

The TCGA-PAAD RNA sequencing data used for survival analysis were obtained by using the R package "RTCGAToolbox" [17]. The TCGA-PAAD data was used for survival analysis by reducing 6 other pancreatic cancer, namely 4 pancreas colloid (mucinous non-cystic) carcinoma, 1 undifferentiated carcinoma, 1NA. We evaluated the association between the overall survival (OS) and the activated PSC signatures, as well as between the OS and progression related genes. Hierarchical clustering of the expression of the genes divided the sample into two groups, and this division

was used as a categorical variable to perform survival analysis. Kaplan-Meier survival analysis and the log-rank test were used to evaluate the independence of the prognostic value of the gene signatures. The variables were tested by COX regression analysis, including year, gender, TNM (tumor–node–metastasis) stage and activated PSC gene signature. The variable which was significant was used for multi-variable COX analysis. $P < 0.05$ was considered significant for all these analyses.

2.6. Experimental model and histopathology assessment

C57BL/6 strain mice received 7 hourly intra-peritoneal injections of cerulein (50 μg/kg) to induce AP. Ten days after modeling, another episode of AP injection was given to induce RAP model. Animals were sacrificed at 12 h, 24 h, 4 days, 7 days and 10 days after the first injection of this repeat episode. All methods and animal experiments were performed in accordance with guidelines approved by Sichuan University Animal Ethics Committee. All experiments were performed in accordance with relevant guidelines and regulations.

Samples of pancreatic tissue for histopathology assessment were fixed in 10% formalin at room temperature for 24 h before embedded in paraffin. Sections were stained with hematoxylin and eosin (H&E). Ten fields were systematically selected at 200 × magnification and graded in a blinded fashion. The subjective grading score gave equal weight (from 0 to 4) for edema, inflammatory infiltration, and necrosis, as previously described [18]. Then were combined as summary score. Quantification of acinar-to-ductal metaplasia (ADM) area (%) in pancreas tissue was performed at 100 × magnification by two blinded independent investigators.

2.7. Quantitative reverse transcriptase PCR analysis

Total RNA was extracted from pancreatic tissue with TRIzol according to the manufacturer's protocol. Reverse transcription was performed with the iScript Reverse Transcription Supermix using 1 μg of total RNA, and the synthesized cDNA samples were used as templates for quantitative real-time PCR (qPCR) analysis. A CFX 96 RT-qPCR instrument (2000 Alfred Nobel Drive, Hercules, California 94547, USA) was applied to support the reaction run at one cycle of 95 °C for 5 min, followed by 40 cycles of 95 °C for 20 s, 60 °C for 20 s and 72 °C for 20 s. The gene-specific primers used are listed in Table 2. Relative transcript levels were calculated using the comparative $2-\Delta\Delta Ct$ method and normalized to the housekeeping gene, 18S rRNA.

2.8. Statistical analysis

R software (version 3.6.2) was used to perform all statistical analyses using the Student's *t* test. Pearson's correlation coefficients were used to assess the relationship between continuous variables. *T* test and One-way ANOVA was was performed using Graphpad Prism version 9. To conduct a survival analysis, the Kaplan-Meier approach was used, with the subsequent findings compared using the log-rank test (*P-value < 0.05; **P-value < 0.01; ***P-value < 0.005; ****P-value < 0.0001).

Table 2
Primer Sequences for qRT-PCR.

Target	Forward (5′ to 3′)	Reverse (5′ to 3′)
Sparc	CCACACGTTTCTTTGAGACC	GATGTCCTGCTCCTTGATGC
Acta2	GTTCAGTGGTGCCTCTGTCA	ACTGGGACGACATGGAAAAG
Collangen I	TAGGCCATTGTGTATGCAGC	ACATGTTCAGCTTTGTGGACC
Fibronectin	CCTTACACGGTTTCCCATTA	TTGTCATGGCACCATTTAGA
18 s rRNA	AGTCCCTGCCCTTTGTACACA	CGATCCGAGGGCCTCACTA

3. Results

3.1. Development of the PSC signatures

The study was conducted according to the flow chart shown in Fig. 1A. Details of the GEO datasets used to derive the PSC signatures are shown in Table 1. To identify the genes highly expressed in PSC, expression profiles representing all pancreas and pancreatic cell subsets were subtracted from mouse-based transcriptome profiles. In total, 45 samples (GEO samples, GSM) were derived from 7 datasets with same platform (GPL1261) covering 12 pancreatic cell and tissue types. The heatmap showed that the same kinds of cells expressed similar gene levels within each study or dataset, supporting the existence of robust transcriptome programs specific to each pancreas cell type (Fig. 1B). A total of 68 genes were identified as highly and uniquely expressed in PSC (Fig. 1C). PCA (principle component analysis) was also conducted by using the expression profile of all genes and PSC signature genes during 11 different cell types (supplementary Fig. 1A and B). A list of the PSC signature genes is provided in Supplementary Table 1. Several canonical PSC markers (including Collagen I, Acta2) were highly expressed in the PSC samples. We then performed functional annotation analyses by using DAVID Gene Ontology (GO) analyses. As shown, GO molecular function analysis revealed that the PSC signatures participated in extracellular matrix (ECM) structural constituent, collagen binding, ECM binding and other molecular functions important for cell motility, cell adhesion and formation of a fibrotic tissue microenvironment (Fig. 1D). Biological process showed that genes were highly enriched for ECM-related functions (Fig. 1E). Cellular components of the identified PSC genes were shown in Supplementary Fig. 1C.

3.2. Upregulation of the PSC signatures in AP

GEO datasets used for subset analysis in this study are shown in Supplementary Table 2. To validate the PSC signature genes for their relevance to CP and PDAC, we tested for PSC signatures enrichment in datasets (GSE77858, GSE123375) that examined the transcriptional profile in CP and PDAC patients. As expected, GSEA analysis showed PSC signatures were enriched in CP and PDAC. GSVA analysis showed that PSC signature genes' expressions were significantly induced in CP and PDAC, compared with healthy normal pancreas (Fig. 2A-D). To explore the relevance of PSC gene signatures to AP phenotypes, we used GSEA and GSVA to test the PSC signatures and disease relevance in two AP datasets (GSE65146, GSE109227 [19]). Our PSC gene expression signatures were strikingly enriched in AP, compared with healthy normal pancreas (Fig. 2E, F).

Since activated PSC is responsible for producing ECM components during pancreatic diseases, we identified differentially expressed genes between the activated and un-activated PSC. A comparison of the transcriptomes of un-activated (3-day culture) and activated (7-day culture) PSC was performed by Sherman et al. [15] We used this dataset (GSE43770) and found that there were 2462 genes upregulated in the activated PSC with the fold change cutoff at 1.5. When we intersected the PSC signatures with activated PSC gene ontologies, twenty-six genes were recognized (Fig. 3A). GO analysis of the activated and un-activated PSC genes among our PSC gene signature outlined the ones for corresponding biological functions (Fig. 3B). As expected, our PSC signatures with fibrotic phenotype were significantly enriched with functions such as ECM organization, extracellular organization, and TGFβ related

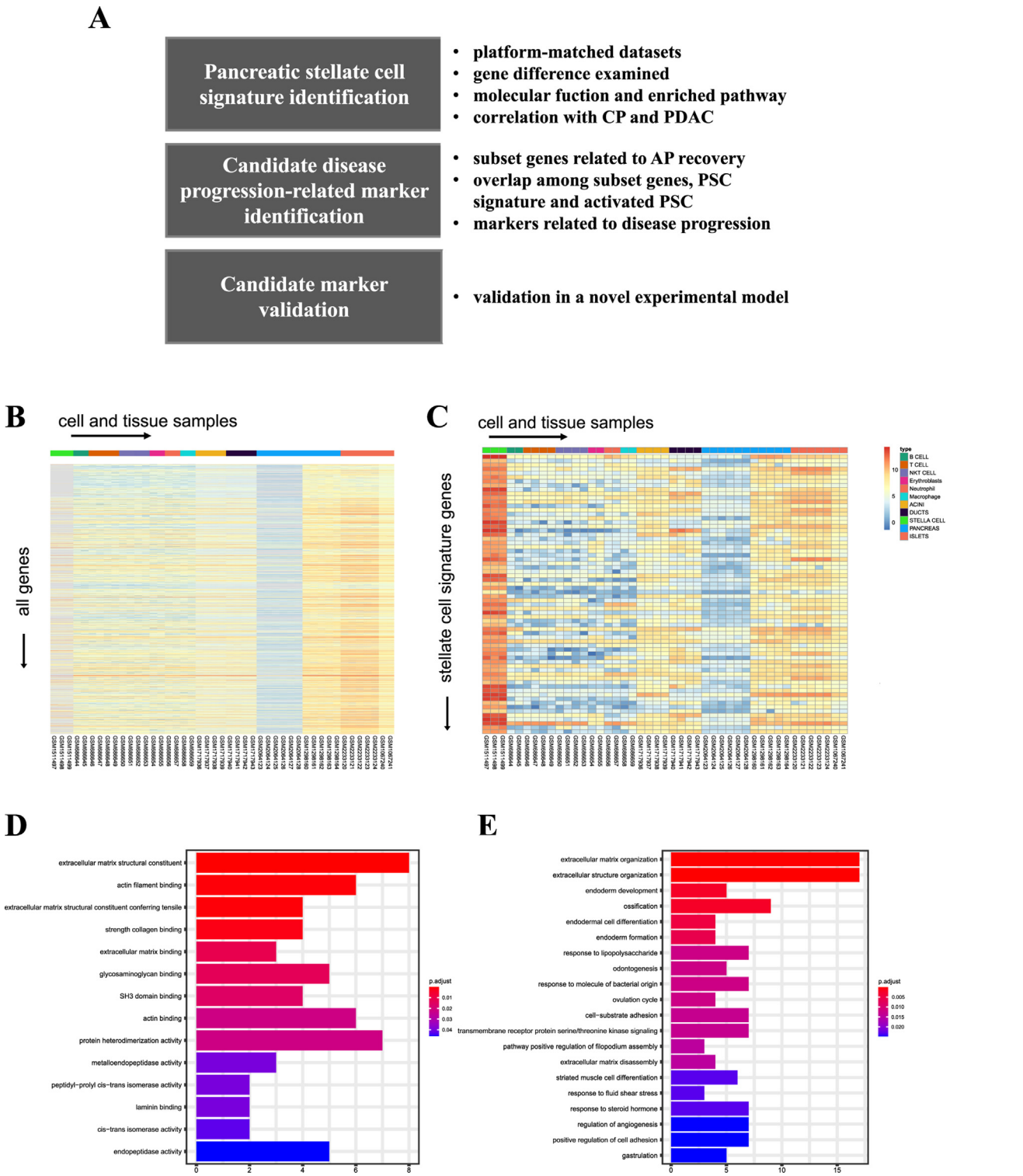

Fig. 1. Identification of genes highly enriched in PSC. (A) Flow chart of study design. (B) Expression of all genes examined, across all tissue and cell types examined. The heat map illustrating degree of similarity. Highly expressed genes were presented for each cell and tissue type, but there was considerable expression overlap between different cells and tissue types. (C) Expression of genes in the PSC signature. Only genes that were uniquely and highly expressed in PSC were included in the signature. The Heatmap showed the expression pattern among different cells. (D) GO molecular function Terms Enriched in the PSC Signatures. (E) GO showed ontologies describing biological process where there was strong enrichment of ECM process.

pathway. In contrast, in genes constituting our signature with unactivated phenotype of PSC, limited GO terms were enriched. Thus, we focused on these 26 activated PSC signature genes to explore their functions in pancreatic disorders. Indeed, 21 out of 26 genes had strong correlation with AP in the GSE65146 dataset (Fig. 3C). Then, we explored the prognostic importance of these activated PSC signature genes by utilizing TCGA-PAAD data. We divided PDAC patients according to high and low activated PSC enrichment. As shown, patients with high PSC signature expression had significantly worse survival (P = 0.0046) (Fig. 3D, E).

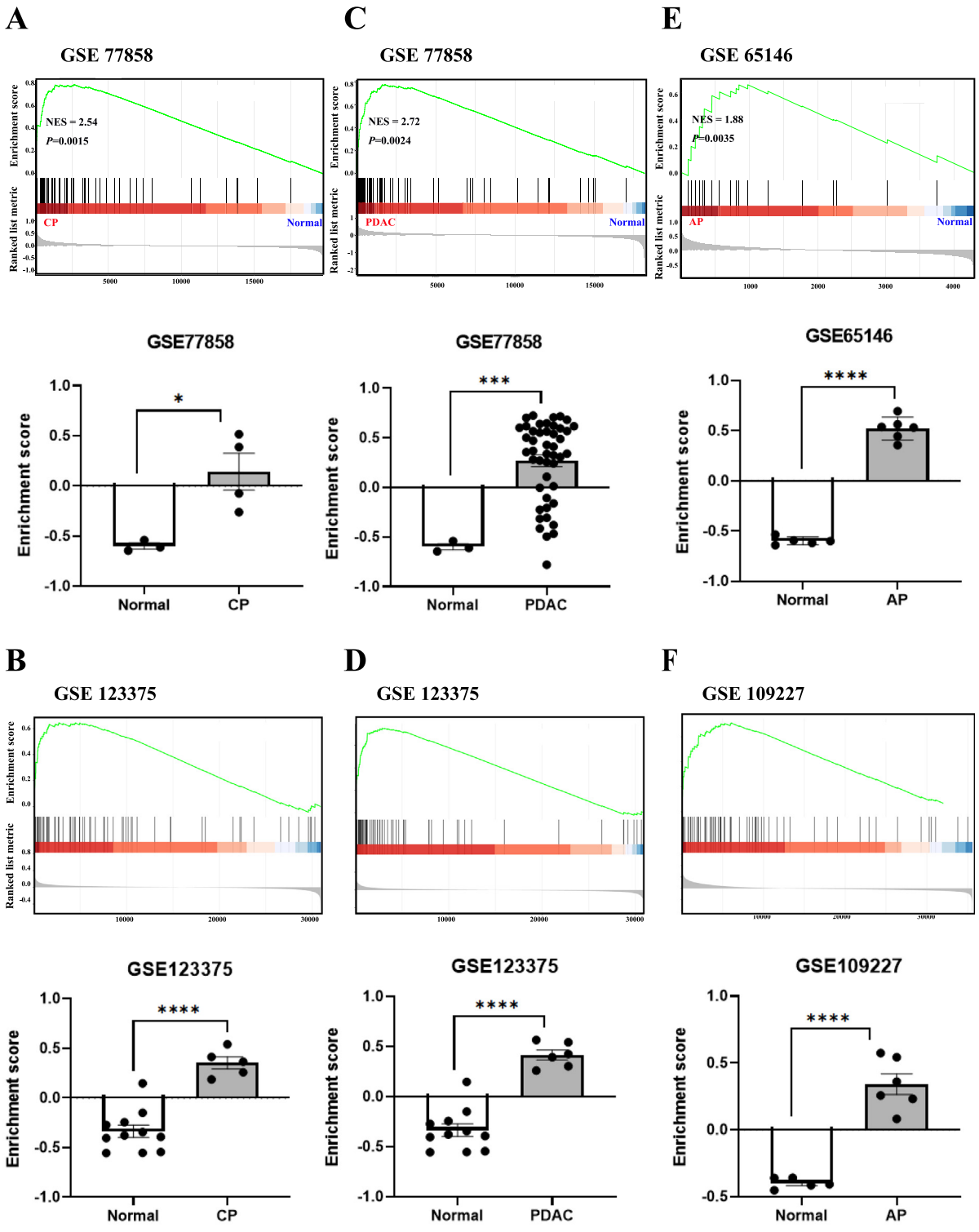

Fig. 2. The PSC signatures are enriched in pancreatic disease. (A-D) GSEA and GSVA of the PSC signature genes were confirmed in CP and PDAC cohorts. Array genes are ordered from the highest in diseased pancreas (left side) to the highest in normal pancreas (right side). Vertical black bars indicate the location of each gene in the PSC signatures. The normalized enrichment score and statistical significance were included. The PSC signature genes' expression were assessed for each sample and assigned numeric value by GSVA method. Mean enrichment scores and SEM were shown for each group. (E,F) GSEA and GSVA of the PSC signatures were confirmed in AP cohorts. Two different GEO datasets were analyzed. The PSC gene signatures were enriched in AP.

3.3. A PSC marker related to pancreatic disease progression

To validate our hypothesis that activated PSC might be sensitive and unresolved (persistently activated) after a first attack of AP, and therefore potentially responsible for progression of AP to CP, we discovered progression/unresolved disease-related genes during AP recovery from the published dataset (GSE99774), which provided mouse pancreatic transcriptomic profile at baseline (non-injured), day 7 (post AP injury), and day 14 (post AP injury). A correlation analysis was conducted between gene expression

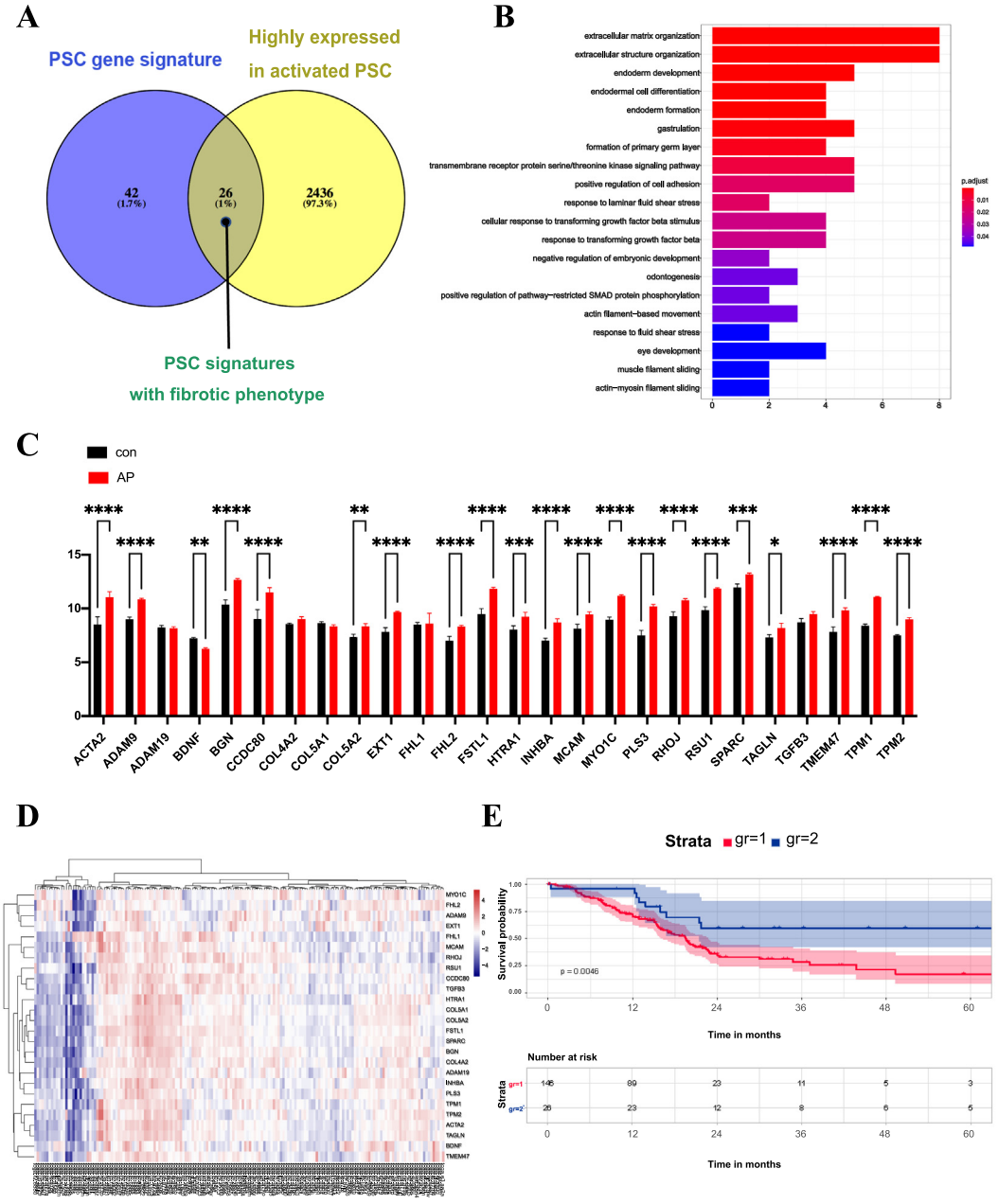

Fig. 3. The PSC signatures with fibrotic phenotype have strong correlation with AP. (A) Overlap between PSC signatures and activated PSC gene ontologies. (B) Gene Ontology Terms Enriched in the PSC Signatures with fibrotic phenotype. (C) Elevated genes expression of PSC signature with fibrotic phenotype were confirmed in previously reported AP array data. (D) Heatmaps of PDAC patients from TCGA PSC signature expression. Patients were divided to higher PSC enrichment and lower PSC signature enrichment. Each column was one patient and each row is one PSC gene. (E) There was robust separation between each of the risk groups for survival. Patients with high PSC signature expression had significantly worse survival.

levels and time (days after AP). Positive related genes according to Pearson correlation index $\geq$ 0.4 and $P < 0.05$ were identified as recovery related genes. In total, 2793 recovery related genes were identified. A Venn diagram was performed to intersect these recovery related genes with PSC signatures with the activated PSC phenotype (Fig. 4A). Ten genes were recognized, including ADAM9, COL4A2, COL5A1, COL5A2, EXT1, FSTL1, MCAM, RHOJ, SPARC and TPM1. Next, we examined the expression pattern of these genes during AP recovery, CP and PDAC. As shown in Fig. 4B-D, most of the genes were overexpressed during AP, CP and PDAC compared to the normal control. We also performed survival analysis for correlation of the selected 10-gene with PDAC patients. Patients with high expression of genes of this set had significantly worse survival ($P = 0.016$) (Fig. 4E, F). This approach highlighted Secreted Protein Acidic and Rich in Cysteine (SPARC), an extracellular glycoprotein secreted by fibroblasts, as a candidate marker for the disease progression, which manifested as persistently differentially expressed even five days (120 h) after AP injury. Furthermore, SPARC was uniquely and highly related with CP and PDAC in two independent clinical datasets (GSE143754 [20], GSE123375 [21]).

3.4. Validation of SPARC and canonical PSC markers' expression in AP/RAP mice model

To validate our findings that PSC contributes to the progression of the continuum of pancreatic disorders, we established a preclin-

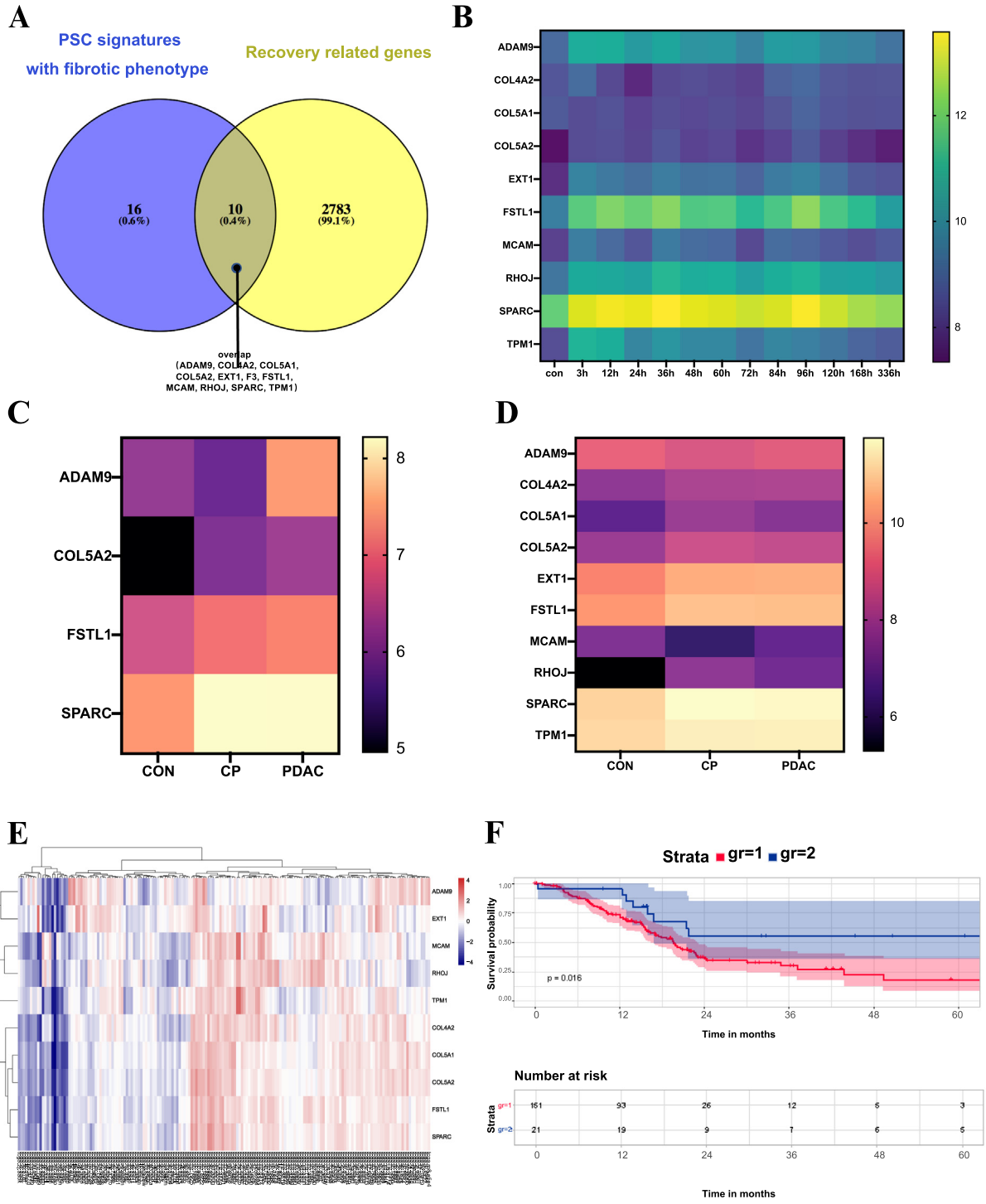

Fig. 4. SPARC is a candidate PSC marker related to pancreatic disease progression. (A) Overlap between PSC signatures with fibrotic phenotype, and recovery related genes. (B) Of the ten markers, we tested gene expression pattern during AP recovery dataset, confirming a gradually increasing expression during time after AP occurrence. (C, D) Of the ten markers, we tested gene expression pattern in CP and PDAC, showed an increase pattern in CP and PDAC compared to normal pancreas. (E) Heatmaps of PDAC patients from TCGA with selected 10-gene set. (F) There was robust separation between each of the risk groups for survival. The survival analysis showed patients with high expression of genes of this set have significantly worse survival.

ical model for AP and RAP as outlined in Fig. 5A by repeating induction of cerulein AP on day 10 after a first attack. These models were based on the general idea that murine pancreas was completely recovered by 7 days post injury on a histological level [16]. We found more severe pancreatic edema, necrosis and inflammatory infiltration at the acute phase of RAP in comparison with AP group. Furthermore, long term observation indicated an unresolving pathology at the recovery phase of RAP, while selective formation of morphologic features, resembles acinar-to-ductal metaplasia (ADM) in this group of mice (Fig. 5B).

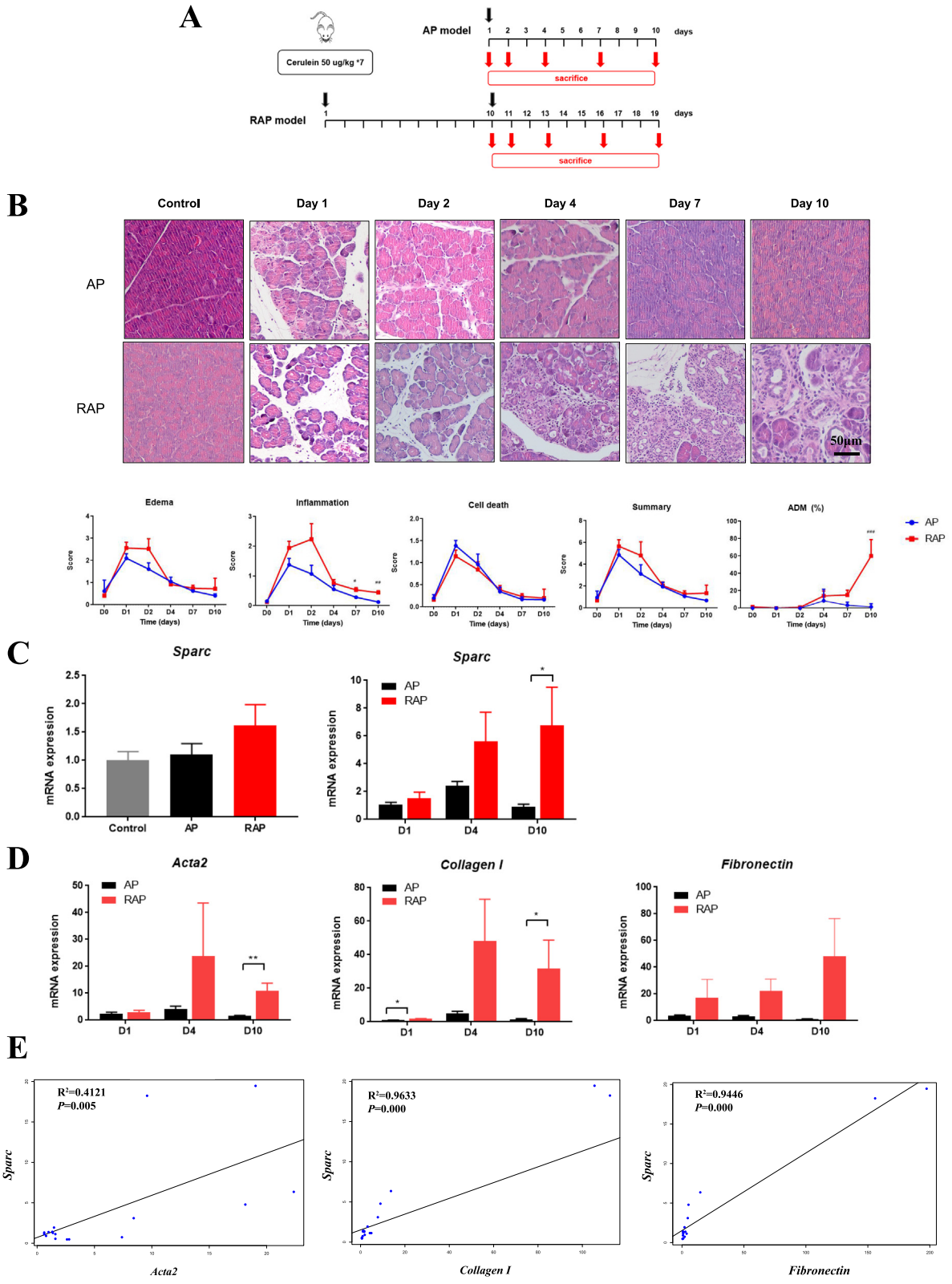

Fig. 5. Increased SPARC and canonical PSC markers during AP/RAP recovery phase. (A) The schematic representation of AP and RAP models. C57BL/6 mice were intraperitoneally injected with 50 μg/kg cerulein (AP) or saline as control (D0) hourly for 7 times. Ten days after modeling, another episode of AP was given to induce RAP model. As indicated by black arrows. Mice were sacrificed on Day 1, 2, 4, 7, 10 post AP induction (red arrows). (B) Representative H&E pancreas images with histopathology scores (edema, inflammatory cell infiltration, acinar cell necrosis, their sum value, and ADM). Scale bar, 50 μm. (C) Expression of Sparc mRNA was measured in mice pancreatic tissues. (D) Genes related to PSC activation and fibrotic phenotype were measured in mice pancreatic tissues. (E) The expression of Sparc was positively correlated with canonical PSC genes, namely Acta2, Collagen I, Fibronectin with R^2 = 0.41, 0.96, 0.94 respectively. Datas were expressed as means ± SEM from 2 independent experiments (n = 8–12 for each group). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ for RAP versus AP group at each time point. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Activated PSC specifically express α-SMA and secrete collagen I, collagen III, fibronectin, and other ECM components to promote the formation of pancreatic fibrosis. We next performed gene expression analysis of Sparc, and fibrotic phenotype markers (Acta2, Collagen I, Fibronectin) on this RAP model in comparison with AP model. As shown, these genes were all significantly increased in RAP than those in the AP group, particularly at the recovery phases (Fig. 5C, D). The expression of Sparc and the Acta2, Collagen I, Fibronectin showed strong correlations after 10 days recovery, respectively (R^2 = 0.4121, 0.9633, 0.9446; P = 0.005, 0.000, 0.000) (Fig. 5E, Supplementary Fig. 2). Correlation between Sparc and canonical PSC markers on day 1 and day 4 were shown in Supplementary Fig. 3.

4. Discussion

Patients with CP suffer from chronic pain, exocrine and endocrine pancreatic insufficiency, reduced quality of life, and have a shorter life expectancy [22–25], with increased risk of developing of PDAC [3]. However, early recognition of CP is still lacking. Growing evidence confirms that AP is a 'sentinel event' of the subsequent pancreatic disorders, in which repetitive inflammation may subsequently lead to the development of CP [10,26–28]. These findings present an opportunity for early detection of CP, as well as prevention of AP-RAP progression to devastating pancreatic disorders. Several studies focused on the role of immune mediators. Park *et al.* [29] found four cytokines (Resistin, SCF, MIP-1a, and IL-17F) that increasingly expressed in CP patients compared to RAP patients which might be potential immune signatures for CP. Wu *et al.* [30] illustrated a dynamic phenotype and regulatory mechanism of macrophages during AP repair/regeneration, which provided macrophage-based therapies to treat/prevent the related disorders post AP injury.

It is established that activated, myofibroblast-like PSC modulate disease states and is major contributor to the extensive pancreatic desmoplasia and fibrosis observed in CP and PDAC [31]. Persistent activation of PSC leads to an imbalance between ECM synthesis and degradation, eventually resulting in pancreatic fibrosis, a cardinal feature of CP [32–33]. During tumorigenesis, PSC could create a favorable tumor microenvironment and facilitate cancer progression by increasing proliferation, invasiveness and inducing treatment resistance of pancreatic cancer cells [34–36]. In PDAC, the presence of PSC correlates with poor prognosis [37,38]. It has also been shown in experimental AP models, that there may be PSC activation and early fibrosis even after the first episode of AP [39]. In view of these concepts, we hypothesized that, PSC might be primed to be more sensitive and persistently activated after first episode of AP. Recurrent stimuli might make PSC reach a certain threshold after which the process becomes irreversible. We provide here for the first time, a bioinformatics analysis to identify a group of genes that are uniquely expressed in PSC. Furthermore, we show that these genes correlate with the progression from AP/RAP to CP.

The transcriptome profiling datasets of whole pancreas, acinar cells, islets, PSC, and immune cells were derived from a series of mouse-based studies. As we expected, the genes in the PSC signature were highly enriched with ECM related functions and pathways. Since a high level of PSC activity (dependent on the expression of αSMA, Collagen I, Fibronectin) is responsible for producing the stromal fibrosis during pancreatic diseases, we identified a set of signature genes related to activated PSC, and explored whether genes in the PSC signature with activated phenotype would have correlation with AP. Interestingly, we found that most of the genes in activated PSC signatures are strongly correlated with AP. The highly activated PSC signature genes were also associated with reduced survival in PDAC cohort. Multivariate Cox (Supplementary Table 3) proportional hazards regression analysis validated the activated PSC signature gens as an independent prognostic factor.

We further explored the termination of PSC activation following AP, and its role in the progression of the continuum of pancreatic disorders. It has been reported by profiling the transcriptome of the recovering pancreas that although there is pathological resolution one week after cerulein induced AP in mice, there are still numerous differentially genes expressed one and even two weeks post injury [16]. We discovered recovery related genes in AP and intersected these genes with PSC signature with activated phenotype. Surprisingly, we identified SPARC as a candidate marker, showing persistent and differential pattern of expression during AP recovery. In addition, SPARC was highly related with CP and PDAC in two dependent clinical datasets.

SPARC is an extracellular glycoprotein involved in ECM assembly and cell-matrix communication during tissue remodeling, embryonic development, and tumor progression [40]. It has been reported to bind numbers of ECM components including collagenase and fibronectin with good affinity [41]. Upregulated expression of SPARC is often associated with the stages of fibrogenesis in several diseases [42–45]. SPARC downregulation with adenovirus or gene knock-out ameliorates liver fibrosis [45,46]. Studies on hepatic stellate cell (HSC) reveals that SPARC expression dramatically increases in activated HSC, and exerts its profibrotic effects partially through mediating profibrogenic activities of TGFβ and PDGF-BB [43,47]. PSC and HSC have similar morphological and functional features. Studies on PSC reveals that SPARC expression in PSC is necessary for ECM remodeling [48]. SPARC upregulation in PSC is governed by YAP, which is crucial in driving and sustaining the activation of PSC [49]. In both rat and human CP, SPARC gene expression is massively upregulated in pancreatic tissue [50,51]. Its high expression in peritumoral fibroblasts has been reported associated with a poor prognosis in PDAC patients [52]. Moreover, SPARC is proposed as a negative predictive factor for the treatment with gemcitabine [53]. Although SPARC has been investigated in CP and PDAC, little is known about its expression during AP/RAP, as well as its correlation with pancreatic disease progression. Here, we provided a preclinical model of AP/RAP, and validated our hypothesis on animal models. We found that our candidate gene (Sparc), PSC activation (Acta2) and fibrotic phenotype markers (Collagen I, Fibronectin) were dramatically higher increased in RAP than AP group during recovery phases. These results substantiate our method for identifying PSC-related disease progression markers, making Sparc an interesting molecule not only as a potential target for therapy, but also as a possible marker of pancreatic disease progression.

There are clear limitations to this study, including lack of functional insights into how PSC influence the transition of AP/RAP to CP. This should be addressed by quantitative estimation and genetic modification in future studies. Second, this pilot study mostly relied on reported dataset is lack of clinical or pre-clinical sampling to assess temporal progression from AP/RAP to CP.

5. Conclusions

Our work establishes a novel and clinically relevant platform for discovery of PSC genes, and provide evidence that PSC plays a role in the progression from AP to CP. This creates a unique opportunity to systematically uncover fibrosis biology, identify candidate drug targets and define candidate biomarkers that monitor disease progression from AP to CP.

6. Ethics statement

Animal studies were performed in accordance with guidelines approved by Sichuan University Animal Ethics Committee.

Author contributions

CH, LY, AL, SJP, LD and QX participated in the design and coordination of the study; CH, LY, ZC, YL and XZ performed experiments, data acquisition, analysis and interpretations; CH, LY, LD and QX participated in study concept; CH, LY, RTW, AL, LW, YH, SJP and LD participated in manuscript writing. LD, QX, AL and SJP reviewed and approved the final manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by the National Science Foundation of China (No. 82104715, CH; No. 82103413, LY; No. 82074230, LD); Major Difficult Diseases and Science and Technology Key Applied Basic Research Projects of Sichuan Province (No. 2020YJ0235, CH); the China Postdoctoral Science Foundation (No. 2019M663520, CH); the Post-Doctor Research Project, West China Hospital, Sichuan University (No. 2020HXBH131, LY; 2019HXBH026, CH).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully thank research assistants (Mingjie Xu, Xiao Xiao) from West China Hospital for kindly assistance. These authors also thank West China Hospital owned public research platforms including Experimental Animal Center (Xiaoting Chen), Laboratory of Pathology (Li Li, Chunjuan Bao, and Fei Chen).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.11.031.

References

[1] Beyer G, Habtezion A, Werner J, Lerch MM and Mayerle J. Chronic pancreatitis. *Lancet* 2020; 396: 499-512. 2020/08/18. DOI: 10.1016/S0140-6736(20)31318-0.
[2] Gardner TB, Adler DG, Forsmark CE, Sauer BG, Taylor JR and Whitcomb DC. ACG Clinical Guideline: Chronic Pancreatitis. *Am J Gastroenterol* 2020; 115: 322-339. 2020/02/06. DOI: 10.14309/ajg.0000000000000535.
[3] Raimondi S, Lowenfels AB, Morselli-Labate AM, Maisonneuve P and Pezzilli R. Pancreatic cancer in chronic pancreatitis; aetiology, incidence, and early detection. *Best Pract Res Clin Gastroenterol* 2010; 24: 349-358. 2010/06/01. DOI: 10.1016/j.bpg.2010.02.007.
[4] Yadav D and Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology* 2013; 144: 1252-1261. 2013/04/30. DOI: 10.1053/j.gastro.2013.01.068.
[5] Takeyama Y. Long-term prognosis of acute pancreatitis in Japan. *Clin Gastroenterol Hepatol* 2009; 7: S15-17. 2009/12/05. DOI: 10.1016/j.cgh.2009.08.022.
[6] Lankisch PG, Breuer N, Bruns A, Weber-Dany B, Lowenfels AB and Maisonneuve P. Natural history of acute pancreatitis: a long-term population-based study. *Am J Gastroenterol* 2009; 104: 2797-2805; quiz 2806. 2009/07/16. DOI: 10.1038/ajg.2009.405.
[7] Nojgaard C, Becker U, Matzen P, Andersen JR, Holst C and Bendtsen F. Progression from acute to chronic pancreatitis: prognostic factors, mortality, and natural course. *Pancreas* 2011; 40: 1195-1200. 2011/09/20. DOI: 10.1097/MPA.0b013e318221f569.
[8] Yadav D, O'Connell M and Papachristou GI. Natural history following the first attack of acute pancreatitis. *Am J Gastroenterol* 2012; 107: 1096-1103. 2012/05/23. DOI: 10.1038/ajg.2012.126.
[9] Petersen OH, Gerasimenko JV, Gerasimenko OV, Gryshchenko O, Peng S. The roles of calcium and ATP in the physiology and pathology of the exocrine pancreas. Physiol Rev 2021;101(4):1691–744. https://doi.org/10.1152/physrev.00003.2021.
[10] Witt H, Apte MV, Keim V and Wilson JS. Chronic pancreatitis: challenges and advances in pathogenesis, genetics, diagnosis, and therapy. *Gastroenterology* 2007; 132: 1557-1573. 2007/05/01. DOI: 10.1053/j.gastro.2007.03.001.
[11] Whitcomb DC. Genetic risk factors for pancreatic disorders. *Gastroenterology* 2013; 144: 1292-1302. 2013/04/30. DOI: 10.1053/j.gastro.2013.01.069.
[12] Gu Z, Eils R and Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics (Oxford, England)* 2016; 32: 2847-2849. 2016/05/22. DOI: 10.1093/bioinformatics/btw313.
[13] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, 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):15545–50. https://doi.org/10.1073/pnas.0506580102.
[14] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC bioinformatics 2013;14:7. https://doi.org/10.1186/1471-2105-14-7.
[15] Sherman MH, Yu RT, Engle DD, Ding N, Atkins AR, Tiriac H, et al. Vitamin D receptor-mediated stromal reprogramming suppresses pancreatitis and enhances pancreatic cancer therapy. *Cell* 2014; 159: 80-93. 2014/09/27. DOI: 10.1016/j.cell.2014.08.007.
[16] Boggs K, Wang T, Orabi AI, Mukherjee A, Eisses JF, Sun T, et al. Pancreatic gene expression during recovery after pancreatitis reveals unique transcriptome profiles. *Sci Rep* 2018; 8: 1406. 2018/01/25. DOI: 10.1038/s41598-018-19392-0.
[17] Samur MK. RTCGAToolbox: a new tool for exporting TCGA Firehose data. *PloS one* 2014; 9: e106397. 2014/09/03. DOI: 10.1371/journal.pone.0106397.
[18] Zhou X, Xue C. Ghrelin inhibits the development of acute pancreatitis and nuclear factor kappaB activation in pancreas and liver. Pancreas 2009;38:752–7. https://doi.org/10.1097/MPA.0b013e3181a86b74.
[19] Norberg KJ, Nania S, Li X, Gao H, Szatmary P, Segersvärd R, et al. RCAN1 is a marker of oxidative stress, induced in acute pancreatitis. *Pancreatology : official journal of the International Association of Pancreatology (IAP) [et al]* 2018; 18: 734-741. 2018/08/25. DOI: 10.1016/j.pan.2018.08.005.
[20] Chhatriya B, Mukherjee M, Ray S, Saha B, Lahiri S, Halder S, et al. Transcriptome analysis identifies putative multi-gene signature distinguishing benign and malignant pancreatic head mass. *J Transl Med* 2020; 18: 420. 2020/11/09. DOI: 10.1186/s12967-020-02597-1.
[21] Barrera LN, Evans A, Lane B, Brumskill S, Oldfield FE, Campbell F, et al. Fibroblasts from Distinct Pancreatic Pathologies Exhibit Disease-Specific Properties. *Cancer research* 2020; 80: 2861-2873. 2020/05/13. DOI: 10.1158/0008-5472.Can-19-3534.
[22] Drewes AM, Olesen AE, Farmer AD, Szigethy E, Rebours V and Olesen SS. Gastrointestinal pain. *Nat Rev Dis Primers* 2020; 6: 1. 2020/01/08. DOI: 10.1038/s41572-019-0135-7.
[23] Schuetz P, Fehr R, Baechli V, Geiser M, Deiss M, Gomes F, et al. Individualised nutritional support in medical inpatients at nutritional risk: a randomised clinical trial. Lancet 2019;393(10188):2312–21. https://doi.org/10.1016/S0140-6736(18)32776-4.
[24] Hart PA, Bellin MD, Andersen DK, Bradley D, Cruz-Monserrate Z, Forsmark CE, et al. Type 3c (pancreatogenic) diabetes mellitus secondary to chronic pancreatitis and pancreatic cancer. *Lancet Gastroenterol Hepatol* 2016; 1: 226-237. 2017/04/14. DOI: 10.1016/S2468-1253(16)30106-6.
[25] Zhu X, Liu D, Wei Q, Lin H, Zhi M, Chen Y, et al. New-Onset Diabetes Mellitus After Chronic Pancreatitis Diagnosis: A Systematic Review and Meta-analysis. *Pancreas* 2019; 48: 868-875. 2019/07/04. DOI: 10.1097/MPA.0000000000001359.
[26] Hegyi PJ, Soos A, Toth E, Ebert A, Venglovecz V, Marta K, et al. Evidence for diagnosis of early chronic pancreatitis after three episodes of acute pancreatitis: a cross-sectional multicentre international study with experimental animal model. *Sci Rep* 2021; 11: 1367. 2021/01/16. DOI: 10.1038/s41598-020-80532-6.
[27] Masamune A, Kikuta K, Nabeshima T, Nakano E, Hirota M, Kanno A, et al. Nationwide epidemiological survey of early chronic pancreatitis in Japan. *J Gastroenterol* 2017; 52: 992-1000. 2017/01/29. DOI: 10.1007/s00535-017-1311-8.
[28] Masamune A, Nabeshima T, Kikuta K, Hamada S, Nakano E, Kume K, et al. Prospective study of early chronic pancreatitis diagnosed based on the Japanese diagnostic criteria. *J Gastroenterol* 2019; 54: 928-935. 2019/07/05. DOI: 10.1007/s00535-019-01602-9.
[29] Park WG, Li L, Appana S, Wei W, Stello K, Andersen DK, et al. Unique circulating immune signatures for recurrent acute pancreatitis, chronic pancreatitis and pancreatic cancer: A pilot study of these conditions with and without diabetes. Pancreatology 2020;20(1):51–9. https://doi.org/10.1016/j.pan.2019.11.008.
[30] Wu J, Zhang Li, Shi J, He R, Yang W, Habtezion A, et al. Macrophage phenotypic switch orchestrates the inflammation and repair/regeneration following acute pancreatitis injury. EBioMedicine 2020;58:102920. https://doi.org/10.1016/j.ebiom.2020.102920.
[31] Apte MV, Wilson JS, Lugea A and Pandol SJ. A starring role for stellate cells in the pancreatic cancer microenvironment. *Gastroenterology* 2013; 144: 1210-1219. 2013/04/30. DOI: 10.1053/j.gastro.2012.11.037.

[32] Lee AT, Xu Z, Pothula SP, Patel MB, Pirola RC, Wilson JS, et al. Alcohol and cigarette smoke components activate human pancreatic stellate cells: implications for the progression of chronic pancreatitis. *Alcohol Clin Exp Res* 2015; 39: 2123-2133. 2015/10/16. DOI: 10.1111/acer.12882.
[33] Xue J, Zhao Q, Sharma V, Nguyen LP, Lee YN, Pham KL, et al. Aryl Hydrocarbon Receptor Ligands in Cigarette Smoke Induce Production of Interleukin-22 to Promote Pancreatic Fibrosis in Models of Chronic Pancreatitis. *Gastroenterology* 2016; 151: 1206-1217. 2016/11/05. DOI: 10.1053/j.gastro.2016.09.064.
[34] Sousa CM, Biancur DE, Wang X, Halbrook CJ, Sherman MH, Zhang L, et al. Pancreatic stellate cells support tumour metabolism through autophagic alanine secretion. *Nature* 2016; 536: 479-483. 2016/08/12. DOI: 10.1038/nature19084.
[35] Xu Z, Vonlaufen A, Phillips PA, Fiala-Beer E, Zhang X, Yang L, et al. Role of pancreatic stellate cells in pancreatic cancer metastasis. *Am J Pathol* 2010; 177: 2585-2596. 2010/10/12. DOI: 10.2353/ajpath.2010.090899.
[36] Fujita H, Ohuchida K, Mizumoto K, Egami T, Miyoshi K, Moriyama T, et al. Tumor-stromal interactions with direct cell contacts enhance proliferation of human pancreatic carcinoma cells. *Cancer Sci* 2009; 100: 2309-2317. 2009/09/09. DOI: 10.1111/j.1349-7006.2009.01317.x.
[37] Erkan M, Michalski CW, Rieder S, Reiser–Erkan C, Abiatari I, Kolb A, et al. The activated stroma index is a novel and independent prognostic marker in pancreatic ductal adenocarcinoma. Clin Gastroenterol H 2008;6(10):1155–61. https://doi.org/10.1016/j.cgh.2008.05.006.
[38] Fujita H, Ohuchida K, Mizumoto K, Nakata K, Yu J, Kayashima T, et al. alpha-Smooth Muscle Actin Expressing Stroma Promotes an Aggressive Tumor Biology in Pancreatic Ductal Adenocarcinoma. *Pancreas* 2010; 39: 1254-1262. 2010/05/15. DOI: 10.1097/MPA.0b013e3181dbf647.
[39] Elsasser HP, Adler G and Kern HF. Time course and cellular source of pancreatic regeneration following acute pancreatitis in the rat. *Pancreas* 1986; 1: 421-429. 1986/01/01. DOI: 10.1097/00006676-198609000-00006.
[40] Brekken RA and Sage EH. SPARC, a matricellular protein: at the crossroads of cell-matrix communication. *Matrix Biol* 2001; 19: 816-827. 2001/02/27. DOI: 10.1016/s0945-053x(00)00133-5.
[41] Sasaki T, Hohenester E, Gohring W and Timpl R. Crystal structure and mapping by site-directed mutagenesis of the collagen-binding epitope of an activated form of BM-40/SPARC/osteonectin. *EMBO J* 1998; 17: 1625-1634. 1998/05/02. DOI: 10.1093/emboj/17.6.1625.
[42] Zhou X, Tan FK, Guo X and Arnett FC. Attenuation of collagen production with small interfering RNA of SPARC in cultured fibroblasts from the skin of patients with scleroderma. *Arthritis Rheum* 2006; 54: 2626-2631. 2006/07/28. DOI: 10.1002/art.21973.
[43] Nakatani K, Seki S, Kawada N, Kitada T, Yamada T, Sakaguchi H, et al. Expression of SPARC by activated hepatic stellate cells and its correlation with the stages of fibrogenesis in human chronic hepatitis. *Virchows Arch* 2002; 441: 466-474. 2002/11/26. DOI: 10.1007/s00428-002-0631-z.
[44] Mazzolini G, Atorrasagasti C, Onorato A, Peixoto E, Schlattjan M, Sowa JP, et al. SPARC expression is associated with hepatic injury in rodents and humans with non-alcoholic fatty liver disease. *Sci Rep* 2018; 8: 725. 2018/01/18. DOI: 10.1038/s41598-017-18981-9.
[45] Atorrasagasti C, Peixoto E, Aquino JB, Kippes N, Malvicini M, Alaniz L, et al. Lack of the matricellular protein SPARC (secreted protein, acidic and rich in cysteine) attenuates liver fibrogenesis in mice. *PLoS One* 2013; 8: e54962. 2013/02/15. DOI: 10.1371/journal.pone.0054962.
[46] Camino AM, Atorrasagasti C, Maccio D, Prada F, Salvatierra E, Rizzo M, et al. Adenovirus-mediated inhibition of SPARC attenuates liver fibrosis in rats. *J Gene Med* 2008; 10: 993-1004. 2008/07/11. DOI: 10.1002/jgm.1228.
[47] Atorrasagasti C, Aquino JB, Hofman L, Alaniz L, Malvicini M, Garcia M, et al. SPARC downregulation attenuates the profibrogenic response of hepatic stellate cells induced by TGF-beta1 and PDGF. *Am J Physiol Gastrointest Liver Physiol* 2011; 300: G739-748. 2011/02/12. DOI: 10.1152/ajpgi.00316.2010.
[48] Tanaka HY, Kitahara K, Sasaki N, Nakao N, Sato K, Narita H, et al. Pancreatic stellate cells derived from human pancreatic cancer demonstrate aberrant SPARC-dependent ECM remodeling in 3D engineered fibrotic tissue of clinically relevant thickness. *Biomaterials* 2019; 192: 355-367. 2018/11/27. DOI: 10.1016/j.biomaterials.2018.11.023.
[49] . Xiao Y, Zhang H, Ma Q, Huang R, Lu J, Liang X, et al. YAP1-mediated pancreatic stellate cell activation inhibits pancreatic cancer cell proliferation. *Cancer Lett* 2019; 462: 51-60. 2019/07/29. DOI: 10.1016/j.canlet.2019.07.015.
[50] Friess H, Ding J, Kleeff J, Liao Q, Berberat PO, Hammer J, et al. Identification of disease-specific genes in chronic pancreatitis using DNA array technology. *Ann Surg* 2001; 234: 769-778; discussion 778-769. 2001/12/01. DOI: 10.1097/00000658-200112000-00008.
[51] Reding T, Wagner U, Silva AB, Sun LK, Bain M, Kim SY, et al. Inflammation-dependent expression of SPARC during development of chronic pancreatitis in WBN/Kob rats and a microarray gene expression analysis. *Physiol Genomics* 2009; 38: 196-204. 2009/05/14. DOI: 10.1152/physiolgenomics.00028.2009.
[52] Infante JR, Matsubayashi H, Sato N, Tonascia J, Klein AP, Riall TA, et al. Peritumoral fibroblast SPARC expression and patient outcome with resectable pancreatic adenocarcinoma. J Clin Oncol 2007;25(3):319–25. https://doi.org/10.1200/JCO.2006.07.8824.
[53] Sinn M, Sinn BV, Striefler JK, Lindner JL, Stieler JM, Lohneis P, et al. SPARC expression in resected pancreatic cancer patients treated with gemcitabine: results from the CONKO-001 study. *Ann Oncol* 2014; 25: 1025-1032. 2014/02/25. DOI: 10.1093/annonc/mdu084.