Molecular Biotechnology (2025) 67:1188–1200
https://doi.org/10.1007/s12033-024-01118-5

ORIGINAL PAPER

Identification of Potential Hub Genes Related to Acute Pancreatitis and Chronic Pancreatitis via Integrated Bioinformatics Analysis and In Vitro Analysis

Lu Yuan[1] · Yiyuan Liu[1] · Lingyan Fan[2] · Cai Sun[1] · Sha Ran[1] · Kuilong Huang[1] · Yan Shen[1]

Received: 7 November 2023 / Accepted: 2 February 2024 / Published online: 23 March 2024

Abstract

Acute pancreatitis (AP) and chronic pancreatitis (CP) are considered to be two separate pancreatic diseases in most studies, but some clinical retrospective analyses in recent years have found some degree of correlation between the two in actual treatment, however, the exact association is not clear. In this study, bioinformatics analysis was utilized to examine microarray sequencing data in mice, with the aim of elucidating the critical signaling pathways and genes involved in the progression from AP to CP. Differential gene expression analyses on murine transcriptomes were conducted using the R programming language and the R/Bioconductor package. Additionally, gene network analysis was performed using the STRING database to predict correlations among genes in the context of pancreatic diseases. Functional enrichment and gene ontology pathways common to both diseases were identified using Metascape. The hub genes were screened in the cytoscape algorithm, and the mRNA levels of the hub genes were verified in mice pancreatic tissues of AP and CP. Then the drugs corresponding to the hub genes were obtained in the drug-gene relationship. A set of hub genes, including *Jun, Cd44, Epcam, Spp1, Anxa2, Hsp90aa1*, and *Cd9*, were identified through analysis, demonstrating their pivotal roles in the progression from AP to CP. Notably, these genes were found to be enriched in the Helper T-cell factor (Th17) signaling pathway. Up-regulation of these genes in both AP and CP mouse models was validated through quantitative real-time polymerase chain reaction (qRT-PCR) results. The significance of the Th17 signaling pathway in the transition from AP to CP was underscored by our findings. Specifically, the essential genes driving this progression were identified as *Jun, Cd44, Epcam, Spp1, Anxa2, Hsp90aa1, and Cd9*. Crucial insights into the molecular mechanisms underlying pancreatitis progression were provided by this research, offering promising avenues for the development of targeted therapeutic interventions.

Keywords Acute pancreatitis · Chronic pancreatitis · Helper T-cell factor signaling pathway · Hub genes

Introduction

Chronic pancreatitis (CP) is characterized as a chronic inflammatory ailment that leads to irreversible alterations in both pancreatic tissue and function [1]. These changes are instigated by a complex interplay of genetic, environmental, and various other factors. The primary pathological features include the atrophy of acinar cells, chronic inflammation, and the development of pancreatic fibrosis. Globally, CP exhibits an annual incidence rate of 9.62 cases per 100,000 individuals, with a corresponding mortality rate of 0.09 cases per 100,000 per year [2]. CP is recognized for its long duration and persistent symptoms, which seriously affects patients' quality of life and may increase their risk of developing pancreatic cancer. [3].

CP's underlying mechanisms are multifaceted, involving acinar cell impairment, acinar stress response, ductal dysfunction, sustained or modified inflammatory pathways, and potential neuroimmune communication. [4]. For instance, the depletion of Treg cells triggers a type 2 immune response, which leads to significant activation of stellate cells and the production of extracellular matrix proteins,

Lu Yuan and Yiyuan Liu have equal contribution to this work.

✉ Yan Shen
shenbmy@126.com

[1] School of Pharmacy and Bioengineering, Chongqing University of Technology, Chongqing 400054, China

[2] Qingdao Central Hospital, University of Health and Rehabilitation Sciences (Qingdao Central Medical Group), Qingdao 266042, China

thereby exacerbating CP [5]. An initial event in CP is cell death resulting from the abnormal activation of trypsinogen in acinar cells [6]. Similarly, the activation of digestive enzymes in the pancreas is the underlying cause of acute pancreatitis (AP) [7]. AP, a prevalent pancreatic disease worldwide, can sometimes be a life-threatening condition [8]. Factors such as persistent inflammation, recurrent episodes, continuous pancreatic injury, genetic predisposition, autoimmune responses, ductal changes, smoking, or specific medications may all contribute to the progression of AP to CP [9–11]. Repeated episodes of AP can worsen pancreatic tissue damage, increasing the risk of sustained inflammation and scar formation. While not all cases of AP develop into CP, the mechanisms governing the transition from AP to CP remain unclear, the factors driving this transition are poorly understood, and there are currently no specific drugs or effective methods for the treatment and prevention of AP and CP. Hence, a comprehensive understanding of the molecular mechanisms is crucial for identifying potential therapeutic interventions.

In this study, differentially expressed genes in AP and CP were identified through the dataset by conducting differential gene analysis. Subsequently, gene function enrichment analysis of the pathway was performed to discover genes with significant differences. The expression levels of mRNA of these genes were verified in vitro. The relationship between genes and drugs was also explored to predict the role of treatment in the development of AP to CP. Through bioinformatics analysis, new insights into the molecular mechanisms underlying Caerulein-induced AP (CER-AP) and CP development and progression were aimed to be provided. This allowed for the identification of new targets for the diagnosis and treatment of AP and CP, contributing to advancements in the understanding of these pancreatic disorders and potentially opening doors to innovative approaches for their clinical management.

Methods

Data Collection

The National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database was employed to access gene expression profiles pertinent to AP (GSE109227, GSE3644) and CP (GSE41418). The GSE109227 microarray dataset was generated utilizing the [MoGene-1_0-st] Affymetrix Mouse Gene 1.0 ST Array platform (GPL6246) and encompassed five wild-type mice as control subjects injected with sodium chloride, along with six mice subjected to intraperitoneal CER injection to induce experimental AP [12]. The GSE3644 microarray data, on the other hand, was derived from the [MOE430A] Affymetrix Mouse Expression 430A Array platform (GPL339) and included three mice injected with saline as controls and three mice administered CER to induce AP [13]. The GSE41418 dataset was established based on the mouse430a2 platform (GPL1261) and comprised six pancreatic tissue samples from CP, juxtaposed with six normal tissue samples as controls [14]. These datasets were harnessed as valuable resources for our study, enabling us to investigate gene expression patterns associated with AP and CP, facilitating a deeper understanding of the molecular mechanisms underpinning these pancreatic disorders.

Data Process and Identification of DEGs

The R package "limma" was employed for the analysis of differentially expressed genes (DEGs) within pancreatic tissues pertaining to both AP and CP [15]. The criteria for defining DEGs were established with an adjusted p-value threshold of less than 0.05. This rigorous statistical approach was utilized to discern significant gene expression variations in order to further our understanding of the molecular differences associated with AP and CP.

Gene Function and Pathway Enrichment Analysis of DEGs

In this study, DEGs identified were subjected to comprehensive functional and biochemical pathway analysis utilizing Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases [16]. The GO terms encompassed three fundamental categories: biological process (BP), cellular component (CC), and molecular function (MF). For this analysis, the Database for Annotation, Visualization, and Integrated Discovery was utilized, which is accessible at "https://metascape.org/" [17]. GO and KEGG pathway enrichment analysis of DEGs, employing a significance threshold of $P < 0.05$, was carried out through Meatscape. This approach allowed for the systematic exploration of the biological processes, cellular components, and molecular functions associated with the identified DEGs, shedding light on the underlying molecular mechanisms implicated in AP and CP.

Protein–Protein Interaction (PPI) Network Construction of DEGs and Module Selection

In this study, PPI analysis was conducted to elucidate interactions among the up-regulated and down-regulated DEGs. [18]. This analysis was executed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database. Subsequently, hub genes were identified from the up-regulated and down-regulated DEGs employing the maximal clique centrality algorithm, which is part of the

CytoHub plug-in integrated within Cytoscape. To comprehensively identify hub genes, various algorithms, including maximal clique centrality, closeness, edge percolated component, and betweenness ranking methods, were employed. These diverse approaches enhanced our ability to pinpoint hub genes, shedding light on their pivotal roles within the network of interactions among the differentially expressed genes.

Assessment of Hub Genes Expression on Single-Cell RNA-seq Data

To further validate the differential expression of hub genes in CP, a single-cell RNA-seq analysis was performed using the GSE182971 dataset. The GSE182971 data were generated using the Illumina HiSeq 4000 platform (GPL21103) and included three wild-type mice that underwent intraperitoneal CER injections for 28 days to induce CP. Following the methodology outlined by Lin et al. [19], our analysis revealed twelve distinct individual clusters, each characterized by specific marker genes, including pancreatic stellate cells (PSCs) *(Pdgfra*, *Col1a1)*, acinar cells (*Ctrb1*, *Cpa1*, *Try4*), duct cells (*Krt8*, *Krt18*), endothelial cells (*Pecam1*, *Tek*), pericytes (*Mylk*, *Acta2*), neutrophils (*Ngp*, *S100a8/9*, *Camp*), macrophages (*Lyz2*, *Csf1r*), T cells (*Cd3d*, *Cd3e*), B cells (*Cd79a*, *Igkc*), NK cells (*Ncr1*, *Nkg7*), dendritic cells (*H2-Aa*, *Rel*), and group 2 innate lymphoid cells (ILC2) (*Il1rl1*, *Gata3*, *Il7r*). This single-cell RNA-seq analysis provided a detailed and comprehensive characterization of distinct cell clusters within the pancreatic tissue, enabling a more precise assessment of gene expression patterns in the context of CP.

Drug-Gene Interaction and Functional Analysis of Potential Genes

In order to elucidate interactions between genes and existing drugs and explore potential applications of new drug indications for pancreatitis, the Drug-Gene Interaction Database (DGIdb) [20]. DGIdb provides comprehensive functionality, including searching, browsing, and filtering information pertaining to drug-gene interactions, drawing upon data from over thirty trusted sources. To identify potential targets for drug intervention, we inputted the module genes into the drug-gene database, seeking to uncover any existing drugs or compounds associated with these genes. The Interaction Score remains constant across searches, as it is a static score. It relies on the evidence of an interaction and is similar to the Query Score, but without dependence on queried gene or drug sets. Instead, it utilizes evidence scores, the ratio of average known gene partners for all drugs to the known partners for the specific drug, and the ratio of average known drug partners for all genes to the known partners for the specific gene [21]. All variables referred to in the following formulas are numerical, with no associated units and the formulas are as follows:

$$\text{relative drug specificity} = \frac{\text{average known gene partners for all drugs}}{\text{known gene partners for drug } d} \quad (1)$$

$$\text{relative gene specificity} = \frac{\text{average known drug partners for all genes}}{\text{known drug partners for gene } g} \quad (2)$$

$$\text{evidence score} = \text{publication count} + \text{source count} \quad (3)$$

The Query Score for gene search is determined by multiplying three sets of data: the "evidence score (Eq. 3)", "queried genes interacting with drug d" and "relative drug specificity (Eq. 1)". Similarly, the Query Score for drug search is calculated by multiplying three sets of data: the "evidence score (Eq. 3)", "queried drugs interacting with gene *g*" and "relative gene specificity (Eq. 2)". The Interaction Score, without a specific query, is derived by multiplying three sets of data: the "evidence score (Eq. 3)", "relative drug specificity (Eq. 1)" and "relative gene specificity (Eq. 2)". It's important to note that all variables are numerical with no associated units. Subsequently, the potential genes for which matched drugs were found were extracted and subjected to functional enrichment analysis. This multifaceted approach allowed us to explore promising avenues for drug repurposing and the development of novel therapeutic strategies for pancreatitis.

Animal Procedures

Male C57BL/6J mice weighing 20–25 g were procured from Hunan SJA Laboratory Animal Co., Ltd. (Chongqing, China). The mice underwent a one-week acclimatization period before the commencement of the experiment and were housed under controlled conditions at 24 ± 2 °C, with a 12-h light–dark cycle and provided a standard diet. Ethical approval for the experimental procedures was obtained from the Ethics Committee of Chongqing University of Technology (Approval Number: 202329).

Eighteen mice were randomly allocated to one of three groups: the Control group, the AP group, and the CP group. The AP group was induced through intraperitoneal injection of 50 μg/kg of CER every hour for a duration of 9 h [22]. Subsequently, mice were euthanized 12 h after the initiation of the first intraperitoneal injection. On the other hand, the CP group was induced through recurrent episodes of AP. Mice in this group received intraperitoneal injections of 50 μg/kg CER every hour for 6 h, 3 days per week, over a total period of 3 weeks [23]. These mice were then sacrificed 3 days following the final CER injection. Pancreatic tissue samples were collected for subsequent analysis.

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

Total RNA extracted from pancreatic tissues was prepared using Trizol reagent. Subsequently, 1 μg of isolated RNA was reverse transcribed into cDNA utilizing the SweScript All-in-One First-Strand cDNA Synthesis SuperMix kit [24]. Amplification of the samples was achieved by incorporating Blue SYBR Green qPCR Master Mix into the Light Cycler 480 Real-Time PCR System.

To determine the relative quantity of the target mRNA, normalization to the expression level of the reference gene β-actin was performed. Data analysis was carried out using the $2^{-\Delta\Delta Ct}$ method. The gene-specific primers used are listed in Table 1.

Statistical Analysis

Statistical analysis was carried out using GraphPad Prism 8.0 software. The data were presented as mean ± standard error of the mean (SEM). Statistical significance was defined as a *P*-value less than 0.05 ($P < 0.05$).

Results

Identification of DEGs in AP and CP

Raw microarray data were sourced from the GEO database and processed through R language programming. The adjusted *P*-value serves to control the false discovery rate, minimizing the chances of identifying false positives. The logFC represents the multiplicative relationship between the control group and the model's expression after logarithmic transformation. The cut-off logFC is determined by the logFC of genes in various datasets, leading to variations in the cut-off logFC across different datasets. A total of 984 DEGs were extracted from GSE109227, including 813 up-regulated genes and 171 down-regulated genes ($logFC > 1.26$, $P < 0.05$) in the pancreas (Fig. 1A, D). The DEGs underwent further clustering analysis (Fig. 1B, E). Employing the same procedure, 292 up-regulated genes and 235 down-regulated genes ($logFC > 1.437$, $P < 0.05$) were identified through differential gene analysis of the GSE3644 dataset, representing a separate cohort of AP. Subsequently, the R language's intersect function was utilized to pinpoint genes with shared differential expression in both datasets, reflecting commonalities in acute pancreatitis. For the CP disease model, the relevant dataset was selected from the database, and limma differential analysis was applied to obtain 655 DEGs, comprising 40 up-regulated genes and 615 down-regulated genes ($logFC > 1.635$, $P < 0.05$) (Fig. 1C, F). Following these analyses, Venn diagrams (Fig. 1G) were generated to identify the set of genes differentially expressed in common across all three datasets, resulting in a total of 49 shared differential genes. Subsequently, protein interaction network maps were constructed using the STRING website and were also imported into Cytoscape software for visualization. This approach facilitated insights into the interconnected relationships among these common differential genes at a protein level.

Enrichment Analysis of DEGs

The genes enclosed within the boxes were submitted to the STRING website to generate protein interaction network maps. Additionally, these genes were imported into Cytoscape software to generate PPI network maps. The varying sizes of nodes in the figure indicate the degree of interaction, with the node size being directly proportional to the degree of interaction. However, it's important to note that the remaining 18 genes are not part of the main interaction network, resulting in only 31 genes being depicted. (Fig. 2A). Among them, *Hsp90aa1*, *CD44*, *Anxa2*, *Jun*, and *Epcam* showed more potential interactions with other proteins in the network. Furthermore, they were entered into the Meatscape website to procure network maps depicting GO and KEGG enrichment analysis results (Fig. 2B). The analysis suggests that significantly enriched on KEGG are Salmonella infections and focal adhesion as well as regulation of wound healing and positive regulation of cell motility on GO.

Table 1 Sequences of primers used in this study

Genes	Forward (5′ to 3′)	Reverse (5′ to 3′)
Jun	TGAAGGAAGAGCCGCAGACC	CCTTGATCCGCTCCTGAGACTC
Cd9	TGGTTTCCTGGGCTGCTGTG	GGCGGCGGCTATCTCAATGG
Cd44	AGTGCGAACCAGGACAGTGG	CAGAGCCAGTGCCAGGAGAG
Spp1	AGAGCGGTGAGTCTAAGGAGTCC	TGGCTGCCCTTTCCGTTGTTG
Epcam	CCGAAGGGGCGATCCAGAAC	GGACTCCGGCGGTGTTGAC
Hsp90aa1	CCAGCAAACAGGACCGAACC	GCTTTGGTGCCCGACTTGG
β-actin	GGCTGTATTCCCCTCCATCG	CCAGTTGGTAACAATGCCAT

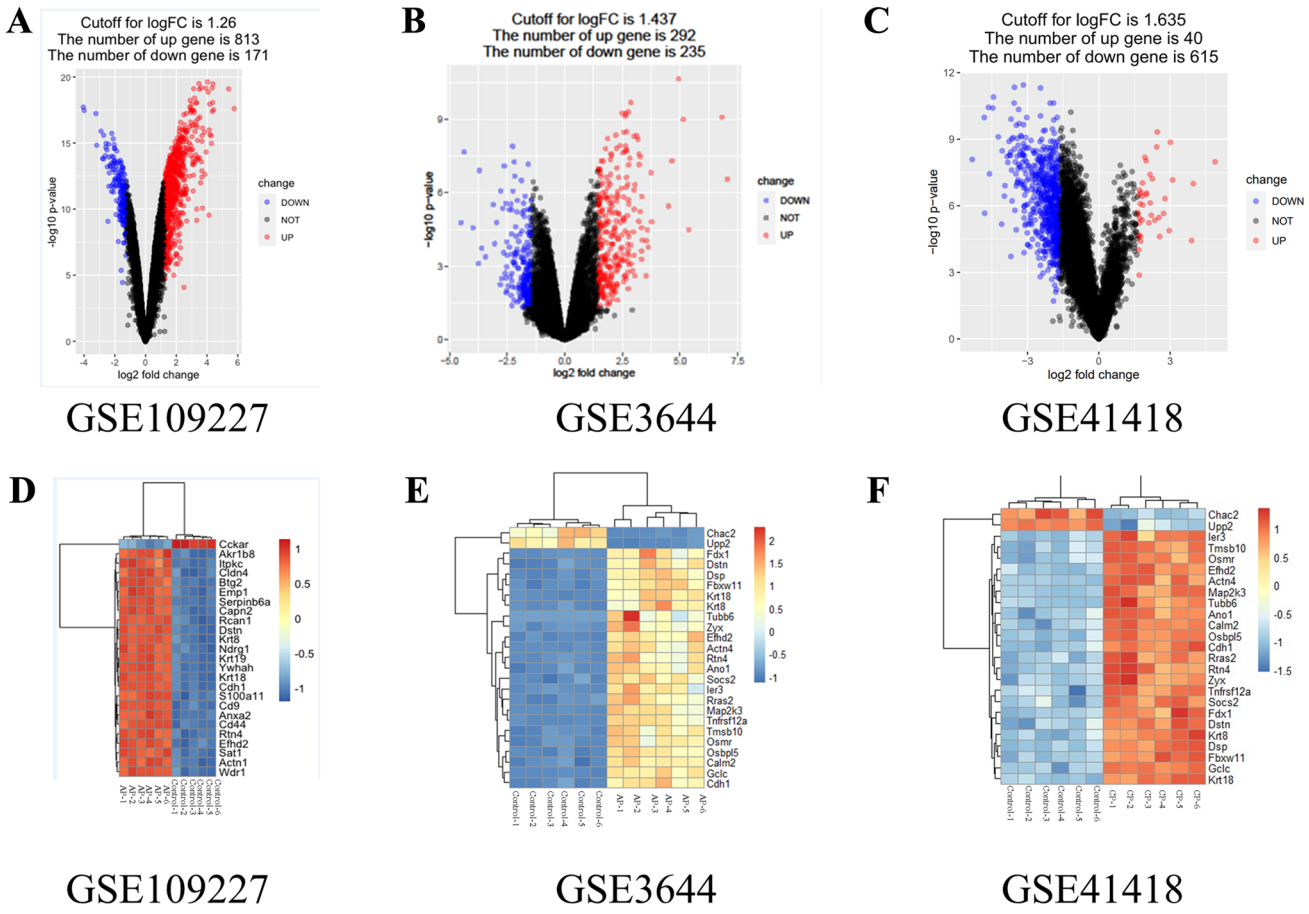

Fig. 1 Data processing and identification of DEGs. **A–C** Volcano plot DEGs in GSE109227, GSE3644 and GSE41418 datasets. Red showed up-regulated genes and blue showed down-regulated genes. **D–F** Heat maps of DEGs in GSE109227, GSE3644 and GSE41418 datasets. **G** Venn diagram of DEGs among three GEO datasets. Identified overlapping DEGs to all three datasets

Determine and Analyze hub Genes

Critical genes were identified utilizing the Cytohubba plugin within Cytoscape software, employing both ranking and meso ranking methods. Four distinct algorithms, including betweenness, closeness, edge percolated component, and maximal clique centrality, were used to calculate the top ten genes exhibiting the highest correlations (Fig. 3A–D). The

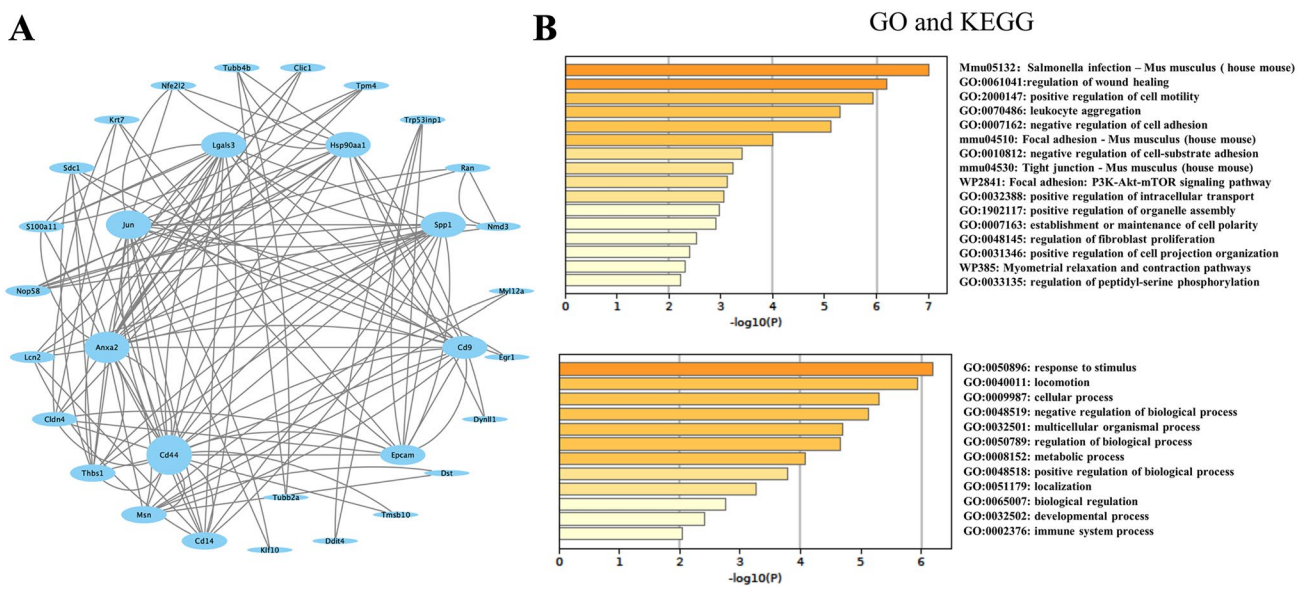

Fig. 2 DEGs functional enrichment analysis in AP and CP datasets. **A** PPI of the overlapping DEGs. **B** GO annotation and KEGG pathway enrichment analysis of the overlapping DEGs

top-scoring ten genes from each sorting method were designated as the hub genes. Using the interlinked relationship scoring plugin in Cytoscape software, we identified seven critical genes that play a role in the progression from AP to CP. These genes are *Jun*, *Cd44*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1*, and *Cd9* (Fig. 3E). Expression levels of these genes were plotted in each of the three datasets, and significant differences were observed (Fig. 3G–I).

Subsequently, GO enrichment analysis of the aforementioned genes was conducted using the Meatscape website. The results were compared with the KEGG database, revealing functional enrichment in the immune system, notably highlighting enrichment in Helper T-cell factor (Th17) cells, which distinguished it from the PI3K-Akt signaling pathway (Fig. 3F). Th17 cells represent a subpopulation of T helper cells known for their production of IL-17 and are recognized as an inflammatory subset that contributes to chronic tissue inflammation and organ dysfunction [25]. Th17 cells produce signature cytokines, including IL-17A, IL-17F, and IL-22, which induce chemokine production to recruit neutrophils during inflammation [26].

Single-Cell RNA-seq Analysis Revealed the Patterns of *Jun*, *Cd44*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1*, *Cd9* in CP Databases

To investigate the expression patterns of *Jun*, *Cd44*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1*, and *Cd9* within CP datasets, we conducted a single-cell RNA-seq analysis on GSE182971 dataset. Following quality control, our analysis revealed eleven distinct individual clusters that were projected onto uniform manifold approximation and projection (UMAP) plots (Fig. 4A), which was used to downscale the data and visualize the cell clusters. These clusters encompassed PSCs marked by *Pdgfra* and *Col1a1*, acinar cells expressing *Ctrb1, Cpa1*, and *Try4*, duct cells with *Krt8* and *Krt18* as markers, endothelial cells marked by *Pecam1* and *Tek*, pericytes with *Mylk* and *Acta2*, neutrophils expressing *Ngp*, *S100a8/9* and *Camp*, macrophages characterized by *Lyz2* and *Csf1r*, T cells expressing *Cd3d* and *Cd3e*, B cells marked by *Cd79a* and *Igkc*, NK cells with *Ncr1* and *Nkg7* as markers, dendritic cells marked by *H2-Aa* and *Rel*, and ILC2 identified by *Il1rl1*, *Gata3*, and *Il7r*. Subsequently, Subsequently, individual genes were visualized by processing the data to obtain the expression levels of individual genes within the cell. The expression patterns of *Jun*, *Cd9*, *Anxa2*, and *Hsp90aa1* were found to be widespread across most cell types (Fig. 4B, C, F, G). Conversely, *Epcam* exhibited higher expression in acinar and ductal cells (Fig. 4D), while *Spp1* and *Cd44* showed elevated expression in ductal cells, PSCs, and immune cells (Fig. 4E, H). In conclusion, these results indicate that *Jun*, *Cd44*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1*, and *Cd9* are enriched in CP cells and exhibit distinct expression patterns across various cell types. These findings provide valuable insights that can guide further in-depth analyses and investigations.

Validation of Hub Genes

To validate the expression stability of the hub genes, AP and CP mouse models were induced with CER. RNA was extracted from the pancreatic tissues of normal mice, AP

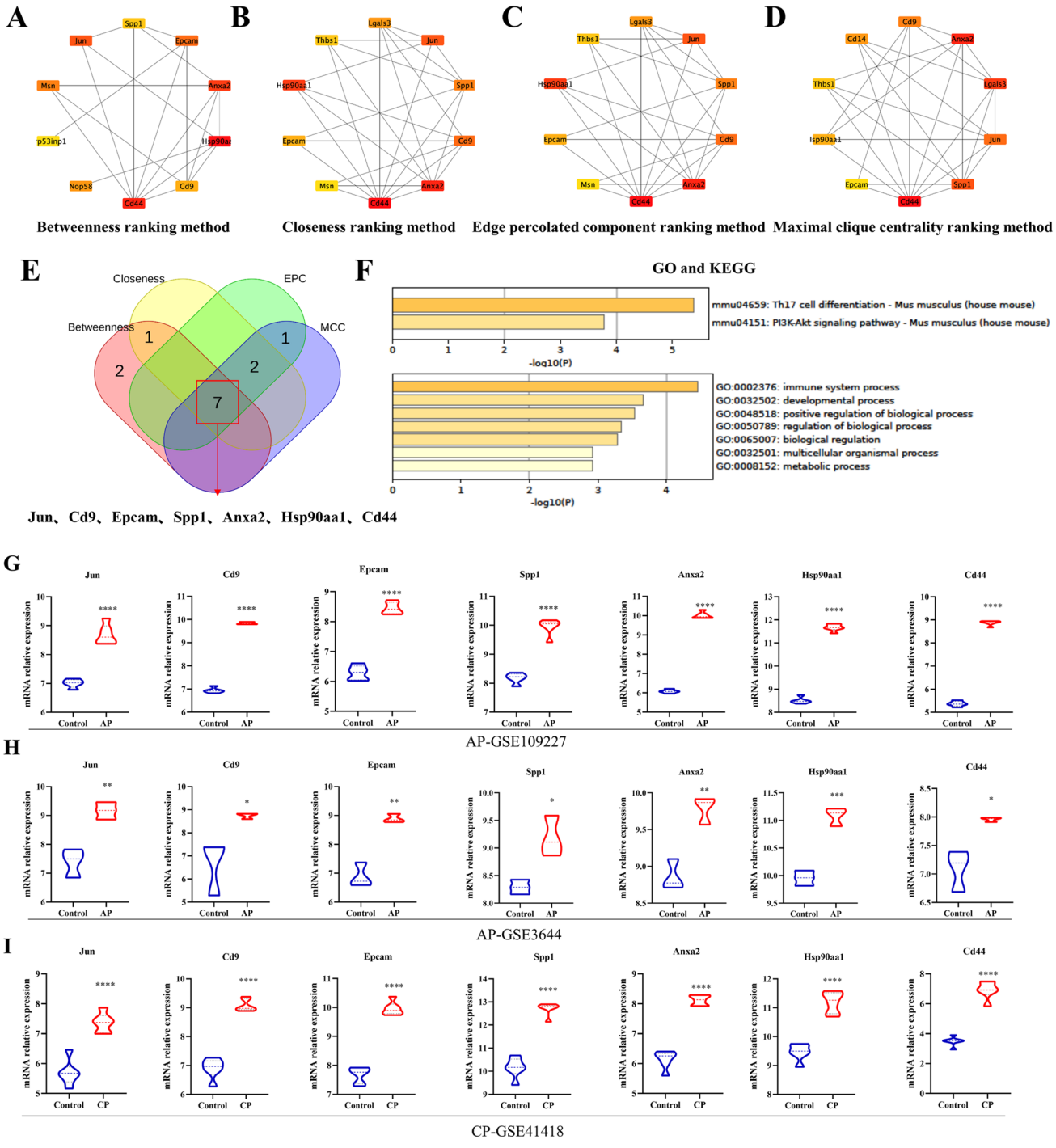

Fig. 3 Identification of hub genes in overlapping DEGs among three GEO datasets. **A** Betweenness ranking method for hub genes identification. **B** Closeness ranking method for hub genes identification. **C** Edge percolated component ranking method for hub genes identification. **D** Maximal clique centrality ranking method for hub genes identification. **E** Venn diagram for identifying hub genes among betweenness, closeness, edge percolated component and axial clique centrality ranking method. **F** GO annotation and KEGG pathway enrichment analysis of hub genes. **G** The expression of *Jun*, *Cd9*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1* and *Cd44* in GSE109227. **H** The expression of *Jun*, *Cd9*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1* and *Cd44* in GSE3644. **I** The expression of *Jun*, *Cd9*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1* and *Cd44* in GSE41418. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$

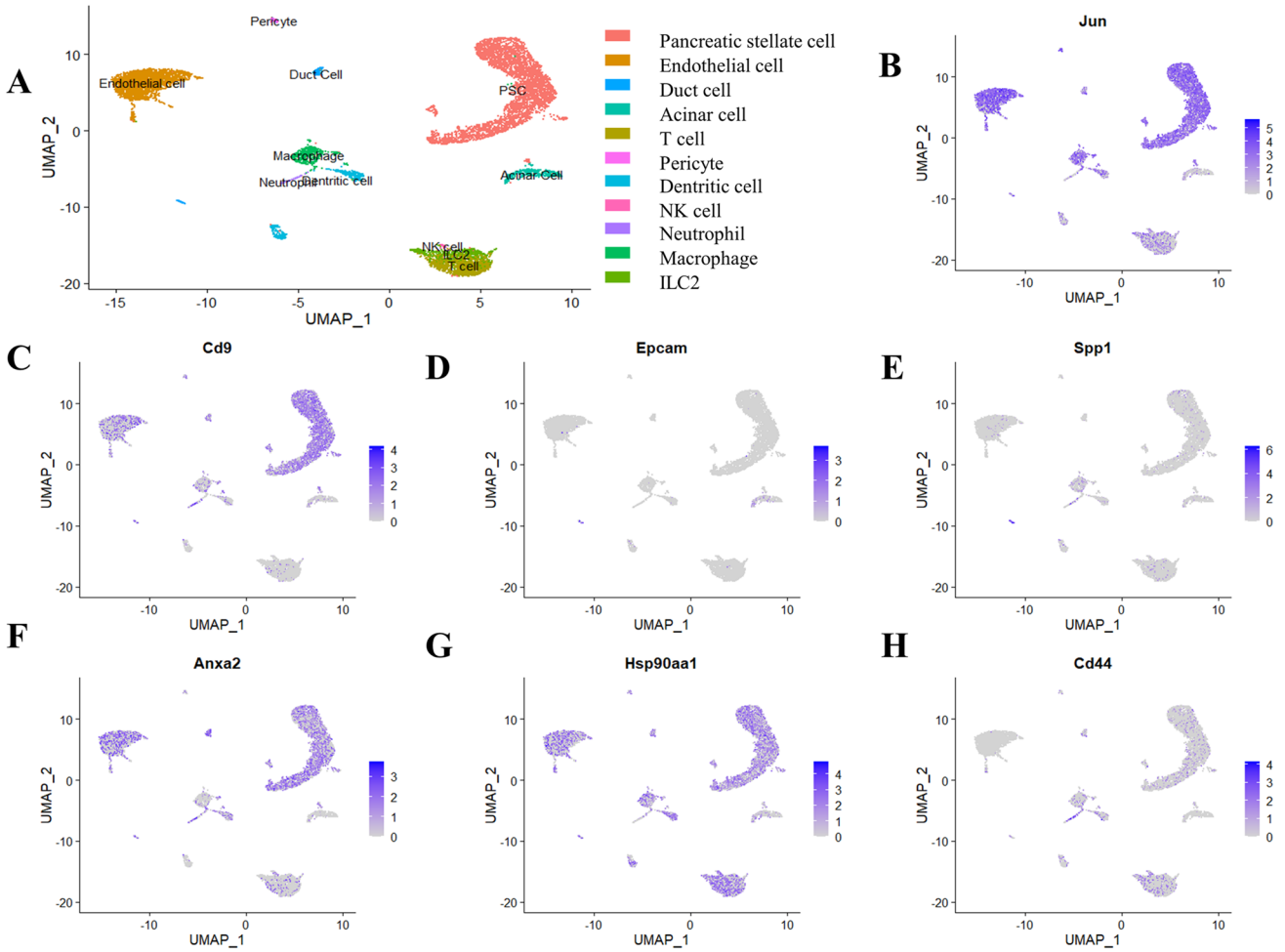

Fig. 4 Verification of hub genes expression pattern on single-cell RNA-seq analysis. **A** UMAP depicts identity of clusters of CP in GSE182971. **B**–**H** Distribution of *Jun*, *Cd9*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1* and *Cd44* in CP. Purple color represents hub genes expressed in the corresponding cells. The purple and white scale on the right represents the level of expression, with blue being high expression

mice, and CP mice, respectively. Subsequent qRT-PCR experiments revealed that the mRNA expression levels of the hub genes, including *Jun*, *Cd44, Epcam*, *Anxa2*, *Hsp90aa1*, and *Cd9*, were significantly upregulated when compared to the normal group, consistent with the data predictions (Fig. 5A). Furthermore, the results indicated that the mRNA expression levels of *Cd9*, *Spp1*, *Anxa2*, and *Cd44* were elevated in CP mouse tissues (Fig. 5B). Overall, these findings demonstrate a reasonable level of consistency with the bioinformatics analyses.

Analysis of Gene-Drug Interactions

Currently, there is no drug available for the prevention of CP, and based on the results of the above experiments, genes critical in the pathologic process of AP progressing to CP were identified. Methods to utilize these genes as potential drug targets, allowing the prediction of drugs that interact with them, are currently available [27]. Subsequently, the seven identified genes were analyzed on the website to obtain drugs targeting the genes, respectively. The hub genes were input into the online DGIdb database to screen for potential drugs that may interact with them. Drug candidates targeting *Jun*, *Cd44*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1* and *Cd9* were identified individually.

Specifically, ten drug candidates were found for *Jun*, three for *Cd44*, nine for Epcam, six for *Spp1*, two for *Anxa2*, and ten for *Hsp90aa1*, respectively. However, *Cd9* did not yield any matching drug candidates (Table 2). The selection of these drug candidates was based on the interaction scoring, and the top five drugs with the highest scores were chosen.

It's worth noting that the term "antagonist interaction" refers to a scenario in which a drug blocks or attenuates agonist-mediated responses rather than eliciting a biological

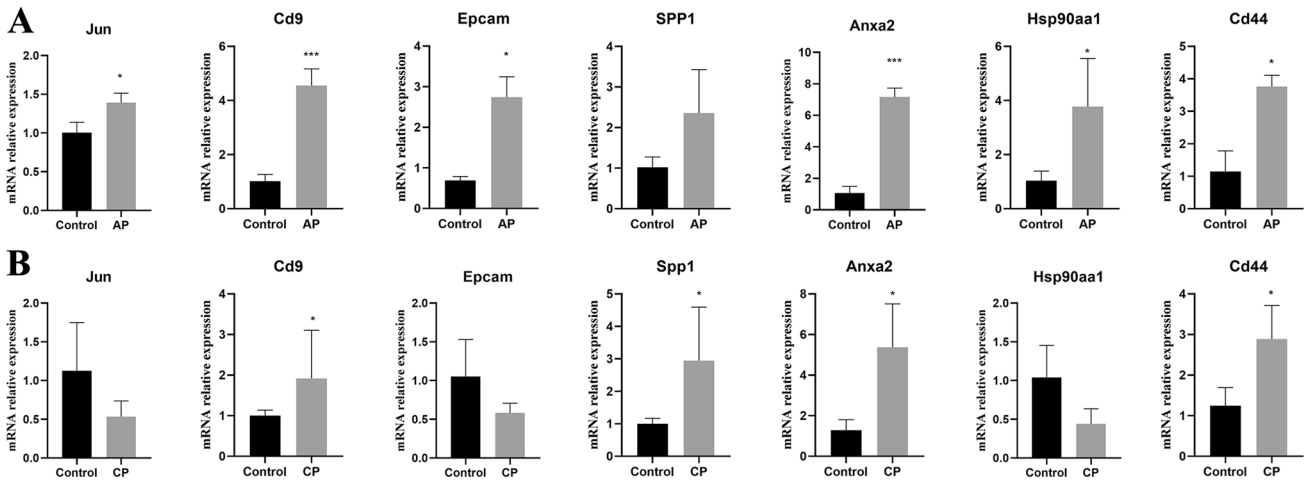

Fig. 5 Validation of the mRNA expression levels of hub genes using qRT-PCR. **A** The expression of hub genes in normal mice pancreas tissues and AP pancreas tissues. **B** The expression of hub genes in normal mice pancreas tissues and CP pancreas tissues. $n \geq 3$ per group. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$

Table 2 Hub gene-drug interactions

Target of hub genes	Drug	Interaction type and directionality	Interaction score	Approved
Jun	Bruceantin	n/a	2.81	No
	Sergeolide	n/a	2.81	No
	Irisolidone	n/a	2.81	No
	Neochamaejasmin a	n/a	2.81	No
	Chembl477052	n/a	2.81	No
Cd44	Hyaluronan	n/a	6.87	No
	Gentamicin	n/a	1.96	No
	Progesterone	n/a	0.62	No
Epcam	ING-1	n/a	23.19	No
	Adecatumumab	Inhibitor	15.46	No
	Oportuzumab monatox	n/a	15.46	No
	Citatuzumab monatox	n/a	15.46	No
	Tucotuzumab celmoleukin	n/a	7.73	No
	Catumaxomab	n/a	5.8	No
Spp1	ASK-8007	Inhibitor	10.3	No
	Calcitonin	n/a	3.43	No
	Gentamicin	n/a	0.98	No
	Wortmannin	n/a	0.74	No
	Alteplase	n/a	1.08	Yes
Anxa2	n-Ethylmaleimide	n/a	12.37	No
	Withaferin a	n/a	5.15	No
Hsp90aa1	Chembl 1,834,096	n/a	0.91	No
	Argenteoside	n/a	0.91	No
	Luminespib	n/a	0.91	No
	Bllb021	Inhibitor	0.45	No
	Alvespimycin	Inhibitor	0.45	No

response itself upon binding to a target receptor. This interaction type is characterized by its inhibitory directionality. In cases where the reporting source did not specify the interaction type, it was labeled as "N/A" in DGIdb. The magnitude of the interaction score is influenced by factors such as the evidence score (including publication and source credibility), the search set, and the specific drug interaction search [28]. Therefore, genes and drugs with numerous interactions tend to receive higher rankings. It is noteworthy, however, that drugs or genes involved in numerous interactions typically receive lower scores.

As of now, Alteplase is the only drug approved for marketing, specifically for thrombolytic therapy targeting Spp1.

Discussion

A retrospective analysis revealed that the initial episode of AP resulted in recurrent pancreatitis in 17% of patients, and nearly 8% of patients progressed to CP within 5 years [29]. This retrospective study also identified pancreatic necrosis as an independent risk factor for the development of recurrent pancreatitis (RP) and CP [29]. Furthermore, it was found that both alcoholism and smoking independently played key roles in the pathogenesis of both CP and AP [30]. Some studies have put forward the hypothesis of a 'necrosis-fibrosis' connection when AP is present with CP, highlighting a strong association between these conditions [31].

In an effort to enhance the reliability of our data for subsequent analyses, we aimed to minimize the occurrence of false positive genes by expanding the dataset for AP. Following this, we performed differential gene analysis on the microarray sequencing data of mouse AP and CP, identifying genes with significant differential expression. The shared gene percentage in the GSE109227 dataset and the GSE3644 dataset is 24.4% and 44.1%, respectively. The variation in this percentage might be attributed to inconsistencies in mold-making sampling, experimental extraction operations, and the probe platforms analyzed. Notably, in the two AP datasets, GSE109227 and GSE3644, the DEGs consistently exhibit either upregulation or downregulation. However, when comparing the CP dataset GSE41418 with either the AP datasets GSE3644 or GSE109227, we observe instances where commonly differentially expressed genes demonstrate inconsistent expression patterns. This discrepancy could arise from disparities in the pathophysiology, immune regulation, and etiology between AP and CP. Utilizing this data analysis approach, we identified key hub genes, including *Jun*, *Cd44*, *Epcam*, *Spp1*, *Anxa1*, *Hsp90aa1*, and *Cd9*. To validate these findings, qRT-PCR testing was performed, revealing that the mRNA expression levels of these genes were significantly elevated in the pancreas tissues of AP mice when compared to healthy mice controls, aligning with the results of our analysis. Additionally, the qRT-PCR results showed that the mRNA expression levels of *Cd9*, *Spp1*, *Anxa2*, and *Cd44* were increased in the pancreatic tissues of CP mice compared to those of healthy mouse controls. This discrepancy may be attributed to variations in the timing and severity of CP induced by CER in mice, as well as the relatively small sample size of chronic pancreatitis tissue, which could introduce inconsistencies between the validation and analysis results. Future research endeavors should focus on validating the expression levels of these hub genes at the protein level using larger sample sizes. Notably, *CD9* and *CD44* are recognized as markers for tumor stem cells in the pancreas [32–34]. The elevated mRNA expression levels of *CD9* and *CD44* in the pancreatic tissues of both AP and CP mice may be linked to the progression from AP to CP, eventually culminating in pancreatic cancer [35, 36]. Despite the lack of existing literature regarding the role of *Anxa2* in pancreatitis, studies related to nonalcoholic steatohepatitis (NASH) have indicated a positive correlation between *Anxa2* expression and the development of hepatocyte pyroptosis and fibrosis. In vitro and in vivo experiments have provided evidence that *Anxa2* promotes focal death of hepatocytes and contributes to liver fibrosis [37]. Additionally, in a model of bleomycin-induced pulmonary fibrosis, *Anxa2* was found to interact with bleomycin and serve as a specific target for bleomycin. Genetic deletion of *Anxa2* in mice resulted in the attenuation of bleomycin-induced pulmonary fibrosis. The binding of bleomycin to *Anxa2* disrupted TFEB-induced autophagic flux, ultimately leading to the development of pulmonary fibrosis [38]. The elevated expression of *Anxa2* in CP tissues suggests its potential role in promoting pancreatic fibrosis. Hendley et al. conducted research on pancreatic ductal cells and identified and validated the epithelial-mesenchymal transition (EMT) axis using Monocle 3, highlighting *Spp1* as one of the key regulators of the epithelial-to-mesenchymal transition in ductal cells [39]. Deletion of *Spp1* was found to protect human pancreatic duct cells from adopting a tumorigenic phenotype [40].

Furthermore, it was found that these genes were enriched in the signaling pathway of Th17. These results suggest that the relationship between AP and CP may be mediated through the Th17 signaling pathway. Th17 is amplified by IL-23 through the induction of IL-6 differentiation, TGF-β, and IL-1β, as well as through the STAT3 signaling pathway [41]. The signaling pathway of Th17-associated cytokines may be a key factor in adaptive immunity. During the evolution of AP from early localized inflammation to systemic disease, proinflammatory factors disrupt the intestinal barrier, allowing bacteria and endotoxins from the gut to enter the bloodstream, triggering a severe systemic inflammatory response [42]. In this process, Th17 produces a variety of cytokines, including IL-17A and IFN-g, to regulate inflammatory cell infiltration and tissue destruction [43]. A

significant imbalance in the ratio between Th17 cells and Treg cells has been found in serologic studies of patients with different severity and prognosis of AP [44]. Gut microbes play a role in promoting the differentiation and accumulation of extra-thymic regulatory T cells, and a lack of Treg cells triggers an enhanced specific immune response [45]. Many studies have shown that gastrointestinal flora homeostasis plays an important role in the systemic infectious effects of AP and has become a target for clinical AP therapy [42]. A serologic study found that levels of C–C motif regulator ligand 20 (CCL20) were significantly elevated in both AP and CP compared to controls [46]. CCL20 promotes the recruitment of Th17 and Treg cells to sites of inflammation, suggesting that Th17 cells are involved in the pathogenesis of AP and CP.

The regulation of Th17 in CP is linked to the regulation of the STING, which, when suppressed, promotes Th17 polarization through the upregulation of fibrotic genes, contributing to the manifestation of fibrosis and inflammation [47]. Th17 can shift towards Th1 and Th2 phenotypes through autocrine or paracrine mechanisms, in addition to its own secretion of tumor necrosis factor-alpha (TNF-α) [43]. Among these, Th2 releases IL-13 and IL-4 that bind to innate adaptive immunity during pancreatic stellate cell fibrosis. Treg cell subsets are important in CP to balance the immune response and limit the progression of fibrosis, suppressing the Th2-driven type 2 immune response [5]. The research strongly indicates an increased likelihood that Th17 plays a pivotal role in the progression from AP to CP. In summary, our study has unveiled a common relationship between AP and CP within the Th17 signaling pathway. Considering the existing research landscape, the co-expressed genes identified in this study may serve as target genes within the Th17 pathway for both diseases.

The bioinformatics analysis results come with certain limitations. The singularity of algorithmic conditions fails to adequately simulate the complexity of pathological changes, and there is a lack of experimental validation. Moreover, variations in sequencing methods and sample conditions directly impact the sequencing results, consequently influencing the analysis outcomes. Despite these limitations, such bioinformatics analyses are crucial for investigating the relationship between diseases and genes. They allow for an expanded scope of study while saving both time and cost in research and development.

Conclusion

Our findings emphasize the importance of the Th17 signaling pathway in the transition from AP to CP. Notably, hub genes implicated in driving this progression, such as *Jun*, *Cd44*, *Epcam*, *Spp1*, *Anxa2*, *Hsp90aa1*, and *Cd9*, were identified. This research contributes essential insights into the molecular mechanisms that underlie the progression of pancreatitis, paving the way for potential targeted therapeutic interventions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12033-024-01118-5.

Author Contributions Yan Shen designed this study and critically revised the manuscript. Lu Yuan and Yiyuan Liu drafted the manuscript. Lingyan Fan acquired and analyzed the data. Yiyuan Liu, Lu Yuan, Cai Sun and Sha Ran did literature search and performed the experiments. Kuilong Huang directed the qRT-PCR experiments and analyzed the data. All authors have reviewed and confirmed the final manuscript.

Funding This work was supported by National Natural Science Foundation of China (No. 82100684), Natural Science Foundation of Chongqing China (CSTB2022NSCQ-MSX1493), Scientific Research Foundation of Chongqing University of Technology, National Natural Science Foundation Incubation project of Chongqing University of Technology (2022PYZ037), and the Funding Achievements of the Action Plan for High Quality Development of Graduate Education at Chongqing University of Technology (gzlcx20233384).

Data availability The data in this article was obtained online and can be obtained from publicly available databases. This manuscript is not applicable.

Declarations

Conflict of interest The authors confirm that they do not have any acknowledged competing financial interests or personal affiliations that might have appeared to influence the research documented in this article.

References

1. Beyer, G., Habtezion, A., Werner, J., et al. (2020). Chronic pancreatitis. *Lancet, 396*, 499–512.
2. Singh, V. K., Yadav, D., & Garg, P. K. (2019). Diagnosis and management of chronic pancreatitis: A review. *JAMA, 322*, 2422–2434.
3. Balázs, A., Balla, Z., Kui, B., et al. (2018). Ductal mucus obstruction and reduced fluid secretion are early defects in chronic pancreatitis. *Frontiers in Physiology, 9*, 632.
4. Zhang, Y., Zhang, W. Q., Liu, X. Y., et al. (2023). Immune cells and immune cell-targeted therapy in chronic pancreatitis. *Frontiers in Oncology, 13*, 1151103.
5. Glaubitz, J., Wilden, A., Golchert, J., et al. (2022). In mouse chronic pancreatitis CD25(+)FOXP3(+) regulatory T cells control pancreatic fibrosis by suppression of the type 2 immune response. *Nature Communications, 13*, 4502.
6. Mayerle, J., Sendler, M., Hegyi, E., et al. (2019). Genetics, cell biology, and pathophysiology of pancreatitis. *Gastroenterology, 156*, 1951-1968.e1951.
7. Sidhu, S., Pandhi, P., Malhotra, S., et al. (2011). Rosiglitazone promotes pancreatic regeneration in experimental model of acute pancreatitis. *Fundamental & Clinical Pharmacology, 25*, 237–247.
8. Paragomi, P., Spagnolo, D. M., Breze, C. R., et al. (2020). Introduction and validation of a novel acute pancreatitis digital tool:

Interrogating large pooled data from 2 prospectively ascertained cohorts. *Pancreas, 49*, 1276–1282.
9. Yadav, D., O'Connell, M., & Papachristou, G. I. (2012). Natural history following the first attack of acute pancreatitis. *American Journal of Gastroenterology, 107*, 1096–1103.
10. Liu, J., Gao, M., Nipper, M., et al. (2019). Activation of the intrinsic fibroinflammatory program in adult pancreatic acinar cells triggered by Hippo signaling disruption. *PLoS Biology, 17*, e3000418.
11. Han, S. Y., Conwell, D. L., Diaz, P. T., et al. (2022). The deleterious effects of smoking on the development and progression of chronic pancreatitis. *Pancreatology, 22*, 683–687.
12. Norberg, K. J., Nania, S., Li, X., et al. (2018). RCAN1 is a marker of oxidative stress, induced in acute pancreatitis. *Pancreatology, 18*, 734–741.
13. Kowalik, A. S., Johnson, C. L., Chadi, S. A., et al. (2007). Mice lacking the transcription factor Mist1 exhibit an altered stress response and increased sensitivity to caerulein-induced pancreatitis. *American Journal of Physiology Gastrointestinal and Liver Physiology, 292*, G1123-1132.
14. Ulmasov, B., Oshima, K., Rodriguez, M. G., et al. (2013). Differences in the degree of cerulein-induced chronic pancreatitis in C57BL/6 mouse substrains lead to new insights in identification of potential risk factors in the development of chronic pancreatitis. *American Journal of Pathology, 183*, 692–708.
15. Yuan, Q., Ren, J., Chen, X., et al. (2022). Contributions and prognostic performances of m7G RNA regulators in pancreatic adenocarcinoma. *Chinese Medical Journal (Engl), 135*, 2101–2103.
16. Jin, S., & Hyun, T. K. (2020). Ectopic expression of production of anthocyanin pigment 1 (PAP1) improves the antioxidant and anti-melanogenic properties of ginseng (Panax ginseng C.A. Meyer) Hairy Roots. *Antioxidants (Basel), 9*, 922.
17. Zhou, Y., Zhou, B., Pache, L., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications, 10*, 1523.
18. Alur, V. C., Raju, V., Vastrad, B., et al. (2019). Mining featured biomarkers linked with epithelial ovarian cancerbased on bioinformatics. *Diagnostics (Basel), 9*, 39.
19. Yang, X., Chen, J., Wang, J., et al. (2022). Very-low-density lipoprotein receptor-enhanced lipid metabolism in pancreatic stellate cells promotes pancreatic fibrosis. *Immunity, 55*, 1185-1199.e1188.
20. Shu, J., Ren, Y., Tan, W., et al. (2022). Identification of potential drug targets for vascular dementia and carotid plaques by analyzing underlying molecular signatures shared by them. *Frontiers in Aging Neuroscience, 14*, 967146.
21. Freshour, S. L., Kiwala, S., Cotto, K. C., et al. (2021). Integration of the drug-gene interaction database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Research, 49*, D1144-d1151.
22. Wen, L., Voronina, S., Javed, M. A., et al. (2015). Inhibitors of ORAI1 prevent cytosolic calcium-associated injury of human pancreatic acinar cells and acute pancreatitis in 3 mouse models. *Gastroenterology, 149*, 481-492.e487.
23. Bansod, S., Aslam Saifi, M., Khurana, A., et al. (2020). Nimbolide abrogates cerulein-induced chronic pancreatitis by modulating β-catenin/Smad in a sirtuin-dependent way. *Pharmacological Research, 156*, 104756.
24. Huang, Q. Y., Zhang, R., Zhang, Q. Y., et al. (2023). Disulfiram reduces the severity of mouse acute pancreatitis by inhibiting RIPK1-dependent acinar cell necrosis. *Bioorganic Chemistry, 133*, 106382.
25. Yasuda, K., Takeuchi, Y., & Hirota, K. (2019). The pathogenicity of Th17 cells in autoimmune diseases. *Seminars in Immunopathology, 41*, 283–297.
26. Littman, D. R., & Rudensky, A. Y. (2010). Th17 and regulatory T cells in mediating and restraining inflammation. *Cell, 140*, 845–858.
27. Griffith, M., Griffith, O. L., Coffman, A. C., et al. (2013). DGIdb—Mining the druggable genome. *Nature Methods, 10*, 1209–1210.
28. Freshour, S. L., Kiwala, S., Cotto, K. C., et al. (2020). Integration of the drug-gene interaction database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Research, 49*, D1144–D1151.
29. Ali, U. A., Issa, Y., Hagenaars, J. C., et al. (2016). Risk of recurrent pancreatitis and progression to chronic pancreatitis after a first episode of acute pancreatitis—ScienceDirect. *Clinical Gastroenterology and Hepatology, 14*, 738–746.
30. Alexandre, M., Pandol, S. J., Gorelick, F. S., et al. (2011). The emerging role of smoking in the development of pancreatitis. *Pancreatology, 11*, 469–474.
31. Yadav, D., Hawes, R. H., Brand, R. E., et al. (2009). Alcohol consumption, cigarette smoking, and the risk of recurrent acute and chronic pancreatitis. *Archives of Internal Medicine, 169*, 1035–1045.
32. Li, C., Heidt, D. G., Dalerba, P., et al. (2007). Identification of pancreatic cancer stem cells. *Cancer Research, 67*, 1030–1037.
33. Wang, V. M., Ferreira, R. M. M., Almagro, J., et al. (2019). CD9 identifies pancreatic cancer stem cells and modulates glutamine metabolism to fuel tumour growth. *Nature Cell Biology, 21*, 1425–1435.
34. Chen, K., Wang, Q., Li, M., et al. (2021). Single-cell RNA-seq reveals dynamic change in tumor microenvironment during pancreatic ductal adenocarcinoma malignant progression. *eBioMedicine, 66*, 103315.
35. Singhi, A. D., Pai, R. K., Kant, J. A., et al. (2014). The histopathology of PRSS1 hereditary pancreatitis. *American Journal of Surgical Pathology, 38*, 346–353.
36. Singhi, A. D., Koay, E. J., Chari, S. T., et al. (2019). Early detection of pancreatic cancer: Opportunities and challenges. *Gastroenterology, 156*, 2024–2040.
37. Feng, Y., Li, W., Wang, Z., et al. (2022). The p-STAT3/ANXA2 axis promotes caspase-1-mediated hepatocyte pyroptosis in nonalcoholic steatohepatitis. *Journal of Translational Medicine, 20*, 497.
38. Wang, K., Zhang, T., Lei, Y., et al. (2018). Identification of ANXA2 (annexin A2) as a specific bleomycin target to induce pulmonary fibrosis by impeding TFEB-mediated autophagic flux. *Autophagy, 14*, 269–282.
39. Hendley, A. M., Rao, A. A., Leonhardt, L., et al. (2021). Single-cell transcriptome analysis defines heterogeneity of the murine pancreatic ductal tree. *eLife, 10*, e677776.
40. Adams, C. R., Htwe, H. H., Marsh, T., et al. (2019). Transcriptional control of subtype switching ensures adaptation and growth of pancreatic cancer. *eLife, 8*, e45313.
41. Park, H., Li, Z., Yang, X. O., et al. (2005). A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nature Immunology, 6*, 1133–1141.
42. Cen, M. E., Wang, F., Su, Y., et al. (2018). Gastrointestinal microecology: A crucial and potential target in acute pancreatitis. *Apoptosis, 23*, 377–387.
43. Cosmi, L., Santarlasci, V., Maggi, L., et al. (2014). Th17 plasticity: Pathophysiology and treatment of chronic inflammatory disorders. *Current Opinion in Pharmacology, 17*, 12–16.
44. Guo, J., Li, Z., Tang, D., et al. (2020). Th17/Treg imbalance in patients with severe acute pancreatitis: Attenuated by high-volume hemofiltration treatment. *Medicine (Baltimore), 99*, e21491.
45. Campbell, C., Dikiy, S., Bhattarai, S. K., et al. (2018). Extrathymically generated regulatory T cells establish a niche for intestinal border-dwelling bacteria and affect physiologic metabolite balance. *Immunity, 48*, 1245-1257.e1249.

46. Lee, B., Jones, E. K., Manohar, M., et al. (2023). Distinct serum immune profiles define the spectrum of acute and chronic pancreatitis from the multicenter prospective evaluation of chronic pancreatitis for epidemiologic and translational studies (PROCEED) study. *Gastroenterology, 165*, 173–186.
47. Zhao, Q., Manohar, M., Wei, Y., et al. (2019). STING signalling protects against chronic pancreatitis by modulating Th17 response. *Gut, 68*, 1827–1837.