

RESEARCH

Open Access



# BNIP3+ fibroblasts associated with hypoxia and inflammation predict prognosis and immunotherapy response in pancreatic ductal adenocarcinoma

Bo Gao<sup>1†</sup>, Guohua Hu<sup>1†</sup>, Boshi Sun<sup>2</sup>, Wenqiang Li<sup>3</sup> and Hao Yang<sup>4\*</sup>

## Abstract

**Background** Pancreatic ductal adenocarcinoma (PDAC) is one of the most malignant tumors that lacks effective treatment options. Cancer-associated fibroblasts (CAFs), an important component of the tumor microenvironment, associated with tumor progression, prognosis, and treatment response. This work aimed to explore the novel CAFs-associated target to improve treatment strategies in PDAC.

**Methods** The PDAC single-cell sequencing data (CRA001160,  $n=35$ ) were downloaded and integrated based on GSA databases to classify fibroblasts into fine subtypes. Functional enrichment analysis and coexpression regulatory network analysis were used to identify the functional phenotypes and biological properties of the different fibroblast subtypes. Fibroblast differentiation trajectories were constructed using pseudochronological analysis to identify initial and terminally differentiated subtypes of fibroblasts. The changes in the proportions of different fibroblast subtypes before and after PDAC immunotherapy were compared in responsive and nonresponding patients, and the relationships between fibroblast subtypes and PDAC immunotherapy responsiveness were determined based on GSA and GEO database. Using molecular biology methods to confirm the effects of BNIP3 on hypoxia and inflammation in CAFs. CAFs were co cultured with pancreatic cancer cells to detect their effects on migration and invasion of pancreatic cancer.

**Results** Single-cell data analysis divided fibroblasts into six subtypes. The differentiation trajectory suggested that BNIP3+ Fibro subtype exhibited terminal differentiation, and the expression of genes related to hypoxia and the inflammatory response increased gradually with differentiation time. The specific overexpressed genes in the BNIP3+ Fibro subtype were significantly associated with overall and disease progression-free survival in the patients with PDAC. Interestingly, the greater the proportion of the BNIP3+ Fibro subtype was, the worse the response of PDAC patients to immunotherapy, and the CRTL treatment regimen effectively reduced the proportion of the BNIP3+ Fibro subtype. After knocking out BNIP3, the hypoxia markers and inflammatory factors of CAFs were inhibited. Co-culture of CAFs with pancreatic cancer cells can increase the migration and invasion of pancreatic cancer, but this could be reversed by knocking out BNIP3.

<sup>†</sup>Bo Gao and Guohua Hu have contributed equally to this work.

\*Correspondence:

Hao Yang

yanghao@hrbmu.edu.cn

Full list of author information is available at the end of the article



**Conclusions** This study revealed the BNIP3+ Fibro subtype associated with hypoxia and inflammatory responses, which was closely related to the poor prognosis of patients with PDAC, and identified signature genes that predict the immunotherapy response in PDAC.

**Keywords** Pancreatic ductal adenocarcinoma, BNIP3+ fibroblasts, Prognosis, Immunotherapy

## Introduction

Pancreatic ductal adenocarcinoma (PDAC) accounts for approximately 90% of all pancreatic cancers. Pancreatic cancer is characterized clinically by its gradual onset, high invasiveness, and propensity for easy metastasis [1]. Therefore, the majority of patients with PDAC are diagnosed at an advanced stage. It is extremely malignant and has a 5-year survival rate of less than 10%. According to the American Cancer Statistics Association, pancreatic cancer will become the second leading cause of cancer death worldwide in 2030 [2]. Immunotherapy has seen widespread application and remarkable progress in treating various solid tumors in recent years. However, strategies aimed at deconstructing the surrounding desmoplastic stroma and targeting the immunosuppressive pathways have largely failed in PDAC [3]. Recent study has found that pancreatic cancer tissue contains abnormally abundant tumor stroma, and remodeling or alleviating the tumor microenvironment may make PDAC sensitive to immune checkpoint inhibitors.

A dense stromal tumor microenvironment accounts for 90% of pancreatic tumor masses [4]. Fibroblasts are the main component of the matrix and can secrete large amounts of extracellular components, such as extracellular matrix and matrix metalloproteinases, thereby producing unique PDAC tumors with immunosuppressive effects [5]. Activated cancer-associated fibroblasts (CAFs) can also release collagen, fibronectin and hyaluronic acid. These extracellular matrices infiltrate the tumor microenvironment, increasing tumor density and preventing killer T cells from infiltrating the tumor to a large extent. Moreover, the response of PDAC to immunotherapy is limited [6, 7]. Fibroblast populations exhibit extensive heterogeneity in terms of cell origin and phenotype, which leads to differences in behavior during tumor development, with most fibroblast subtypes acting as tumor promoters, whereas other subtypes simultaneously play a tumor suppressor role. In addition, the mutual transformation between different subtypes also indicates the plasticity and pluripotency of fibroblasts. The proportions of different fibroblast subtypes may change dynamically with the progression of tumors or the progression of immunotherapy [8]. Therefore, identifying reliable and specific cell surface markers is the key to distinguishing different fibroblast subtypes [9, 10].

Single-cell sequencing, capable of quantifying the attributes of individual cells, serves as a powerful tool for examining the cellular components and their interactions within the tumor microenvironment. This study used single-cell sequencing data to divide the patients with PDAC into fine subtypes, identify the biological functions and differentiation trajectories of the fibroblast subtypes, and explore the relationships between the fibroblast subtypes and the prognosis and immunotherapy response of patients with PDAC. We found that the COLEC11+ Fibro and PLA2G2A+ Fibro subtypes were dominant in normal pancreatic tissue. In PDAC tissue, these two subtypes were significantly reduced, and a specific BNIP3+ Fibro subtype differentiated. In the immunotherapy cohort, the greater the proportion of patients in the COLEC11+ Fibro and PLA2G2A+ Fibro subtypes was, the greater the immunotherapy response was, while the proportion of patients in the BNIP3+ Fibro subtype in the poor response group increased, which provides theoretical basis for the selection of immunotherapy options for PDAC patients.

## Materials and methods

### Single-cell data download and preprocessing

The pancreatic cancer single-cell data were downloaded with the data number CRA001160 in the GSA database (<https://ngdc.cncb.ac.cn/gsa/>) [11]. Whole single-cell transcriptome data analysis was completed using the Seurat package (version: 4.0.3) in R language. The CreateSeuratObject function was used for the merged sample to construct a single-cell analysis object. The parameter “min.features=100” was used to ensure that the initialized cells had at least 100 genes. The “min.cells=10” parameter was used to ensure that the genes were present in at least 10 cells. The initial filtering process involves certain expressions.

### Integrated analysis of single-cell data and cell type annotation

After data preprocessing and quality control, we obtained high-quality expression data for 35 samples. To avoid the influence of sample batch effects, we used canonical correlation analysis (CCA) to analyze multiple samples. Unbiased integration was performed. The

FindIntegrationAnchors function was subsequently used to find similarity anchors between pairs of data, and the IntegrateData function was used to integrate multiple sample groups.

After the sample integration step, we performed dimensionality reduction clustering and cell type annotation on the data and used the FeaturePlot function to mark the expression of various cell landmark genes and annotate cell types. Subtypes without marker gene expression, as well as cell groups expressing two or more cell marker genes, were defined as low-quality cells and double cells, respectively. Low-quality cells and double cells were uniformly removed and were not subjected to further analysis. A total of 57,340 high-quality single-cell expression data points were ultimately obtained.

### Single-cell pathway activity and functional enrichment analysis

The R package G SVA (version 1.40.1) was used to implement the gene set variation analysis (GSVA) algorithm. The overall process is divided into three steps: 1. The average expression function in Seurat was used to calculate the number of genes in each cell subtype. Mean expression matrix; 2. The getGmt function constructs the set of pathways and genes to be analyzed. The metabolic pathway genes were obtained from published articles (<https://doi.org/10.1126/sciadv.abd9738>). The Hallmark50 gene set was downloaded from the Molecular Signatures Database (v7.5.1) (<https://www.gsea-msigdb.org/gsea/index.jsp>); 3. The gsva function was used to perform G SVA analysis, and the parameters were set to  $\text{min.sz} = 5$  and  $\text{max.sz} = 500$ .

The R package fgsea (version: 1.18.0) was used to implement the G SEA (Gene Set Enrichment Analysis) algorithm, which is divided into four steps: 1. The Wilcoxon function of the R package presto (version: 1.0.0) was used to calculate the differences in gene expression between the tumor tissue and control tissue; 2. The gmtPathways function was used to construct the pathways and gene sets to be analyzed. The Hallmark50 gene set was downloaded from the Molecular Signatures Database (v7.5.1) (<https://www.gsea-msigdb.org/gsea/index.jsp>); 3. The fgseaMultilevel function performs G SEA analysis and sets the parameters  $\text{minSize} = 10$ ,  $\text{maxSize} = 500$  and  $\text{eps} = 0$ ; 4. The plotEnrichment function was used to plot the gene set enrichment.

### Pseudosequential differentiation trajectory inference and analysis

To conduct an in-depth analysis of the progression and differentiation of different fibroblast subtypes in pancreatic cancer, Monocle2 (version: 2.20.0) software was first used to perform pseudochronological analysis of

pancreatic cancer and adjacent fibroblasts to construct and visualize cell differentiation path. The analysis involved the following four steps: (1) The newCellDataSet function was used to construct a CellDataSet object, and the expressionFamily parameter was set to  $\text{negbinomial.size}()$ ; (2) the estimateSizeFactors and estimateDispersions functions were used to calculate the factors and divergences of gene expression, after which differential-GeneTest was used to select the top 1000 highly variable genes of each subtype for trajectory construction; and (3) the DDRTree method of the reduceDimension function was used to reduce the dimensionality of the gene expression matrix, after which the orderCells function was used to restore the cell differentiation path from the dimensionally reduced data. The position of cells on the differentiation path was fixed. (4) Visualization of cell differentiation trajectories included the following methods. The plot cell trajectory function visualizes the cell differentiation path. The plot genes in pseudotime and heatmap functions visualize a specific gene and multiple genes, respectively, in the pseudotime sequence. The expression change trend, plot genes branched pseudotime and plot genes branched heatmap functions were used to visualize the change in gene expression of a specific gene and multiple genes before and after a specific branch point, respectively, in the pseudotime series. The selection of key genes at the bifurcation point was completed by the BEAM algorithm function.

Monocle2 was used to construct the tree differentiation structure of the data. To better reflect the high-dimensional characteristics of the data, Monocle3 (version 1.2.9) was further used to study the differentiation trajectory. The analysis involved the following four steps: (1) The new cell data set function created the Monocle3 analysis object; (2) Based on the Seurat software, the clustering results were mapped to Monocle 3; (3) The learn graph function inferred the cell differentiation trajectory, and the order cell function was sorted according to the order of differentiation; (4) The graph test function was used to perform differential gene calculations based on the graph and then calculate the coexpressed gene module in the high-dimensional space based on Moran's index. The biological function of the module was determined with the clusterProfiler package.

### Identification of subtype transcription factors

The gene regulatory network (GRN) is composed of transcription factors (TFs), cofactors and their regulated target genes (target genes), which together determine the transcriptional status of cells in a certain state. SCENIC software (version 1.2.4) was used to identify transcriptional regulators of different subtypes. The overall SCENIC process includes three steps: (1) GENIE3/

GRNBoost (gradient boosting) was used to infer the coexpression module between transcription factors and candidate target genes based on coexpression; and (2) since the GENIE3 model is based only on coexpression, there will be many false-positives and indirect regulatory effects. To identify direct binding targets, RcisTarget was used to construct cis-regulatory motifs in each coexpression module. TF-motif enrichment analysis was then performed to identify direct targets. Each TF and its potential direct target genes are called a regulon. (3) The AUCell algorithm was used to score the activity of each regulon in each cell. For a regulon, comparing AUCell scores between cells can reveal which cells have significantly greater activity. This will determine in which cells regulon is actively turned on. The reference database used in the above analysis process was downloaded from the website (<https://resources.aertslab.org/cistarget/databases>). The cell type-specific transcription factors were evaluated based on the regulation specificity score (RSS).

#### **Weighted gene coexpression regulatory network analysis**

Weighted gene coexpression regulatory network (WGCNA) is a method used to analyze and describe gene association patterns between different samples. The tool can be used to identify gene sets that change highly collaboratively and to identify key factors based on the interconnectivity of gene sets and the associations between gene sets and phenotypes. Compared with studies that only focus on differentially expressed genes, WGCNA (version 1.71) uses the information of thousands or nearly 10,000 most highly expressed genes or all genes to identify gene sets of interest and performs significant correlation analysis with phenotypes.

#### **Survival analysis**

Survival analysis was completed through the online website GEPIA (<http://gepia.cancer-pku.cn/>), and pancreatic cancer patient survival data from The Cancer Genome Atlas were selected for analysis. Patients were sorted according to gene expression levels and divided into high- and low-expression groups using upper and lower quartiles. Overall survival and progression-free survival were analyzed. The log rank  $p$  value was calculated, and a log rank  $p$  value  $< 0.05$  was considered to indicate a significant difference in survival.

#### **Immunotherapy data download and preprocessing**

The data number GSE202051 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE202051>) was downloaded comprehensively used single-cell transcriptome sequencing technology to analyze single-cell data before and after neoadjuvant treatment for pancreatic cancer

and the response of patients to different treatment regimens [12]. The data included data from 43 pancreatic cancer patients, 25 of whom were untreated tumor samples and 18 of whom had undergone neoadjuvant treatment. We selected all the samples for subsequent integrated analysis and research. The data preprocessing procedure was consistent with the aforementioned method, and a total of 224,988 high-quality single-cell expression data points were ultimately obtained.

#### **Immunotherapy data integrated analysis and cell type annotation**

In order to compare neoadjuvant treatment data with previous data, we performed data integration and cell type identification based on comparison of reference data and query data. First, the previously annotated fibroblast subpopulation was used as a reference data set, and the overall fibroblasts after neoadjuvant treatment were extracted as a query data set. FindTransferAnchors function was used to calculate the anchor points between the two data sets. Then TransferData function was carried out to convert and annotate the calculated anchor points. Finally, MapQuery function was used to display the query data set in the UMAP space of the reference data set.

#### **Analysis of cell state transition based on cell rank**

CellRank software was used to calculate the pseudotime of differentiation to estimate the possibility of various cells differentiating into stable cell types [13]. Moreover, the authors estimated which cells were at the starting point and end of differentiation based on the transfer characteristics of cells on the Markov chain. To further determine the initial and terminal differentiation subtypes of fibroblast subtypes, CellRank software was used to infer the states of the cells based on Pseudotime in Monocle3.

#### **Cell culture**

Pancreatic cancer cells (Panc-1, BxPC-3 and Panc02) were purchased from Procell Life Science & Technology (Wuhan, China), and the cells were cultured according to the manual instructions. The cell lines were cultured in RPMI-1640 medium (Gibco, USA) supplemented with 10% fetal bovine serum (FBS) (Gibco, USA) and 1% penicillin/streptomycin (Gibco, USA).

#### **Extraction of pancreatic cancer-associated fibroblasts through the collagenase method**

Surgically-resected PDAC tissues and paracancerous normal pancreatic tissues were collected and immediately placed in DMEM culture medium supplemented with a penicillin–streptomycin dual antibiotic solution. An electronic balance was used to weigh 40 mg

of type II collagenase (ThermoFisher, USA). A total of 10 ml of DMEM culture medium was added to the collagenase and mixed by pipetting until it was fully dissolved. This yielded a 0.4% type II collagenase solution, which was then passed through a 0.22  $\mu\text{m}$  filter and stored in a refrigerator at 4 °C. Next, ophthalmic scissors were used to remove adipose tissue and other tissues from around the PDAC and normal pancreatic tissues. The specimens were cut into 1–2  $\text{mm}^3$ -sized pieces, placed in a 0.1% type II collagenase solution, with a volume approximately five times that of the specimen, and shaken for mixing. Then, the specimens were left to digest at 37 °C in an incubator containing 5% CO<sub>2</sub> for 4 h and were retrieved and shaken every 20 min. Next, the mixture was passed through a 4- $\mu\text{m}$  cell strainer to remove tissue debris, and the filtrate was transferred to a 15-ml centrifuge tube, centrifuged at 1500 rpm for 5 min, and the supernatant was discarded. The cells were washed twice, transferred to a T25 culture flask, and cultured in an incubator. Non-adherent cells were discarded after 24 h. The medium was changed every 2 days. When the cells reached 80% confluence, EDTA-trypsin was used to passage the cells at a 1:2 ratio. The cells were ready for use in cytological experiments after 3–6 passages.

CAF were cultured in 24-well plates and transfected with previously constructed RNA interference lentiviral vectors (Genechem, China) or a negative control (empty plasmid) for 24 h. Interference sequences of lentiviral are shown in Additional file 1. The medium was changed to complete medium, and cell culture was continued for 1 week. The medium was then changed to complete medium containing puromycin. After 72 h, the fluorescence intensity was observed under a fluorescence microscope, and the visible fluorescence of the cells indicated that the transfection was successful. The lentivirus was resistant to puromycin, and the stable expression lentiviral cell lines were screened by adding puromycin in the medium. In the process of culture, the cells were overgrown in 24-well plates, and gradually passed into 12-well plates and 6-well plates.

#### **Co-culture of pancreatic cancer cells and cancer-associated fibroblasts**

Pancreatic cancer cells (Panc-1, BxPC-3) and CAFs were adjusted to a concentration of  $1 \times 10^6$  cells/ml using DMEM complete medium. Then, the pancreatic cancer cells and CAFs were co-cultured in a six-well plate cell culture chamber (pore size: 0.4  $\mu\text{m}$ ; LAB-SELECT, China) (Fig. 7E). A total of 200  $\mu\text{l}$  of CAF

suspension was added to the chamber and 800  $\mu\text{l}$  of pancreatic cell suspension was added to the plate.

#### **Colony formation assay**

Cells were cultured following the method described above for the co-culture of CAFs and pancreatic cancer cells. A total of  $1 \times 10^3$  pancreatic cancer cells (Panc-1, BxPC-3) were seeded onto a six-well plate and colony formation was observed with the naked eye 1 week after incubation. Then, after fixing with 4% paraformaldehyde for 15 min, the cells were stained with 0.5% crystal violet for 15 min.

#### **Wound-healing assay**

Cells were cultured following the method described above for the co-culture of CAFs and pancreatic cancer cells. When pancreatic cancer cells filled the six-well plate, the medium in the chamber and six-well plate was substituted with DMEM supplemented with 1% fetal bovine serum. The tip of a 10- $\mu\text{l}$  pipette was used to scratch the pancreatic cancer cell layer on the six-well plate to form a wound. Then, the cells were further incubated in an incubator containing 5% CO<sub>2</sub> at 37 °C. Images were captured using a microscope at 0 and 24 h to determine the degree of wound healing.

#### **Transwell assays**

Pancreatic cancer cells were seeded at a concentration of  $2 \times 10^5$  cells per well in a transwell chamber (pore size: 8  $\mu\text{m}$ ; Corning, USA) containing 200  $\mu\text{l}$  of serum-free DMEM, while CAFs were seeded at a concentration of  $1 \times 10^6$  cells per well in a lower culture plate containing 800  $\mu\text{l}$  of DMEM complete medium. After incubation for 24 h, the cells were stained with crystal violet for 15 min. Five random fields of view were selected under a microscope for capturing images and performing cell counts.

#### **Western blotting**

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, Epizyme, China) was used for western blotting of gastric cancer cells. After lysing cells, the lysate was subjected to electrophoresis, membrane transfer, and blocking of non-specific antigens. This was then incubated overnight at 4 °C with primary antibodies specific for anti-BNIP3 (Abcam, USA), anti-HIF- $\alpha$  (Affinity, USA), anti-GLUT1 (Affinity, USA), anti-IL1 (Affinity, USA), anti-IL6 (Affinity, USA), anti-IL8 (Affinity, USA), anti-TNF- $\alpha$  (Affinity, USA) and GAPDH (ImmunoWay, USA). The following day, the membrane was incubated with secondary antibodies for 1 h at room temperature. After visualization of protein bands, grayscale analysis

was performed using the ImageJ software (version 1.8.0). The grayscale of the target protein was divided by the grayscale of Actin to obtain the relative amount of the target protein in each protein sample. Then GraphPad Prism 8.0 software was used for statistical analysis of target protein levels between samples.

### Immunofluorescence staining

Immunofluorescence staining was performed using primary antibodies anti-BNIP3 and anti-FAP. Corresponding Alexa Fluor dyes were used for fluorescent detection. DAPI was used for nuclear counter staining. Images were captured on the Zeiss LSM780 laser scanning confocal microscope. Quantitative analysis of immunohistochemistry and immunofluorescence images was performed using ImageJ software by measuring mean optical density values, and differences were analyzed using the *t* test.

### Statistics

Data were analyzed and visualized using the GraphPad Prism 8.0 software. The Student's *t* test was used to compare means between two groups, and one-way ANOVA was conducted to determine the significance of differences among multiple groups (>2). Subcutaneous tumor growth curves were analyzed using two-way ANOVA.  $p < 0.05$  was considered statistically significant.

## Results

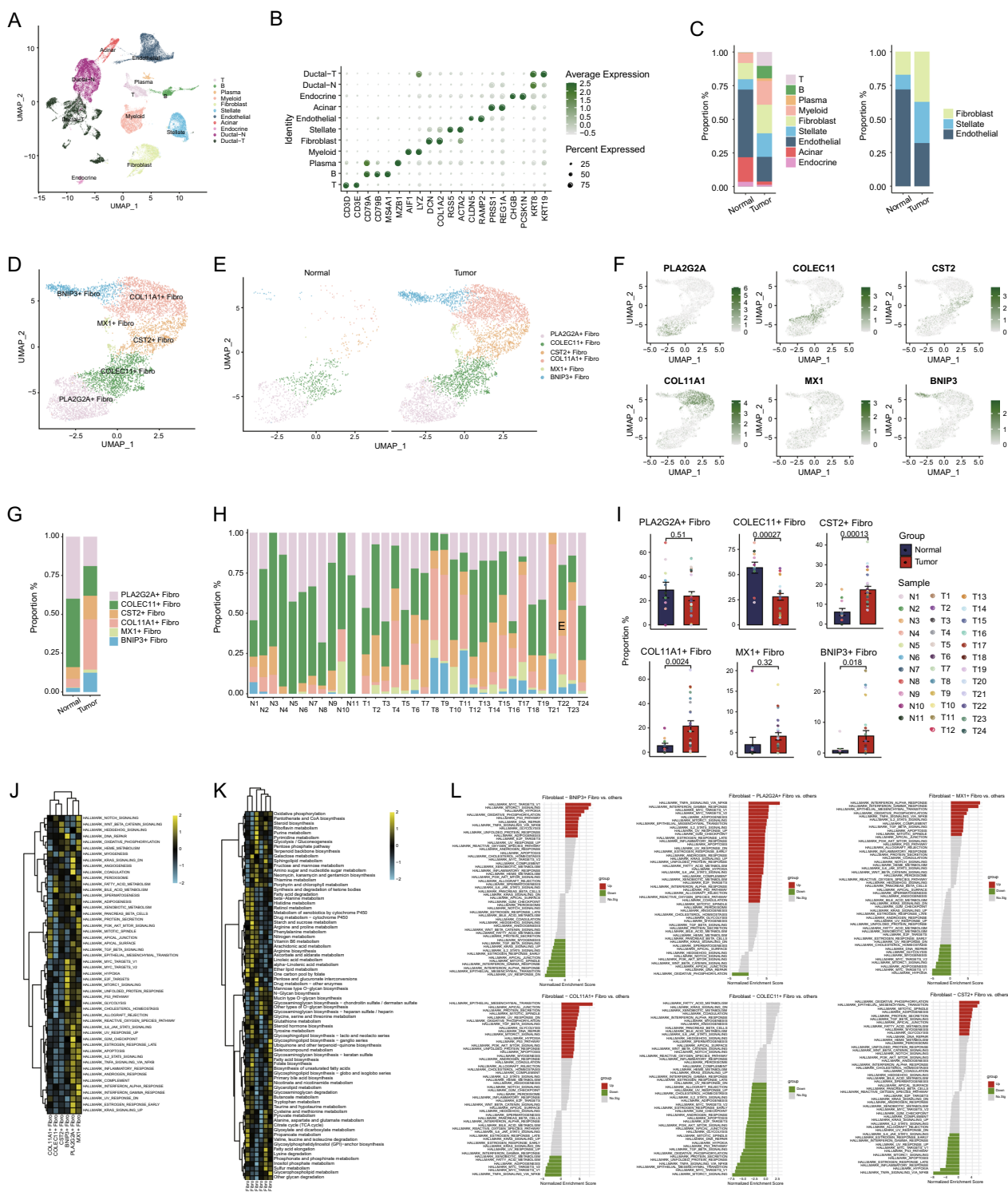
### Single-cell data analysis and fibroblast subtypes

Single-cell sequencing data on pancreatic cancer were downloaded and organized through the use of the GSVA database. The data were obtained from 35 patients with untreated pancreatic cancer, including 24 tumor samples and 11 adjacent healthy control samples. After data preprocessing and quality control, 57,340 high-quality single-cell expression data points were obtained from

35 samples. The FeaturePlot function was used to determine the expression of various cell landmark genes, and the cell types were annotated; these genes were divided into 11 cell clusters (Fig. 1A; Fig. S1A, B). The 11 cell clusters included T cells (CD3D and CD3E) [14], B cells (CD79A, CD79B and MS4A1) [15], plasma cells (CD79A and MZB1) [16], myeloid cells (AIF1 and LYZ) [17], fibroblasts (DCN and COL1A2) [18], stellate cells (RGS5 and ACT2), endothelial cells (CLDN5 and RAMP2) [19], acinar cells (PRSS1 and REG1A), endocrine cells (CHGB and PCSK1N), Ducatl-N cells (KRT8), and Ducatal-T cells (KRT8 and KRT9) (Fig. 1B; Fig. S1C, D). The percentage chart shows that fibroblasts, stellate cells and endothelial cells account for the highest proportions of cells (Fig. 1C). Fibroblast and stellate cell counts were greater in tumor samples, while the endothelial cell count was greater in normal samples. Dynamic remodeling of the TME plays a key role in tumor progression, and fibroblasts are considered key stromal cells involved in TME remodeling [20, 21]. However, the high heterogeneity of fibroblast populations often results in functional diversity [10, 22]. We further divided fibroblasts into fine subtypes and clustered them into 6 subtypes (Fig. 1D–F). Analysis of the distribution of cell proportions revealed that two subtypes, COLEC11+ Fibro and PLA2G2A+ Fibro, were dominant in the normal group, while both subtypes were significantly reduced in the tumor group. However, four subtypes with specific states were differentiated: CST2+ Fibro, COL11A1+ Fibro, MX1+ Fibro and BNIP3+ Fibro (Fig. 1G–I). It is suggested that there is obvious transcriptional reprogramming of fibroblasts in PDAC, and normal fibroblasts are specifically induced to differentiate into tumor-associated fibroblasts to adapt to the special microenvironment of PDAC growth and metastasis.

(See figure on next page.)

**Fig. 1** Single-cell data analysis of the heterogeneity of fibroblasts in pancreatic cancer. **A** The Seurat package (4.0.3) of the R package was used to construct a UMAP diagram, and all cells were annotated into 11 cell types according to the expression of various cell marker genes. Each point in the image represents a single cell, whose position is derived by the UMAP algorithm. **B** A bubble chart was drawn based on the featureplot function to display the percentage and abundance of the iconic genes expressed in different cell clusters. The horizontal axis represents the marker genes of different cell subsets, and the vertical axis represents the 11 cell subsets. The size of the dots represents the percentage of expression, and the color of the dots represents the abundance of expression. **C** The ggplot2 package was used to construct a percentage histogram showing the proportions of various types of cells in normal and tumor tissues. **D** Based on the FindNeighbors function, the Mutual Nearest Neighbor (MNN) algorithm was used to calculate the similarity between cells, subdivide fibroblasts into 6 subtypes, and annotate the key marker genes of each subtype. **E** Annotation of fibroblast subtypes according to normal and PDAC tissue sources. The dot plots on the left represent normal pancreas samples, and the dot plots on the right represent PDAC tissue samples. **F** UMAP diagram showing the distribution of key marker genes in the fibroblast subtype in two-dimensional space. The color of the dots represents the abundance of expression. **G** Percentage histogram showing the proportions of six fibroblast subtypes in normal and tumor tissues. **H** Percentage histogram showing the proportions of six fibroblast subtypes in different sample sources. **I** COLEC+ Fibro ( $p = 0.00027$ ) and PLA2G2A+ Fibro ( $p = 0.51$ ) subtypes have a higher proportion in Normal, CST2+ Fibro ( $p = 0.00013$ ), COL11A1+ Fibro ( $p = 0.0024$ ), MX1+ Fibro ( $p = 0.32$ ) and The BNIP3+ Fibro ( $p = 0.018$ ) subtype has a higher proportion in tumor. **J, K** The GSVA algorithm was used to evaluate the differences in the gene set variation scores of six fibroblast subtypes in metabolic pathways and tumor hallmark 50 pathways. **L** The GSEA algorithm was used to analyze the enrichment of gene sets in six fibroblast subtypes



**Fig. 1** (See legend on previous page.)

**Functional characteristics of fibroblast subtypes**  
 GSEA and GSEA revealed that different fibroblast subtypes were associated with different biological

functions. In addition, COL11A1+ Fibro is related to tumor EMT, matrix remodeling and other processes; CST2+ Fibro is related to myogenesis and the Wnt

pathway; MX1+ Fibro is related to oxidative phosphorylation; and BNIP3+ Fibro is related to the inflammatory response, hypoxia, and glycolysis. Analysis of pathways related to specific immune and inflammatory responses revealed that PLA2G2A+ Fibro had the strongest correlation, followed by BNIP3+ Fibro (Fig. 1J). However, the BNIP3+ Fibro subtype exhibited stronger responses related to hypoxia and glycolysis (Fig. 1I), and metabolic pathway analysis revealed significant differences in metabolism between the two groups (Fig. 1K). The distributions of these two subtypes in normal and PDAC tissues were also exactly opposite (Fig. 1I). The above results suggest that there are fibroblasts related to specific immune regulation in normal pancreatic tissue and that a group of special fibroblast subtypes related to both inflammatory and hypoxic responses are specifically induced during the progression of PDAC.

Significant correlation analysis between fibroblast subtypes and phenotypes was performed through WGCNA, and 8 was selected as the optimal soft threshold for subsequent analysis (Fig. 2A). The Pheatmap clustering tree clusters highly correlated genes into a single branch. The gray modules are composed of ungrouped genes and should be ignored for all downstream analyses (Fig. 2B). Correlation analysis of the coexpression modules revealed that the yellow module was most strongly correlated with BNIP3+ Fibro, and the brown module was most strongly correlated with PLA2G2A+ Fibro (Fig. 2C, D, G). Figure 2E plots the genes and functional enrichment results that are strongly related to BNIP3+ Fibro. The results showed that the function of this subtype of genes was closely related to hypoxia and glycolysis (in response to hypoxia and canonical glycolysis). Figure 2F plots the genes and functional enrichment results that are strongly associated with PLA2G2A+ Fibro. The results showed that the function of this subtype of genes was closely related to the immune-inflammatory response

(in response to interleukin-1 and inactivation of MAPK activity).

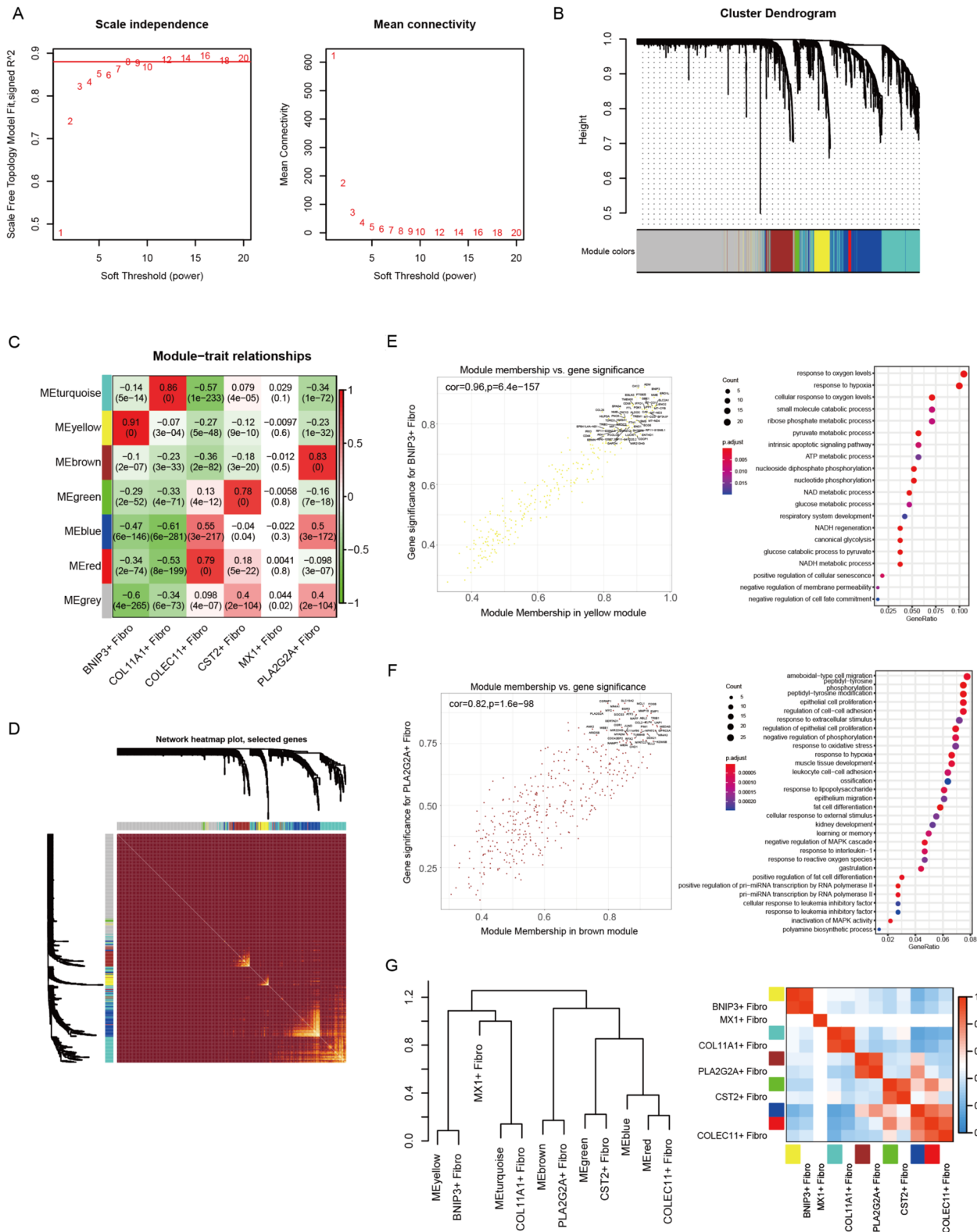
### Pseudochronological differentiation trajectory inference and analysis of fibroblast subtypes

To study the differentiation relationships between different fibroblast subtypes, we evaluated the expression patterns of key genes based on the gene expression of each single cell to sort the cells and deduce the differentiation trajectory of the fibroblast subtypes. The results showed that, starting from COLEC11+ Fibro, the cells gradually differentiated into two branches over time (Fig. 3A, B). The entire trajectory involved a differentiation process from normal to PDAC (Fig. 3C). The two subtypes COLEC11+ Fibro and PLA2G2A+ Fibro, which were more abundant in the normal group, were enriched in the early stage of the differentiation trajectory, while the subtype with a greater proportion in the PDAC group was enriched in the later stage of the differentiation trajectory (Fig. 3D, E). The BNIP3+ Fibro subtype, which we are particularly interested in, differentiates independently, exists on an upward branching branch and is in a terminally differentiated state (Fig. 3D). The genes that play major roles in pseudochronological changes were clustered into four clusters, among which were BNIP3, hypoxia, and immune- and inflammation-related genes (ADM, CCL26, HILPDA, and VEGFA) [23–26]. The genes were highly expressed in Cluster 4 (Fig. 3F). Therefore, it can be inferred that BNIP3+ fibroblasts are a subtype of fibroblasts in the terminal differentiation state, and their biological functions are related mainly to hypoxia and immune-inflammatory responses. In single-cell pseudochronological analysis, there are several branch points, and these branch nodes usually represent programmed changes in cells. The BEAM algorithm was used to identify key genes at the bifurcation point. The heatmap shows that ADM, HILPDA, and CCL26 play key roles in the differentiation of prebranch cells to the cell fate 1 state (Fig. 3G). Representative gene expression

(See figure on next page.)

**Fig. 2** WGCNA modular analysis of the subtypes of fibroblasts. **A** Finding the optimal soft threshold makes the constructed network more consistent with the scale-free topology. The left picture shows the scale-free fit indices (y-axis) under different soft thresholds (x-axis). The red line indicates the selected scale-free fit index value. The picture on the right shows the network connectivity under different soft thresholds. **B** Hierarchical clustering diagram of gene coexpression modules. The upper part is the hierarchical clustering dendrogram of genes, and the lower part is the coexpression gene modules. The same color indicates the same module, and gray indicates that it does not belong to any module. **C** Heatmap of the correlations between different gene coexpression modules and different fibroblast subtypes. **D** Based on the topological overlap matrix, a correlation heatmap was drawn between genes. The brighter the color is, the stronger the interaction between genes, and the same module generally has the strongest correlation. **E, F** The scatter plot on the left plots the correlation between the correlation between genes and modules (module membership, MM) and the correlation between genes and traits (gene significance, GS). Genes that are highly correlated are located in the upper right corner and are marked. The right panel shows the functional enrichment results of the significantly associated genes. **G** Both the hierarchical clustering on the left and the correlation heatmap on the right represent the close relationships between different modules and subtypes





**Fig. 2** (See legend on previous page.)

profiles suggest that the BNIP3+ Fibro subtype and the ADM, CCL26, HILPDA, and VEGFA genes are highly expressed on the upwardly bifurcated branches. Moreover, COL11A1+ Fibro and EMT, matrix remodeling and other functional genes (C1QTNF3, FNDC1, MFAP5, and SDC1) are highly expressed on downwardly branching branches, which is consistent with previous biological function analysis results (Fig. 3H) [27, 28]. In addition to using branched linear inference, we also use Monocle3 (version: 1.2.9) analyzes spatial trajectory differentiation based on UMAP [29]. It was also found that, starting from COLEC11+ Fibro, the tumors gradually differentiated into terminal BNIP3+ Fibro (Fig. 4A, B). We subsequently identified several genes related to differentiation trajectories, such as genes related to hypoxia and immune inflammation, whose expression increased gradually with differentiation time (Fig. 4C). Through spatial gene similarity module mining, several modules related to BNIP3+ Fibro were identified (Modules 10, 19, 20, and 31), whose functions still support special functions such as hypoxia (Fig. 4D–F). We also identified modules (Modules 5, 26, 27, and 35) related to PLA2G2A+ Fibro, whose functions are related to immune regulation, the cell matrix, etc. (Fig. 4D, G, H).

#### Identification of transcription factors in fibroblast subtypes

The phenotype of cells is shaped by the temporal and dynamic expression of different genes. The expression of genes is regulated by transcription factors. The differentiation of different fibroblast subtypes must be regulated by transcription factors. Transcription factor analysis was performed through SCENIC (1.2.4), and highly enriched transcription factors in different fibroblast subtypes were identified from the expression abundance of transcription factors (Fig. 5A–D). The results showed that MXI1, the CEBPG transcription factor, had the highest specificity and transcription factor activity among the BNIP3+ Fibro subtype. Among the PLA2G2A+ Fibro subtype, the EGR3 and CHD1 transcription factors have the highest specificity and transcription factor activity. Moreover, we identified 4 transcription factors (ATF5, CEBPG, DDIT3,

and MXI1) that are closely related to BNIP3+ Fibro differentiation and plotted the area under the curve distribution diagram of these 4 transcription factors in all the cell data, which was greater than the specificity of this transcription factor. Cells with a sexual threshold had higher corresponding regulon activity (Fig. 5E). The UMAP map showed that four transcription factors were highly expressed in the BNIP3+ Fibro subtype (Fig. 5F). Further evaluation of the overall expression of genes regulated by these transcription factors revealed that the overall expression of genes regulated by these transcription factors was significantly greater in BNIP3+ Fibro (Fig. 5G).

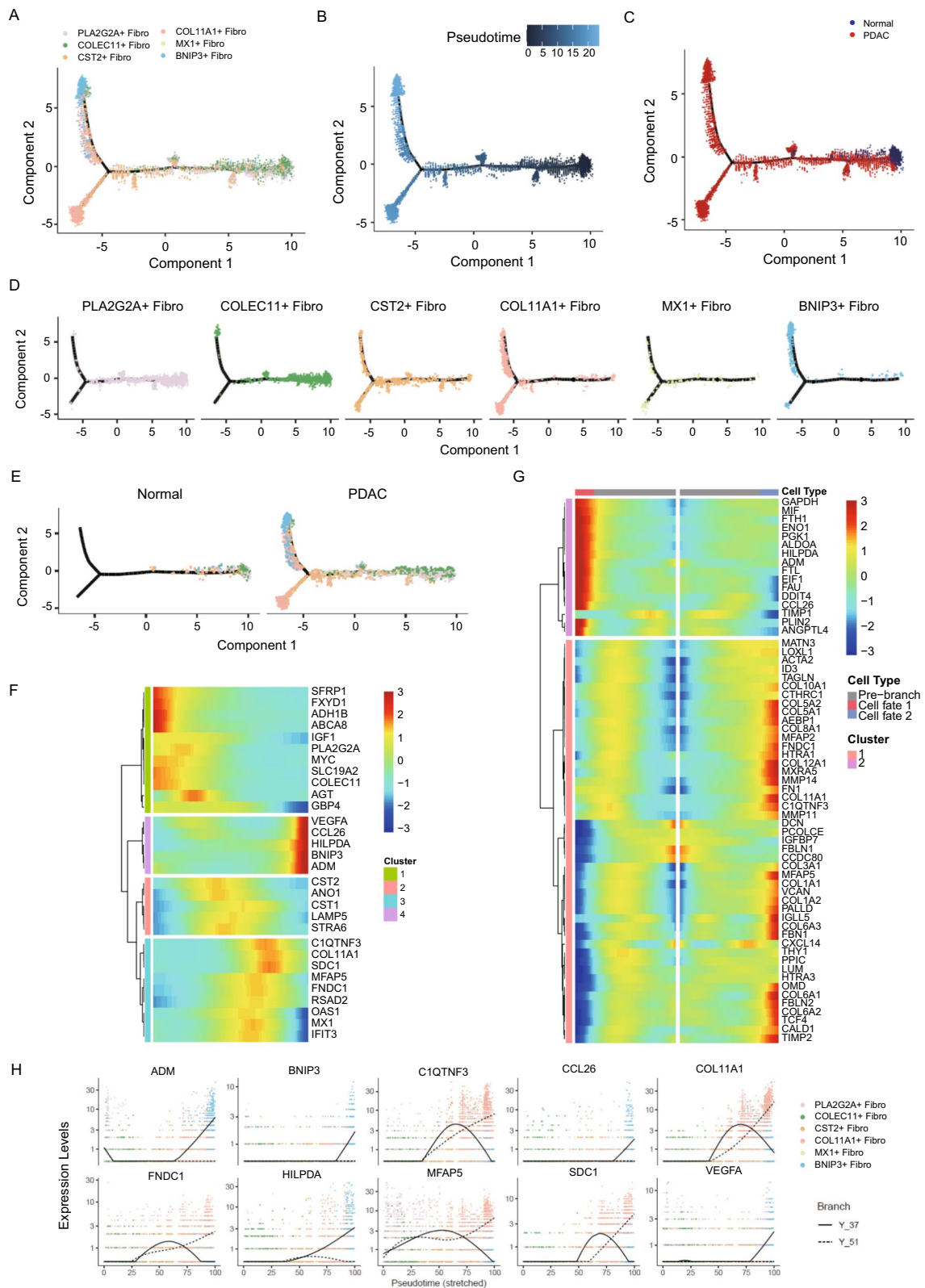
#### Analysis of the effect of fibroblast subtypes on the survival and immunotherapy response of PDAC patients

Based on the TCGA database, we analyzed the survival of PDAC patients. The top 5 genes (ADM, TGFBI, TMEM158, SLC16A3, and ERO1L) that are specific and highly expressed in the BNIP3+ Fibro were identified. The violin plot shows the expression of these genes in all fibroblast subtypes. The results showed that these genes were expressed at high levels in the BNIP3+ Fibro but were expressed at lower levels in the other subtypes (Fig. 6A). PDAC patients were divided into a high-expression group and a low-expression group according to the expression level of the landmark genes. The results showed that the overall survival rate and disease-free survival time of the high-expression group were significantly lower than those of the low-expression group (Fig. 6B, C). Moreover, we screened the top 50 genes according to their expression levels in the BNIP3+ Fibro subtype to determine the expression profile of the BNIP3+ Fibro in the transcriptome dataset. The results showed that the overall survival and disease-free survival of the high-50 signature group were significantly lower than those of the low-50 signature group (Fig. 6D, E). The above analysis revealed that the BNIP3+ Fibro subtype was significantly related to the prognosis of PDAC patients.

To further analyze the impact of the fibroblast subtype on the response of PDAC patients to

(See figure on next page.)

**Fig. 3** Monocle2 (2.20.0) was used to perform pseudochronological analysis of fibroblasts to construct and visualize cell differentiation pathways. **A–C** Distribution of fibroblast subtypes; pseudochronological time; and disease groups, differentiation trajectories. **D,E** Pseudochronological differentiation trajectories are displayed separately according to fibroblast subtype and disease group. **F** Representative heatmap of the changes in gene expression levels with pseudotime. The heatmap is arranged from left to right as the pseudotime increases. The colours from blue to red indicate that the gene expression values ranged from low to high. **G** Heatmap of gene expression patterns that play a major role at the bifurcation point of the cell differentiation trajectory. The gray symbol represents the gene expression pattern before the bifurcation point of the trajectory. Red represents the gene expression pattern of cells on the upward bifurcation trajectory (cell fate 1). Blue represents the gene expression pattern. (Cell fate 2) Gene expression patterns in cells with downward bifurcating trajectories. **H** Linear fitting graph of the change trend of representative gene expression levels with pseudotime. The solid line Y\_37 represents the branch that bifurcates upward (cell fate 1), and the dotted line Y\_51 represents the branch that bifurcates downward (cell fate 2)



**Fig. 3** (See legend on previous page.)

immunotherapy, we downloaded and integrated the largest single-cell dataset of PDAC immunotherapy, which included 43 untreated and immunotherapy PDAC patients (Fig. S2). The data of patients who received immunotherapy and treatment plan information in this dataset were integrated separately (Fig. S3). All the fibroblasts in this dataset were mapped into five previously annotated subtypes (Fig. S4A). We also focused on the differences in the proportions of patients in each subtype before and after immunotherapy, between responding patients and nonresponding patients, and between different immunotherapy regimens. Patient information, immunotherapy response, and immunotherapy regimen were mapped to fibroblasts (Fig. S4B–F). We found that the two subtypes, PLA2G2A+ Fibro and COLEC11+ Fibro, increased significantly after immunotherapy, and the better the patient responded to immunotherapy was, the greater the proportion of these two subtypes (Fig. 7A, B, D). In contrast, for the BNIP3+ Fibro, COL11A1+ Fibro and CST2+ Fibro subtypes, the greater the response to immunotherapy was, the lower the proportion was (Fig. 7D). This result is opposite to what was observed for the normal and tumor tissues in the CRA001160 dataset, reflecting the effectiveness of immunotherapy in changing the composition of the fibroblast subtype. The better the treatment effect is, the closer the sample is to the composition state of normal. Interestingly, the BNIP3+ Fibro subtype increased significantly in the poor response group, indicating that this subtype has a clear correlation with the response to immunotherapy. In patients who had a low response to treatment, this group of cells increased instead. Leading to poor prognosis (Fig. 7A, B, D). The treatment plans for different patients were integrated and mainly divided into CRT, CRTL and other treatment plans (Fig. S5A). Interestingly, the CRT treatment regimen effectively increased the proportion of PLA2G2A+ Fibro and COLEC11+ Fibro cells and reduced the proportion of COL11A1+ Fibro and CST2+ Fibro cells but did not effectively reduce the proportion of BNIP3+ Fibro cells (Fig. S5B, D). The CRTL treatment regimen effectively increased the proportion of PLA2G2A+ Fibro cells and reduced

the proportions of BNIP3+ Fibro, COL11A1+ Fibro and CST2+ Fibro cells (Fig. S5B, D). It is suggested that different treatment options have different effects on different patient subtypes, and appropriate immunotherapy needs to be selected to target different fibroblast subtypes.

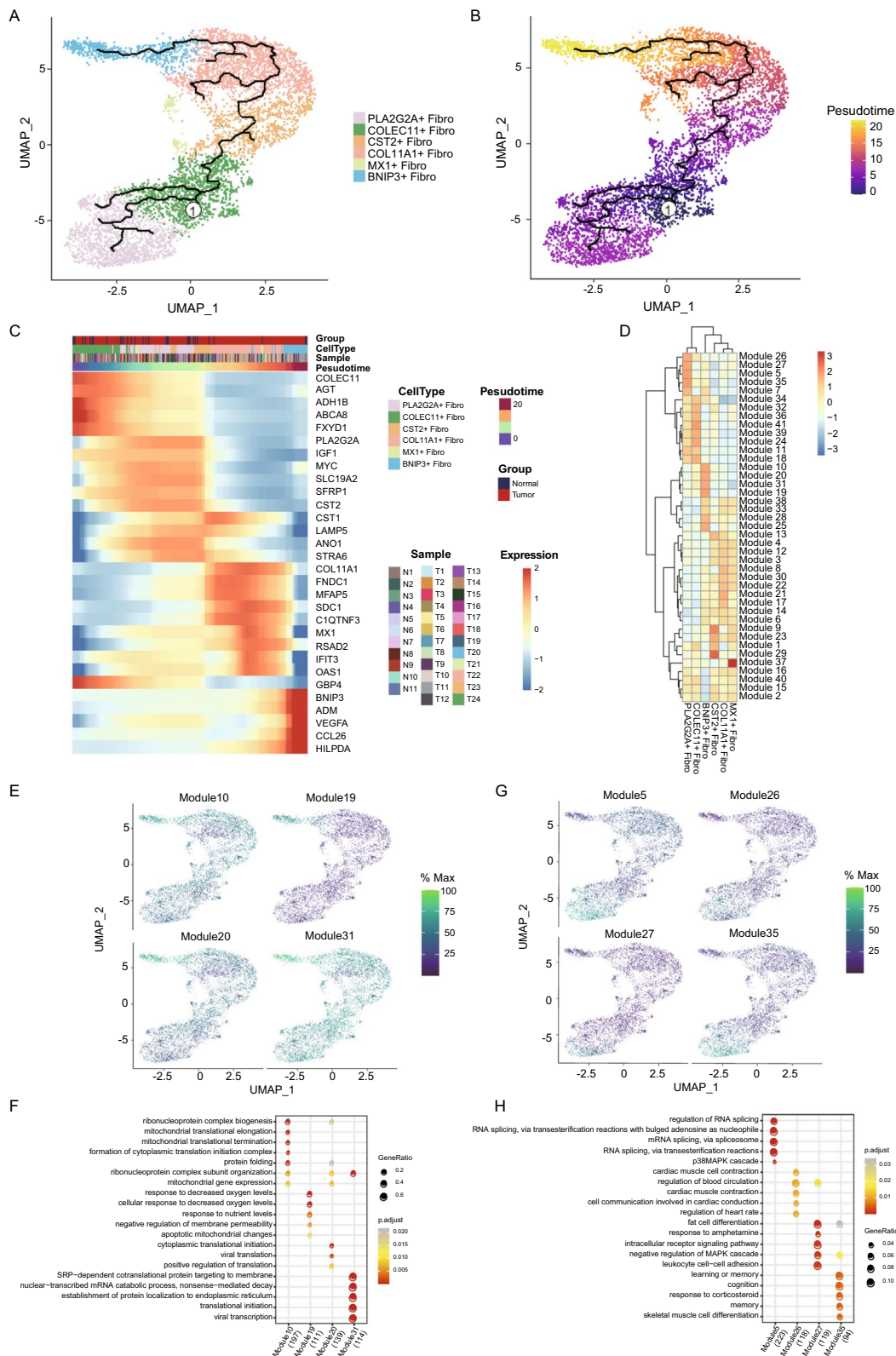
To understand the cell state transformation in the fibroblast differentiation pathway, we used the CellRank algorithm to determine the initial and terminal states of cells in the cell population through the similarity of cells combined with the RNA rate and plotted the cell differentiation trajectory [13, 30, 31]. Figure 9A shows the differentiation relationships between different fibroblast subtypes. COLEC11+ Fibro cells were in the early stage of differentiation, and BNIP3+ Fibro cells were in the late stage of differentiation (Fig. 8B, C). Figure 8D shows that random seed cells were used to mark the starting point and end point of cell differentiation. The results showed that cells at the starting point of differentiation were mainly concentrated in the COLEC11+ Fibro subtype (Fig. 8E), while cells at the end point of differentiation were concentrated in the PLA2G2A+ Fibro and BNIP3+ Fibro subtypes (Fig. 8F). The UMAP diagram shows the two-dimensional spatial distribution of six representative genes in fibroblasts. These genes were mainly concentrated in COL11A1+ Fibro and BNIP3+ Fibro (Fig. 8G). The pseudochronal differentiation expression trend chart showed that the expression of representative genes in the terminal stage of differentiation was significantly increased in the BNIP3+ Fibro subtype, while the increase in the PLA2G2A+ Fibro subtype was not obvious (Fig. 8H).

#### BNIP3+ CAFs enhanced pancreatic cancer cell proliferation, migration and invasion

CAFs were isolated from surgically removed pancreatic cancer tissues and paracancerous tissues. Immunofluorescence results showed that the expression of BNIP3 and FAP in fibroblasts isolated from tumor tissues was significantly higher than that in paracancerous tissues (Fig. 9A). Then BNIP3 was knocked down in CAFs by lentiviral transfection (Fig. 9B). The expression of hypoxia marker proteins HIF1A and GLUT1 was significantly decreased in

(See figure on next page.)

**Fig. 4** Monocle3 (1.2.9) plots fibroblast subtype differentiation trajectories and calculates coexpressed gene modules in high-dimensional space. **A,B** distribution of cell subtypes and pseudotimes in UMAP plots; curves represent inferred differentiation pathways. **C** Representative heatmap of the changes in gene expression levels with pseudotime. The heatmap is arranged from left to right as the pseudotime increases. The colours from blue to red indicate the gene expression values from low to high. **D** Heatmap representing the expression of different gene modules in different subtypes based on Moran's I calculation. **E,F** UMAP plot showing the expression distribution of four gene modules related to the BNIP3+ Fibro subtype (left panel) and the PLA2G2A+ Fibro subtype (right panel) in all fibroblasts. **G,H** Bubble heatmap showing the functional enrichment of four gene modules related to the BNIP3+ Fibro subtype (left picture) and PLA2G2A+ Fibro subtype (right picture)



**Fig. 4** (See legend on previous page.)

the BNIP3-KD group (Fig. 9C). At the same time, multiple inflammatory-related factors (IL1, IL6, IL8, TNF- $\alpha$ ) were also inhibited to varying degrees in the BNIP3-KD group (Fig. 9D). After indirect co-culture of CAFs and pancreatic cancer cells (PANC-1, BxPC-3) (Fig. 9E), the proliferation, migration and invasion abilities of pancreatic cancer cells were significantly enhanced. However, after knocking down BNIP3 in CAFs, this enhancement was weakened to varying degrees (Fig. 9F–H).

## Discussion

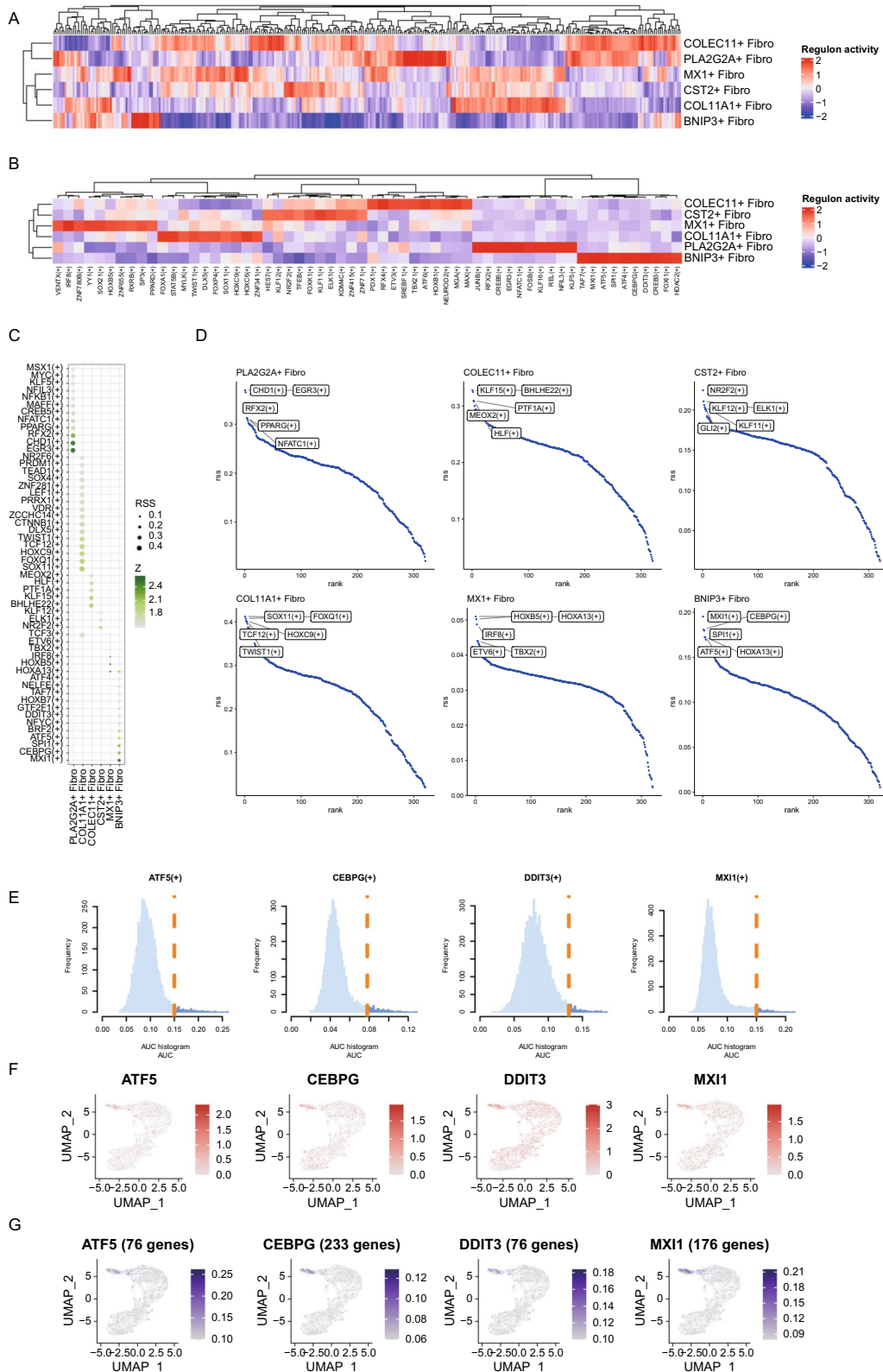
CAFs are the predominant cell type in the stroma of pancreatic cancer, known for producing extensive extracellular matrix components that contribute to stromal fibrosis. They serve as a crucial source of the dense tissue architecture characteristic of PDAC [32]. CAFs in pancreatic cancer exhibit a variety of sources, phenotypes, and functions, which dynamically shift as the tumor progresses [33]. Studies have shown that CAFs engage in extensive and intricate interactions with tumor cells and other stromal cells, playing an active role in the malignant advancement of tumors [34]. However, the specifics of how these interactions govern pancreatic cancer and the ways in which CAFs can be targeted for effective treatment are still not fully understood. CAFs have the dual ability to both stimulate and inhibit tumor growth, suggesting that functionally distinct fibroblast populations may coexist within the tumor microenvironment [35, 36]. The high heterogeneity of CAFs, attributable to their origins from diverse fibroblast lineages and cell types, may be the reason for their complex nature, rendering them challenging to precisely define and characterize [37].

As an emerging sequencing technology, scRNA-seq has shown irreplaceable advantages in revealing the heterogeneity of fibroblasts and mapping cell differentiation trajectories [38]. We divided fibroblasts into 6 subtypes through cell type annotation and found that COLEC11+ Fibro was highly expressed in normal pancreatic tissue, while the BNIP3+ Fibro subtype was highly expressed in PDAC tissue. PDAC usually exhibits a high degree of fibrosis because fibroblasts accumulate in the matrix of the tissue microenvironment, which also leads to the hypovascularization and hypoxic phenotype of PDAC.

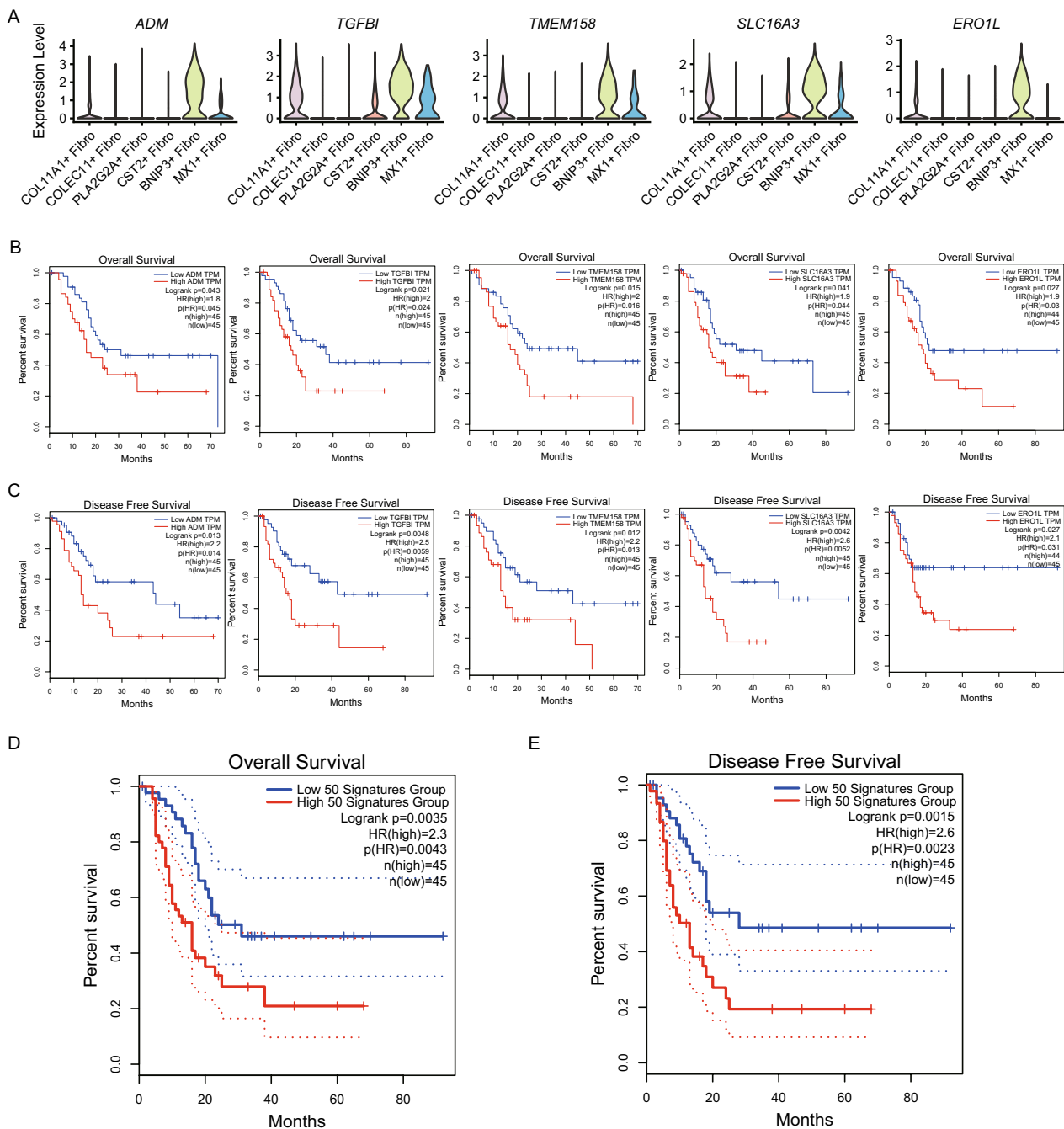
These conditions create infiltration barriers for T cells and myeloid cells in the PDAC tumor microenvironment. By analyzing the enrichment of functional features and differentiation trajectories, we found that hypoxia pathway and glycolysis pathway were significantly enriched in the BNIP3+ Fibro subtype and in the hypoxia-inducible genes BNIP3 [39], ADM [23], and VEGFA [26]. Similarly, the expression of HILDPA in this subtype showed a clear upward trend with differentiation time [40]. Hypoxia can lead to the activation of ATF transcription factors, thereby upregulating the expression of LC3B and ATG5 and maintaining high levels of autophagic degradation. These molecular mechanisms promote the autophagy-mediated survival of tumor cells under hypoxic conditions [41]. It has been reported that activating the transcription factor ATF3 played an important role in KRAS-mediated malignant process in PDAC. Knocking out ATF3 may prevent the occurrence of PDAC [42]. Gene regulatory network analysis revealed that the transcriptional activity of ATF5 was significantly increased in the BNIP3+ Fibro subtype, suggesting that activation of ATF5 under hypoxia may promote the progression of PDAC. We found in vitro experiments that knocking out BNIP3 in CAF resulted in a decrease in the expression of hypoxia markers, demonstrating the positive correlation between BNIP3 and hypoxia. Moreover, a variety of inflammation-related pathways, such as the oxidative phosphorylation, IL6/JAK/STAT3, and TNFA/NFKB pathways, were also enriched in the BNIP3+ Fibro subtype, reflecting the inflammatory characteristics of this subtype. After knocking out BNIP3 in vitro experiments, the expression of inflammatory factors in CAF decreased, demonstrating the pro-inflammatory effect of BNIP3. TGF- $\beta$  is the most important and critical signaling factor that promotes the activation and conversion of normal long-fiber cells into CAFs [43]. CAFs usually transdifferentiate through TGFB1 signaling, produce contractile cells, express  $\alpha$ -SMA, secrete collagen-rich extracellular matrix, and promote the development of various malignant tumors [44]. In this study, TGFB1, a signature gene of the BNIP3+ Fibro tissue, was highly expressed and significantly associated with OS and DFS in PDAC patients. A large number of studies have shown that fibroblasts

(See figure on next page.)

**Fig. 5** Fibroblast subtype transcription factor identification and regulon activity scoring. **A** The activity distribution of all transcription factors. **B** The activity distribution of the top 10 transcription factors highly expressed in each subtype. **C** Bubble heatmap showing the activity of specific transcription factors in each fibroblast subtype. The size of the dot indicates the degree of specificity (RSS: regulation specificity score), and the color of the dot indicates the level of activity (RAS: regulon activity score). **D** Scatter plot showing the specific ranking of transcription factors in each fibroblast subtype. The horizontal axis represents the ranking position, and the vertical axis represents the specificity. **E** Distribution diagram of AUC values of four representative transcription factors screened in the BNIP3+ Fibro subtype in all cells. The dotted line, which is the AUC value, is the threshold for determining that cells are specific for this transcription factor. **F** Expression distribution of transcription factors on the UMAP plot. **G** Expression distribution of transcription factors and target gene sets on the UMAP plot



**Fig. 5** (See legend on previous page.)

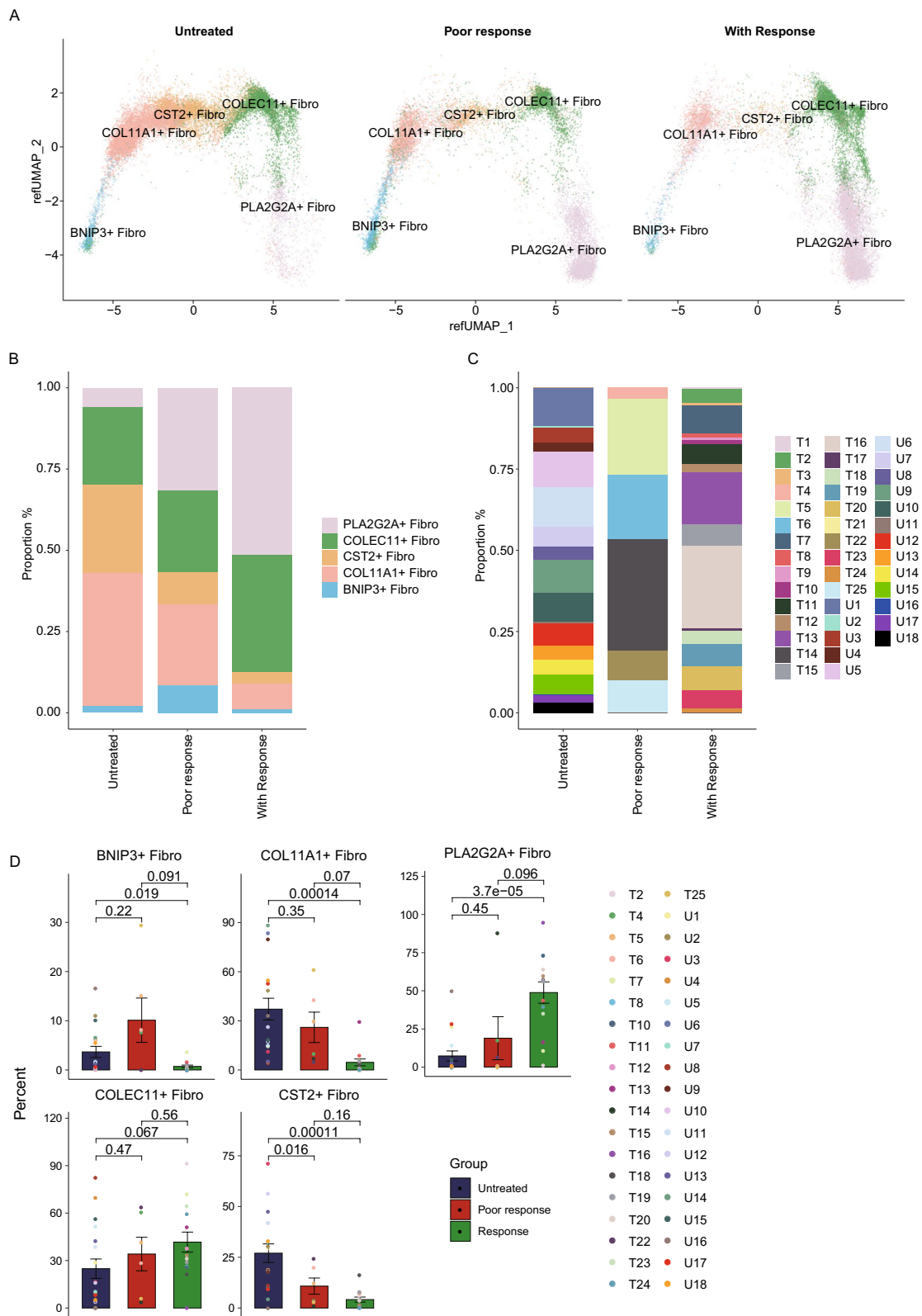


**Fig. 6** Impact of the BNIP3+ Fibro subtype on the survival of PDAC patients. **A** Violin plot plots showing the expression of specific genes in fibroblast subtypes. **B,C** Kaplan–Meier curves showing the impact of specific gene expression levels on patient overall survival and disease-free survival. **D,E** Kaplan–Meier curves showing overall survival and disease-free survival of PDAC patients in the low- and high-grade Top 50 signature groups

(See figure on next page.)

**Fig. 7** Correlation between the fibroblast subtype and immunotherapy response in PDAC patients. **A** UMPA diagram showing the two-dimensional spatial distribution of different fibroblast subtypes in the untreated, poorly responsive and with responsive groups. **B** Percentage histogram showing the proportions of different fibroblast subtypes in the untreated, poorly responsive and with responsive groups. **C** Percentage bar chart showing the sample origin of fibroblasts in the untreated, poorly responsive and with responsive groups. **D** Differential expression of different fibroblast subtypes in the untreated, poor response and with response groups





were highly plastic and multipotent [45]. This implies that within the dynamic regulation of the tumor microenvironment, a fibroblast subtype can transform into another subtype through transdifferentiation, leading to alterations in the fibroblast's function and characteristics [46]. This study analyzed the state transition of fibroblasts based on the CellRank algorithm and revealed that COLEC11+ fibroblasts were the starting point of differentiation. Most of the normal parts differentiate into PLA2G2A+ Fibro, and abnormal pathological conditions differentiate into CST2+ Fibro, MX1+ Fibro, COL11A1+ Fibro and BNIP3+ Fibro. Among them, PLA2G2A+ Fibro and BNIP3+ Fibro were used as the end points of differentiation. Metabolic pathway analysis revealed that these two terminally differentiated subtypes not only had significant differences in metabolic functions but also exhibited complete resistance to immunotherapy in PDAC patients.

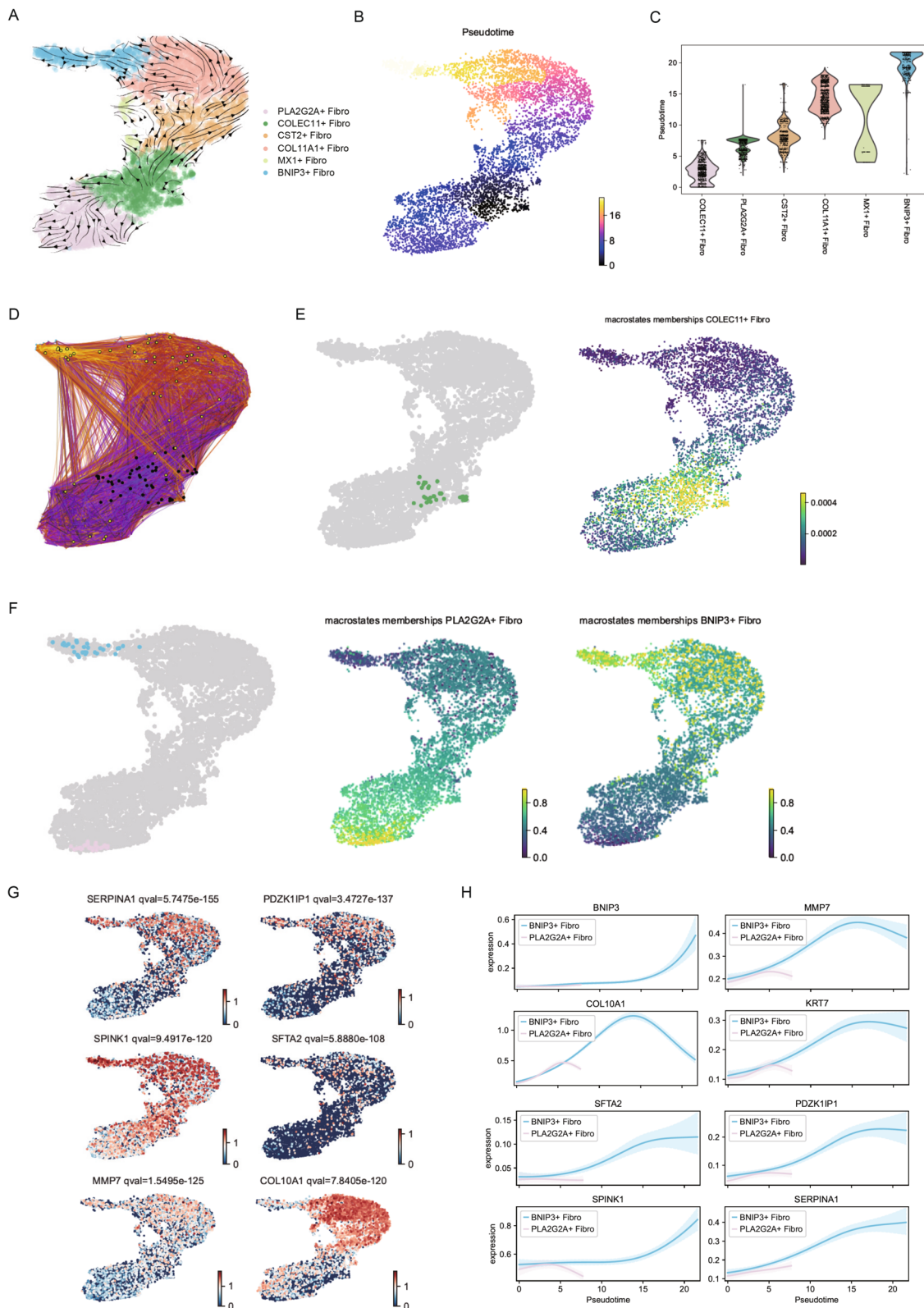
CAFs are the major source of extracellular matrix in the tumor microenvironment. There is evidence that a dense collagen matrix is resistant to PDAC treatment [47]. It has been reported that in samples from patients with PDAC who underwent neoadjuvant therapy, there was an increase in inflammation-related CAF levels. Particularly in patients treated with a combination of gemcitabine and the protein paclitaxel, there was an observed upregulation of the metallothionein gene in iCAFs. This variation could be associated with resistance to chemotherapy [48]. Our study demonstrated that the

composition of fibroblast subtypes underwent changes following neoadjuvant therapy in patients with PDAC. BNIP3+ Fibro expression was significantly increased in the poor response group, while the proportion was significantly decreased in the with response group. After we co cultured CAFs with pancreatic cancer cells indirectly, the proliferation, migration and invasion of pancreatic cancer cells were enhanced. However, after knocking out BNIP3 in CAF, this trend was reversed. Moreover, the expression of MMP7 in the BNIP3+ Fibro subtype also gradually increased with differentiation time. Study revealed that matrix metalloproteinases inhibitors could be effectively used to treat pancreatitis-related malignant process [49]. Moreover, it has been found that TGF- $\beta$ -driven LRRC15+ CAFs were associated with low responsiveness to anti-PD-L1 immunotherapy in PDAC [7]. This work revealed that understanding the characteristics of specific CAF subtypes was critical for improving immunotherapy responsiveness in patients with PDAC. In addition, although BNIP3+ Fibro may lead to a low responsiveness to immunotherapy, it does show a better response to CRTL treatment regimens. Responsiveness also provides a theoretical basis for the selection of clinical immunotherapy options.

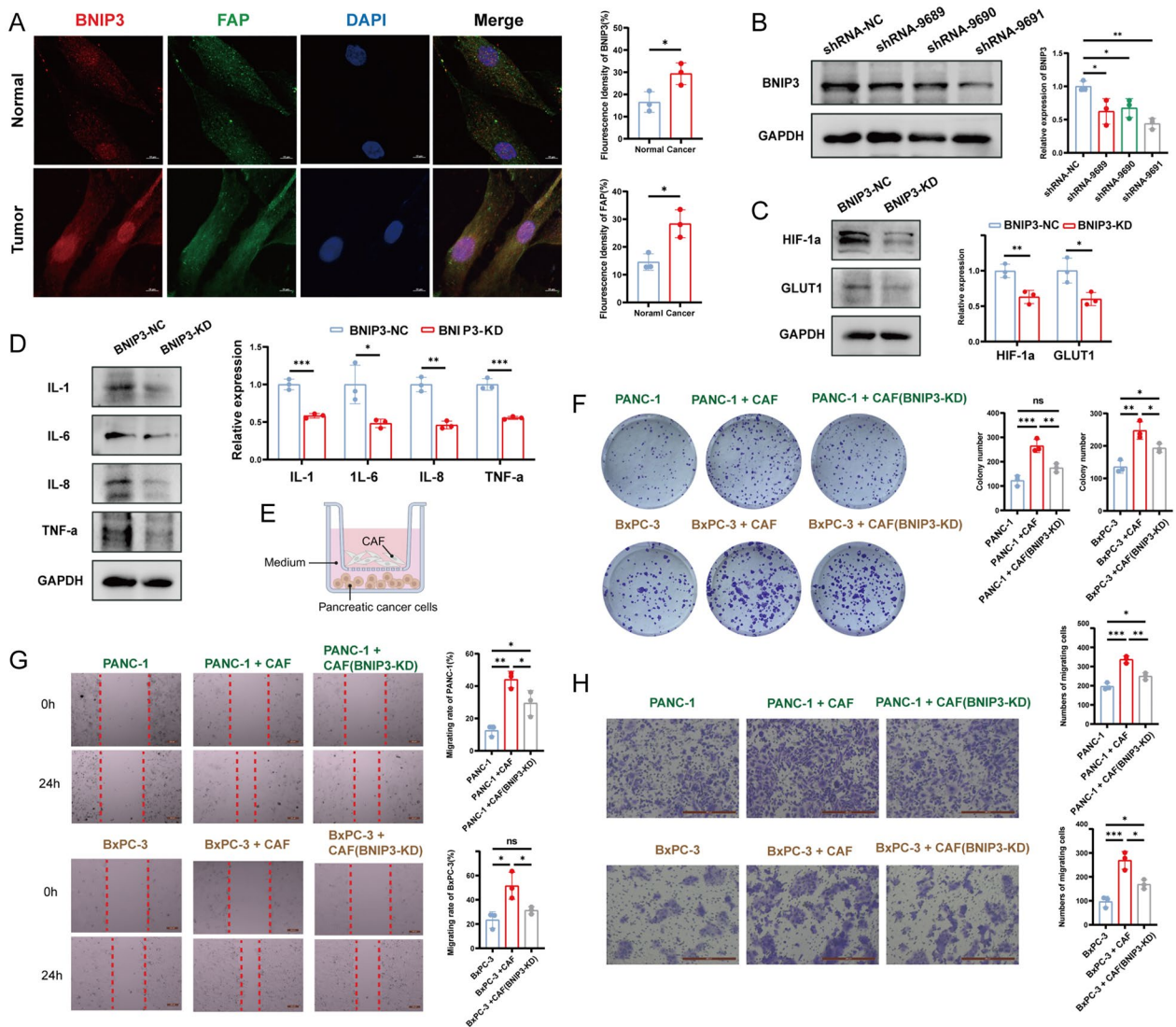
In conclusion, this study characterized the heterogeneity and pluripotency of fibroblasts in PDAC and identified a terminally differentiated BNIP3+ Fibro subtype. The gene expression pattern of this subtype can affect the responsiveness of PDAC patients to immunotherapy and

(See figure on next page.)

**Fig. 8** Cell state transition analysis based on CellRank. **A** Differentiation trajectories of different fibroblast subtypes. The arrow indicates the fibroblast differentiation trajectory predicted by RNA rate analysis. **B** Color depth shows the differentiation period. **C** The differentiation time for each cell was calculated, and the overall differentiation time of the different fibroblast subtypes was determined, which is equivalent to the quantitative statistics of **A** and **B**. **D** The starting point and end point of differentiation are shown for the randomly seeded cells based on differentiation time. The black dots indicate cells at the starting point, and the yellow dots indicate the positions of these cells from the starting point to the terminal state. **E** Representative starting cells and inferred origins. **F** Representative endpoint cells and inferred endpoints. The F plot infers two end points. **G** Expression of representative genes in the UMAP plot. **H** Expression trends of representative genes in two different endpoint state cell populations over differentiation time



**Fig. 8** (See legend on previous page.)



**Fig. 9** In vitro and in vitro experiments of BNIP3+ CAFs and pancreatic cancer cells. **A** Immunofluorescence experiments of fibroblasts derived from pancreatic cancer tissues and adjacent tissues. **B** WB evaluation of BNIP3 lentiviral knockdown efficiency. **C** Changes in the levels of HIF- $\alpha$  and GLUT1 after BNIP3 knockdown. **D** Changes in the levels of inflammatory factors (IL1, IL6, IL8, TNF- $\alpha$ ) after BNIP3 knockdown. **E** Schematic diagram of indirect culture of CAFs and pancreatic cancer cells. **F–H** Plate cloning experiment, scratch experiment, Transwell experiment results and statistical analysis. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; \*\*\*\*  $p < 0.0001$

lead to poor prognosis by forming a hypoxic and inflammatory tumor microenvironment.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-024-05674-x>.

**Additional file 1: Fig. S1.** Cell type annotation and visual display of single-cell data. **A** UMAP plot showing single-cell data; each point represents a cell, and the colours correspond to different patient sample sources. **B** UMAP diagram showing the distribution of 11 cell clusters in normal and tumor samples. **C** The level of expression of the iconic genes in 11 cell clusters in the form of a violin plot. **D** UMAP diagram showing

the 9 most highly expressed landmark genes in the single-cell data. **E** Based on the sample source, a UMAP diagram was generated to display the distribution of fibroblasts in two-dimensional space. **F, G** Bubble charts and heatmaps showing the expression of signature genes in 6 fibroblast subtypes.

**Additional file 2: Fig. S2.** Cell type annotation of single-cell data for PDAC immunotherapy. **A** Unsupervised clustering of all cells from 43 patients with pancreatic cancer. The UMAP diagram shows their two-dimensional spatial distribution. Different cell clusters are represented by different colors. **B** UMAP graph showing the expression of landmark genes in all the cell data. **C** Bubble chart showing the percentage and abundance of signature genes expressed in different cell clusters.

**Additional file 3: Fig. S3.** Annotations based on patient immunotherapy response and immunotherapy regimen in the dataset. **A** Cell

class annotation of single-cell data based on the expression of landmark genes. **B** All the cell data were labeled according to the source of the cell sample. **C, D** All cells are labeled according to the patient's response to immunotherapy. **E, F** All cells are labeled according to the patient's choice of immunotherapy regimen. **G** UMAP graph showing the four most highly expressed landmark genes in the single-cell data.

**Additional file 4: Fig. S4.** Fine subtype division of fibroblasts and annotation according to immunotherapy status. **A** Fibroblasts were extracted from the subtype and integrated with the subtype of the CRA001160 single-cell dataset. **B** Fibroblast subtypes are annotated according to the source of the cell samples. **C, D** Fibroblast subtypes are annotated according to the patient's choice of immunotherapy regimen. **E, F** Fibroblast subtypes are annotated according to the patient's choice of immunotherapy regimen.

**Additional file 5: Fig. S5.** Correlations between fibroblast subtypes and immunotherapy regimens in PDAC patients. **A** UMPA diagram showing the two-dimensional spatial distribution of different fibroblast subtypes in the untreated, CRT, CRTL and other groups. **B** Percentage histogram showing the proportions of different fibroblast subtypes in the Untreated, CRT, CRTL and other groups. **C** Percentage bar graph showing the sample origin of fibroblasts in the untreated, CRT, CRTL and other groups. **D** Differential analysis of the expression of different fibroblast subtypes in the Untreated, CRT, CRTL and other groups.

**Additional file 6: Table. S1.** Interference sequences of Vector.

**Additional file 7.**

#### Acknowledgements

All authors thank open-source TCGA, GSA, and GEO databases for their platforms and contributors for uploading their meaningful datasets.

#### Author contributions

H.Y. and B.G. contributed to conception and designed the experiments; H.Y., B.S., and W.L. collected and analyzed the data; H.Y. and B.S. conducted the molecular biology experiments; H.Y., B.G., B.S. and W.L. drafted a manuscript; B.G. and G.H. revised the manuscript; all authors read and approved the manuscript. H.Y. is responsible for the overall content as guarantor.

#### Funding

This work was supported by the Beijing Natural Science Foundation (No. 7232194) and Peking University People's Hospital Research and Development Funds (No. RDJP2022-13).

#### Data availability

The original contributions presented in the study are publicly available.

#### Declarations

#### Competing interests

The authors have declared that no competing interest exists.

#### Author details

<sup>1</sup>Department of Hernia and Abdominal Wall, Peking University People's Hospital, Beijing, China. <sup>2</sup>Department of General Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, China. <sup>3</sup>Department of General Surgery, Jinchuan Hospital of Fudan University, Shanghai, China. <sup>4</sup>Department of Oncology Surgery, Harbin Medical University Cancer Hospital, Harbin, China.

Received: 15 March 2024 Accepted: 4 September 2024

Published online: 14 October 2024

#### References

- Lee Y, Ju A, Choi H, Kim J, Kim E, Kim T, et al. Rationally designed redirection of natural killer cells anchoring a cytotoxic ligand for pancreatic cancer treatment. *J Control Release*. 2020;326:310–23.

- Siegel R, Miller K, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7–30.
- Ho W, Jaffee E, Zheng L. The tumour microenvironment in pancreatic cancer—clinical challenges and opportunities. *Nat Rev Clin Oncol*. 2020;17(9):527–40.
- Halbrook C, Lyssiotis C, Pasca di Magliano A, Maitra A. Pancreatic cancer: advances and challenges. *Cell*. 2023;186(8):1729–54.
- Velez-Delgado A, Donahue K, Brown K, Du W, Irizarry-Negron V, Menjivar R, et al. Extrinsic KRAS signaling shapes the pancreatic microenvironment through fibroblast reprogramming. *Cell Mol Gastroenterol Hepatol*. 2022;13(6):1673–99.
- Ren B, Cui M, Yang G, Wang H, Feng M, You L, et al. Tumor microenvironment participates in metastasis of pancreatic cancer. *Mol Cancer*. 2018;17(1):108.
- Dominguez C, Müller S, Keerthivasan S, Koeppen H, Hung J, Gierke S, et al. Single-cell RNA sequencing reveals stromal evolution into LRRC15 myofibroblasts as a determinant of patient response to cancer immunotherapy. *Cancer Discov*. 2020;10(2):232–53.
- Caligiuri G, Tuveson D. Activated fibroblasts in cancer: perspectives and challenges. *Cancer Cell*. 2023;41(3):434–49.
- Liu H, Shi Y, Qian F. Opportunities and delusions regarding drug delivery targeting pancreatic cancer-associated fibroblasts. *Adv Drug Deliv Rev*. 2021;172:37–51.
- Biffi G, Tuveson D. Diversity and biology of cancer-associated fibroblasts. *Physiol Rev*. 2021;101(1):147–76.
- Peng J, Sun B, Chen C, Zhou J, Chen Y, Chen H, et al. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res*. 2019;29(9):725–38.
- Hwang W, Jagadeesh K, Guo J, Hoffman H, Yadollahpour P, Reeves J, et al. Single-nucleus and spatial transcriptome profiling of pancreatic cancer identifies multicellular dynamics associated with neoadjuvant treatment. *Nat Genet*. 2022;54(8):1178–91.
- Lange M, Bergen V, Klein M, Setty M, Reuter B, Bakhti M, et al. Cell Rank for directed single-cell fate mapping. *Nat Methods*. 2022;19(2):159–70.
- Gao S, Yan L, Wang R, Li J, Yong J, Zhou X, et al. Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. *Nat Cell Biol*. 2018;20(6):721–34.
- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411–20.
- Smirnova N, Riemondy K, Bueno M, Collins S, Suresh P, Wang X, et al. Single-cell transcriptome mapping identifies a local, innate B cell population driving chronic rejection after lung transplantation. *JCI Insight*. 2022;7(18): e156648.
- Devlin J, Axelrad J, Hine A, Chang S, Sarkar S, Lin J, et al. Single-cell transcriptional survey of ileal-anal pouch immune cells from ulcerative colitis patients. *Gastroenterology*. 2021;160(5):1679–93.
- Valenzi E, Bahudhanapati H, Tan J, Tabib T, Sullivan D, Nouraei M, et al. Single-nucleus chromatin accessibility identifies a critical role for TWIST1 in IPF myofibroblast activity. *Eur Respir J*. 2023;62(1):2200474.
- Paik D, Tian L, Lee J, Sayed N, Chen I, Rhee S, et al. Large-scale single-cell RNA-seq reveals molecular signatures of heterogeneous populations of human induced pluripotent stem cell-derived endothelial cells. *Circ Res*. 2018;123(4):443–50.
- de Visser K, Joyce J. The evolving tumor microenvironment: from cancer initiation to metastatic outgrowth. *Cancer Cell*. 2023;41(3):374–403.
- Mao X, Xu J, Wang W, Liang C, Hua J, Liu J, et al. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. *Mol Cancer*. 2021;20(1):131.
- Kanzaki R, Pietras K. Heterogeneity of cancer-associated fibroblasts: opportunities for precision medicine. *Cancer Sci*. 2020;111(8):2708–17.
- Han J, Wan Q, Seo G, Kim K, El Baghdady S, Lee J, et al. Hypoxia induces adrenomedullin from lung epithelia, stimulating ILC2 inflammation and immunity. *J Exp Med*. 2022;219(6): e20211985.
- Guttman-Yassky E, Bissonnette R, Ungar B, Suárez-Fariñas M, Ardeleanu M, Esaki H, et al. Dupilumab progressively improves systemic and cutaneous abnormalities in patients with atopic dermatitis. *J Allergy Clin Immunol*. 2019;143(1):155–72.
- Povero D, Johnson S, Liu J. Hypoxia, hypoxia-inducible gene 2 (HIG2)/HILPDA, and intracellular lipolysis in cancer. *Cancer Lett*. 2020;493:71–9.

26. Zhou W, Liu K, Zeng L, He J, Gao X, Gu X, et al. Targeting VEGF-A/VEGFR2 Y949 signaling-mediated vascular permeability alleviates hypoxic pulmonary hypertension. *Circulation*. 2022;146(24):1855–81.
27. Micallef P, Vujičić M, Wu Y, Peris E, Wang Y, Chanclón B, et al. C1QTNF3 is upregulated during subcutaneous adipose tissue remodeling and stimulates macrophage chemotaxis and M1-like polarization. *Front Immunol*. 2022;13: 914956.
28. Duan Y, Zhang X, Ying H, Xu J, Yang H, Sun K, et al. Targeting MFAP5 in cancer-associated fibroblasts sensitizes pancreatic cancer to PD-L1-based immunotherapy via remodeling the matrix. *Oncogene*. 2023;42(25):2061–73.
29. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim D, Hill A, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566(7745):496–502.
30. Bergen V, Lange M, Peidli S, Wolf F, Theis F. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol*. 2020;38(12):1408–14.
31. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. *Nature*. 2018;560(7719):494–8.
32. Neuzillet C, Tijeras-Raballand A, Ragulan C, Cros J, Patil Y, Martinet M, et al. Inter- and intra-tumoural heterogeneity in cancer-associated fibroblasts of human pancreatic ductal adenocarcinoma. *J Pathol*. 2019;248(1):51–65.
33. Helms E, Berry M, Chaw R, DuFort C, Sun D, Onate M, et al. Mesenchymal lineage heterogeneity underlies nonredundant functions of pancreatic cancer-associated fibroblasts. *Cancer Discov*. 2022;12(2):484–501.
34. Wang Y, Liang Y, Xu H, Zhang X, Mao T, Cui J, et al. Single-cell analysis of pancreatic ductal adenocarcinoma identifies a novel fibroblast subtype associated with poor prognosis but better immunotherapy response. *Cell Discov*. 2021;7(1):36.
35. McAndrews K, Chen Y, Darpolor J, Zheng X, Yang S, Carstens J, et al. Identification of functional heterogeneity of carcinoma-associated fibroblasts with distinct IL6-mediated therapy resistance in pancreatic cancer. *Cancer Discov*. 2022;12(6):1580–97.
36. Murray E, Menezes S, Henry J, Williams J, Alba-Castellón L, Baskaran P, et al. Disruption of pancreatic stellate cell myofibroblast phenotype promotes pancreatic tumor invasion. *Cell Rep*. 2022;38(4): 110227.
37. Luo H, Xia X, Huang L, An H, Cao M, Kim G, et al. Pan-cancer single-cell analysis reveals the heterogeneity and plasticity of cancer-associated fibroblasts in the tumor microenvironment. *Nat Commun*. 2022;13(1):6619.
38. Lavie D, Ben-Shmuel A, Erez N, Scherz-Shouval R. Cancer-associated fibroblasts in the single-cell era. *Nat Cancer*. 2022;3(7):793–807.
39. Scharping N, Rivadeneira D, Menk A, Vignali P, Ford B, Rittenhouse N, et al. Mitochondrial stress induced by continuous stimulation under hypoxia rapidly drives T cell exhaustion. *Nat Immunol*. 2021;22(2):205–15.
40. Barkley D, Moncada R, Pour M, Liberman D, Dryg I, Werba G, et al. Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat Genet*. 2022;54(8):1192–201.
41. Xia H, Green D, Zou W. Autophagy in tumour immunity and therapy. *Nat Rev Cancer*. 2021;21(5):281–97.
42. Azizi N, Toma J, Martin M, Khalid M, Mousavi F, Win P, et al. Loss of activating transcription factor 3 prevents KRAS-mediated pancreatic cancer. *Oncogene*. 2021;40(17):3118–35.
43. Zhang X, Peng L, Luo Y, Zhang S, Pu Y, Chen Y, et al. Dissecting esophageal squamous-cell carcinoma ecosystem by single-cell transcriptomic analysis. *Nat Commun*. 2021;12(1):5291.
44. Ford K, Hanley C, Mellone M, Szyndralewicz C, Heitz F, Wiesel P, et al. NOX4 inhibition potentiates immunotherapy by overcoming cancer-associated fibroblast-mediated CD8 T-cell exclusion from tumors. *Can Res*. 2020;80(9):1846–60.
45. Sahai E, Astsaturov I, Cukierman E, DeNardo D, Egeblad M, Evans R, et al. A framework for advancing our understanding of cancer-associated fibroblasts. *Nat Rev Cancer*. 2020;20(3):174–86.
46. Chen Y, McAndrews K, Kalluri R. Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol*. 2021;18(12):792–804.
47. Kim D, Jeong J, Lee D, Hyeon D, Park G, Jeon S, et al. PD-L1-directed PlGF/VEGF blockade synergizes with chemotherapy by targeting CD141 cancer-associated fibroblasts in pancreatic cancer. *Nat Commun*. 2022;13(1):6292.
48. Cui Zhou D, Jayasinghe R, Chen S, Herndon J, Iglesia M, Navale P, et al. Spatially restricted drivers and transitional cell populations cooperate with the microenvironment in untreated and chemo-resistant pancreatic cancer. *Nat Genet*. 2022;54(9):1390–405.
49. Liou G, Döppler H, Necela B, Krishna M, Crawford H, Raimondo M, et al. Macrophage-secreted cytokines drive pancreatic acinar-to-ductal metaplasia through NF- $\kappa$ B and MMPs. *J Cell Biol*. 2013;202(3):563–77.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.