



Universiteit  
Leiden  
The Netherlands

## **Serum levels of iCAF-derived osteoglycin predict favorable outcome in pancreatic cancer**

Dings, M.P.G.; Manoukian, P.; Waasdorp, C.; Quik, J.S.E.; Strijker, M.; Lodestijn, S.C.; ... ; Bijlsma, M.F.



### **Citation**

Dings, M. P. G., Manoukian, P., Waasdorp, C., Quik, J. S. E., Strijker, M., Lodestijn, S. C., ... Bijlsma, M. F. (2022). Serum levels of iCAF-derived osteoglycin predict favorable outcome in pancreatic cancer. *International Journal Of Cancer*, 152(3), 511-523.  
doi:10.1002/ijc.34276

Version: Publisher's Version  
License: [Creative Commons CC BY-NC 4.0 license](#)  
Downloaded from: <https://hdl.handle.net/1887/3566997>

**Note:** To cite this publication please use the final published version (if applicable).

# Serum levels of iCAF-derived osteoglycin predict favorable outcome in pancreatic cancer

Mark P. G. Dings<sup>1,2,3</sup>  | Paul Manoukian<sup>1,2,3</sup> | Cynthia Waasdorp<sup>1,2,3</sup> |  
 Judith S. E. Quik<sup>1</sup> | Marin Strijker<sup>4</sup> | Sophie C. Lodestijn<sup>1,2,3</sup> |  
 Sanne M. van Neerven<sup>1,2,3</sup> | Leandro F. Moreno<sup>1,2,3</sup> | Rodrigo Leite de Oliveira<sup>1,2,5</sup> |  
 Bert A. Bonsing<sup>6</sup> | Marco J. Bruno<sup>7</sup> | Olivier R. Busch<sup>4</sup> | Michael Doukas<sup>8</sup> |  
 Casper H. van Eijck<sup>9</sup> | Nadia Haj Mohammad<sup>10</sup> | Ignace H. de Hingh<sup>11</sup> |  
 Quintus I. Molenaar<sup>12</sup> | Marc G. Besselink<sup>4</sup> | Louis Vermeulen<sup>1,2,3</sup> |  
 Jan Paul Medema<sup>1,2,3</sup> | Hanneke W. M. van Laarhoven<sup>1,3,13</sup> |  
 Maarten F. Bijlsma<sup>1,2,3</sup> 

<sup>1</sup>Center for Experimental and Molecular Medicine, Laboratory of Experimental Oncology and Radiobiology, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands

<sup>2</sup>Oncode Institute, Amsterdam, The Netherlands

<sup>3</sup>Cancer Center Amsterdam, Imaging and Biomarkers, Amsterdam, The Netherlands

<sup>4</sup>Department of Surgery, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands

<sup>5</sup>CRISPR Expertise Center, Cancer Center Amsterdam, Amsterdam University Medical Center, Amsterdam, The Netherlands

<sup>6</sup>Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands

<sup>7</sup>Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>8</sup>Department of Pathology, Erasmus MC-University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>9</sup>Department of Surgery, Erasmus MC-University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>10</sup>Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

<sup>11</sup>Department of Surgery, Catharina Hospital, Eindhoven, The Netherlands

<sup>12</sup>Department of Surgery, Regional Academic Cancer Center Utrecht, University Medical Center Utrecht and St Antonius Hospital, Nieuwegein, The Netherlands

<sup>13</sup>Department of Medical Oncology, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands

## Correspondence

Maarten F. Bijlsma, Laboratory for Experimental Oncology and Radiobiology, Center for Experimental and Molecular Medicine, Amsterdam UMC, Cancer Center Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands.  
 Email: [m.f.bijlsma@amsterdamumc.nl](mailto:m.f.bijlsma@amsterdamumc.nl)

## Abstract

Pancreatic ductal adenocarcinoma (PDAC) is characterized by abundant stroma, the main cellular constituents of which are cancer-associated fibroblasts (CAFs). Stroma-targeting agents have been proposed to improve the poor outcome of current treatments. However, clinical trials using these agents showed disappointing results.

**Abbreviations:** CA19.9, carbohydrate antigen 19-9; CAF, cancer-associated fibroblast; CXCL12, platelet factor 12; ELISA, enzyme-linked immunosorbent assay; FABP4, fatty acid binding protein 4; GEO, Gene Expression Omnibus; hg38, homo sapiens (human) genome assembly GRCh38; HR, hazard ratio; IL6, interleukin-6; KD, knock-down; MEF, mouse embryonic fibroblast; MHC-II, antigen-presenting major histocompatibility complex class II; mm10, *Mus musculus* (house mouse) genome assembly GRCm38; OGN, human osteoglycin protein (gene; OGN); Ogn, mouse osteoglycin protein (gene; Ogn); PDAC, pancreatic ductal adenocarcinoma; PDGFR, platelet-derived growth factor receptor; PDX, patient-derived xenograft; RT-qPCR, quantitative reverse transcription-PCR.

Paul Manoukian and Cynthia Waasdorp shared equally the second authorship position.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *International Journal of Cancer* published by John Wiley & Sons Ltd on behalf of UICC.

**Funding information**

KWF Kankerbestrijding; Oncode Institute; PRECODE, Grant/Award Number: H2020-MSCA-ITN 861196; The AMC Foundation; The CCA Foundation

Heterogeneity in the PDAC CAF population was recently delineated demonstrating that both tumor-promoting and tumor-suppressive activities co-exist in the stroma. Here, we aimed to identify biomarkers for the CAF population that contribute to a favorable outcome. RNA-sequencing reads from patient-derived xenografts (PDXs) were mapped to the human and mouse genome to allocate the expression of genes to the tumor or stroma. Survival meta-analysis for stromal genes was performed and applied to human protein atlas data to identify circulating biomarkers. The candidate protein was perturbed in co-cultures and assessed in existing and novel single-cell gene expression analysis from control, pancreatitis, pancreatitis-recovered and PDAC mouse models. Serum levels of the candidate biomarker were measured in two independent cohorts totaling 148 PDAC patients and related them to overall survival. Osteoglycin (OGN) was identified as a candidate serum prognostic marker. Single-cell analysis indicated that *Ogn* is derived from a subgroup of inflammatory CAFs. *Ogn*-expressing fibroblasts are distinct from resident healthy pancreatic stellate cells and arise during pancreatitis. Serum OGN levels were prognostic for favorable overall survival in two independent PDAC cohorts (HR = 0.47,  $P = .042$  and HR = 0.53,  $P = .006$ ). Altogether, we conclude that high circulating OGN levels inform on a previously unrecognized subgroup of CAFs and predict favorable outcomes in resectable PDAC.

**KEYWORDS**

CAF subtypes, liquid biopsy center, pancreatic ductal adenocarcinoma, stroma

**What's new?**

The heterogeneity of cancer-associated fibroblasts (CAFs) in pancreatic ductal adenocarcinoma (PDAC) with both tumor-promoting and tumor-suppressive properties has recently been recognized. Our study identifies osteoglycin as a distinct marker for an inflammatory and tumor-suppressive CAF subpopulation. The favorable prognostic significance of high serum osteoglycin levels was observed in a cohort of 40 patients with resected PDAC and further validated in a nationwide multicenter cohort of 108 patients. The findings point to osteoglycin as a potential noninvasive biomarker that not only predicts favorable clinical outcomes but also distinguishes a novel CAFs subtype.

**1 | INTRODUCTION**

Pancreatic ductal adenocarcinoma (PDAC) has one of the worst prognoses among solid cancers. Late detection and particularly aggressive biology that contributes to early metastatic dissemination are major hurdles to curative therapies. Median survival after diagnosis is about 6 months, and even after surgical resection and adjuvant systemic therapy, patients frequently recur.<sup>1</sup>

PDAC tumors are characterized by an abundance of stroma, which is now well-recognized as a key contributor to the distinct tumor biology of this disease. The stroma consists of various cell types, such as endothelial and immune cells. However, its main cellular constituents are cancer-associated fibroblasts (CAFs), which are either recruited from distant sites or derived from the activation of resident fibroblasts or stellate cells during disease progression. CAFs participate in active crosstalk with tumor cells and produce a dense extracellular matrix that hampers drug delivery and foils treatment approaches. This led to the

concept of stroma-targeted therapies to improve the outcome of current treatment regimens.<sup>2</sup> However, despite many attempts, stroma-targeted approaches have been reported to yield disappointing clinical outcomes; this was subsequently investigated by preclinical work that provided explanations behind this phenomenon.<sup>3-6</sup>

A recent noteworthy discovery is that CAFs are not a homogenous population. Heterogeneity in the PDAC CAF population has been described in the literature as supported by the identification of distinct CAF subgroups that co-exist in the PDAC microenvironment and partake in either tumor-promoting and tumor-suppressive activities.<sup>7,8</sup> It is now thought that effective targeting of the stroma in PDAC should be specific against the tumor-promoting fraction of CAFs, thereby sparing tumor-suppressive CAFs, or tumor-suppressive activities of CAFs in general. In PDAC, the most notable CAF subtypes are myfibroblastic CAFs (myCAF) and inflammatory CAFs (iCAF). MyCAF are thought to produce and remodel the extracellular matrix and are located in close proximity to the malignant cells to form a physical barrier and restrain

invasive growth. In contrast, iCAFs are postulated to be engaged mostly in paracrine signaling, thereby possibly promoting tumor growth.

Current clinical and pathological metrics do not consider CAF heterogeneity and biomarkers for CAF subtypes are not available. Biomarkers, preferably noninvasive, may allow us to predict tumor composition and characteristics at initial diagnosis and to make informed decisions with respect to treatment. As such, we believe that noninvasive biomarkers may be worth incorporating into clinical and histopathological routines. Stromal biomarkers can be used to select patients for treatment strategies that are of limited efficacy in unselected populations. For instance, this could apply to stroma-targeting agents that have recently failed in unstratified phase III trials, or new clinical trials using novel stroma-targeting agents.<sup>9-12</sup> Here, we aimed to identify noninvasive biomarkers for molecular features in the CAF population that contribute to favorable outcome.

## 2 | METHODS

### 2.1 | Gene expression analysis of bulk tissue and meta-analysis

Sequence reads from PDX samples (accession number: E-MTAB-6830) were mapped to the mouse (mm10) and human genomes (hg38) and assigned to either one using Xenofilter.<sup>13</sup> The mapping pipeline was based on and described in more detail in Dijk et al.<sup>14</sup> High confidence stromal genes were identified by calculating mouse-over-human fold change and were cut-off by the following statistical criteria: fold-change >2 or <-2, and  $P < .01$ . Multiple hypothesis testing was adjusted using the Benjamini-Hochberg procedure. For the meta-analysis, data from publicly available transcriptomic (microarray or RNA-Seq) data sets (E-MTAB-6830,<sup>14</sup> GSE36924,<sup>15</sup> GSE15471,<sup>16</sup> GSE71729,<sup>17</sup> GSE21501<sup>18</sup> and TCGA-PAAD<sup>19</sup>) with survival data were downloaded from the data repository Gene Expression Omnibus (GEO). Each data set was scaled to a mean of zero, with a SD of one. For E-MTAB-6830, samples from patients who received adjuvant gemcitabine were excluded. Univariate Cox proportional hazard regression models were evaluated for the top 1000 genes which showed the highest mouse over human fold change. The Metafor R package<sup>20</sup> was used to perform the meta-analysis using a random effect model. Multiple hypothesis testing was adjusted using the Benjamini-Hochberg procedure. Proteomics data from plasma were downloaded from The Human Plasma Proteome Draft of 2017.<sup>21</sup> Micro-dissected PDAC proteomics dataset PXD011289<sup>22</sup> and laser capture micro-dissection PDAC RNAseq dataset GSE93326<sup>23</sup> were downloaded and  $\log_2(\text{OGN} + 1)$  gene expression was analyzed.

### 2.2 | Cell culture

Two mouse embryonic fibroblast (MEF) lines, NIH/3T3 (CVCL\_0594<sup>24</sup>) and C3H/10T1/2 (RRID: CVCL\_0190<sup>25</sup>), as well as the human pancreatic cancer cell line PANC-1 (RRID: CVCL\_0480) were utilized in our

study. All cell lines (ATCC, Manassas, VA) were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 8% FCS, L-glutamine (2 mM), penicillin (100 units/mL) and streptomycin (500  $\mu\text{g}/\text{mL}$ ), in adherence to routine cell culture practices and cultured under normal culture conditions (37°C, 5%  $\text{CO}_2$ ). All human cell lines were STR profiled within the last 3 years. All experiments were performed with mycoplasma-free cells as determined by monthly PCR.

### 2.3 | Lentiviral gene silencing

*Ogn* knockdowns (KDs) of both MEF lines were generated by short hairpin RNA (shRNA)-mediated gene silencing. Lentivirus was produced by transfecting HEK293T cells with Mission TRC library pLKO transfer plasmids together with the envelope/packaging plasmids *pMD2.G*, *pMDL* and *pRSV-REV* using Lipofectamine 2000 transfection reagent (Thermo Fisher). Subsequently, virus-containing supernatant was harvested (at 48 and 72 hours), passed through 0.45  $\mu\text{m}$  filters (Millipore, Billerica, MA), concentrated using Amicon Ultra-15 100K centrifugal filter units and used for MEF transduction. The following TRC clones were used for KD generation: TRCN0000109865, TRCN0000109866, TRCN0000109867, TRCN0000109868 and TRCN0000109869. The *shc002* scrambled sequence was used as a control. In short, MEFs (60% confluence) were transduced overnight with concentrated lentivirus and 8  $\mu\text{g}/\text{mL}$  polybrene (Sigma, St. Louis, MO). Two days after transduction, MEFs were selected and maintained in a culture medium containing 4  $\mu\text{g}/\text{mL}$  puromycin (Sigma) for at least two passages before subsequent experimentation was performed. *Ogn* expression was measured by qRT-PCR to validate knockdown and to determine the most optimal conditions for downstream experiments.

### 2.4 | Co-culture experiments

Co-cultures were performed in 12-well plates under normal cell culture conditions. PANC-1 cells were seeded with *Ogn*-KD MEFs or control MEFs at a 2:1 ratio with a total density of 20 000 cells/ $\text{cm}^2$  and cultured for 72 hours. We opted for clone TRCN0000109865 and clone TRCN0000109869 (later referred to as sh*Ogn*#1 and sh*Ogn*#2, respectively) for co-culture experiments, as these KD MEFs showed the most efficient *Ogn* silencing.

### 2.5 | Quantitative RT-PCR

RNA isolation was performed with RNeasy kit (Macherey Nagel, Düren, Germany) and cDNA was synthesized using Superscript III (Invitrogen). Quantitative PCR (qRT-PCR) was performed using SYBR green reagent and the Lightcycler LC480II system (both from Roche, Basel, Switzerland) following manufacturer guidelines. The comparative threshold cycle (Cp) method was used to calculate transcript levels, which were normalized to *B2m*, and data were represented as

fold change to respective scrambled control. The primer pairs used are listed in the following table.

Mouse primers	Fw	Rv
<i>Ogn</i>	AACGACCTGGAATCTGTGCCTC	TCGCTCCCGAATGTAACGAGTG
<i>B2m</i>	CTTCAGTCGTCAGCATGG	GTTCTTCAGCATTGGATTTC
<i>Acta2</i>	CTGACAGAGGCCACCACTGAA	AGAGGCATAGAGGGACAGCA
<i>Sparc</i>	CCAGGCAAAGGAGAAAGAAG	TTCAGACC GCCAGA AACTCTT
<i>Fn1</i>	TTCAAGTGTGATCCCCATGAAG	CAGGTCTACGGCAGTTGTCA
<i>Cxcl12</i>	CGGTTCTTCGAGAGCCACAT	TCAGCCGTGCAACAATCTGA
Human primers	Fw	Rv
<i>VIM</i>	TCCGCACATTCGAGCAAAGA	ATTCAAGTCTCAGCGGGCTC
<i>B2M</i>	GTCITTCAGCAAGGACTGGTC	CTTCAAACCTCCATGATGC

## 2.6 | FACS analysis

After 72 hours of co-culturing, cells were dissociated using Trypsin-EDTA and resuspended in FACS buffer (PBS with 2% FBS). Cells were stained for N-cadherin surface marker related to EMT, as well as for viability, for 30 minutes at 4°C. Zombie Red was used to discriminate dead cells from live cells and only the latter fraction was considered for subsequent analyses. A gating strategy was used to differentiate between the pancreatic cancer cells, PANC1 cells and the NIH3T3 MEFs; the analysis shown in the figure reflects the PANC1 cell population. The readout was performed on a CytoFlex flow cytometer (Beckman Coulter) and data was processed using FlowJo v.10. The following antibodies/stains were used in this experiment: PE/Cyanine7 anti-human CD325 (N-Cadherin); Zombie Red Fixable Viability Kit (all from BioLegend).

## 2.7 | Cerulein-induced acute pancreatitis

Methods are described in detail in Lodestijn et al.<sup>26</sup> In brief, acute pancreatitis was induced in adult mice (60–70 days old), which received intraperitoneal injections with cerulein (Sigma-Aldrich, C9026) in PBS (50 µg/kg) every hour seven times a day for 2 consecutive days. On the days of cerulein treatment, Metamizol (Sigma-Aldrich, 46232) in PBS (200 mg/kg) was three times orally administered as analgesia. Mice were sacrificed and pancreata were isolated either 9 days (pancreatitis) or 38 days (recovered postpancreatitis) after the last cerulein injection.

## 2.8 | Tissue isolation and preparation for single-cell RNA-seq

To isolate single pancreatic cells for single-cell RNA seq analysis freshly isolated pancreata of control (n = 4; GSE171731), pancreatitis

(n = 2; GSE171731) and postpancreatitis (n = 2), were first minced in small pieces and immediately transferred to 1 mL cold digestion medium (0.5 mg/mL Collagenase P, 10 mM HEPES, 5 mM glucose in HBSS). After processing, the samples were incubated in a water bath at 37°C for 5 minutes at 37°C for digestion. To stop the digestion process 10 mL cold 1% FCS DMEM was added to the samples. Cells were further dissociated by first pipetting them up and down with a P1000 pipet followed by filtering (70 µm filter). After this, cells were centrifuged at 1200 RPM for 3 minutes and the supernatant was removed, the pellet was resuspended in 1 mL cold 1%FCS DMEM.

## 2.9 | Single-cell RNA-seq library preparation

Single cells were sorted (FACS Aria, DB), spun down at 720×g for 5 min and supernatant was removed. Cells were diluted in 0.5% BSA/PBS to 500 to 2000 cells/µL. Directly after this, libraries were prepared according to Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (PN-1000121). Briefly, Gel Beads-in-emulsion (GEMs) were generated by combining barcoded Single Cell 3' v3.1 Gel Beads, a Master Mix containing cells and Partitioning Oil onto Chromium Next GEM Chip. Immediately after GEM generation, the Gel Bead was dissolved, primers were released, and any co-partitioned cell was lysed. Primers were mixed with the cell lysate and a mastermix containing reverse transcription (RT) reagents. Next, Silane magnetic beads were used to purify the first-strand cDNA from the post-GEM-RT reaction mixture. The full-length cDNA was amplified via PCR to generate sufficient mass for library construction. After end repair, A-tailing, adaptor ligation and PCR amplification, the final libraries were ready for sequencing (HiSeq).

## 2.10 | Single-cell analysis

Fibroblast-enriched single-cell gene expression from KPC mice was downloaded from GSE129455. Our earlier published single-cell gene expression data, available at GSE171731, were utilized and supplemented with novel postpancreatitis data (GSE197908). The gene expression matrices were imported into R and further processed by the Seurat R package<sup>27</sup> version 3.2.2. For GSE129455, we in silico isolated CAFs by their markers and subsequently assigned the CAF subtype by ConsensusClusterPlus<sup>28</sup> using previously reported CAF subtype markers. To identify subclusters in iCAFs, we subset the iCAFs and applied Seurat's FindClusters with a resolution = 0.1. For GSE171731 and GSE197908, additional steps were applied and later analyzed as one data set. In brief, we filtered <10% of mitochondrial gene expression and >1000 unique gene counts (nFeature RNA) < 4000, and normalized by SCTransform<sup>29</sup> with regression for nFeature RNA and the percent mitochondrial genes. In addition, we integrated and harmonized the datasets using the function FindIntegrationAnchors and IntegrateData from the Seurat package. Fibroblast subset was identified based on clusters positive for *Sparc*. *Ogn*<sup>+</sup> fibroblasts were defined as fibroblasts with >0 expression of

the *Ogn* gene. The sequencing coverage and quality statistics for each sample are summarized in Table S6.

## 2.11 | Survival analysis with CAF signature

Markers for the iCAF subgroups were identified by FindMarkers from the Seurat package, and retained only genes from stromal origin (all genes with a positive mouse over human fold change). To obtain a signature score, all genes were z-score normalized and averaged per sample. We applied the signature to the TCGA PAAD ( $n = 178$ ),<sup>19</sup> and retained only the samples with histologically confirmed PDAC and survival data ( $n = 130$ ). The median summary score was used as a cutoff to generate two groups for survival analysis.

## 2.12 | Collection of blood samples

In the BioPAN cohort,<sup>30</sup> serum samples were obtained preoperatively from 40 patients undergoing resection and additional clinicopathological data were obtained from medical records. Samples were collected in the period between November 2011 and December 2014. In the Pancreas-Parel cohort,<sup>31</sup> patients who underwent resection because of nonmetastatic PDAC (or chronic pancreatitis) in the period between March 2015 and December 2017 and whose serum sample was collected less than 30 days before resection were included. Only patients with a tumor located in the pancreas and histologically confirmed PDAC were included in the analysis. Patients with 60-day mortality were excluded, as it was assumed that these deaths were due to postoperative complications.

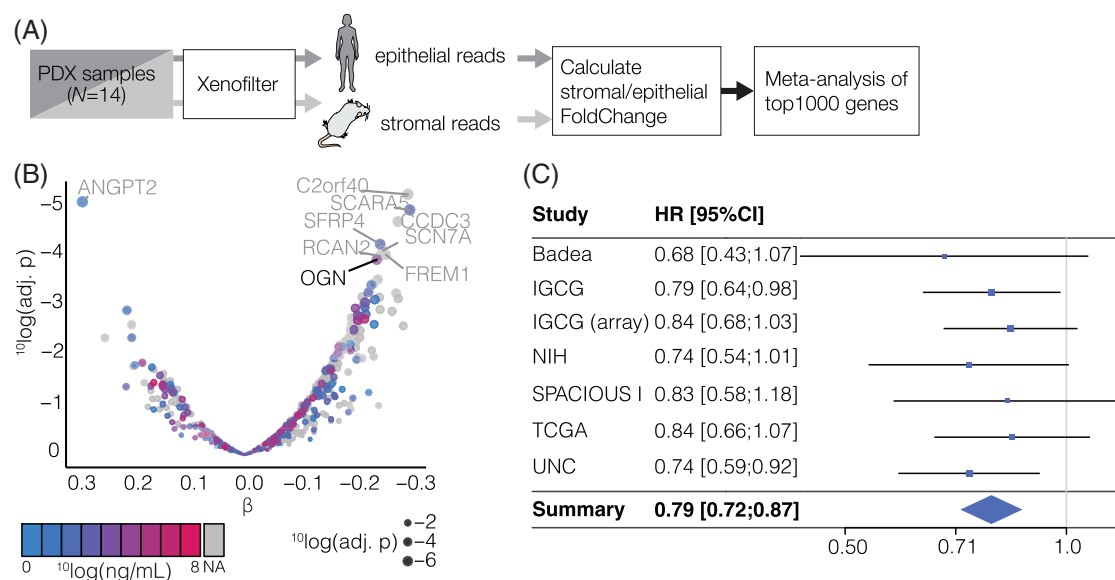
From The Liquid Biopsy Center biobank, a biobank containing all types of biofluids from patients diagnosed with cancer and healthy controls, blood samples of 22 nonage-matched healthy individuals without any indication of malignancy were collected as a control group.

## 2.13 | Enzyme-linked immunosorbent assay

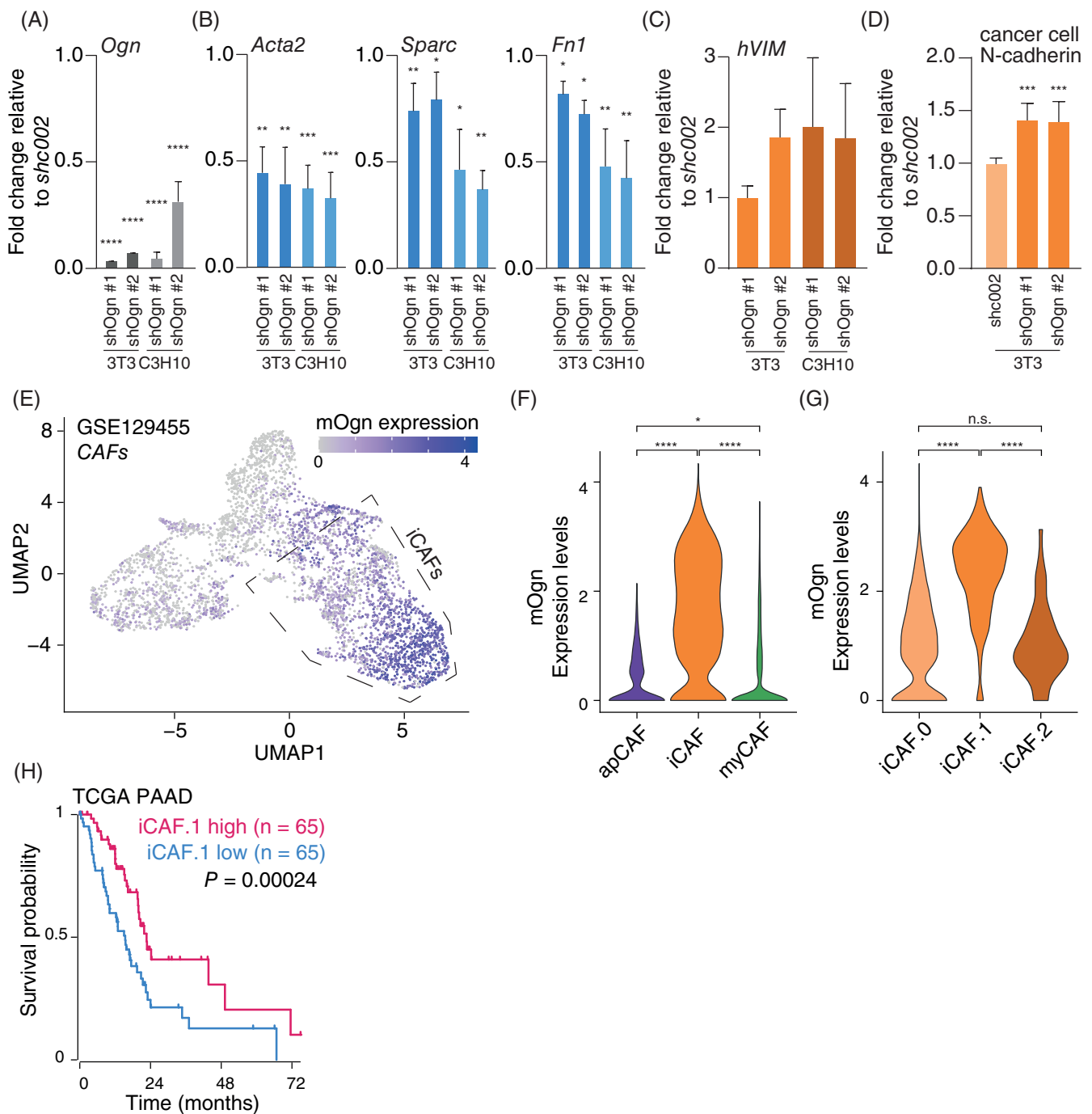
Serum was obtained by centrifugation of blood according to biobank-specific protocol and stored at  $-80^{\circ}\text{C}$  until analysis. Samples from the DPCG and LBC cohort were  $5\times$  diluted prior to use. OGN levels were determined using a human Osteoglycin ELISA kit according to manufacturer's instructions (E-EL-H2163, Elabscience Biotechnology, Houston, MA). Briefly, samples were added to precoated micro-ELISA wells and incubated for 90 minutes at  $37^{\circ}\text{C}$ , followed by a 60 minutes incubation with a biotinylated detection antibody at  $37^{\circ}\text{C}$ . After washing steps, Avidin-Horseradish Peroxidase (HRP) conjugate was added for 30 minutes at  $37^{\circ}\text{C}$ , followed by the addition of stop solution. The optical density was measured at 450 and 570 nm using a microplate reader (BioTek Instruments, Inc. Winooski, VT). Prior to analysis, a wavelength correction was performed by subtracting the 570 nm readings from the 450 nm values. OGN serum concentrations were calculated using a 4 Parameter Logistic (4PL) nonlinear regression model.

## 2.14 | Statistics

In vitro experimental and in vivo CAF marker comparisons were evaluated by unpaired two-tailed Student's *t* test. The correlation between



**FIGURE 1** Identification of stromal genes associated with favorable prognosis in PDAC. (A) Diagram indicating the analysis workflow to interrogate gene expression specific to the mouse host (stroma), and the grafted human tumor cells. (B) Survival meta-analysis was performed on the top 1000 most stromal-associated genes using a random-effect model. The volcano plot shows the effect size on the x-axis and *P*-value on the y-axis for each biomarker. The colors indicate the concentrations found in blood by the Human Protein Atlas.<sup>46</sup> (C) Forest plot of hazard ratios (HRs) for the prognostic value of OGN across PDAC gene expression dataset. All seven data sets show a decreased hazard ratio with increasing OGN expression levels



**FIGURE 2** OGN is required for fibroblast activation and marks a distinct iCAF population. (A) Quantification of *mOgn* expression by qPCR in mouse embryonic fibroblasts (C3H/10T1/2, NIH/3T3) after knockdown. Values were normalized to scrambled control (mean  $\pm$  SD,  $n = 3$  biological replicates per group). (B) Quantification of stromal activation markers in *mOgn*-KD MEFs after coculture with human tumor cells using qPCR. Values were normalized to scrambled control (mean  $\pm$  SD,  $n = 3$  biological replicates per group). (C) As for experiment shown in panel B, showing expression of human *Vimentin* (*hVIM*). Data are mean  $\pm$  SD,  $n = 3$  biological replicates per group, normalized to scrambled control (*shc002*). (D) As for panels A to C, using PANC-1 cells and NIH/3T3 fibroblasts. Cells were processed for flow cytometry using an antibody against human N-cadherin. In addition, cancer cells were identified on scatter profile. Shown are gMFI, normalized to scrambled control (included in the panel). xData are technical triplicates. (E) UMAP embedding of CAFs from KPC mice, showing heterogeneous expression of *mOgn*. (F,G) Violin plot of *mOgn* expression for each CAF subtype, and for each iCAF subcluster, respectively. (H) Kaplan-Meier survival analysis of TCGA PAAD patients based on *iCAF.1* signature as identified by stromal gene ranking analysis. The samples were dichotomized by the median. n.s. = not significant, \* $P < .05$ , \*\* $P < .01$ , \*\*\* $P < .001$  and \*\*\*\* $P < .0001$ .

clinical pathological characteristics and gene expression with overall survival was evaluated by Kaplan-Meier and log-rank test. The prognostic value of serum OGN was tested with univariate and multivariate Cox regression analysis. CA19.9 was nonlinear in both cohorts and was therefore log-transformed. CA19.9 was dichotomized based on the median. The proportional hazard assumption was checked by Schoenfeld residuals and was met in all Cox regression analyses. A *P*-value of  $<.05$  was considered statistically significant.

### 3 | RESULTS

#### 3.1 | Stroma-derived OGN predicts favorable outcomes in PDAC

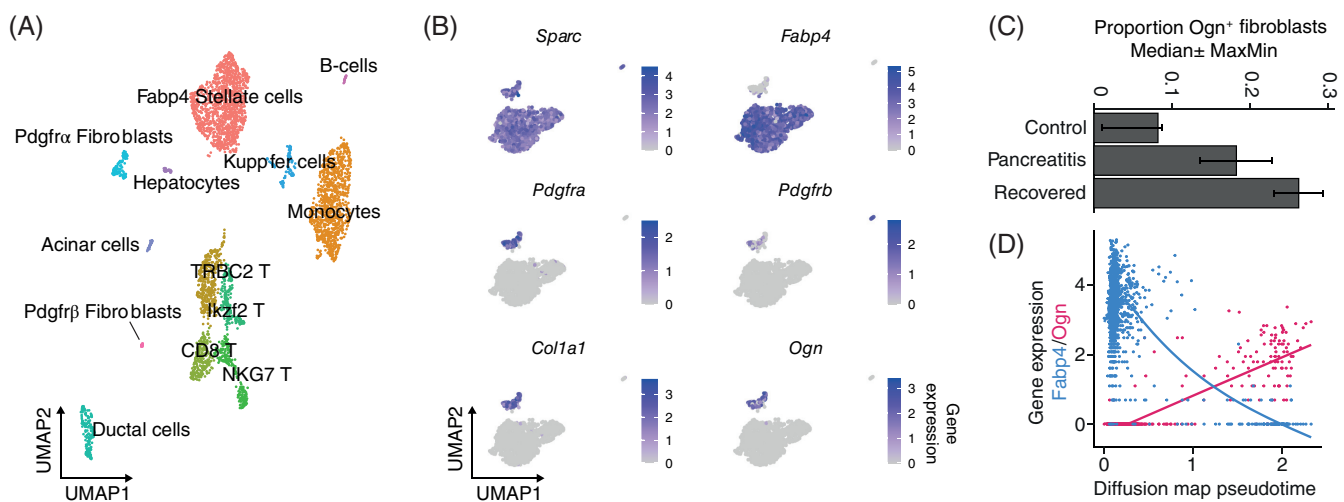
To identify CAF biomarkers, an unbiased gene expression meta-analysis was performed using stromal genes. To obtain genes that derive exclusively from the stroma, we mapped RNA-sequencing reads from patient-derived xenografts (PDXs) to the mouse and human genomes (Figure 1A). This segregates the gene expression to the stroma or tumor, respectively. By using stringent statistical criteria, 5360 genes were found to be differentially expressed between the stromal and the tumor compartment, including 1940 genes uniquely expressed in the stroma and 3420 uniquely expressed in the tumor (Figure S1A).

The top 1000 genes associated with stroma were then processed in a survival meta-analysis. The meta-analysis included patients from six cohorts of which gene expression and survival data were available. This revealed 64 genes that were significantly associated with favorable prognosis, and five associated with poor outcomes (Figure 1B). The observation that stromal genes appear to mostly predict favorable

rather than poor outcomes underlines the notion that the stroma does not harbor purely tumor-promoting activities. To prepare for clinical translation, we set out to identify a blood-borne stromal biomarker and cross-referenced the results from the meta-analysis with the Human Protein Atlas<sup>21</sup> for protein expression found in blood plasma (indicated by colors in Figure 1B). We reasoned that the likelihood of clinical implementation would be greatest if the biomarker can distinguish a set of patients with a favorable outcome, in whom providing additional treatment may be beneficial. In these analyses, Osteoglycin (OGN), also known as Mimecan, showed consistent prognostic power towards favorable outcomes across all six cohorts and was reported to be detectable in serum (Figure 1C). We, therefore, considered OGN (HR = 0.79,  $P = 8.6 \times 10^{-5}$  from the meta-analysis) a likely candidate to follow-up as CAF biomarker. To further validate the stromal origin of OGN, we interrogated laser capture microdissected gene expression and proteomics datasets from PDAC patients (Figure S1B,C). This clearly demonstrated that OGN expression was indeed strictly confined to stromal compartments.

#### 3.2 | OGN is associated with fibroblast activation

OGN is a member of the small leucine-rich proteoglycan family. It exerts various regulatory functions in the extracellular matrix, through both structural and cellular processes, in inflammation and fibrosis.<sup>32</sup> To investigate the role of OGN in the tumor microenvironment, we generated genetic knockdown models for OGN (*shOgn*) in two mouse fibroblast lines. We co-cultured *shOgn* fibroblasts with PANC-1 PDAC cells and measured stromal activation markers by species-specific qRT-PCR. This revealed a substantial and consistent reduction of alpha-smooth muscle actin (*Acta2*), secreted protein acidic and cysteine rich (*Sparc*) and

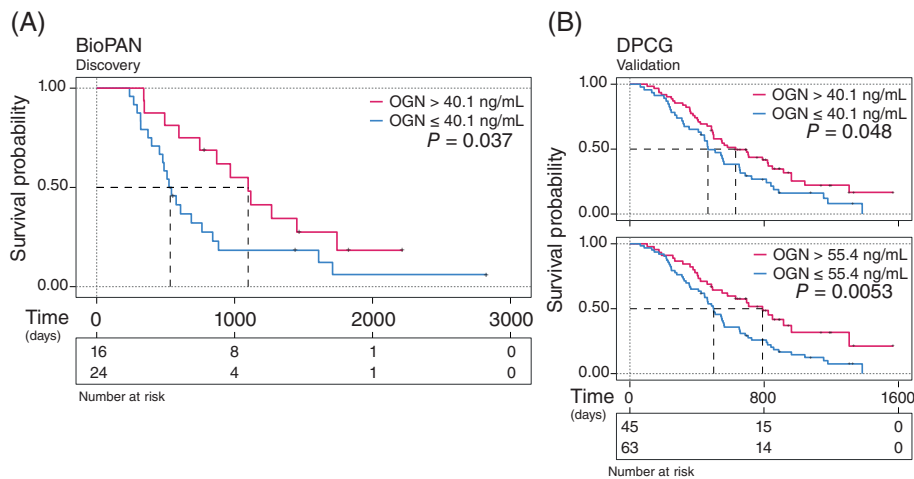


**FIGURE 3** OGN<sup>+</sup> fibroblasts pre-exist in the nonmalignant pancreas. (A) UMAP embedding of 3691 cells from eight mice across three experimental conditions. Cell types identified based on marker expression are indicated. (B) Expression of selected stellate cell and fibroblast markers using the UMAP representation from in silico subset fibroblasts. (C) Proportion of Ogn<sup>+</sup> fibroblasts from the total fibroblast population, grouped by experimental condition (median ± minimum/maximum). (D) Feature plot of expression distribution for *Fabp4* (red) and *Ogn* (blue) across pseudotime. *Fabp4* associates with the resident pancreatic stellate cell population. *Ogn* is part of an activated fibroblast population, which increases in pseudo-time



fibronectin (*Fn1*) in the fibroblasts in response to *Ogn* knockdown (Figure 2B). Similar results were obtained in fibroblast monoculture (Figure S2A). In the cocultured PDAC cells (PANC-1), an upregulation of the EMT markers Vimentin and N-cadherin was observed in response to

*Ogn* knockdown, implying that stromal *Ogn* expression may contribute to maintaining epithelial cancer cell states in PDAC (Figure 2C,D). We conclude from this that OGN is not only associated with stroma, but that it also contributes to shaping CAF activation status and phenotype.



**FIGURE 4** Serum OGN levels correlate with favorable outcomes in resectable PDAC patients. (A) Kaplan-Meier survival analysis of AMC PDAC patients (BioPAN,  $n = 75$ ) using a cut-off value of 40.1 ng/mL, determined by ROC. (B) Validation of the prognostic power of serum OGN in an independent cohort of PDAC patients (DPCG,  $n = 108$ ). Upper panel shows the validation of the 40.1 ng/mL threshold. The lower panel shows the most optimal cut-off (55.4 ng/mL) for this cohort determined by ROC in this cohort

Characteristics	HR (univariable)	HR (multivariable)
OGN, log-scale	0.73 (0.26-2.02, $P = 0.543$ )	—
OGN		
Low	—	—
High	<b>0.47 (0.23-0.97, <math>P = .042</math>)</b>	1.06 (0.19-5.85, $P = .947$ )
CA19.9, log-scale	1.40 (0.95-2.06, $P = .087$ )	—
CA19.9		
Low	—	—
High	<b>4.66 (1.42-15.28, <math>P = .011</math>)</b>	4.83 (0.97-24.14, $P = .055$ )
Bilirubin	1.29 (0.99-1.67, $P = .055$ )	—
Sex		
Female	—	—
Male	1.54 (0.75-3.15, $P = .238$ )	—
Age	1.02 (0.98-1.06, $P = .244$ )	—
BMI	0.96 (0.85-1.09, $P = .564$ )	—
ECOG		
0	—	—
1	0.91 (0.17-5.05, $P = .918$ )	—
Adjuvant therapy		
No	—	—
Yes	0.64 (0.27-1.49, $P = .301$ )	—
Radicality		
$R_0$	—	—
$R_1$	1.95 (0.95-4.02, $P = .068$ )	—
Tumor size	1.10 (0.88-1.39, $P = .400$ )	—
Location		
Head	—	—
Corpus/tail	1.76 (0.75-4.08, $P = .192$ )	—
Head and tail	—	—

**TABLE 1** Univariate and multivariate survival analysis, BioPAN. A  $P$ -value of  $<.05$  was considered statistically significant and is indicated by bold markup

**TABLE 2** Univariate and multivariate survival analysis, DPCG. A *P*-value of <.05 was considered statistically significant and is indicated by bold markup

Characteristics	HR (univariable)	HR (multivariable)
OGN, log-scale	<b>0.51 (0.31-0.85, <i>P</i> = .010)</b>	—
OGN		
Low	—	—
High	<b>0.53 (0.33-0.83, <i>P</i> = .006)</b>	<b>0.56 (0.35-0.91, <i>P</i> = .019)</b>
CA19.9, log-scale	<b>1.03 (1.01-1.05, <i>P</i> = .002)</b>	—
CA19.9		
Low	—	—
High	<b>1.55 (1.00-2.40, <i>P</i> = .047)</b>	1.40 (0.89-2.19, <i>P</i> = .148)
Bilirubin, per 100	1.00 (0.78-1.28, <i>P</i> = .999)	—
Sex		
Female	—	—
Male	1.16 (0.74-1.82, <i>P</i> = .524)	—
Age	<b>1.03 (1.00-1.05, <i>P</i> = .021)</b>	1.02 (0.99-1.04, <i>P</i> = .131)
BMI	0.97 (0.91-1.03, <i>P</i> = .288)	—
ECOG		
0	—	—
1	1.37 (0.86-2.18, <i>P</i> = .190)	—
2	1.07 (0.42-2.75, <i>P</i> = .881)	—
3	1.55 (0.47-5.07, <i>P</i> = .467)	—
Adjuvant therapy		
No	—	—
Yes	0.69 (0.43-1.11, <i>P</i> = .129)	—
Adjuvant therapy, type		
Chemotherapy	—	—
Chemoradiotherapy	1.82 (0.65-5.10, <i>P</i> = .255)	—
Radiotherapy	—	—
Neoadjuvant therapy		
No	—	—
Yes	0.86 (0.48-1.56, <i>P</i> = .625)	—
Neoadjuvant therapy, type		
Chemotherapy	—	—
Chemoradiotherapy	0.40 (0.12-1.35, <i>P</i> = .140)	—
Radiotherapy	—	—
Radicality		
R <sub>0</sub>	—	—
R <sub>1</sub>	<b>1.69 (1.08-2.65, <i>P</i> = .022)</b>	1.58 (1.00-2.51, <i>P</i> = .050)
Tumor size	1.02 (0.99-1.04, <i>P</i> = .266)	—
Location		
Head	—	—
Corpus/tail	0.87 (0.47-1.60, <i>P</i> = .648)	—

### 3.3 | OGN is an iCAF marker

To relate *Ogn* expression and CAF heterogeneity, we queried a published single-cell RNA sequencing data set from KPC mice.<sup>33</sup> Three different CAF subtypes were apparent, in line with previous class discovery efforts (Figure S2B). *Ogn* was significantly higher expressed in

iCAFs than in myCAFs or antigen-presenting CAFs (apCAFs; Figure 2F). However, even within the iCAF population, *Ogn* expression was markedly heterogeneous (Figure 2E). This prompted us to assess whether iCAFs can be further divided into subgroups, and unsupervised clustering of the iCAF population yielded three subclusters (Figure S2C,D). In one of the iCAF subclusters (iCAF.1), *Ogn*

expression was most prominent (Figure 2G, Tables S1 and S2). Given that OGN was found to associate with favorable prognosis in patients, we reasoned that—in contrast with the current paradigm—the iCAF.1 subcluster may contribute to favorable outcome. Applying the top 50 stromal marker genes of iCAF.1 to the TCGA-PAAD dataset, a survival difference was observed that supports a tumor-restrictive activity for this subcluster (Figure 2H). To exclude OGN as a colinear factor, we also omitted OGN from the analysis and the results were similar (Figure S2E).

Because OGN is required for fibroblast activation *in vitro*, we wondered if OGN is involved in such dynamic processes *in vivo*. This hypothesis is bolstered by earlier findings from single-cell sequencing studies investigating fibroblasts in a nonmalignant setting, in which it was reported that OGN is an upregulated marker preceding myofibroblast differentiation.<sup>34–36</sup> Given the association of OGN with the development of fibrosis, we speculated that OGN could be detected as early as the pancreatitis stage. During pancreatitis, pancreatic stellate cells become activated and transit from a quiescent to a proliferative, migratory and fibrogenic cell state. To demonstrate the association between OGN and fibrosis, we queried our single-cell RNA sequencing data set from mice, which contained pancreata from control and cerulein-induced pancreatitis mice, which were supplemented with pancreatitis-recovered samples. For downstream analysis, we analyzed the data sets as one merged set. Unsupervised clustering identified 11 cell clusters (Figures 3A and S3A–C). Using the differentially expressed gene signatures, we attributed clusters to their putative identities. By comparing *Ogn* expression across all different cell types, we confirmed again that *Ogn* is solely of stromal origin (Figure S3D). This quality control step illustrates the robustness of dissecting stromal genes using gene expression from PDXs. We observed three distinct fibroblast populations: *Fabp4*<sup>+</sup> pancreatic stellate cells, *Pdgfra*<sup>+</sup> fibroblasts and *Pdgfrβ*<sup>+</sup> fibroblasts (Figure 3B). PDGFRα signaling is generally considered as a driver of fibrosis in multiple other tissues. This is also confirmed by its increased *Col1a1* expression in the PDGFRα cluster. Within that same cluster, *Ogn* expression was detected, supporting the idea that OGN is involved in cellular contributions to the extracellular matrix. Considering the OGN<sup>+</sup> population kinetics across experimental conditions, we observed a 2-fold increase of *Ogn*-expressing fibroblasts after pancreatitis induction, and persisted in the days after pancreatitis (Figure 3C) in which tissue architecture was largely restored.<sup>26</sup> When ordering cells in pseudotime, we observed an anticorrelation of *Fabp4* and *Ogn*, of which *Ogn* is solely expressed at the end of the trajectory (Figure 3D). In the dataset, no CAF subtypes could be observed by clustering the signature genes. Marker-wise, we observed only two iCAF markers (*Col14a1*, *Clec3b* and not *Has1* and *Il6*) present in the premalignant stages, while myCAF and apCAF markers were absent (Figure S4A–D). These data indicate that OGN plays a pivotal role in fibroblast activation and differentiation, and possibly contributes to the early stages of tissue remodeling to restrain injury to the pancreas. Therefore, it can be considered that OGN<sup>+</sup> fibroblasts are a distinct subpopulation of CAFs that in fact exist before the malignant transformation of the pancreatic epithelium.

### 3.4 | Serum levels of OGN are prognostic

To investigate whether OGN can be used as a noninvasive biomarker for the abundance of this tumor-restraining CAF population, an ELISA for OGN was performed on 40 preoperatively collected serum samples from patients with resectable PDAC (Table S3). A threshold of 40.1 ng/mL was determined by AUC of the ROC curve to dichotomize high and low serum expression of OGN, based on survival status at the time of analysis (Figure 4A). The analysis of clinical data identified low serum OGN level as a factor for survival in univariate analysis, which was not found to be independent in multivariate analysis (Table 1). In a larger cohort from the nationwide pancreatic biobank with 108 PDAC patients, we validated the prognostic power of serum OGN (Table S4). In addition, a more optimal cutoff point in this cohort was found at 55.4 ng/mL (Figure 4B). The multivariate Cox regression, revealed that reduced OGN is an independent factor for survival (HR = 0.56, *P* = .019, Table 2) after adjustment for other significant preoperatively available predictors. OGN serum levels were also compared to 22 healthy controls (Table S5). Mean serum OGN did not differ significantly between PDAC patients and healthy controls (73.4 ng/mL vs 57.1 ng/mL, respectively). However, some PDAC patients showed very high values; up to eight times the mean value of healthy controls. Together, we take the above to show that serum OGN is not a PDAC-detecting biomarker, but that its presence does inform the abundance of a particular subset of CAFs with importance for PDAC prognosis.

## 4 | DISCUSSION

PDAC is a highly aggressive malignancy against which limited therapeutic options exist. Failed clinical trials with stroma targeting agents forced the research community to revise the prevailing view of the stroma as strictly supportive of neoplastic growth and therapy resistance. In the presented work, we identified a noninvasive iCAF biomarker, OGN, that is able to predict favorable outcomes. Notably, OGN was not only prognostic at the mRNA level, but we also demonstrated in two independent cohorts, totaling 148 patients, that high serum levels of OGN were associated with a favorable prognosis. Increased serum OGN levels are likely derived from a specific iCAF subcluster, which already arises after premalignant injuries to the pancreas.

iCAFs have been suggested to harbor mostly tumor-promoting properties.<sup>37</sup> This idea is largely based on experiments in genetically engineered tumor-bearing mice from which α-SMA<sup>+</sup> fibroblasts were depleted, which resulted in shortened survival and poorly differentiated tumors.<sup>4</sup> This was in contrast to the long-standing suspicion that CAFs in general were tumor-supportive. Deductive reasoning led to the conclusion that an opposing CAF population must then be pro-tumor. Indeed, studies confirmed pro-tumorigenic properties on CAF subpopulations, and associated them with iCAFs by either CXCL12 or IL6 positivity.<sup>38,39</sup> In our study, however, we found evidence that iCAFs may harbor both tumor-supportive and tumor-restrictive

properties. We identified three subpopulations of iCAFs, one of which is marked by *OGN* expression and is most associated with favorable prognosis.

*OGN* is a small leucine-rich proteoglycan that is involved in various cellular processes, such as development, cell proliferation by expression-based regulation and collagen fibrinogenesis via direct binding and cross-linking of collagen fibers.<sup>32</sup> Several interaction partners for *OGN* have been reported, for instance from proteomics analysis of binding partners but many of these are for instance intracellular proteins that would not be expected reflect a biologically meaningful interaction.<sup>40</sup> In the InnateDB database, CellphoneDB includes a reported interaction with the *HLA-DRB1* gene, a beta chain of the antigen-presenting major histocompatibility complex class II (MHCII) molecule.<sup>41,42</sup> Based on this, *OGN* may be expected to be involved in immune responses. However, given the effects observed in our co-culture experiments (from which immune cells are entirely absent; Figure 2A-D), a nonimmune related interaction of *OGN* should exist.

In novel single-cell sequencing datasets containing fibroblasts, *OGN* is increasingly detected as a marker for a fibroblast subpopulation. In human intestinal development, *OGN*-high cells are annotated as fibroblast progenitors which later differentiate into pericytes, smooth muscle cells and myofibroblasts.<sup>35</sup> Slightly more related to PDAC stroma, in human kidney fibrosis it was found that the presence of myofibroblast differentiation is preceded by *OGN*-high fibroblasts.<sup>34</sup> Similar results were found in murine skin wounds, which revealed extensive fibroblast heterogeneity and that *OGN*-positive cells precede those expressing contractile genes such as *Acta2* and *Tagln* in pseudo-time.<sup>36</sup> Our data support the observed *OGN* dynamics. In the pancreatitis mouse model, we indeed saw an increased population of *OGN*<sup>+</sup> fibroblasts that lacked any actual CAF subtype markers, suggesting that *OGN* may be a sort of CAF precursor marker that persists after the cells adopt a CAF phenotype. This trend was also found in data from KPC mice, in which we detected a specific subcluster of iCAFs which were positive for *OGN*. Merging the observations, this suggests that myCAFs, apCAFs and the remaining iCAF subclusters differ from these *OGN*<sup>+</sup> fibroblasts. Although a link between *OGN* and TGF- $\beta$  signaling has been suggested<sup>32</sup> (and could be inferred from the effects of *shOgn* on *Acta2*, a TGF- $\beta$  responsive gene), we were unable to find consistent correlations between the two in bulk and single-cell gene expression data (data not shown).

Mice genetically deficient for *Ogn* were generated two decades ago and are used to study various diseases.<sup>43</sup> Although these mice were never studied in a cancer context, they could provide clues and help generate hypotheses as to why stromal *OGN* contributes to favorable prognosis, and whether this is driven by its association with a particular iCAF subcluster. *Ogn*-null mice do not show developmental or other phenotypic abnormalities to their wild-type counterparts. However, reduced skin fragility was detected by altered fibrinogenesis in *Ogn*-null mice.<sup>32</sup> In a myocardial infarction model, *Ogn*-null mice prevented cardiac dilatation and dysfunction through infarct collagen strengthening.<sup>44</sup> This suggests that *OGN* is mostly required in disease processes to which fibroblast activation contributes. Our data in *shOgn* fibroblasts complement this idea in which *Ogn* depletion leads

to fibroblast deactivation. Altogether, a bigger picture emerges in which *OGN* functions as a structural regulator and affects the stromal composition by tissue remodeling to ultimately influence the outcome of PDAC.

The results of our study should, however, be interpreted in light of several limitations. First of all, lineage connectivity between CAF subtypes and *OGN*<sup>+</sup> fibroblasts was implicated by the results of the two analyzed data sets, but future lineage tracing experiments are required to map the CAF differentiation trajectory. Second, we did not use a chronic pancreatitis mouse model, which is the main risk factor for PDAC, but an acute pancreatitis model. Long-term inflammation can possibly affect *OGN* dynamics. However, *OGN*<sup>+</sup> fibroblasts persist during cancer which makes it likely that acute pancreatitis suffices as a model to study *OGN* dynamics. Third, we failed to identify a standardized cut-off value for *OGN* serum, which prevents definitive conclusions on how to use *OGN* as a biomarker for prognostication. Strengths of our study that mitigate some of these limitations, however, include (a) the use of a comprehensive meta-analysis approach to identify a novel prognostic biomarker with fidelity, (b) in combination with experiments and analyses to explain *OGN*'s biological function in various steps of PDAC development and (c) the ability to validate the prognostic power of serum *OGN* in an independent cohort.

To sum up, with the use of a meta-analysis approach we identified *OGN* as a favorable outcome biomarker in patients with resectable PDAC, both in tissue gene expression as well as serum samples. *OGN* is a distinct marker for a particular subset of PDAC iCAFs and plays a critical role in fibroblast activation and differentiation in both the premalignant and malignant pancreas, and may function to limit injury to the pancreas by tissue remodeling in stages as early as pancreatitis. In the future, *OGN*-lineage tracing experiments should rule out whether *OGN* fibroblasts form the common cell of origin for various CAF subtypes. For future clinical application, *OGN* should be investigated in prospective studies to understand its role in specific stages of PDAC, and in relation to different treatment strategies. As a proxy for a tumor-restraining constituent in the stroma, it should also be considered to stratify patients in future stroma-targeting randomized controlled trials.

#### AUTHOR CONTRIBUTIONS

The work reported in the study has been performed by the authors unless clearly specified in the text. **Mark P. G. Dings, Maarten F. Bijlsma:** Design research. **Mark P. G. Dings, Cynthia Waasdorp, Paul Manoukian, Judith S. E. Quik, Marin Strijker, Sophie C. Lodestijn, Sanne M. van Neerven, Leandro F. Moreno, Rodrigo Leite de Oliveira, Bert A. Bonsing, Marco J. Bruno, Olivier R. Busch, Michael Doukas, Casper H. van Eijck, Nadia Haj Mohammad, Ignace H. de Hingh, I. Quintus Molenaar, Marc G. Besselink, Louis Vermeulen, Jan Paul Medema, Hanneke W. M. van Laarhoven, Maarten F. Bijlsma:** Data interpretation, analysis, generation and/or collection. **Jan Paul Medema, Hanneke W. M. van Laarhoven, Maarten F. Bijlsma.** Wrote the paper: **Mark P. G. Dings, Maarten F. Bijlsma:** Supervised the project.

## ACKNOWLEDGEMENTS

The authors would like to thank the patients for participating, and to the Liquid Biopsy Center, Mai Tran, Melissa Fidder, Dr Michiel Pegtel for sample acquisition and logistical support. Funding was provided by KWF Dutch Cancer Society, the AMC Foundation, the CCA Foundation, H2020-MSCA-ITN 861196 PRECODE and Oncode.

## CONFLICT OF INTEREST

Maarten F. Bijlsma has received research funding from Celgene, Frame Therapeutics and LeadPharma, and has acted as a consultant to Servier. Hanneke W. M. van Laarhoven has served as a consultant for BMS, Drag-onfly, Lilly, Merck, Nordic Pharma and Servier, and has received unrestricted research funding and/or study medication from Bayer, BMS, Celgene, Janssen, Incyte, Lilly, Merck, Nordic Pharma, Philips, Roche and Servier, and has been a speaker for Astellas and Novartis. Jan Paul Medema has acted as a consultant to AbbVie. Louis Vermeulen received consultancy fees from Bayer, MSD, Genentech, Servier and Pierre Fabre. None of these parties were involved in the design of our study or drafting of the study. The other authors have no conflict of interest to declare.

## DATA AVAILABILITY STATEMENT

The gene expression data set from PDAC PDX are available at E-MTAB-6830. The single-cell RNA-sequencing data from KPC mice are available GSE129455. The single-cell RNA-sequencing data from control and cerulein-induced pancreatitis data are available at GSE171731. The novel postpancreatitis single-cell RNA-sequencing data generated in our study is available in GEO under accession number GSE197908. All other data from the current study are available from the authors on reasonable request.

## ETHICS STATEMENT

For the BioPAN and the Liquid Biopsy Center, the collection of material was approved by the Ethics Committee of the Amsterdam UMC, location Academic Medical Center (AMC 2014\_181, AMC 2018\_167, respectively). The study was approved by the Dutch Pancreas Biobank scientific committee. For the PancreasParel, the use of residual material from an earlier issuance for which local approvals were obtained<sup>31</sup> was accepted by the Dutch Pancreatic Cancer Group committee. In all cases, written informed consent was received from all participants. The REMARK criteria were followed in the writing of this tumor marker prognostic study.<sup>45</sup>

## ORCID

Mark P. G. Dings  <https://orcid.org/0000-0001-8509-0006>

Maarten F. Bijlsma  <https://orcid.org/0000-0001-6627-3229>

## REFERENCES

- Mizrahi JD, Surana R, Valle JW, Shroff RT. Pancreatic cancer. *Lancet*. 2020;395:2008-2020.
- Olive KP, Jacobetz MA, Davidson CJ, et al. Inhibition of hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science (80-)*. 2009;324:1457-1461.
- Van Mackelenbergh MG, Strees CI, Spijker R, et al. Clinical trials targeting the stroma in pancreatic cancer: a systematic review and meta-analysis. *Cancers (Basel)*. 2019;11:588.
- Özdemir BC, Pentcheva-Hoang T, Carstens JL, et al. Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell*. 2014;25:719-734.
- Lee JJ, Perera RM, Wang H, et al. Stromal response to hedgehog signaling restrains pancreatic cancer progression. *Proc Natl Acad Sci USA*. 2014;111:E3091-E3100.
- Rhim AD, Oberstein PE, Thomas DH, et al. Stromal elements act to restrain, rather than support, pancreatic ductal adenocarcinoma. *Cancer Cell*. 2014;25:735-747.
- Öhlund D, Handy-Santana A, Biffi G, et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J Exp Med*. 2017;214:jem.20162024.
- Biffi G, Oni TE, Spielman B, et al. IL1-induced JAK/STAT signaling is antagonized by TGF $\beta$  to shape CAF heterogeneity in pancreatic ductal adenocarcinoma. *Cancer Discov*. 2019;9:282-301.
- Hingorani SR, Zheng L, Bullock AJ, et al. HALO 202: randomized phase II study of PEGPH20 plus nab-paclitaxel/gemcitabine versus nab-paclitaxel/gemcitabine in patients with untreated, metastatic pancreatic ductal adenocarcinoma. *J Clin Oncol*. 2018;36:359-366.
- De Jesus-Acosta A, Sugar EA, O'Dwyer PJ, et al. Phase 2 study of vismodegib, a hedgehog inhibitor, combined with gemcitabine and nab-paclitaxel in patients with untreated metastatic pancreatic adenocarcinoma. *Br J Cancer*. 2020;122:498-505.
- Picozzi VJ, Pishvaian MJ, Mody K, et al. Effect of anti-CTGF human recombinant monoclonal antibody pamrevlumab on resectability and resection rate when combined with gemcitabine/nab-paclitaxel in phase 1/2 clinical study for the treatment of locally advanced pancreatic cancer patients. *J Clin Oncol*. 2018;36:4016.
- Murphy JE, Wo JY, Ryan DP, et al. Total Neoadjuvant therapy with FOLFIRINOX in combination with losartan followed by Chemoradiotherapy for locally advanced pancreatic Cancer: a phase 2 clinical trial. *JAMA Oncol*. 2019;5:1020-1027.
- Kluin RJC, Kemper K, Kulman T, et al. Xenofilter: computational deconvolution of mouse and human reads in tumor xenograft sequence data. *BMC Bioinformatics*. 2018;19:366.
- Dijk F, Veenstra VL, Soer EC, et al. Unsupervised class discovery in pancreatic ductal adenocarcinoma reveals cell-intrinsic mesenchymal features and high concordance between existing classification systems. *Sci Rep*. 2020;10:337.
- Bailey P, Chang DK, Nones K, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*. 2016;531:47-52.
- Badea L, Herlea V, Dima SO, Dumitrascu T, Popescu I. Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia. *Hepatogastroenterology*. 2015; 55:2016-2027.
- Moffitt RA, Marayati R, Flate EL, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet*. 2015;47:1168-1178.
- Stratford JK, Bentrem DJ, Anderson JM, et al. A six-gene signature predicts survival of patients with localized pancreatic ductal adenocarcinoma. *PLoS Med*. 2010;7:e1000307.
- Raphael BJ, Hruban RH, Aguirre AJ, et al. Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell*. 2017;32: 185-203.e13.
- Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36:1-48.
- Schwenk JM, Omenn GS, Sun Z, et al. The human plasma proteome draft of 2017: building on the human plasma PeptideAtlas from mass spectrometry and complementary assays. *J Proteome Res*. 2017;16: 4299-4310.
- le Large TYS, Mantini G, Meijer LL, et al. Microdissected pancreatic cancer proteomes reveal tumor heterogeneity and therapeutic targets. *JCI Insight*. 2020;5:e138290.

23. Maurer C, Holmstrom SR, He J, et al. Experimental microdissection enables functional harmonisation of pancreatic cancer subtypes. *Gut*. 2019;68:1034-1043.
24. Jainchill JL, Aaronson SA, Todaro GJ. Murine sarcoma and leukemia viruses: assay using clonal lines of contact-inhibited mouse cells. *J Virol*. 1969;4:549-553.
25. Reznikoff CA, Brankow DW, Heidelberger C. Establishment and characterization of a cloned line of C3H mouse embryo cells sensitive to Postconfluence inhibition of division. *Cancer Res*. 1973;33:3231-3238.
26. Lodestijn SC, van den Bosch T, Nijman LE, et al. Continuous clonal labeling reveals uniform progenitor potential in the adult exocrine pancreas. *Cell Stem Cell*. 2021;28:2009-2019.e4.
27. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell*. 2019;177:1888-1902.e21.
28. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26:1572-1573.
29. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol*. 2019;20:296.
30. Damhofer H, Ebbing EA, Steins A, et al. Establishment of patient-derived xenograft models and cell lines for malignancies of the upper gastrointestinal tract. *J Transl Med*. 2015;13:115.
31. Strijker M, Gerritsen A, van Hilst J, et al. The Dutch pancreas biobank within the PARELSNOER institute: a Nationwide biobank of pancreatic and Periampullary diseases. *Pancreas*. 2018;47:495-501.
32. Deckx S, Heymans S, Papageorgiou A. The diverse functions of osteoglycin: a deceitful dwarf, or a master regulator of disease? *FASEB J*. 2016;30:2651-2661.
33. Elyada E, Bolisetty M, Laise P, et al. Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting Cancer-associated fibroblasts. *Cancer Discov*. 2019;9:1102-1123.
34. Kuppe C, Ibrahim MM, Kranz J, et al. Decoding myofibroblast origins in human kidney fibrosis. *Nature*. 2021;589:281-286.
35. Fawcner-Corbett D, Antanaviciute A, Parikh K, et al. Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell*. 2021;184:810-826.e23.
36. Guerrero-Juarez CF, Dedhia PH, Jin S, et al. Single-cell analysis reveals fibroblast heterogeneity and myeloid-derived adipocyte progenitors in murine skin wounds. *Nat Commun*. 2019;10:650.
37. Hosein AN, Brekken RA, Maitra A. Pancreatic cancer stroma: an update on therapeutic targeting strategies. *Nat Rev Gastroenterol Hepatol*. 2020;17:487-505.
38. Feig C, Jones JO, Kraman M, et al. Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proc Natl Acad Sci USA*. 2013;110:20212-20217.
39. Djurec M, Graña O, Lee A, et al. Saa3 is a key mediator of the protumorigenic properties of cancer-associated fibroblasts in pancreatic tumors. *Proc Natl Acad Sci USA*. 2018;115:E1147-E1156.
40. Conrad AH, Zhang Y, Tasheva ES, Conrad GW. Proteomic analysis of potential keratan sulfate, chondroitin sulfate A, and hyaluronic acid molecular interactions. *Invest Ophthalmol Vis Sci*. 2010;51:4500-4515.
41. Breuer K, Foroushani AK, Laird MR, et al. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res*. 2013;41:D1228-D1233.
42. Rienks M, Papageorgiou A, Wouters K, et al. A novel 72-kDa leukocyte-derived osteoglycin enhances the activation of toll-like receptor 4 and exacerbates cardiac inflammation during viral myocarditis. *Cell Mol Life Sci*. 2017;74:1511-1525.
43. Tasheva ES, Koester A, Paulsen AQ, et al. Mimecan/osteoglycin-deficient mice have collagen fibril abnormalities. *Mol Vis*. 2002;8:407-415.
44. Van Aelst LNL, Voss S, Carai P, et al. Osteoglycin prevents cardiac dilatation and dysfunction after myocardial infarction through infarct collagen strengthening. *Circ Res*. 2015;116:425-436.
45. Sauerbrei W, Taube SE, Mcshane LM, Cavenagh MM, Altman DG, Sauerbrei W. Reporting recommendations for tumor marker prognostic studies (REMARK): an abridged explanation and elaboration. *J Natl Cancer Inst*. 2018;110:88-811.
46. Uhlen M, Fagerberg L, Hallstrom BM, et al. Tissue-based map of the human proteome. *Science (80- )*. 2015;347:1260419.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Dings MPG, Manoukian P, Waasdorp C, et al. Serum levels of iCAF-derived osteoglycin predict favorable outcome in pancreatic cancer. *Int J Cancer*. 2023;152(3):511-523. doi:[10.1002/ijc.34276](https://doi.org/10.1002/ijc.34276)