



## OPEN Evaluation of tumor targets selected from public genomic databases for imaging of pancreatic ductal adenocarcinoma

Nada Badr<sup>1</sup> , Luca Ten Elshof<sup>1</sup>, Ruben D. Houvast<sup>1</sup>, Lysanne D. A. N. de Muynck<sup>1</sup>, A. Stijn L. P. Crobach<sup>2</sup>, Gerard J. P. van Westen<sup>3</sup>, Ronald L. P. van Vlierberghe<sup>1</sup>, J. Sven D. Mieog<sup>1</sup>, Alexander L. Vahrmeijer<sup>1</sup> & Peter J. K. Kuppen<sup>1</sup>

Pancreatic ductal adenocarcinoma (PDAC) is an aggressive malignancy with a 5-year survival rate of approximately 5–7%, and complete surgical resection remains the only curative treatment but is often unfeasible. Fluorescence-guided surgery (FGS) using tumor-targeted probes may improve tumor visualization and facilitate complete resection. This study aimed to identify and validate tumor targets for FGS during PDAC resection procedures. RNA expression data from over 4000 cell surface genes, obtained from public genomic databases, were analyzed to identify genes encoding PDAC-associated proteins. Eleven potential tumor targets were identified, including CEACAM5, TMPRSS4, COL17A1, CLDN18, and AQP5. Protein expression was evaluated by immunohistochemistry (IHC) in tissues from 44 PDAC and 7 chronic pancreatitis (CP) patients. All targets, except COL17A1, showed significantly higher expression in PDAC tissue compared to healthy pancreatic, CP, and duodenal tissue ( $p < 0.001$ ), as well as in tumor-positive versus tumor-negative lymph nodes. Especially CEACAM5, TMPRSS4, and AQP5 were identified as the most promising targets for distinguishing PDAC from healthy tissues and detecting lymph node metastasis during FGS. The development of probes targeting multiple markers, such as AQP5 with CEACAM5 and/or TMPRSS4, may help overcome interpatient variability and enhance detection across patients.

**Keywords** Pancreatic cancer, Tumor-target selection, Fluorescence-guided surgery, Data-driven approach, Molecular imaging

### Abbreviations

CP	Chronic pancreatitis
CSPA	Cell-surface protein atlas
DAB	3,3'-diaminobenzidine
Duo	Duodenal tissue
FFPE	Formalin-fixed paraffin-embedded
FGS	Fluorescence-guided surgery
GTex	Genotype-tissue expression
HE	Hematoxylin and eosin
HPA	Human Protein Atlas
HRP	Horseradish peroxidase
ICG	Indocyanine green
IHC	Immunohistochemistry
MSI	Mean staining intensity
NAT	Neoadjuvant therapy
NIR	Near-infrared
Panc	Pancreatic tissue
PBS	Phosphate-buffered saline

<sup>1</sup>Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands. <sup>2</sup>Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands. <sup>3</sup>Drug Discovery and Safety, Leiden Academic Centre of Drug Research, Leiden, The Netherlands. ✉email: n.badr@lumc.nl

PDAC	Pancreatic ductal adenocarcinoma
PE	Protein expression
R0	Complete resections
R1	Incomplete resections
ROI	Regions of interest
RT	Room temperature
TCGA	The Cancer Genome Atlas
TNLN	Tumor-negative lymph node
T/N	Tumor-to-normal ratio
TPLN	Tumor-positive lymph node
TPM	Transcript per million

Pancreatic ductal adenocarcinoma (PDAC) accounts for 90% of pancreatic cancers and is known for its poor clinical prognosis with a 5-year survival rate of around 5–7%<sup>1</sup>. Surgery with pre-operative chemo(radio)therapy is the only potentially curative treatment approach. However, survival remains poor, and local recurrence rates are high, mainly due to incomplete resections (R1). Complete resections (R0) have been linked to improved patient survival<sup>2</sup>. R1 resections are often caused by inadequate tumor localization during surgery<sup>3</sup>. Existing preoperative imaging methods may present limitations in accurately distinguishing between healthy tissue and vital tumor tissue, chronic pancreatitis (CP), or fibrosis<sup>3,4</sup>. Consequently, approximately 7% of pancreatic surgeries are performed unnecessarily for benign diseases<sup>5</sup>. In addition, the effects of neoadjuvant therapy (NAT) further complicate adequate tumor recognition before and during surgery. Preoperative imaging techniques may, moreover, fail to detect small metastatic lesions<sup>6</sup>. Lastly, intraoperative assessment of tumor margins through visual inspection, palpation, or imaging techniques is difficult due to the complex spread pattern of PDAC and the presence of inflammatory tissue<sup>7</sup>.

Real-time fluorescence-guided surgery (FGS) allows surgeons to accurately visualize and fully remove tumors. Additionally, FGS assists surgeons in identifying metastatic lymph nodes during the procedure, which is crucial, as the removal of metastatic lymph nodes is associated with improved patient survival<sup>1</sup>. Indocyanine green (ICG) is an example of a clinically approved non-targeted near-infrared (NIR) dye that accumulates in areas with increased blood flow such as tumors or inflamed tissue. However, its effectiveness in pancreatic cancer is limited due to the lack of an enhanced permeability and retention effect, a consequence of the highly stromal nature of PDAC<sup>8</sup>. Tumor-targeted probes consisting of a targeting moiety conjugated to a NIR fluorescent dye have been proposed for real-time FGS in PDAC and their feasibility has been demonstrated<sup>9</sup>. Favorable properties of a target include an extracellular location, and homogeneous overexpression in tumor tissue and metastases compared to normal surrounding (inflammatory) tissue<sup>10</sup>.

Some previously developed tracers for pancreatic cancer have targeted CEACAM5<sup>11</sup>, EGFR<sup>12</sup>, and VEGFR/VEGF-A<sup>13</sup>, among others. Due to high intra and inter-patient heterogeneity, the currently available tracers do not cover all PDAC patients<sup>14,15</sup>. Moreover, several of these tracers fail to distinguish between CP and PDAC, or cannot detect micro-metastases<sup>15,16</sup>. In this study, we use a fast and unbiased method to identify tumor targets that are highly expressed in primary tumor tissue compared to surrounding tissue. In addition, we investigate the protein expression of these targets in surrounding lymph nodes to determine whether they could be used to detect lymph node metastasis. Like in several other studies, we have investigated the effect of NAT on target expression<sup>3,4,15</sup>. Currently, potential tumor targets are often discerned based on a literature search on known (PDAC-associated) biomarkers<sup>3,17</sup>. However, this can lead to a biased search in which potential targets may be overlooked<sup>18</sup>. Euretos (Utrecht, The Netherlands) offers a data-driven pre-clinical discovery tool that considers omics data of the Genotype-Tissue Expression (GTEx) and The Cancer Genome Atlas (TCGA)<sup>19</sup>. Its feasibility to identify novel targets has been demonstrated in several recent studies<sup>18,20</sup>. Moreover, this tool enables the differentiation of gene expression levels of targets in tumor tissue, selected healthy tissue, and cell types like immune cells. In this study, we used this data-driven approach to identify PDAC tumor targets that may be suitable for FGS. This search engine specifically visualizes RNA expression levels in selected tissue, rather than protein expression. The protein expression of selected targets was therefore evaluated using immunohistochemistry.

## Results

### PDAC target identification using Euretos

We started by identifying genes that are potential biomarkers for targeting PDAC, based on RNA expression. Using Euretos, we identified membranous targets that are highly expressed in PDAC compared to surrounding tissues, including healthy pancreatic, CP, and duodenal tissues. A gene set consisting of 4673 genes likely to be expressed on the cell surface was used for the data-driven search using the Euretos search engine. After application of the expression thresholds, 11 potential targets remained (Table 1). CEACAM5 and MSLN are well-known and extensively studied PDAC targets<sup>4,14</sup>, and Euretos showed that both had high PDAC TPM rates, along with low TPM rates in healthy pancreatic tissues. As the aim of this study was to investigate lesser-known potential PDAC targets, we decided to include only CEACAM5 as a reference target. MUCL3 was not selected for IHC evaluation due to high expression in selected normal tissue as determined from a preliminary experiment (data not shown). PSCA has been shown to have low specificity and was therefore also excluded<sup>21</sup>. VSIG2 and FXYD3 showed high protein expression in normal tissue (Table 1). We used a cutoff tumor-to-normal (T/N) TPM ratio of 10, as previously reported<sup>10</sup>, to limit the list to a feasible number of targets to be further investigated. Therefore, GJB3 was not selected. Thus, CEACAM5, TMPRSS4, COL17A1, CLDN18, and AQP5 were selected as candidate genes for further investigation.

For effective visualization of PDAC in FGS and clear differentiation from CP, a high expression of tumor targets in epithelial cells combined with low expression in immune cells is ideal, as CP is characterized by

Potential target	Predicted location	T/N ratio (TCGA)	TPM PDAC (TCGA)	TPM Panc (TCGA)	PE Panc	PE Duo	PE Liver	3D crystal structure
CEACAM5	Membrane	353.2	257.8	0.7	ND	ND	Low	Yes
MUCL3	Membrane/cytoplasm	152.0	57.8	0.4	ND	ND	ND	No
TMPRSS4	Membrane/intracellular	131.0	72.1	0.6	Medium	Medium	Low	No
COL17A1	Cell junction, desmosome, membrane	96.8	78.4	0.8	ND	ND	ND	No
MSLN	Membrane	67.5	272.7	4.0	ND	ND	ND	No
CLDN18	Membrane	64.4	133.3	2.1	ND	ND	ND	No
PSCA	Membrane	51.8	135.2	2.6	Low	ND	ND	No
AQP5	Membrane, cytoplasmic vesicle membrane	37.9	68.9	1.8	Low	Low	Low	Yes
VSIG2	Membrane	18.9	159.0	8.4	High	High	Medium	No
FXYD3	Membrane	17.4	236.8	13.6	ND	Medium	ND	No
GJB3	Membrane, cell junction	9.3	56.0	6.0	ND	ND	ND	No

**Table 1.** Characteristics of genes of interest obtained from the Euretos-based search, ranked by T/N ratio in TCGA. The median of the transcripts per million (TPM) values is given. Protein expression (PE) levels are as given by the Human Protein Atlas (HPA), with ND: Not Detected. PDAC: Pancreatic Ductal Adenocarcinoma. Panc: healthy pancreatic tissue. Duo: Duodenal tissue. T/N: PDAC-to-Panc ratio. TCGA: The Cancer Genome Atlas.

extensive immune cell infiltration. The ‘Single Cell Study Viewer’ tool of Euretos was used to further investigate in which types of cells these targets are highly expressed. As shown in Figure S1, the RNA expression levels of the selected genes were all higher in tumor epithelial cells compared to immune cells, including macrophages, mast cells, and natural killer cells.

In addition, a volcano plot analysis was performed to visualize the overlapping genes between the surfaceome gene set and the genes from TCGA that are correlated with PDAC. In this analysis, the focus was solely on the upregulation and downregulation of specific genes in PDAC, without considering healthy pancreas or surrounding tissue. The genes of interest, shown in Table 1, are highlighted in the volcano plot. Each of these genes was significantly upregulated in PDAC (Supplementary Figure S2).

### Patient selection and clinicopathological characteristics

Tissue samples were selected from 51 patients who underwent pancreatic resection (Table 2). In the studied cohort, a total of 44 patients were diagnosed with PDAC, and 7 patients were ultimately identified as tumor-free and diagnosed with CP. In the PDAC group, 17 patients received NAT, of whom 12 patients received chemoradiotherapy and 5 patients received chemotherapy. Samples included healthy pancreas tissue from 31 patients, PDAC from 33 patients, Tumor-positive lymph nodes (TPLN) from 26 patients, Tumor-negative lymph nodes (TNLN) from 25 patients, CP from 7 patients, and duodenum from 11 patients.

### Evaluation of CEACAM5, TMPRSS4, COL17A1, CLDN18 and AQP5 protein expression

The protein expressions of CEACAM5, TMPRSS4, COL17A1, CLDN18, and AQP5 in representative PDAC, CP, pancreatic, duodenal, and liver tissue sections are shown in Figure 1.

In PDAC tissue, high levels of CEACAM5 expression were observed, particularly associated with the cell membrane, showing substantial inter-patient variability across the cohort. Protein expression of CEACAM5 was relatively low across healthy pancreatic tissues, including exocrine acinar cells, ducts, and stroma, as well as in duodenal, CP, and liver tissues. TMPRSS4 expression was increased in PDAC tissue, predominantly in epithelial ductal cells, compared to selected healthy tissues, including the pancreas, duodenum, and CP (Fig. 1).

Expression levels of COL17A1 and CLDN18 were moderate and high in PDAC, respectively, and heterogeneous among patients. Moreover, high staining intensity of both proteins was observed in tumor-negative tissues. These proteins were present in pancreatic tissues, specifically within exocrine cells, islets of Langerhans, and epithelial cells lining the ducts. In duodenal tissue, the expression of both targets was predominantly localized to the mucosal layer, with a particular presence in the intestinal glands. In CP tissues, COL17A1 staining showed low intensity, whereas CLDN18 staining marked positive epithelial cells within ducts and regions of fibrotic tissue. In addition, CLDN18 was highly expressed in healthy liver tissue, in which COL17A1 appeared to be minimally expressed.

AQP5 expression was higher in PDAC tissue, particularly in epithelial ductal cells, whereas it remained low in healthy pancreatic, CP, and liver tissues. Notably, AQP5 showed high expression in the intestinal glands of duodenal tissue.

### Protein expression quantification using ImageJ

The mean staining intensity (MSI) was quantified for each relevant tissue type (PDAC, pancreas, CP, and duodenum). Figure 2 shows the MSIs of all targets in relevant tissues. For CEACAM5, TMPRSS4, CLDN18, and AQP5, the MSIs in PDAC were significantly higher (all  $p < 0.001$ ) compared to those in pancreatic, CP, and

Clinical pathological characteristics	Total (n = 51)	PDAC (n = 44)	PDAC no NAT (n = 27)	PDAC NAT (n = 17)	CP (n = 7)
Age (years), mean (SD)	62.7 (11.1)	64.9 (9.8)	66.7 (9.5)	62.1 (9.8)	49 (9.4)
Gender, n (%)					
Male	26 (51.0)	20 (45.5)	13 (48.1)	7 (41.2)	6 (85.7)
Female	25 (49.0)	24 (54.5)	14 (51.9)	10 (58.8)	1 (14.3)
Surgery type, n (%)					
Pancreaticoduodenectomy	36 (70.6)	33 (75.0)	23 (85.2)	13 (76.5)	3 (42.9)
Pancreatic corpus/tail resection	12 (23.5)	8 (18.2)	2 (7.4)	3 (17.6)	4 (57.1)
Total pancreatectomy	3 (5.9)	3 (6.8)	2 (7.4)	1 (5.9)	0 (0)
Tumor differentiation, n (%)					
Good	–	5 (11.4)	3 (11.1)	1 (5.9)	–
Moderate	–	11 (25.0)	10 (37.0)	1 (5.9)	–
Poor	–	14 (31.8)	11 (40.7)	2 (11.8)	–
Missing	–	14 (31.8)	3 (11.1)	13 (76.5)	–
Primary tumor, n (%)					
pT0	–	2 (4.5)	1 (3.7)	1 (5.9)	–
pT1	–	13 (29.5)	6 (22.2)	7 (41.2)	–
pT2	–	23 (52.3)	15 (55.6)	8 (47.1)	–
pT3	–	6 (13.6)	5 (18.5)	1 (5.9)	–
Regional lymph nodes, n (%)					
pN0	–	11 (25.0)	4 (14.8)	7 (41.2)	–
pN1	–	19 (43.2)	11 (40.7)	8 (47.1)	–
pN2	–	14 (31.8)	12 (44.4)	2 (11.8)	–
NAT, n (%)					
No	34 (66.7)	27 (61.4)	27 (100.0)	0	7 (100.0)
Yes, chemoradiotherapy	12 (23.5)	12 (27.3)	0 (0)	12 (70.6)	0 (0)
Yes, chemotherapy	5 (9.8)	5 (11.4)	0 (0)	5 (29.4)	0 (0)
Tumor size (mm), mean (SD)	–	26.7 (11.5)	30.0 (11.9)	21.6 (8.8)	–

**Table 2.** Clinicopathological characteristics of PDAC patients ( $n = 44$ ) and CP patients ( $n = 7$ ) in the immunohistochemistry cohort. PDAC patients are categorized into two groups: NAT patients and no NAT patients. *PDAC* pancreatic ductal adenocarcinoma, *CP* chronic pancreatitis, *NAT* neoadjuvant therapy, *SD* standard deviation.

duodenal tissues. For COL17A1, no significant difference was observed between the MSIs of PDAC tissue and pancreatic or duodenal tissues. Expression of COL17A1 was only significantly higher in PDAC compared to CP tissue ( $p < 0.05$ ).

To account for inter-patient heterogeneity in target expression, MSIs of PDAC and relevant tissues were collected and visualized for each patient, as shown in Figure 3. In all patients, CEACAM5, TMPRSS4, and AQP5 showed significantly higher MSIs in PDAC compared to pancreatic tissues (all  $p < 0.001$ ). CLDN18 expression was significantly higher in PDAC tissue compared to healthy pancreatic tissue ( $p < 0.001$ ). However, due to the high average MSI level in pancreatic tissue, the differences in expression were less pronounced in some patients. The difference in MSI between PDAC and pancreatic tissue stained for COL17A1 was not significant. Additionally, COL17A1 expression showed a heterogeneous pattern across patients, with some pancreatic tissues exhibiting higher protein levels than those observed in PDAC tissues.

In Figure 3f, a heat map of T/N ratios for individual patients illustrates the variability of each target across the patient cohort. The mean T/N ratios for CEACAM5, TMPRSS4, COL17A1, CLDN18, and AQP5 were 7.7, 3.2, 1.2, 1.5, and 17.3, respectively. A T/N ratio greater than 10 is considered sufficient for the selection of potential targetable biomarkers for imaging<sup>10</sup>. For CEACAM5 expression, 6 of 19 patients had a T/N ratio greater than 10. In contrast, none of the patients showed a T/N ratio above 10 for TMPRSS4, COL17A1, or CLDN18. However, for AQP5 expression, 16 of 19 patients had a T/N ratio of at least 10. Among the three patients with a T/N ratio below 10 for AQP5, two had T/N ratios higher than those of CEACAM5.

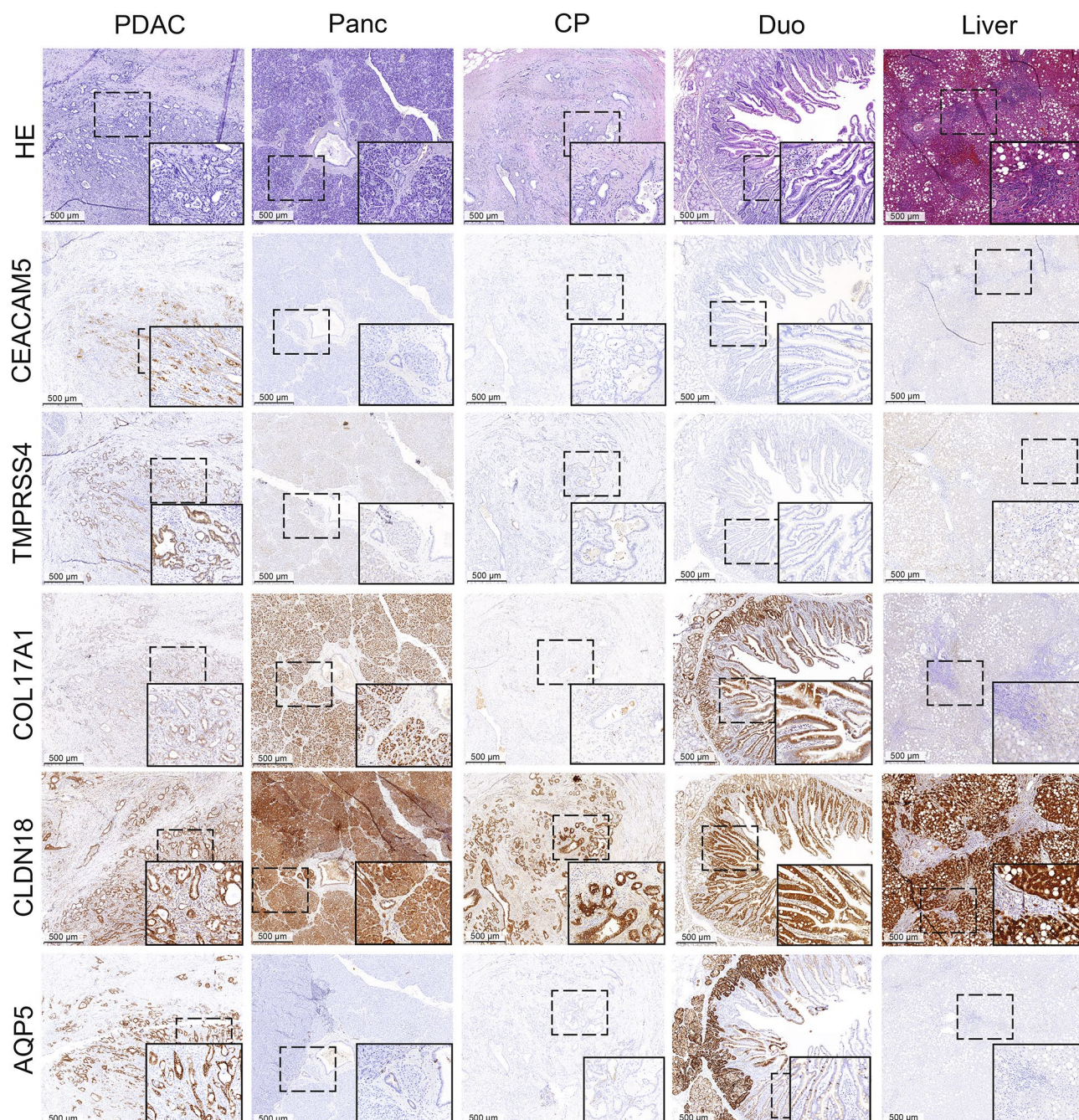
### Target expression after neoadjuvant(chemo)radiotherapy

To determine whether NAT affects tumor target expression levels, MSIs of CEACAM5, TMPRSS4, COL17A1, CLDN18, and AQP5 were compared between patients who received NAT and those who did not receive NAT, as shown in Supplementary Figure S3. Statistical analysis showed no significant differences in MSI values between the NAT and no-NAT groups for any of the investigated target proteins.

### Detection of lymph node metastases with targets of interest

The potential of the selected targets in detecting lymph node metastases was evaluated by immunohistochemical staining of TPLN and TNLN tissues from the same patients (Fig. 4). As shown in Figure 4a, high levels of



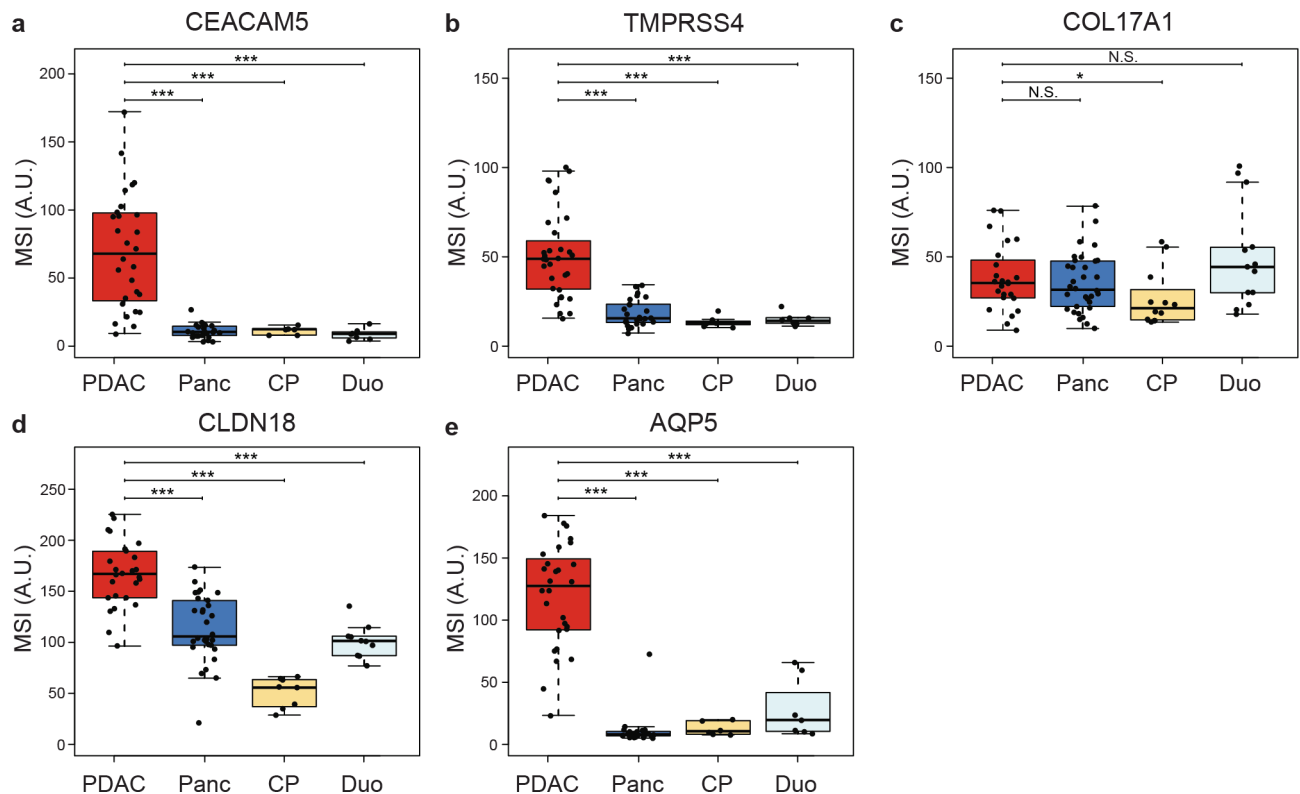


**Fig. 1.** Representative immunohistochemical staining images of tissue sections stained with anti-CEACAM5, anti-TMPRSS4, anti-COL17A1, anti-CLDN18, and anti-AQP5. Scale bars represent 500  $\mu$ m. HE: hematoxylin and eosin staining. Tissue types: PDAC: pancreatic ductal adenocarcinoma, Panc: pancreas, CP: chronic pancreatitis, Duo: duodenal tissue, Liver: liver tissue.

CEACAM5 expression were observed in TPLN within PDAC epithelial cells, whereas CEACAM5 expression was low in TNLN. A similar pattern was observed for TMPRSS4, and COL17A1 showed this to a lesser extent. Conversely, CLDN18 had higher expression levels in TNLN compared to TPLN. Lastly, AQP5 expression was low in TNLN and high in TPLN, as shown in Figure 4a. Staining intensities of all targets were significantly higher in TPLN tissue (all  $p < 0.001$ , except for CLDN18 for which  $p < 0.05$ ) compared to TNLN tissue as illustrated in Figures 4b-f.

In addition, a heat map of TPLN/TNLN T/N ratios for individual patients (Fig. 4g) shows the heterogeneity of each target across the patient population. The mean T/N ratio for CEACAM5, TMPRSS4, COL17A1, CLDN18, and AQP5 is 9.7, 5.9, 1.6, 1.8, and 9.7, respectively. For CEACAM5 expression, 6 of 17 patients had a T/N ratio above 10. TMPRSS4 showed this level of expression in 1 of 16 patients, while none of the patients exceeded this





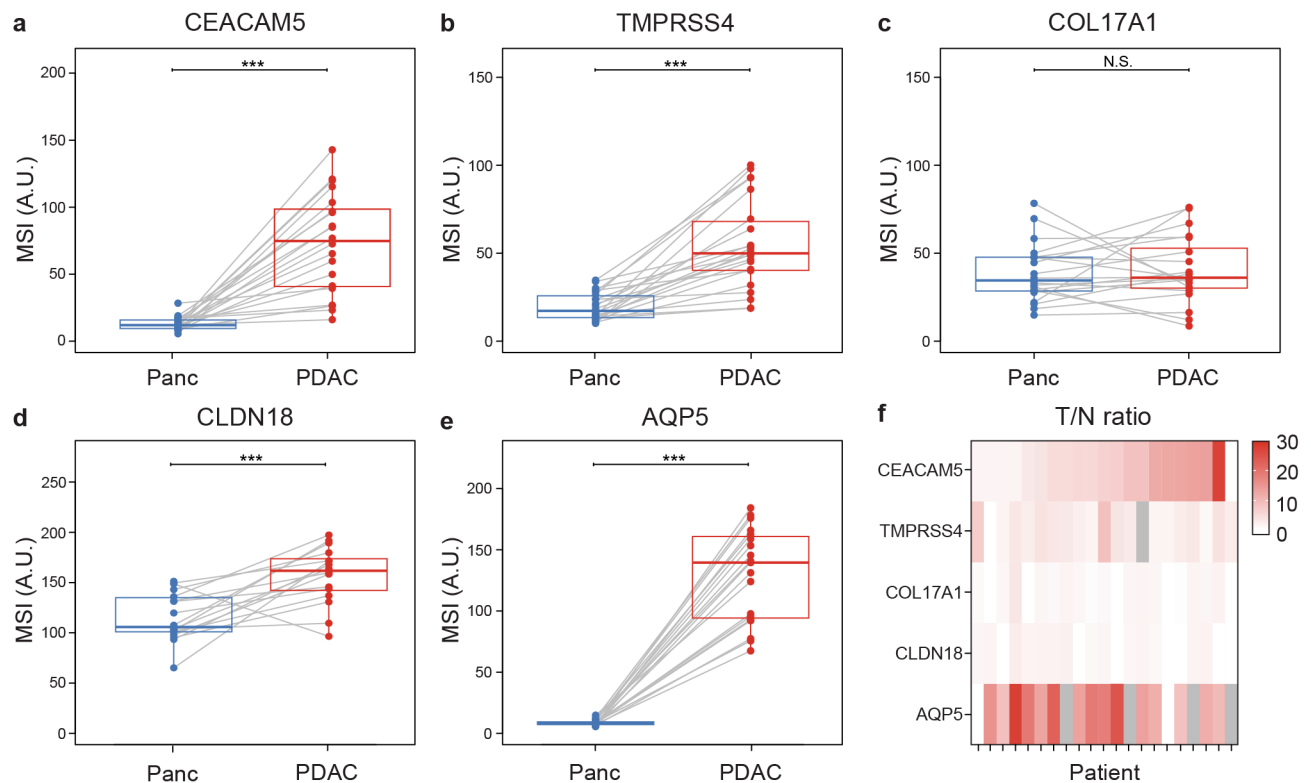
**Fig. 2.** Mean staining intensity (MSI) in tissue sections immunohistochemically stained with anti-CEACAM5, anti-TMPRSS4, anti-COL17A1, anti-CLDN18, and AQP5. **(a)** MSI for CEACAM5 expression. **(b)** MSI for TMPRSS4 expression. **(c)** MSI for COL17A1 expression. **(d)** MSI for CLDN18 expression. **(e)** MSI for AQP5 expression. Thick lines indicate medians, boxes represent the interquartile ranges (IQRs), whiskers indicate the 95% confidence intervals (CI), and circles data from individual patients. Horizontal bars indicate significance levels (Mann-Whitney U-test: \*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ ; N.S.; not significant). Tissue types: PDAC: pancreatic ductal adenocarcinoma, Panc: pancreas, CP: chronic pancreatitis, DUO: duodenal tissue.

threshold for COL17A1 and CLDN18. In contrast, AQP5 protein expression was increased in the TPLN tissue of 13 of 18 patients, with T/N ratios reaching at least 10.

## Discussion

FGS is useful when tumor identification is limited. Untargeted moieties, e.g., ICG, are recognized to lack the specificity to guide surgeons in pancreatic cancer and therefore achieve complete tumor and metastatic lymph node resection<sup>8</sup>. In this study, a data-driven approach was used to identify upregulated genes encoding for cell-surface proteins that can be targeted during FGS of PDAC. CEACAM5, TMPRSS4, COL17A1, CLDN18, and AQP5 were selected for further evaluation based on their RNA expression and predicted protein expression. Analysis of IHC stains revealed high protein expression levels of CEACAM5, TMPRSS4, and AQP5 in PDAC compared to healthy pancreatic tissue. Moreover, these tumor targets were significantly higher expressed in PDAC than in CP, supporting their use as intraoperative biomarkers to differentiate PDAC from CP. Although COL17A1 and CLDN18 were initially selected as potential targets based on their RNA expression and predicted protein levels, high protein expression was also detected in normal healthy pancreatic tissue. This highlights the importance of evaluating protein expression to refine target selection. The study also investigated several other factors relevant to potential FGS targets for pancreatic cancer. To assess the suitability of these targets for distinguishing locally advanced pancreatic head carcinomas with duodenal invasion<sup>15</sup>, protein expression was also evaluated in healthy duodenal tissue. CEACAM5, TMPRSS4, CLDN18, and AQP5 expressions were significantly higher in PDAC than in duodenal tissue. Moreover, tumor target expression remained consistent in stains of both NAT-treated and untreated patients for all targets, suggesting that neoadjuvant therapy does not affect target expression in PDAC tumors. Furthermore, CEACAM5 and AQP5 were significantly higher expressed in tumor-positive compared to tumor-negative lymph nodes and demonstrated particularly high T/N ratios, indicating their potential to detect metastatic lymph nodes for FGS and staging<sup>15</sup>.

Our target selection was based on a data-driven search in Euret. Among the top selected genes of interest, CEACAM5, PSCA, and MSLN were also included, all of which are already well-studied targets for the treatment and imaging of PDAC<sup>4,14,22,23</sup>. The IHC staining results of CEACAM5 showed high protein expression in PDAC but with substantial inter-patient heterogeneity. Heterogeneity of CEACAM5 expression has been observed

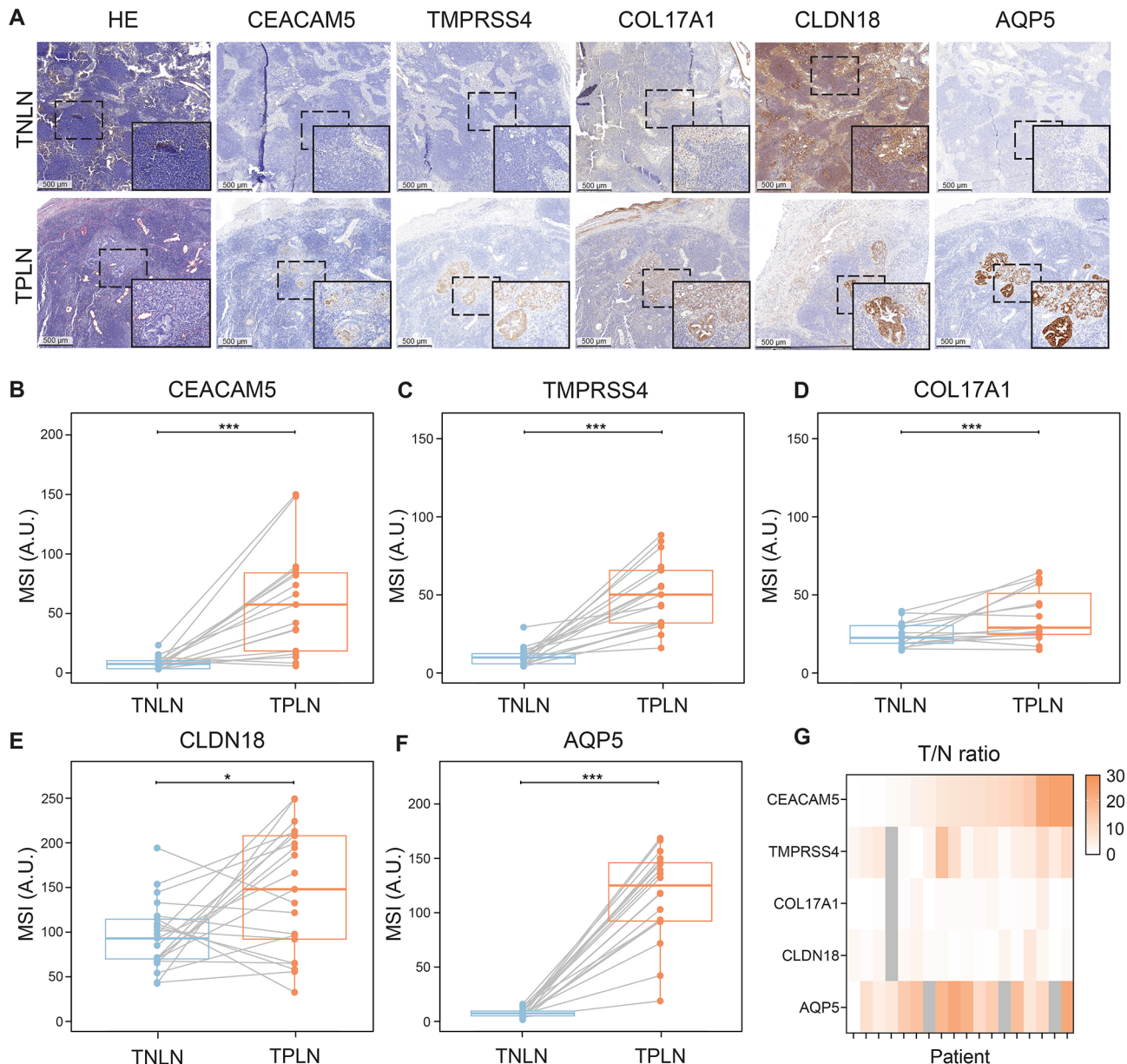


**Fig. 3.** Paired analysis of the mean staining intensities (MSIs) of targets in tissue sections from patients with both pancreatic ductal adenocarcinoma (PDAC) and pancreas (Panc) tissue available. **(a)** MSI for CEACAM5 expression. **(b)** MSI for TMPRSS4 expression. **(c)** MSI for COL17A1 expression. **(d)** MSI for CLDN18 expression. **(e)** MSI for AQP5 expression. Dots and lines represent data from individual patients. Thick lines indicate medians and boxes represent interquartile ranges (IQRs). Horizontal bars indicate significance levels (Wilcoxon-signed rank test: \*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ ; N.S.; not significant). **(f)** Heat map showing the T/N ratios (MSI PDAC/MSI Panc) for patients with available PDAC and Panc data. Gray regions indicate missing data.

before and may be associated with neoadjuvant therapy<sup>3,4</sup>. In our study, the protein expression of the selected targets, including CEACAM5, was similar between patients who received NAT and those who did not. However, our sample size for these particular analyses was relatively limited, which may have affected the results. TMPRSS4 protein expression in PDAC has been studied previously and has been identified as a promising therapeutic target<sup>24</sup>. Similar IHC staining patterns observed in this study further support its potential applicability for FGS. COL17A1, in contrast to previous IHC findings<sup>25</sup>, was not found to be a suitable target due to its high protein expression in both PDAC and healthy pancreatic tissues. CLDN18 has previously been identified as a biomarker for PDAC using a data-driven approach<sup>26</sup>. The utility of CLDN18 for image-guided surgery has been described and two CLDN18-specific imaging probes were able to facilitate the resection of tumor tissue in orthotopic mouse tumor models<sup>27</sup>. We observed high protein expression levels of CLDN18 in normal pancreatic tissues, adding to conflicting findings reported in the literature about CLDN18 expression<sup>28,29</sup>. The conflicting findings reflect the broader inconsistency in the literature regarding protein expression patterns of COL17A1 and CLDN18. These differences may partially result from variability in the affinity of the antibodies or epitope recognition. As such, future studies employing multiple antibodies per target could provide additional clarity and aid in distinguishing between technical variation and biological differences.

AQP5 and COL17A1 expression have recently been linked to PDAC progression<sup>25,30</sup>. In addition, AQP5 expression has been related to PDAC tumor differentiation<sup>31</sup>. In our study, AQP5 T/N ratios were consistently high over all patients relative to the other studied targets. AQP5 did exhibit expression in normal ductal epithelial cells, consistent with previous research<sup>32</sup>. The absence in the top-ranked results of other widely investigated targets, such as EGFR, can be attributed to their relatively high or ambiguously reported expression levels in healthy pancreatic tissues<sup>33</sup>. This led to their exclusion during the selection process. It must be noted that the found T/N ratios may not reflect clinically feasible T/N ratios to the target. In addition, some already-known targets have T/N ratios of less than 10 and are used as biomarkers for FGS<sup>10</sup>.

Traditional target screening methods, including literature searches, may be subject to several kinds of bias, such as publication bias. By combining omics data with artificial intelligence to handle various data inputs, data-driven tools facilitate the unbiased discovery of therapeutic targets that traditional techniques may fail to detect<sup>34</sup>. In this study, we used Euretos as its potential had been demonstrated previously<sup>18,20</sup>. Other data-driven



**Fig. 4.** (a) Representative immunohistochemical (IHC) staining of tissue sections stained with anti-CEACAM5, anti-TMPRSS4, anti-COL17A1, anti-CLDN18, and anti-AQP5. Scale bars represent 500  $\mu$ m. HE: hematoxylin and eosin staining. Tissue types: TNLN: Tumor-negative lymph node, TPLN: Tumor-positive lymph node. (b–f) Paired analysis of the mean staining intensities (MSIs) of targets in tissue sections from patients with both TNLN and TPLN tissues available. (b) MSI for CEACAM5 expression. (c) MSI for TMPRSS4 expression. (d) MSI for COL17A1 expression. (e) MSI for CLDN18 expression. (f) MSI for AQP5 expression. Dots and lines represent data of individual patients. Thick lines indicate medians and boxes represent interquartile ranges (IQRs). Horizontal bars indicate significance levels (Wilcoxon-signed rank test: \*\*\*,  $p < 0.001$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ ; N.S. = not significant). (g) Heat map showing the T/N ratio (MSI TPLN / MSI TNLN) for patients in whom TNLN and TPLN data were available. Gray regions indicate missing data.

approaches have been shown to be successful in identifying PDAC targets as well<sup>26,35,36</sup>. However, a limitation of the use of bulk RNA sequencing data is that samples contain a mixture of different cell types, which can be misleading when a biomarker is only present in a subset of cells<sup>37</sup>. Single-cell RNA sequencing overcomes this limitation by analyzing expression at the individual cell level. Euretos' 'Single Cell Study Viewer' tool was used in this study to ensure that promising targets were also all highly expressed in epithelial cells and minimally in immune cells. An additional limitation may be the predominant use of TCGA data, which mainly includes cancer and less healthy tissue samples, and may be subject to batch effects or other biases<sup>35</sup>. In this study, GTEx data were not included as Euretos does not currently apply a normalization strategy compatible with combining GTEx and TCGA datasets. Integrating these datasets could provide valuable insights by enabling more accurate



comparison between normal and tumor tissue expression<sup>38</sup>. Moreover, some potential targets could not be retrieved in EuretOS. Of one known target,  $\alpha\text{v}\beta 6$ <sup>33,39</sup>, only the gene encoding for the subunit B6 (ITGB6) could be found. Prior research indicates a relatively high expression of this gene in normal tissue<sup>3,4</sup>. In addition, EuretOS may not recognize potential targets that originate from the same gene but have different isoforms. For instance, although CD44 has multiple isoforms, the CD44 variant 6 (v6) has been specifically identified and used as a target for FGS in head and neck squamous cell carcinoma<sup>40</sup>. Nevertheless, a combination of search strategies, e.g. adding literature screening, could aid the selection of targets<sup>20</sup>. Lastly, one of the major limitations of the use of omics search tools for tumor target identification is that RNA expression does not always correlate with protein expression<sup>41</sup>, as was shown for COL17A1 and CLDN18. Post-transcriptional modifications can affect the stability, location, and translation efficiency. Moreover, protein turnover dynamics determine the efficiency of the complex networks in which proteins function. This was addressed by evaluating protein expression through IHC, which is a key strength of this research. Data-driven approaches specifically focusing on protein expression and taking into account isoforms or multimers would help identify novel targets.

A key challenge in tumor target identification is to find targets that are consistently highly expressed across all patients and within patients. In this study, we focused on inter-patient heterogeneity, showing that target expression in PDAC is subject to substantial variability between patients, as previously indicated<sup>14</sup>. EuretOS helps address this challenge by enabling the user to visualize patient-specific data. For future research, it will be important to also consider intra-patient heterogeneity, particularly when translating PDAC tumor targets into clinical applications. Developing FGS probes that are robust to inter- and intra-patient heterogeneity, including multi-target tracers, may enhance the effectiveness of PDAC resection surgeries<sup>33</sup>. However, it may be challenging to design a tracer that can simultaneously bind to multiple targets with high affinity. Another interesting development to handle target heterogeneity is multiplex imaging, that is, injecting an array of probes targeting several tumor biomarkers and potentially using multiple imaging modalities<sup>42</sup>. Based on our study's findings, imaging of tumors using AQP5 combined with CEACAM5 and/or TMPRSS4 may enhance the detection of PDAC across a broader patient population compared to any of the targets alone. The availability of the protein crystal structures of AQP5 and CEACAM5 also makes these promising candidates for further probe development using computational screening programs<sup>39</sup>.

The data-driven selection method using EuretOS proved to be a valuable tool for identifying imaging targets in pancreatic cancer. While experimental validation of protein expression is still necessary, this approach allows for rapid and unbiased selection of potential tumor targets for FGS. Four of the five candidate targets selected showed significantly higher expression in PDAC than surrounding tissues and could detect lymph node metastases. Among the potential targets for FGS, AQP5 emerged as the most promising, alongside CEACAM5 and TMPRSS4, due to its especially low expression in healthy tissue. A combination of AQP5 with CEACAM5 and/or TMPRSS4 holds the most promise for addressing the interpatient heterogeneity of PDAC.

## Methods

### Patient and tissue selection

Formalin-fixed, paraffin-embedded (FFPE) tissue blocks from 51 patients who underwent pancreatic resection between 2011 and 2020 at the Leiden University Medical Centre (LUMC) and were diagnosed with PDAC or CP were selected for this study. NAT was administered to 17 of these patients. The tissue blocks contained PDAC, Panc, CP, TPLN, TNLN, and/or duodenal tissue. Medical and pathology records were retrospectively collected. This study was approved by the Gastroenterology Biobank Review Committee (protocol reference: 2020-16), and the local medical ethics review committee (protocol reference: B20.052). The study was conducted in compliance with the Dutch code of conduct for the responsible use of human tissue in medical research.

### Biomarker selection

Using the EuretOS search engine (accessed in May 2023; Utrecht, The Netherlands), we compared mRNA levels in transcripts per million (TPM) in Pancreatic Ductal Adenocarcinoma (PDAC) tissue with healthy tissue, obtained from TCGA. The surfaceome gene set was used to select genes that are predicted to encode proteins expressed on the cell surface. This gene set is based on four sources, including two computational surfaceome lists, mass-spectrometry data obtained from Cell Surface Protein Atlas (CSPA), and an EuretOS surfaceome list<sup>43–46</sup>. Expression thresholds were established for normal pancreas tissue (< 15 TPM), surrounding tissue (bile duct and liver, < 30 TPM), and PDAC (> 50 TPM). The tumor-to-normal (T/N) ratio calculation for the ranking of genes was performed using the TCGA dataset exclusively. We sorted the resulting targets by their T/N ratio. In addition, the selection process took into account the following criteria: (1) expression in normal tissues should not be higher than 50 TPM, (2) protein should be expressed on the membrane according to UniProt and the Human Protein Atlas (HPA), (3) protein expression in normal tissue according to the HPA should be low, and (4) a 3D crystal structure or predicted AlphaFold 2 (DeepMind, London, UK) structure with high confidence should be available to enable computational probe development. The gene set was additionally validated through the utilization of the 'Single Cell Study Viewer' tool in EuretOS<sup>47</sup>. This tool allows the estimation of expression levels per cell type in PDAC patients. To illustrate the RNA expression of the selected targets, a volcano plot with  $\text{Log}_2$ -fold changes was created using VolcanoSeR<sup>48</sup>. The input data consisted of genes from the surfaceome gene set and genes from TCGA associated with PDAC, filtered using the Xena Functional Genomics Explorer (University of California, Santa Cruz). cBioPortal was used to filter gene data specifically from PDAC patients.

### Hematoxylin and Eosin (HE) and immunohistochemistry (IHC) staining

FFPE tissue blocks were cut into sections of 4  $\mu\text{m}$  thickness. For hematoxylin and eosin staining, slides were twice incubated in xylene for 5 min, then rehydrated through a graded ethanol series (100%, 70%, 50%) and demineralized water. Sections were stained in Mayer's Hematoxylin (Klinipath, Duiven, The Netherlands) for

4 min, rinsed in tap water (10 min) and demineralized water (1 min), and then incubated in eosin (Dako, Glostrup, Denmark) for 2 min. After thorough rinsing in demineralized water, slides were dehydrated with ethanol concentrations of 50%, 70%, and 100% respectively. Lastly, the slides were incubated twice in xylene for 5 min and mounted in Pertex (Histolab, Askim, Sweden).

For immunohistochemical staining, tissue sections were deparaffinized in a series of xylene and rehydrated in descending ethanol solutions of 100%, 70%, and 50%. After washing with demineralized water, endogenous peroxidase was blocked with 0.3% hydrogen peroxide (Merck, New Jersey, USA) for 20 min. Depending on the antibody of interest, different antigen retrieval conditions were used in EnVision Flex Target Retrieval Solution using the PT-Link module (Dako, Glostrup, Denmark). The antigen retrieval consisted of heating at 95 °C for 10 min with different pH levels, which are shown in supplementary Table S1. After cooling and washing three times with phosphate-buffered saline (PBS), the slides were incubated with a solution of primary antibody in PBS with 1% BSA. The primary antibody concentrations and characteristics can be found in Supplementary Table S1. The incubation step took place in a humidified closed chamber at room temperature (RT) overnight. In the specific case of TMRSS4, a blocking step was performed using 5% normal goat serum for 10 min before the incubation with the primary antibodies. Next, the slides were washed three times with PBS and incubated with horseradish peroxidase (HRP)-labeled secondary antibody (anti-rabbit or anti-mouse Envision, Dako) for 30 min at RT. After washing three times with PBS, the staining was visualized using 3,3'-Diaminobenzidine (DAB) substrate buffer (Dako) for 10 min. Afterwards, DAB was deactivated by washing twice with demineralized water and counterstained with Mayer's Hematoxylin for 15 s. Finally, the slides were rinsed with tap water for 5 min, dehydrated at 37 °C for one hour, and mounted in Pertex.

### IHC staining scoring: semi-automated imaging analysis

The slides were scanned using the PANNORAMIC 250 Flash III scanner (3DHISTECH, Budapest, Hungary) at 20x magnification. Multiple regions of interest (ROIs) of PDAC, pancreatic, TPLN, TNLN, and/or duodenal tissues were selected per patient. In healthy pancreatic tissue, representative regions were selected that included exocrine cells, ducts, islets of Langerhans, and other pancreatic structures. In PDAC samples, only tumor ductal epithelial cells were selected as ROIs to assess target protein expression specifically within the malignant cell population. The regions were annotated by an experienced hepato-pancreato-biliary pathologist (A.S.L.P.C.). The mean staining intensity (MSI) of an annotated region was determined using ImageJ (1.54d, Java 1.8.0\_345).

### Statistical analysis

Statistical analysis was performed using RStudio (4.2.1). Wilcoxon signed-rank tests and Mann-Whitney U tests were used to compare paired and unpaired data, respectively, with a significance threshold of  $\alpha = 0.05$ .

### Data availability

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 16 January 2025; Accepted: 29 April 2025

Published online: 16 May 2025

### References

- Adamska, A., Domenichini, A. & Falasca, M. Pancreatic ductal adenocarcinoma: current and evolving therapies. *Int. J. Mol. Sci.* **18** <https://doi.org/10.3390/ijms18071338> (2017).
- Tummers, W. S. et al. Impact of resection margin status on recurrence and survival in pancreatic cancer surgery. *Br. J. Surg.* **106**, 1055–1065. <https://doi.org/10.1002/bjs.11115> (2019).
- Tummers, W. S. et al. Selection of optimal molecular targets for tumor-specific imaging in pancreatic ductal adenocarcinoma. *Oncotarget* **8**, 56816–56828. <https://doi.org/10.18632/oncotarget.18232> (2017).
- Vuijk, F. A. et al. Molecular targets for diagnostic and intraoperative imaging of pancreatic ductal adenocarcinoma after neoadjuvant FOLFIRINOX treatment. *Sci. Rep.* **10**, 16211. <https://doi.org/10.1038/s41598-020-73242-6> (2020).
- Gerritsen, A. et al. Preoperative characteristics of patients with presumed pancreatic cancer but ultimately benign disease: a multicenter series of 344 pancreatoduodenectomies. *Ann. Surg. Oncol.* **21**, 3999–4006. <https://doi.org/10.1245/s10434-014-3810-7> (2014).
- Verbeek, F. P. et al. Image-guided hepatopancreatobiliary surgery using near-infrared fluorescent light. *J. Hepatobiliary Pancreat. Sci.* **19**, 626–637. <https://doi.org/10.1007/s00534-012-0534-6> (2012).
- Tummers, W. S. et al. Advances in diagnostic and intraoperative molecular imaging of pancreatic Cancer. *Pancreas* **47**, 675–689. <https://doi.org/10.1097/MPA.0000000000001075> (2018).
- de Muynck, L. et al. Consensus Statement on the Use of Near-Infrared Fluorescence Imaging during Pancreatic Cancer Surgery Based on a Delphi Study: Surgeons' Perspectives on Current Use and Future Recommendations. *Cancers (Basel)*. **15**. <https://doi.org/10.3390/cancers15030652> (2023).
- de Valk, K. S. et al. First-in-Human assessment of cRGD-ZW800-1, a zwitterionic, Integrin-Targeted, Near-Infrared fluorescent peptide in Colon carcinoma. *Clin. Cancer Res.* **26**, 3990–3998. <https://doi.org/10.1158/1078-0432.CCR-19-4156> (2020).
- van Oosten, M., Crane, L. M., Bart, J., van Leeuwen, F. W. & van Dam, G. M. Selecting potential targetable biomarkers for imaging purposes in colorectal Cancer using target selection criteria (TASC): A novel target identification tool. *Transl. Oncol.* **4**, 71–82. <https://doi.org/10.1593/tlo.10220> (2011).
- Hoogstins, C. E. S. et al. Image-Guided surgery in patients with pancreatic cancer: first results of a clinical trial using SGM-101, a novel carcinoembryonic Antigen-Targeting, Near-Infrared fluorescent agent. *Ann. Surg. Oncol.* **25**, 3350–3357. <https://doi.org/10.1245/s10434-018-6655-7> (2018).
- Lu, G. et al. Tumour-specific fluorescence-guided surgery for pancreatic cancer using panitumumab-IRDye800CW: a phase 1 single-centre, open-label, single-arm, dose-escalation study. *Lancet Gastroenterol. Hepatol.* **5**, 753–764. [https://doi.org/10.1016/S2468-1253\(20\)30088-1](https://doi.org/10.1016/S2468-1253(20)30088-1) (2020).
- Mulder, B. G. S. et al. Intraoperative molecular fluorescence imaging of pancreatic Cancer by targeting vascular endothelial growth factor: A multicenter feasibility Dose-Escalation study. *J. Nucl. Med.* **64**, 82–89. <https://doi.org/10.2967/jnumed.121.263773> (2023).

14. Houvast, R. D. et al. Prediction of biomarker expression on primary pancreatic ductal adenocarcinoma tissues using Fine-Needle biopsies: paving the way for a Patient-Tailored molecular imaging approach. *Mol. Diagn. Ther.* **27**, 261–273. <https://doi.org/10.1007/s40291-022-00635-w> (2023).
15. Houvast, R. D. et al. An immunohistochemical evaluation of Tumor-Associated glycans and mucins as targets for molecular imaging of pancreatic ductal adenocarcinoma. *Cancers (Basel)*. **13**. <https://doi.org/10.3390/cancers13225777> (2021).
16. van Dam, M. A. et al. Overview and future perspectives on Tumor-Targeted positron emission tomography and fluorescence imaging of pancreatic Cancer in the era of neoadjuvant therapy. *Cancers (Basel)*. **13**. <https://doi.org/10.3390/cancers13236088> (2021).
17. Harsha, H. C. et al. A compendium of potential biomarkers of pancreatic cancer. *PLoS Med.* **6**, e1000046. <https://doi.org/10.1371/journal.pmed.1000046> (2009).
18. Tummers, F. et al. Biomarker identification for endometriosis as a target for real-time intraoperative fluorescent imaging: A new approach using transcriptomic analysis to broaden the search for potential biomarkers. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **288**, 114–123. <https://doi.org/10.1016/j.ejogrb.2023.07.007> (2023).
19. *Euretos - Data drive disease insight*, <https://www.euretos.com> (2023).
20. Meijer, R. P. J. et al. Data-Driven identification of targets for Fluorescence-Guided surgery in Non-Small cell lung Cancer. *Mol. Imaging Biol.* **25**, 228–239. <https://doi.org/10.1007/s11307-022-01791-5> (2023).
21. Dim, D. C. et al. The usefulness of S100P, mesothelin, fascin, prostate stem cell antigen, and 14-3-3 Sigma in diagnosing pancreatic adenocarcinoma in cytological specimens obtained by endoscopic ultrasound guided fine-needle aspiration. *Diagn. Cytopathol.* **42**, 193–199. <https://doi.org/10.1002/dc.21684> (2014).
22. Li, T. et al. Research progress and design optimization of CAR-T therapy for pancreatic ductal adenocarcinoma. *Cancer Med.* **8**, 5223–5231. <https://doi.org/10.1002/cam4.2430> (2019).
23. Almeida, P. P., Cardoso, C. P. & de Freitas, L. M. PDAC-ANN: an artificial neural network to predict pancreatic ductal adenocarcinoma based on gene expression. *BMC Cancer*. **20**, 82. <https://doi.org/10.1186/s12885-020-6533-0> (2020).
24. Gu, J. et al. TMRSS4 promotes cell proliferation and inhibits apoptosis in pancreatic ductal adenocarcinoma by activating ERK1/2 signaling pathway. *Front. Oncol.* **11**, 628353. <https://doi.org/10.3389/fonc.2021.628353> (2021).
25. Mao, F. et al. High Expression of COL17A1 Predicts Poor Prognosis and Promotes the Tumor Progression via NF-kappaB Pathway in Pancreatic Adenocarcinoma. *J. Oncol.* **8868245**, (2020). <https://doi.org/10.1155/2020/8868245> (2020).
26. Long, N. P. et al. An integrative data mining and Omics-Based translational model for the identification and validation of oncogenic biomarkers of pancreatic Cancer. *Cancers (Basel)*. **11** <https://doi.org/10.3390/cancers11020155> (2019).
27. Zhao, C. et al. Targeting Claudin 18.2 using a highly specific antibody enables Cancer diagnosis and guided surgery. *Mol. Pharm.* **19**, 3530–3541. <https://doi.org/10.1021/acs.molpharmaceut.1c00947> (2022).
28. Woll, S. et al. Claudin 18.2 is a target for IMAB362 antibody in pancreatic neoplasms. *Int. J. Cancer*. **134**, 731–739. <https://doi.org/10.1002/ijc.28400> (2014).
29. Maeda, K., Uehara, T., Hosoda, W., Kuraishi, Y. & Ota, H. Expression profiles of Cadherin 17 and Claudin 18.2 in comparison with peptide hormonal expression in pancreatic neuroendocrine tumours: implications for targeted immunotherapy. *Pathol. Res. Pract.* **262**, 155537. <https://doi.org/10.1016/j.prp.2024.155537> (2024).
30. Chen, G. et al. AQP5 is a novel prognostic biomarker in pancreatic adenocarcinoma. *Front. Oncol.* **12**, 890193. <https://doi.org/10.3389/fonc.2022.890193> (2022).
31. Direito, I., Paulino, J., Vigia, E., Brito, M. A. & Soveral, G. Differential expression of aquaporin-3 and aquaporin-5 in pancreatic ductal adenocarcinoma. *J. Surg. Oncol.* **115**, 980–996. <https://doi.org/10.1002/jso.24605> (2017).
32. Burghardt, B. et al. Distribution of Aquaporin water channels AQP1 and AQP5 in the ductal system of the human pancreas. *Gut* **52**, 1008–1016. <https://doi.org/10.1136/gut.52.7.1008> (2003).
33. de Geus, S. W. et al. Selecting Tumor-Specific molecular targets in pancreatic adenocarcinoma: paving the way for Image-Guided pancreatic surgery. *Mol. Imaging Biol.* **18**, 807–819. <https://doi.org/10.1007/s11307-016-0959-4> (2016).
34. Ding, M. Q., Chen, L., Cooper, G. F., Young, J. D. & Lu, X. Precision oncology beyond targeted therapy: combining omics data with machine learning matches the majority of Cancer cells to effective therapeutics. *Mol. Cancer Res.* **16**, 269–278. <https://doi.org/10.1158/1541-7786.MCR-17-0378> (2018).
35. Mishra, N. K., Southekal, S. & Guda, C. Survival analysis of Multi-Omics data identifies potential prognostic markers of pancreatic ductal adenocarcinoma. *Front. Genet.* **10**, 624. <https://doi.org/10.3389/fgene.2019.00624> (2019).
36. Jiang, F. et al. Integrated analysis of Multi-Omics data to identify prognostic genes for pancreatic Cancer. *DNA Cell. Biol.* **41**, 305–318. <https://doi.org/10.1089/dna.2021.0878> (2022).
37. Morse, D. L. et al. Identification of novel pancreatic adenocarcinoma cell-surface targets by gene expression profiling and tissue microarray. *Biochem. Pharmacol.* **80**, 748–754. <https://doi.org/10.1016/j.bcp.2010.05.018> (2010).
38. Wang, Q. et al. Unifying cancer and normal RNA sequencing data from different sources. *Sci. Data.* **5**, 180061. <https://doi.org/10.1038/sdata.2018.61> (2018).
39. Ui, T. et al. Development and characterization of a (68)Ga-labeled A20FMDV2 peptide probe for the PET imaging of alphavbeta6 integrin-positive pancreatic ductal adenocarcinoma. *Bioorg. Med. Chem.* **28**, 115189. <https://doi.org/10.1016/j.bmc.2019.115189> (2020).
40. Odenthal, J. et al. Targeting CD44v6 for fluorescence-guided surgery in head and neck squamous cell carcinoma. *Sci. Rep.* **8**, 10467. <https://doi.org/10.1038/s41598-018-28059-9> (2018).
41. Liu, Y., Beyer, A. & Aebersold, R. On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**, 535–550. <https://doi.org/10.1016/j.cell.2016.03.014> (2016).
42. Kim, Y. I. et al. Simultaneous detection of EGFR and VEGF in colorectal Cancer using Fluorescence-Raman endoscopy. *Sci. Rep.* **7**, 1035. <https://doi.org/10.1038/s41598-017-01020-y> (2017).
43. Bausch-Fluck, D. et al. A mass spectrometric-derived cell surface protein atlas. *PLoS One*. **10**, e0121314. <https://doi.org/10.1371/journal.pone.0121314> (2015).
44. Bausch-Fluck, D. et al. The in Silico human surfaceome. *Proc. Natl. Acad. Sci. U S A.* **115**, E10988–E10997. <https://doi.org/10.1073/pnas.1808790115> (2018).
45. Fonseca, A. L. et al. Bioinformatics analysis of the human surfaceome reveals new targets for a variety of tumor types. *Int. J. Genomics*. **2016** (8346198). <https://doi.org/10.1155/2016/8346198> (2016).
46. da Cunha, J. P. et al. Bioinformatics construction of the human cell surfaceome. *Proc. Natl. Acad. Sci. U S A.* **106**, 16752–16757. <https://doi.org/10.1073/pnas.0907939106> (2009).
47. Steele, N. G. et al. Multimodal mapping of the tumor and peripheral blood immune landscape in human pancreatic Cancer. *Nat. Cancer*. **1**, 1097–1112. <https://doi.org/10.1038/s43018-020-00121-4> (2020).
48. Goedhart, J. & Luijsterburg, M. S. VolcanoR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci. Rep.* **10**, 20560. <https://doi.org/10.1038/s41598-020-76603-3> (2020).

## Acknowledgements

The authors would like to thank Arie Baak for his introduction to Euretos and help with data interpretation. Moreover, we thank Noel F. de Miranda for his contribution to collecting datasets of genes associated with PDAC and his assistance with data analysis.



## Author contributions

Conceptualization: Nada Badr, Ruben D. Houvast, J. Sven D. Mieog, Alexander L. Vahrmeijer, and Peter J. K. Kuppen; Methodology: Nada Badr, Ronald L.P. van Vlierberghe, A. Stijn L. P. Crobach, and Peter J. K. Kuppen; Formal analysis & Investigation: Nada Badr, Luca Ten Elshof, Lysanne D. A. N. de Muynck, A. Stijn L. P. Crobach, Ronald L.P. van Vlierberghe, and Peter J. K. Kuppen; Writing – original draft preparation: Nada Badr, Luca Ten Elshof and Lysanne D. A. N. de Muynck; Writing – review and editing: Ruben D. Houvast, Gerard J. P. van Westen, J. Sven D. Mieog, Alexander L. Vahrmeijer, and Peter J. K. Kuppen; Resources: Ronald L.P. van Vlierberghe and A. Stijn L. P. Crobach; Supervision: Gerard J. P. van Westen, J. Sven D. Mieog, Alexander L. Vahrmeijer, and Peter J. K. Kuppen.

## Declarations

### Competing interests

The authors declare no competing interests.

### Ethics approval and consent to participate

All procedures involving human participants complied with the ethical standards of the institutional and/or national research committee and followed the principles outlined in the 1964 Helsinki Declaration, including its amendments or equivalent ethical standards. The Ethics Committee of the Leiden University Medical Center reviewed and approved this study (protocol code B20.052, approval date: 17 December 2020). The requirement for informed consent was waived by the Ethics Committee.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-00517-1>.

**Correspondence** and requests for materials should be addressed to N.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025