Original Research

Targeted and explorative profiling of kallikrein proteases and global proteome biology of pancreatic ductal adenocarcinoma, chronic pancreatitis, and normal pancreas highlights disease-specific proteome remodelling

Janina Werner [a,b,1], Patrick Bernhard [a,c,d,1], Miguel Cosenza-Contreras [a,d], Niko Pinter [a], Matthias Fahrner [a,f], Prama Pallavi [b], Johannes Eberhard [b], Peter Bronsert [a], Felix Rückert [b,e,*], Oliver Schilling [a,f]

[a] *Institute for Surgical Pathology, Medical Center, Faculty of Medicine, University of Freiburg, Germany*
[b] *Department of Surgery, Universitätsmedizin Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany*
[c] *Spemann Graduate School of Biology and Medicine (SGBM), University of Freiburg, Germany*
[d] *Faculty of Biology, University of Freiburg, Freiburg, Germany*
[e] *Surgical Department, Diakonissen-Stiftungs-Krankenhaus Speyer, Paul-Egell-Straße 33, Speyer D-67346, Germany*
[f] *German Cancer Consortium (DKTK) and Cancer Research Center (DKFZ), Freiburg, Germany*

ARTICLE INFO

ABSTRACT

Pancreatic ductal adenocarcinoma (PDAC) represents one of the most aggressive and lethal malignancies worldwide with an urgent need for new diagnostic and therapeutic strategies. One major risk factor for PDAC is the pre-indication of chronic pancreatitis (CP), which represents highly inflammatory pancreatic tissue. Kallikreins (KLKs) are secreted serine proteases that play an important role in various cancers as components of the tumor microenvironment. Previous studies of KLKs in solid tumors largely relied on either transcriptomics or immunodetection. We present one of the first targeted mass spectrometry profiling of kallikrein proteases in PDAC, CP, and normal pancreas. We show that KLK6 and KLK10 are significantly upregulated in PDAC (n=14) but not in CP (n=7) when compared to normal pancreas (n=16), highlighting their specific intertwining with malignancy. Additional explorative proteome profiling identified 5936 proteins in our pancreatic cohort and observed disease-specific proteome rearrangements in PDAC and CP. As such, PDAC features an enriched proteome motif for extracellular matrix (ECM) and cell adhesion while there is depletion of mitochondrial energy metabolism proteins, reminiscent of the Warburg effect. Although often regarded as a PDAC hallmark, the ECM fingerprint was also observed in CP, alongside with a prototypical inflammatory proteome motif as well as with an increased wound healing process and proteolytic activity, thereby possibly illustrating tissue autolysis. Proteogenomic analysis based on publicly accessible data sources identified 112 PDAC-specific and 32 CP-specific single amino acid variants, which among others affect KRAS and ANKHD1. Our study emphasizes the diagnostic potential of kallikreins and provides novel insights into proteomic characteristics of PDAC and CP.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) represents the most common type of pancreatic cancer, accounting for more than 90% of all pancreatic malignancies [1]. It is characterized by a highly aggressive and invasive growth, resulting in an abysmal 5-year survival rate of merely 3–15% [2]. To date, surgical resection is the only option for cure [3]. However, the majority of patients receive palliative treatment, as most patients present with their disease at an already locally advanced or metastatic state [4]. Although the protein carbohydrate antigen 19-9 (Ca 19-9), the only clinically accepted and routinely used biomarker for PDAC, serves as a vital follow-up parameter during cancer progression and response to treatment, it is insufficient for screening and diagnostic among the general population [5,6]. This highlights the lack of diagnostic approaches for PDAC and illustrates the urgent need for the identification of novel biomarkers which enable early diagnosis and subsequent therapy.

Challenges in the early detection of pancreatic cancer further include identifying at-risk individuals among the general population who would particularly benefit from diagnostic biomarkers for PDAC [7]. One of these major risk factors for the development of PDAC represents chronic pancreatitis (CP), which is characterized by recurrent, highly inflammatory episodes in the pancreatic tissue and has been discussed to be a preliminary stage to PDAC [8–10].

Human kallikreins (KLKs) comprise a family of 15 secreted serine proteases. For PDAC, KLK 6 and 10 have been suggested to represent new diagnostic and/or prognostic biomarkers for pancreatic cancer [11–15]. Moreover, KLK3, known as prostate-specific antigen (PSA), is already of great clinical value as a routinely applied diagnostic marker and follow-up parameter in prostate cancer [16]. Some kallikreins have been proposed as therapeutic targets for pancreatic cancer, as elevated levels of KLK6, KLK7, KLK8, KLK10, and KLK11 in PDAC were correlated with a significantly worse overall patient survival and are further suggested to play a crucial role within pancreatic cancer development [11,14,17,18]. Various pro-tumorigenic mechanistic routes have been proposed such as facilitating cancer cell migration by cleavage of cell adhesion and extracellular matrix proteins, suggesting KLKs as important factors in modulating the tumor microenvironment [19–22]. KLK6 levels were found to be significantly elevated in invasive tumor areas compared with non-invasive tumor areas in PDAC tissue sections [17]. Besides the upregulation of kallikreins in PDAC tissues, multiple kallikreins, including KLK5, KLK6, KLK7 and KLK10, were also found to be elevated in cell-conditioned medium of the pancreatic cancer cell line MiaPaCa-2 [23]. However, currently available data on KLK secretion or expression in pancreatic cancer cell lines almost exclusively originates from transcriptomic data or immunodetection such as Western blot analysis, immunohistochemistry, or ELISA, which are often limited to few targets per study [11,14,23].

Gene or protein expression profiling approaches are increasingly used to gain deep insight into pathological traits pancreatic malignancies, as for example moffit and CPTAC. In this study, we applied targeted mass spectrometry (MS), namely parallel reaction monitoring (PRM) to investigate KLK expression in PDAC, ampullary cancer (AMPAC), CP, and normal non-malignant adjacent pancreas. To our knowledge, this is one of the first reports using targeted mass spectrometry for the detection of KLK proteins in these disease conditions. Furthermore, using exploratory proteomics, we performed proteome profiling of PDAC, AMPAC, CP, and normal non-malignant adjacent pancreas to gain better insight into the disease-related proteome biology.

Materials and methods

Cell culture and preparation of Cell-Conditioned Medium (CCM)

The human pancreatic cancer cell line MiaPaCa-2 was cultured in Dulbecco´s modified Eagle´s medium (DMEM, high glucose, GlutaMAX^{TM} supplement; Gibco, Thermo Fisher Scientific), while the human pancreatic cancer cell line AsPC-1 (passage 11) was cultured in RPMI-1640 medium (L-glutamine supplement; Gibco, Thermo Fisher Scientific), both supplemented with 10% (v/v) Fetal bovine serum (FBS, Gibco, Thermo Fisher Scientific), 100 U/mL penicillin and 100 μg/mL streptomycin at 37 °C under a humidified atmosphere containing 5% CO_2. The human primary pancreatic cancer cell lines PaCaDD-159, PaCaDD-165 and MaPaC-107 were cultured in a 2:1 mixture of supplemented DMEM as described above and Keratinocyte-Serum-free medium (KSFM, Thermo Fisher Scientific) supplemented with 2.5 μg Epidermal Growth Factor (EGF, Human Recombinant) and 25 mg Bovine Pituitary Extract (BPE) at 37 °C under a humidified atmosphere containing 5% CO_2. All cell lines were maintained and subcultured at a cell confluency between 50 and 90% and were regularly tested for mycoplasma contaminations (Eurofins Genomics).

For the preparation of cell-conditioned medium (CCM), cells were cultured in 145 mm cell culture dishes (Greine Bio-One) and expanded until 70–90% confluency. Subsequently, cells were incubated for 24 h in serum-free medium (SFM, Gibco, Thermo Fisher Scientific) supplemented with 100 U/mL penicillin, 100 μg/mL streptomycin, 1% (v/v) Minimum Essential Medium (MEM) solution and 1% (v/v) Non-essential Amino Acid solution. Resulting cell-conditioned medium (CCM) was separated and supplemented with 10 mM EDTA and 0.1 mM PMSF followed by centrifugation (4000 rpm, 10 min) and storage of the clear supernatant at - 80 °C until further processing. The corresponding cell number of each CCM dish was determined by harvesting cells using TrypLE^{TM} Express Enzyme trypsin-EDTA solution (Gibco, Thermo Fisher Scientific) and counting cells by using 0.4% (v/v) Trypan blue staining solution (NanoEntek) and an automated cell counter (Eve Automatic Cell Counter, NanoEnTek).

Cohort information

The patient cohort included formalin-fixed paraffin-embedded (FFPE) tissues of PDAC (n=14), chronic pancreatitis (n=7), non-malignant adjacent pancreas (n=11), normal pancreas (n=5) and ampullary cancer (n=3), which were kindly provided by the surgical department of the University of Heidelberg, Medical Faculty Mannheim (PDAC, chronic pancreatitis and non-malignant adjacent pancreas, ampullary cancer), and the University of Freiburg, Institute for Surgical Pathology (normal pancreas). The latter samples were obtained from the pancreas of multiple organ donors (median age 63 years; range 43±84). FFPE blocks of each tissue unit were cut at different thicknesses ranging from 1.5 to 10 μm and dried overnight at 60 °C by using a microtome (HM355 S, Thermo Fisher Scientific). All tissue slices were routinely stained with hematoxylin and eosin (H&E) using the Dako Coverstainer (Agilent Technologies) and were reviewed by an experienced pathologist. The area of interest was then marked on the tissue slices; regions that were not marked on the slide were excluded from further experiments. Written informed consent was obtained from all patients before inclusion into our study. Ethics approval was obtained by local authorities of the Ethics Committee of the University Medical Center Freiburg (file number: 367/14) and the Ethics Committee of the University Medical Center Heidelberg, Faculty of Mannheim (file number: 2012-293N-MA).

Protein extraction and proteomic sample preparation

For sample preparation of cell-conditioned medium, 5 mL CCM of each cell line was transferred to a 15 mL falcon and was supplemented with 1 μg Avidin (Sigma) as internal standard and with 0.2% (v/v) SDS, 35 mM TRIS pH 8.5 and 9 U/mL benzonase for protein extraction, respectively. Samples were incubated for 15 min at ambient temperature before performing cysteine reduction using 2 mM tris(2-carboxyethyl)phosphin-hydrochlorid (TCEP, 30 min, 80 °C water bath) and subsequent alkylation using 5 mM iodoacetamide (30 min, room

temperature, in the dark). Sample pH was adjusted to pH 7 by adding 83 mM 2-Morpholinoethanesulfonic acid monohydrate (MES) before performing protein enrichment and SDS removal by using a previously published sp3-bead protocol [24]. In short, sp3-beads (Sera-Mag Speed-Beads, GE Healthcare) were added to a final concentration of 0.2 μg/μL before adding acetonitrile to a final proportion of 60% (v/v) and incubating for 10 min at room temperature to allow protein binding. Beads were pelleted by centrifugation (800 g, 8 min), resuspended in 80% (v/v) ethanol and transferred to Eppendorf tubes to allow bead separation using a magnetic rack. Beads were successively washed with 80% (v/v) ethanol and acetonitrile before air drying beads. Protein-coupled beads were resuspended in 0.1 M HEPES pH 8.0 containing 0.1% (v/v) of an acid-labile surfactant (sodium 3-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl)methoxy]-1- propanesulfonate) before eluting proteins by sonication (10 min) and heating (95 °C, 10 min). Beads were removed using the magnetic rack only using the clear supernatant for further processing. Protein concentration of each eluate was determined by using the BCA assay (Thermo Fisher Scientific) before subjecting 10–70 μg protein amount to in-solution tryptic digestion. Proteolytic digestion was performed by adding Lysyl Endopeptidase (LysC, Wako Chemicals) in a protease:protein ratio of 1:50 (w/w) and incubating for 2 h at 42 °C, before adding trypsin (Promega) in a protease:protein ratio of 1:50 (w/w) and further incubate samples for 19 h at 37 °C. Digestion was stopped by acidification (2% (v/v) TFA), before incubating (37 °C, 30 min) and centrifuging (20,000 g, 5 min) the samples in order to precipitate acid-labile surfactant. Clear supernatant was used for peptide desalting using iST C18 mixed phase cartridges (PreOmics, Martinsried, Germany) according to manufacturer´s instructions. After determining the peptide concentration *via* the BCA assay (Thermo Fisher Scientific), eluates were vacuum dried and stored at – 80 °C until MS measurement.

For sample preparation of formalin-fixed, paraffin-embedded (FFPE) samples, tissue slices were deparaffinized by successively incubating in increasing concentrations of xylene followed by decreasing concentrations of ethanol (xylol 4x, alcohol absolute, alcohol 96%, alcohol 70%, aqua distilled 2x). Samples (volume ranging from 0.3 to 0.6 mm^3) were macrodissected into Eppendorf tubes and stored at - 80 °C until further processing. For protein extraction, tissue samples were incubated with detergent-containing buffer (0.1 M HEPES pH 8.0, 0.1% (v/v) SDS) before performing heat-induced antigen retrieval (2 h at 95 °C) followed by ultrasonication (Bioruptor, 10 cycles, 45/15 s on/off time, high intensity) and centrifugation (21.000 g, 10 min), thereupon only using the clear supernatant. The protein concentration was determined by using the BCA assay (Thermo Fisher Scientific) before subjecting 100 μg protein amount to in-solution proteomic sample preparation including cysteine reduction with 5 mM DTT (37 °C, 30 min), subsequent alkylation with 15 mM iodoacetamide (30 min, room temperature, in the dark) and protein enrichment and SDS removal by using sp3-beads like described above. On-bead proteolytic digest was performed by adding Lysyl Endopeptidase (LysC, Wako Chemicals) in a protease:protein ratio of 1:100 (w/w) and incubating for 2 h at 42 °C, before adding trypsin (Promega) in a protease:protein ratio of 1:50 (w/w) and further incubate samples for 19 h at 37 °C. Post-digestion processing was performed like described above.

LC-MS/MS measurement

For SRM measurement, vacuum dried peptides were resolubilized in 0.1% (v/v) formic acid to a final concentration of 0.2 μg/μL, sonicated for 5 min and centrifuged at 20,000 g for 10 min before transferring the supernatant to the measurement tube. 1 μg of each sample, together with 200 fmol of indexed retention time (iRT) peptides and 100 fmol of selected isotope-labelled heavy reference peptides, were analyzed using a nanoflow liquid chromatography (LC) system, Proxeon Easy-nLC II (Thermo Fisher Scientific) equipped with a trapping column (Fused Silica Capillary, 5 cm length, 100 μm inner diameter (ID),

VICI Jour, Schenkon, Switzerland) and an analytical column (Selfpack PicoFrit column, 35 cm length, 75 μm ID, New Objective, Woburn, MA) both in-house packed with C18 particles (Dr. Maisch, ReproSil-Pur 120 C18-AQ, 3 μm C18 particle size, 120 Å pore size). Samples were trapped at 220 bars with 100% Buffer A (0.1% (v/v) formic acid) and separated using a flow rate of 250 nL/min with the analytical column temperature set to 60 °C. A 83 min multistep gradient of 8% to 56% (v/v) buffer B (50% (v/v) ACN in 0.1% (v/v) formic acid in H$_2$O) in buffer A was used for separation followed by washing (100% B) and reconditioning of the column to 5% B (Table S1 for detailed gradient overview). The Easy-nLC II system was coupled online to a TSQ Vantage (Thermo Fisher Scientific) triple quadrupole mass spectrometer *via* a Nanopray Flex Ionsource (Thermo Fisher Scientific) with an applied voltage of 2.5 kV for electrospray ionization. The analytical PicoFrit column contained an integrated uncoated pre-cut emitter (Silica Tip, 10 μm tip inner diameter, New Objective). The mass spectrometer was operated in selected reaction monitoring (SRM) mode with scheduled acquisition. SRM acquisition was performed with Q1 and Q3 operated at unit resolution (0.70 m/z half maximum peak width) with 449 transitions and a fixed dwell time of 15 ms per transition. Collision energies (CEs) were calculated in Skyline according to the formulas: CE = 0.03 x m/z + 2.905 and CE = 0.038 x m/z + 2.281 (m/z, mass-to-charge ratio of the precursor ion) for doubly and triply charged precursor ions, respectively. Samples were measured in randomized sample order.

For PRM measurement, vacuum dried peptides were resolubilized like described above. 800 ng of each sample together with 100 fmol of iRT peptides and 100 fmol of selected isotope-labeled heavy reference peptides, were analyzed using a nanoflow LC system, Easy-nLC 1000 (Thermo Fisher Scientific) equipped with a trapping column (50 cm μPac^{TM} trapping column, Thermo Fisher Scientific) and an analytical column (200 cm μPac^{TM} analytical column, Thermo Fisher Scientific) tempered to 45°C. Samples were trapped at 250 bars with 100% buffer A (0.1% v/v formic acid) and separated using a dynamic flow rate of 350–700 nL/min. A 120 min multistep gradient of 8% to 55 % buffer B (80% v/v acetonitrile, 0.1% v/v formic acid) in buffer A was used for separation, followed by washing (100 % B) and reconditioning of the column to 8% B (Table S2 for detailed gradient overview). The Easy-nLC 1000 system was coupled online to a Q-Exactive plus (Thermo Fisher Scientific) mass spectrometer *via* a Nanospray Flex Ionsource (Thermo Fisher Scientific) with an applied voltage of 2.1 kV for electrospray ionization. The analytical column was coupled to a pulled, uncoated ESI emitter (10 μm tip inner diameter, 20 μm inner diameter, 7 cm length, CoAnn Technologies) *via* a μPacTM Flex iON Connect ESI-MS interface (PharmaFluidics). The mass spectrometer was operated in unscheduled parallel reaction monitoring (PRM) acquisition mode. Therefore, MS2 scans (1 μscan) of doubly-charged precursor ions were performed with an isolation window size of 1.2 m/z, MS2 resolution was set to 35,000, automatic gain control (AGC) to 3e6 and maximum injection time was set to 150 ms using stepped normalized collision energy (NCE) of 25 and 30 for fragmentation. Samples were measured in randomized sample order.

For DIA measurement, vacuum dried peptides were resolubilized like described above. 800 ng of each sample together with 100 fmol of iRT peptides, were analyzed using the same nanoflow LC and mass spectrometer system as described for PRM measurements.

For the individual DIA-samples, the mass spectrometer was operated in data independent acquisition (DIA) mode. Therefore, two cycles of 24 m/z broad windows ranging from 400 to 1000 m/z with a 50% shift between the cycles (staggered window scheme) [25] was used, leading to an overall isolation window size of 12 m/z. MS2 resolution was set to 17,500, automatic gain control (AGC) to 1e6 and maximum injection time was set to 80 ms using stepped NCE of 25 and 30. After 25 consecutive MS2 scans, a MS1 survey scan was triggered covering the same range, but with a resolution of 70,000, an AGC target of 3e6 and maximum injection time was set to 50 ms. Samples were measured in randomized sample order.

Table 1

Characteristics of included pancreatic cancer cell lines. The used cell lines vary in their origin, cell type, sex, and age at sampling.

Cell Line	Origin	Cell Type	Sex of Cell	Age at Sampling	Citation
MiaPaCa-2	Primary tumor	Immortalized cell line	Male	65	https://web.expasy.org/cellosaurus/CVCL_0428
AsPC-1	Metastasis (Ascites)	Immortalized cell line	Female	62	https://web.expasy.org/cellosaurus/CVCL_0152
PaCaDD-165	Metastasis (Ascites)	Primary cell line	Male	54	[30]
PaCaDD-159	Primary tumor	Primary cell line	Male	78	[30]
MaPaC-107	Metastasis (Ascites)	Primary cell line	–	–	Provided by Dr. Felix Rückert

For GPF-library measurements to generate an experiment-specific database, similar sample amounts were mixed to generate a master-pool sample, which was subsequently measured. The same settings were applied with the exception that the total mass scan range of 400 to 1000 m/z has been subdivided into 6 individual measurements, each covering a mass scan range of 100 m/z with an overall isolation window size of 4 m/z.

Proteomic data analysis

For SRM and PRM data, proteomic data analysis was performed using the open source software tool Skyline (version 19.1) [26]. All integrated peaks were manually inspected to ensure correct peak detection and integration. The resulting peak areas of KLK proteins were successively normalized to both the internal standard avidin and the total cell number the CCM derived from. For statistical evaluation, a two-sample t-test was performed using Benjamini–Hochberg for multiple testing correction. Resulting adjusted p-values and corresponding fold changes were visualized as volcano plots, with proteins revealing a significant abundance change (adjusted p-value ≤ 0.05) together with an absolute log2 fold change ≥ 1 (corresponds to 2-fold abundance change) were considered as differentially expressed.

For DIA data analysis, DIA-NN (1.7.12) was used with recommended settings [27]. Peptide identification was performed using a human proteome database containing reviewed UniProt sequences without isoforms downloaded from Uniprot on 14th June 2021 (20,856 entries). This database was refined using the GPF-measurements. For library refinement, deep learning-based spectra and RTs prediction was activated. Tryptic cleavage specificity (Trypsin/P) with 1 missed cleavage was applied, while setting an allowed peptide length between 7 and 30 amino acids. The N-terminal M excision, C carbamidomethylation, retention time (RT) profiling and RT-dependent cross-normalization options were enabled. Precursor FDR was set to 1.0%, while decoys of the database entries were automatically created by DIA-NN. For the in silico predicted library search, the same settings were applied with exception of deactivated deep learning-based spectra and RTs prediction option.

The DIA-NN output was further processed in R (V 4.1.0) with RStudio (V 1.3.1093) as an integrated development environment. After filtering for unique proteins, log2-transformation and median centering was performed using the DIANN package (V 1.0.1) [27]. For sparsity reduction, the dataset was filtered for proteins with at least 80% valid values per protein, per condition before performing ComBat batch correction by using the sva (V 3.40.0) package [28].

Partial least squares discriminant analysis (PLS-DA) and extraction of protein representatives was performed by using the mixOmics package (V 6.16.3). Cluster of proteins that are consistently co-expressed were identified by using the Clust algorithm (V 1.10.10) (Clust Citation) in Python (V 3.8.9) using Z-score normalization prior to k-means clustering with seed number k=8 and tightness weight = 0.5. Differential expression analysis was performed using the limma (V 3.48.3) package. Gene Ontology annotation and gene enrichment analysis of protein subsets were performed using the topgo (V 2.44.0), clusterprofiler (V 4.0.5) and ReactomePA (V 1.36.0) packages.

For Semi-Tryptic Peptide Analysis of DIA data, peptide identification was performed like described above but using a human proteome database already containing semi-tryptic peptides (10,514,442 entries)

deriving from the abovementioned human database downloaded from Uniprot on 14th June 2021. After additional peptide-to-protein mapping by using an in-house developed publicly available R-script (publicly available *via* repository: pep2prot (Version 1.2), https://github.com/npinter/pep2prot) and subsequent log2-transformation and median centering like described above, identified peptides were annotated as tryptic or semi-tryptic peptides by using another in-house developed publicly available R-script (publicly available *via* repository: Fragterminomics (Version 1.0.0), https://github.com/MiguelCos/Fragterminomics).

For proteogenomic analysis, the open-source and free-to-use Galaxy platform [29] was used on the European Galaxy server (https://usegalaxy.eu). The tools for generating a proteogenomic FASTA database are connected within a Galaxy workflow that enables a standardized and parallel high-throughput analysis combined with file collections. Selected Paired-end RNA-seq data from the Sequence Read Archive (SRA) database (https://www.ncbi.nlm.nih.gov/sra) were used as input for the Galaxy workflow. To validate the quality of the provided input files, quality control was performed using FastQC (v0.73) and MultiQC (v1.11). First, raw RNA-seq reads were mapped to the human reference genome GRCh38 using HISAT2 (v2.2.1). Subsequently, the output of HISAT2 was deployed for variant calling using FreeBayes (v1.3.1). The resulting variant call file (VCF) and the mapped RNA-seq reads (BAM) were submitted to CustomProDB (v1.22.0) for generation of the proteogenomic FASTA database containing splice isoforms, single amino acid variants (SAAVs) as well as insertion and deletion (InDels) mutations. Further post-processing using Galaxy's standard text manipulation tools yield the final proteoform database containing single amino acid variants (SAAVs). Subsequently, this database was processed locally with Python scripts to generate tryptic variant peptides (10,806 entries). The abovementioned canonical human proteome database (downloaded from Uniprot on 14th June 2021; 20,856 entries) was subsetted (subFASTA proteome) with the protein identifiers (UniProt IDs) from the initial DIA search and utilized for further filtering of those tryptic SAAV peptides to obtain a unique and peptide-centric variant database. The subFASTA proteome (6159 entries) and the generated SAAV database were merged and applied as a reference proteome for the variant peptide search in DIA-NN. Subsequent peptide identification from the acquired DIA data was performed like described above but without refinement using DIA-NN and the combined database (16,965 entries).

Results and discussion

KLK detection in cell-conditioned medium of pancreatic cancer cell lines

As a proof-of-concept system for mass spectrometric detection of secreted kallikrein proteases, we chose cell-conditioned medium (CCM) of cultured pancreatic cancer cells. Different human pancreatic cancer cell lines (MiaPaCa-2, AsPC-1, PaCaDD-165, PaCaDD-159 and MaPaC-107; Table 1) were maintained by standard cell culture before incubating in serum-free medium (SFM) to prepare CCM. Mass spectrometric analysis of this CCM by using selected reaction monitoring (SRM) enabled the targeted detection of secreted kallikreins from the respective cell lines. We probed all 15 human KLKs with a panel comprising 41 peptides, leading to the robust detection of KLK3, KLK5, KLK6, KLK7 and KLK10 in the considered cell lines (Fig. 1). Further KLKs remained below detection.
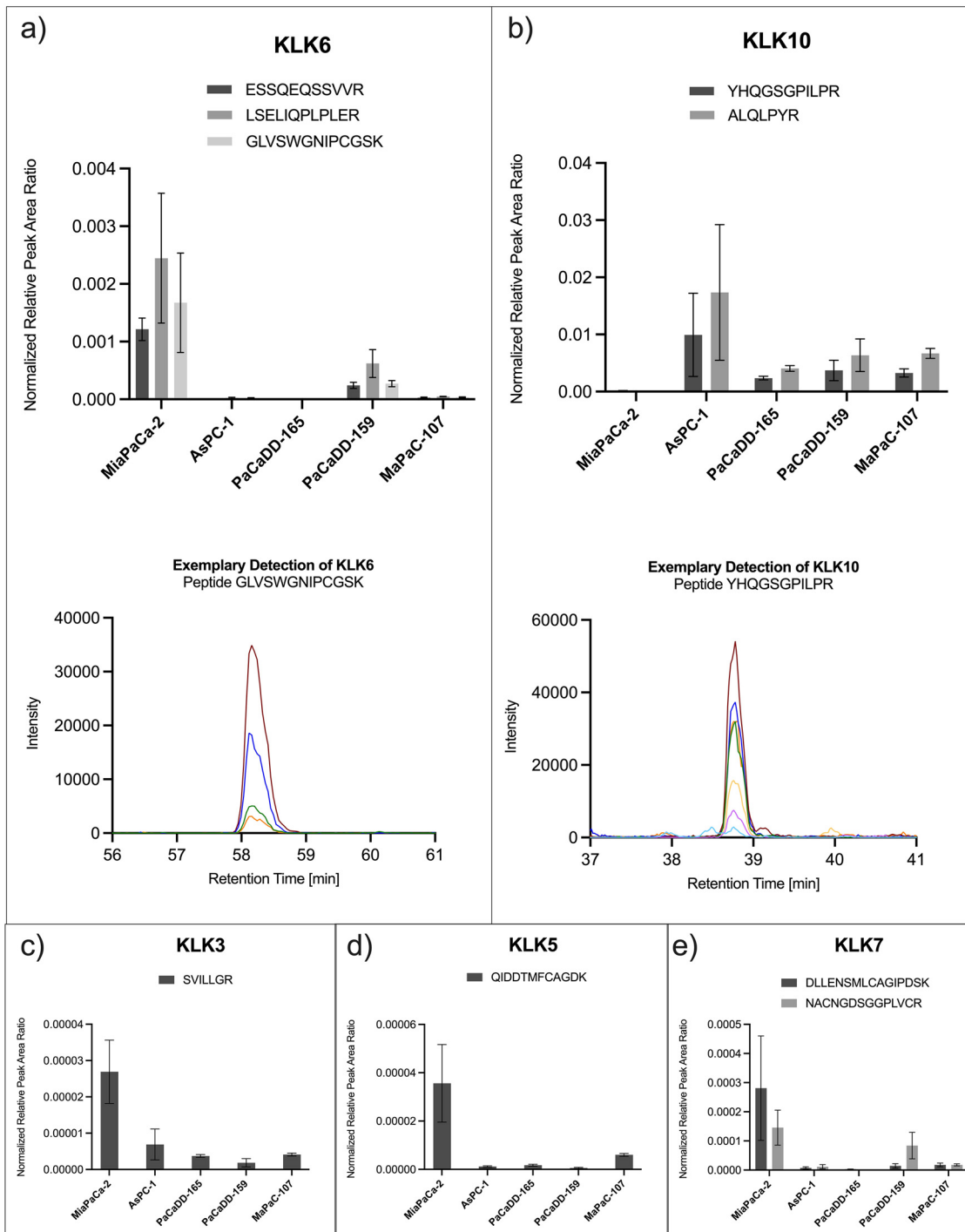
Fig. 1. *Targeted detection of secreted kallikrein proteins in cell-conditioned medium of pancreatic cancer cell lines by selected reaction monitoring.* Serum-free, cell-conditioned medium of five pancreatic cancer cell lines (MiaPaCa-2, AsPC-1, PaCaDD-165, PaCaDD-159, MaPaC-107) were analyzed in a randomized sample order using selected reaction monitoring (SRM) to enable targeted detection and comparison of secreted kallikreins. The assay allowed for the detection of (a) KLK6 with three peptides and (b) KLK10 with two peptides, each of them with several fragments, exemplarily depicted below as co-eluting chromatographic peaks in different colors. Furthermore, (c) KLK3, (d) KLK5 and (e) KLK7 was detected with one or two peptides. Detected peak areas were successively normalized to avidin as internal standard and to the total cell number the CCM derived from. Bar plots show normalized peak area ratios after 2-step normalization for each peptide in each of the investigated cell lines. Error bars correspond to standard deviation.

We have refrained from absolute quantitation of KLKs in CCM due to (a) the proof-of-concept nature of this data and (b) the strong impact of aspects such as cell lysis in SFM on the total protein content of CCM. For normalization purposes, we spiked in 1 µg avidin upon CCM harvesting to correct for potentially occurring technical variance during the sample preparation process. This data serves to illustrate which KLKs are amenable to robust detection by SRM in prototypical CCM of cultured pancreatic cancer cell lines. Of note, this data does not address limits of identification (LOI) or limits of quantitation (LOQ).

KLK6 was strongly detected with three peptides in the CCM of MiaPaCa-2 and the primary cell line PaCaDD-159. In other cell lines, KLK6 transitions remained below detection or were only marginally

present. The inter-peptide differences of normalized peptide abundance, although originating from the same protein, probably derive from varying peptide-specific features like ionization or fragmentation [31].

KLK10 showed a different behavior and was detected with two peptides in the CCM of AsPC-1 and all primary cell lines, but not in MiaPaCa-2 (Fig. 1 b). This is consistent with a previous report, which describes an elevated level of KLK10 in transcriptomic and proteomic data of AsPC-1 compared to MiaPaCa-2 [11].

KLK3/PSA, which is clinically used as a biomarker for prostate cancer [16], was detected in all investigated pancreatic tumor cell lines (Fig. 1 c). KLK5 and KLK7 were strongly detected in CCM of MiaPaCa-2 (Fig. 1 d and e), which, together with KLK6, is consistent with a previous report that investigated KLK protein levels in CCM by ELISA [23].

KLK detection in clinical FFPE tissue samples

Based on the robust, mass spectrometry-based detection of KLKs in CCM of cultured pancreatic cancer cells, we sought to investigate the presence of KLKs in pancreatic FFPE tissue samples using a targeted proteomics approach. We have refrained from absolute quantitation of KLKs in these samples since this data primarily serves to illustrate in which pancreatic pathological condition (PDAC, AMPAC, CP) KLK6, KLK7 or KLK10 are robustly detectable. This approach was motivated by KLKs having escaped detection by explorative mass spectrometry in previous pilot studies (data not shown). Of note, this data does not address LOI or LOQ.

We assembled a cohort including cases of PDAC (n=14), non-malignant adjacent pancreas (n=11), chronic pancreatitis (n=7), normal pancreas (n=5), and ampullary cancer (n=3). Basic patient information is summarized in Table S3. Tissue samples underwent targeted mass spectrometric analysis using parallel reaction monitoring (PRM) focusing on KLK6, KLK7, and KLK10, as these KLKs are considered promising diagnostic and/or prognostic candidates for pancreatic cancer [17] and showed robust detection in our previous cell culture experiment (Fig. 1).

Qualitative evaluation of PDAC tissue data revealed the near-consistent detection of KLK6 (14/17 cases) and KLK10 (15/17 cases) in the malignant tissue entities PDAC and AMPAC. In the entire set of non-malignant samples (total n=23) KLK6 and 10 were absent in all samples except for one case each (Fig. 2 a). The obviously stronger presence of KLK6 and KLK10 in PDAC/AMPAC compared to CP or benign/normal pancreas is in accordance with previous publications [12,14,15,17,22,32,33].

Although detected in our preliminary cell culture experiment (Fig. 1), KLK7 failed to be detected in any of the clinical FFPE tissue samples. This is in contrast to several publications that could show the presence of KLK7 in PDAC tissue either on transcriptomic level or immunodetection [17,20,34].

Although KLK6 and 10 remained below the LOD in benign pancreas (normal pancreas, non-malignant adjacent pancreas, and chronic pancreatitis), chromatographic background signals can be used to assess the extent of their upregulation in PDAC or AMPAC.

This semi-quantitative evaluation of KLK6 and KLK10 in PDAC compared to benign tissue controls showed a significant (p < 0.05, two-sample t-test using Benjamini Hochberg) upregulation of both kallikreins in PDAC compared to all controls (Fig. 2 b). This is consistent with several previous reports [17,32] and has even been associated with shorter patient survival [14,17]. In addition, KLK6 has been reported to be one of the most differentially expressed gene in pancreatic cancer compared with normal controls [32] which further supports our proteomic data, in which KLK6 revealed a more pronounced upregulation compared to KLK10. Furthermore, KLK6 and KLK10 also appeared to be significantly overexpressed in tissue samples of ampullary cancer compared to controls (Fig. S1), which underlines the already suggested similarities between pancreatic and ampullary cancer [33].

It should be noted that our bulk approach is not suited to distinguish intracellular (e.g. in secretory granules) and secreted KLKs, nor do we distinguish zymogens from activated KLKs. For the targeted proteomic array, we have included glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as an endogenous protein for glycolytic energy metabolism. We observed varying abundance of GAPDH, with higher levels in malignant tissues (PDAC and AMPAC) and lower levels in benign tissues (Fig. 2 c; raw peak areas from 1 µg injected tryptic sample). The variability of GAPDH abundance is in line with a study that found GAPDH to be among the genes with the highest expression variability in tumorous compared to normal pancreatic tissue [35], possibly related to the "Warburg effect" in pancreatic cancer [36–38].

Proteome profiling allow differentiation of different pancreatic entities

To obtain a global insight into the proteomic differences between the considered entities, an explorative analysis of clinical FFPE samples was additionally performed using LC-MS/MS operated in data independent acquisition (DIA) mode.

In a recent benchmarking study, we have shown that DIA-type proteomics is a powerful approach for characterizing the proteome biology of diseased or healthy tissue even in the presence of inter-individual heterogeneity [39]. Based on the benchmarking study, we opted for a gas-phase fractionation approach [40] to generate an experiment-specific spectral library as well as using linear model of microarray analysis (limma) to statistically evaluate alterations in the proteome biology of the different entities.

Using the aforementioned approach, we identified a total of 5936 proteins in our cohort. On a per-sample basis, we identified and quantified 8073-31177 peptides and 2527-5023 proteins per sample (Fig. 3 a). After median centering, the protein intensity distribution of the samples is comparable and spans several orders of magnitude (Fig. 3 b). To further improve consistency within the dataset, we focused only on proteins with less than 20% missingness per condition resulting in 2279 remaining proteins with a total of 2.1 % missing values in the overall dataset. The proteome coverage remains below the recently published CPTAC study on PDAC proteomics [41] but is substantially above the proteome coverage reported by non-fractionated DDA-type proteomics of PDAC [42].

We employed sparse partial least squares discriminant analysis (sPLS-DA) to depict the global proteomic divergence or similarity between the considered entities (Fig. S2). In this initial approach we observed near-complete overlap of the normal pancreas samples and the non-malignant adjacent pancreas samples (NMAP). Since highly similar proteomic profiles of normal and non-malignant adjacent pancreatic tissues have been reported previously [43], we merged these two conditions to a joint condition named normal non-malignant control (NNMC) for further analyses (Fig. S2). Further, we noticed that one NNMC sample was located beyond the 95% confidence interval of NNMC (Fig. S3). Since this sample represented an NMAP sample with proximity to macrodissected tumor tissue, we consider the possibility of insufficient microdissection and actual presence of tumor tissue, leading to this control sample being considered inadequate and excluded from further analysis. Furthermore, the same sPLS-DA plot showed a complete overlap of the malign entities PDAC and ampullary cancer. Due to the low AMPAC case number (n=3), we refrained from further data analysis of the AMPAC proteomes.

The remaining dataset was submitted to a second sPLS-DA analysis to globally compare the proteome profiles of PDAC, CP, and NNMC (Fig. 3 c). It is evident that all three entities show distinct clusters, suggesting proteomic differences, however with partial overlapping regions. Since the benign entities NNMC and CP have a slightly greater overlap than the malignant entity PDAC, one might speculate that the proteome difference between malignant and non-malignant entities appears to be more pronounced. The fact that benign pancreatic entities are crucially distinct from malignant tissues of PDAC is in agreement with the literature [43].
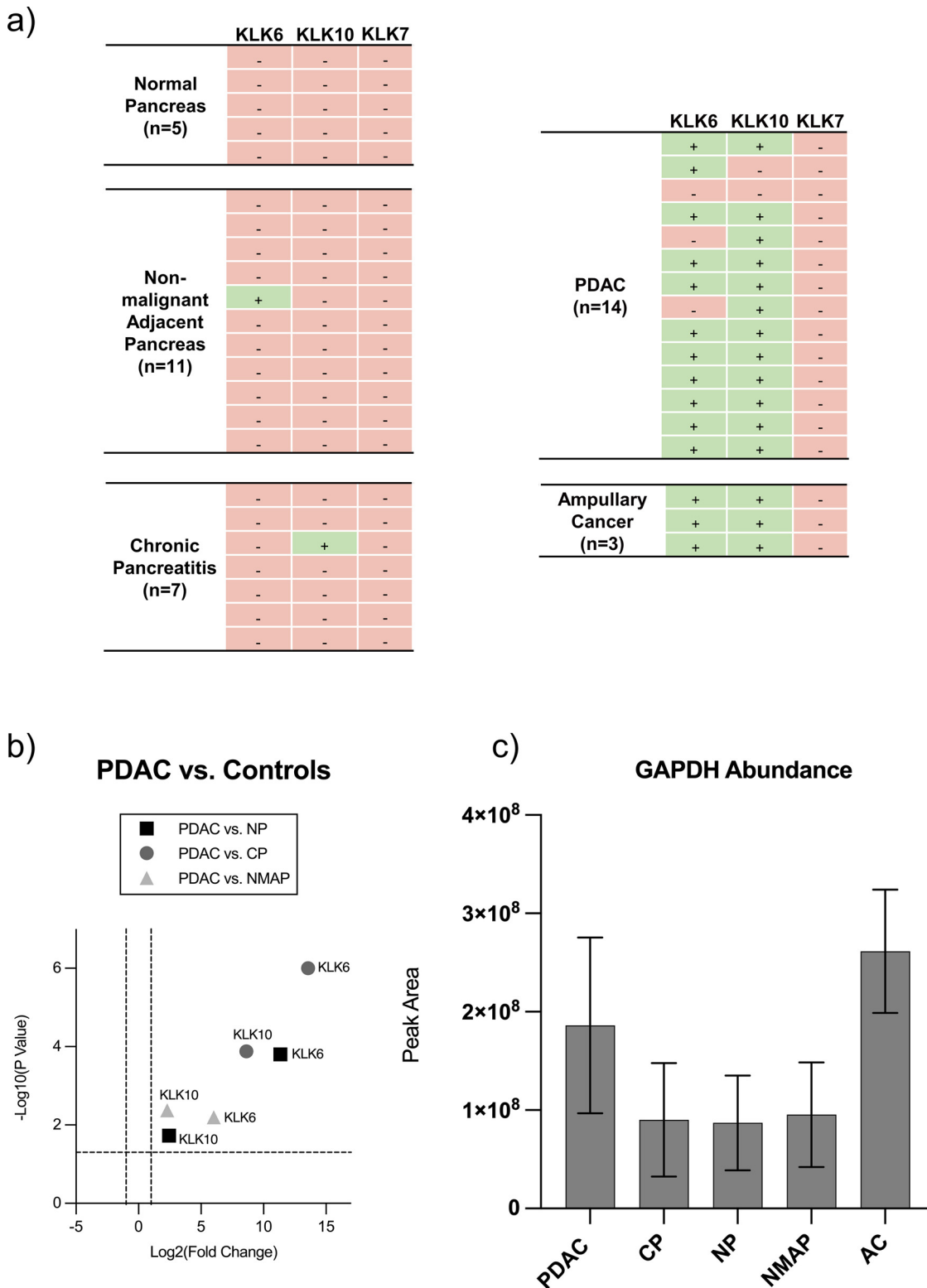
Fig. 2. *Qualitative and quantitative analysis of kallikrein proteins and GAPDH in clinical FFPE tissue samples by targeted mass spectrometry.* Clinical FFPE tissue samples deriving from pancreatic ductal adenocarcinoma (PDAC), non-malignant adjacent pancreas (NMAP), chronic pancreatitis (CP), normal pancreas (NP) and ampullary cancer (AMPAC) were subjected to targeted mass spectrometric analysis using parallel reaction monitoring (PRM). Samples were measured in randomized sample order. (a) Focusing on KLK6, KLK7 and KLK10 enabled the qualitative evaluation of their detectable presence or absence in the respective tissue entities. (b) Results of pairwise quantitative comparisons (two-sample t-test) between PDAC and the respective controls (NP, CP and NMAP) are illustrated as a combined volcano plot depicting KLK6 and KLK10. The log2 fold changes are plotted on the x-axis and corresponding adjusted p-values in -log10 scale are shown on the y-axis. The applied log2 fold change cut-off was set to 2-fold, while the adjusted p-value cut-off was set to 0.05 (1.3 in -log10 scale), each depicted as dashed lines. A log2 fold change > 0 corresponds to an upregulation in the PDAC condition. (c) Evaluation of GAPDH-abundance (represented as peak area) in the respective tissue entities is shown as bar plot. Error bars correspond to standard deviation.
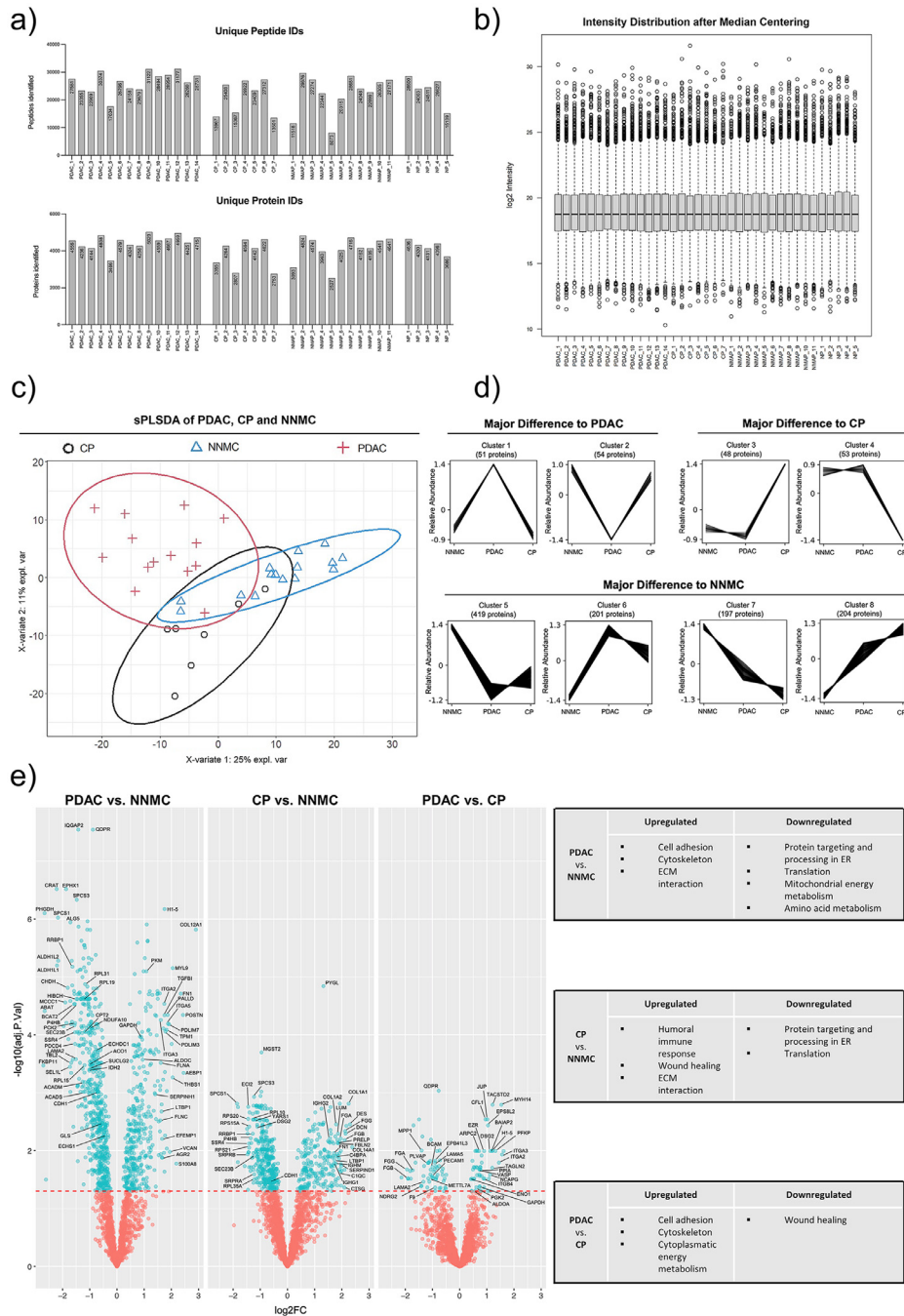
Fig. 3. *Proteomic profiling of different pancreatic tissue entities by explorative mass spectrometry.* Clinical FFPE tissue samples deriving from pancreatic ductal adenocarcinoma (PDAC), non-malignant adjacent pancreas (NMAP), chronic pancreatitis (CP) and normal pancreas (NP) were subjected to explorative mass spectrometry analysis using data-independent acquisition (DIA) mode. Samples were measured in randomized sample order. (a) Resulting numbers of identified peptides and proteins are shown for each sample and tissue entity. (b) Box-and-Whisker-Plot shows the intensity distribution of each sample after median centering. In the following, the conditions NP and NMAP were combined to the condition normal non-malignant control (NNMC). (c) Protein profiles of each sample together with the corresponding condition annotation (CP, NNMC, PDAC) were submitted to sPLS-DA analysis. X- y-axis represent the percentage of explained variance of the respective component, while ellipses represent the 95% confidence intervals. (d) Depending on the measured abundance in different conditions (NNMC, PDAC, CP), the identified proteins were assigned to different co-abundance clusters with a confidence interval of 95% by using the Clust algorithm. Each line represents an individual protein, while the y-axis illustrates the relative abundance change after Z-score normalization. Co-abundance clusters were sorted into three subgroups, depending on the conditions showing an abundance change compared to the normal control. The number of assigned proteins per cluster are shown above each graph. (e) For differential expression analysis, a pairwise multi-group limma approach was used to compare the conditions "PDAC vs. NNMC", "CP vs. NNMC" and "PDAC vs. CP" while results were illustrated as volcano plots. The log2 fold changes (log2FC) are plotted on the x-axis and corresponding adjusted p-values in -log10 scale are shown on the y-axis. The applied adjusted p-value cutoff was set to 0.05 (1.3 in -log10 scale, depicted as horizontal line). Each plot highlights significantly up- or down-regulated proteins (blue). Hereby, a log2FC > 0 corresponds to an upregulation in the first-mentioned condition. Selected proteins are marked with their respective gene names. Gene Ontology (GO) enrichment analysis for the respective up- and down-regulated proteins were performed separately and a selection of the resulting significant GO terms (adjusted p-value < 0.05) are listed for each pairwise comparison.

To identify protein groups exhibiting a comparable abundance behavior across the considered entities, we performed a co-abundance cluster analysis using the publicly available Clust algorithm [44]. The clustering revealed eight distinct abundance-courses, with a total of 1227 assigned proteins (Fig. 3 d). The eight clusters were further classified into three subgroups, depending on the condition showing the major abundance change. Considering the number of proteins assigned to each of these subgroups revealed the majority of proteins showing a similar abundance behavior of PDAC and CP compared to the control condition. This illustrates that both diseases introduce proteomic changes in the pancreas, emphasizing their deviation from the normal physiological state but also suggesting some joint characteristics [8]. To find out whether certain biological processes or functions are particularly represented in the differently identified proteins, we further performed a gene ontology (GO) enrichment analysis (Fig. S5). While we observed a strong presence of proteins involved in cell adhesion and extracellular

matrix (ECM) in PDAC, proteins involved in wound healing were evident in CP. In both PDAC and CP, there appears to be a neglect of normal exocrine pancreatic function.

In order to identify individual proteins that are significantly enriched or depleted in CP or PDAC compared to NNMC, we performed a multi-group limma approach for pairwise statistical testing to identify differentially expressed proteins in each entity. This analysis revealed 980, 588, and 112 significant hits (adjusted p-value ≤ 0.05) for the comparison PDAC vs. NNMC, CP vs. NNMC, and PDAC vs. CP, respectively (Fig. 3 e, Sup.File1). To again evaluate whether certain biological processes or functions are especially represented within the differentially expressed proteins, we performed GO enrichment analysis for the up- and down-regulated proteins separately.

For the comparison of PDAC vs. NNMC, GO terms such as cytoskeleton, ECM interaction, and cell adhesion were found to be significantly upregulated in PDAC (Fig. 3 e). Upregulated cell adhesion in PDAC com-

pared to NNMC is also apparent when considering a heatmap representation of respectively associated proteins (Fig. S4 a), which include e.g. fibronectin, vitronectin, thrombospondin, vinculin, fibulin, filamin, and several integrin subunits. Furthermore, our proteomic data also suggest an increase of cell adhesion-associated proteins in PDAC compared to CP, which could be also visualized *via* heatmap. For most epithelial tumors, downregulation of cell adhesion proteins is considered a hallmark feature. However, for PDAC, further studies support upregulation of cell adhesion proteins as compared to non-malignant control, e.g. for the tight junction protein ZO-1 [45] or the cell surface adhesion protein CEACAM6 [46] as well as unique aspects of laminin and integrin biology [47]. To the best of our knowledge, the counter-intuitive enrichment of cell adhesion proteins in PDAC (as opposed to NNMC) has been largely under-appreciated.

A strong desmoplastic reaction and unique ECM biology is another hallmark feature of PDAC [48]. Our data recapitulates this aspect as PDAC revealed a significant upregulation of proteins associated with ECM interaction, which could be visualized *via* heatmap representation (Fig. S4 b).

Downregulated GO terms in PDAC vs. NNMC comprise protein processing in the endoplasmic reticulum (ER) and mitochondrial energy metabolism (Fig. 3 e). The former likely represents de-differentiation and loss of exocrine protein secretion functionality, whereas the latter may represent the Warburg effect. As such, glycolytic processes in the cytosol often appear to be upregulated, whereas oxidative phosphorylation in mitochondria is reduced, which has been also reported for pancreatic cancer [37]. This effect is reflected by our data, as we observed a significant upregulation of several cytoplasmic glycolytic proteins (pyruvate kinase, PKM; fructose-bisphosphate aldolase C, ALDOC; ATP-dependent 6-phosphofructokinase, PFKP; phosphoglycerate kinase 2, PGK2; fructose-bisphosphate aldolase A, ALDOA; alpha-enolase, ENO1; and glyceraldehyde-3-phosphate dehydrogenase, GAPDH) in PDAC compared to the benign controls. Consistently, we also detected the decrease in the mitochondrial energy metabolism showing a significant downregulation of citrate cycle proteins (isocitrate dehydrogenase, IDH2; cytoplasmic aconitate hydratase, ACO1; succinate–CoA ligase, SUCLG2) in PDAC as well as an enriched GO term for downregulated mitochondrial energy metabolism. Furthermore, we detected a decreased amino acid degradation and downregulation of catabolic amino acid enzymes in PDAC, resulting in an increased availability of amino acids and thus providing an important source for cancer development and growth. Thereby, this data shows the metabolic reprogramming including the Warburg effect in PDAC on proteome level.

In chronic pancreatitis, GO terms like humoral immune response as well as wound healing were found to be significantly upregulated when compared to NNMC and show consistency with the literature [49,50]. As representatives thereof, the complement protein c4b-binding protein alpha chain (C4BPA), several immunoglobulins (IGHG1, IGHG2, IGHM) and several fibrinogens (FGA, FGB, FGG) were found to be significantly upregulated in CP. Since the systemic activation of the complement system in pancreatitis is already known, our data further suggests the triggering of the complement system *via* the classical pathway as the C4BPA protein as well as immunoglobulins are known to be involved in this specific complement pathway [49,51]. Thereby, our data confirms the inflammatory state of this entity on proteome level and highlights the role of the complement system in chronic pancreatitis.

Despite the described differences in the proteomes of PDAC and CP, we also identified several similarities when compared to NNMC. One of them is the common upregulation of ECM interaction and corresponding protein representatives including fibronectin (FN1), vinculin (VCL), fibulin (FBLN1), filamin (FLNA) and integrin subunits (e.g. ITGB1, IT-GAV). The jointly increased abundance of proteins associated with ECM interaction in PDAC and CP is also apparent when considering a more global heatmap representation (Fig. S4 b). This may reflect the role of ECM proteins in tissue remodelling not only during cancer progression but also during inflammatory processes [52,53].

Furthermore, we observed a strong reduction of the exocrine pancreatic function in both PDAC and CP, since GO terms like translation and post-translational processing were found to be significantly downregulated as well as several representative proteins thereof including ribosomal proteins (RPL15, RPL19, RPL31, RPS20, RPS21, RPS15A, RPL10, RPL35A), members of protein export mechanisms (SPCS1, SPCS3, SRPRA, SRPRB) as well as protein processing (RRBP1, P4HB). This highlights the overall decreased gene expression in inflammatory as well as in malign pancreatic tissue.

Of note, we did not detect any KLK proteins in our DIA dataset. This may stem from technical limitations of the Q-Exactive plus mass spectrometer with regard to DIA proteomics, e.g. its limited scan speed. However, the same mass spectrometer was used for the successful PRM analysis of KLK6 and KLK10 described above, which is in line with earlier reports highlighting a much increased sensitivity of this instrumentation for targeted proteomics in the context of complex samples [54]. Hence, approaches that combine targeted and explorative methods may be particularly powerful [55].

Pro-tumorigenic and tumor suppressor proteins in PDAC and CP

Sustaining proliferative signaling and evading growth suppressors are common hallmarks of cancer [56], which is why we specifically evaluated our proteomic dataset with regard to pro-tumorigenic and tumor suppressor proteins in PDAC and CP. As previously outlined, we have reduced our dataset to proteins with a maximum missingness of 20% in PDAC, CP, and NNMC. This strict criterium might obscure tumor-driving or -suppressing proteins which are specifically present in PDAC and CP. Hence, we screened our proteomic data without missingness reduction for common protein representatives of each respective group using a reference list of Bosman *et al.* and further literature [57,58]. This resulted in the identification of ten pro-tumorigenic, and five tumor suppressor proteins in our data (Table 2 a and b).

We noticed enrichment in PDAC for the pro-tumorigenic proteins ADAM9, AKT2, ERBB2, KRAS, MSLN, S100A6 and S100P compared to normal pancreas samples (Table 2 a). The strong presence of the cell surface protease ADAM9 in PDAC and its near-absence in normal pancreas samples is in line with earlier reports [59]. For ADAM9, we have previously shown its involvement in PDAC tumor biology by contributing to cell-matrix adhesion, anchorage independent growth, and neoangiogenesis [60]. The proliferation-regulating protein KRAS has been described as the major driver gene in PDAC, which corroborates our proteomic findings [41].

Interestingly, CP shows an intermediate enrichment for ADAM9, AKT2, KRAS, S100A6 and S100P between PDAC and normal pancreas and thereby supports the hypothesis of CP being a preliminary stage to PDAC [8–10]. However, the receptor tyrosine kinase ERBB2 was not detected in CP and thereby showed an even diminished detection compared to normal pancreas samples.

Although it is expected to find the enrichment of pro-tumorigenic proteins in PDAC accompanied by the depletion of tumor suppressor proteins, we could not observe this for the considered tumor suppressor proteins in our proteomic data (Table 2 b). Generally, tumor-driving proteins have lower missingness in PDAC and CP as opposed to NNMC while tumor-suppressing proteins have similar missingness in these entities (Table 2 a and b).

Semi-tryptic peptide analysis to estimate endogenous proteolytic processing

We have previously shown that semi-tryptic (re-)analysis of proteomic data might yield novel insight into endogenous proteolytic processing prior to sample harvesting and eventual trypsination [61–63]. Naturally, this re-analysis can only shed light on non-tryptic, endogenous proteolytic processing. To further explore the acquired proteomic data, we performed a semi-tryptic peptide analysis to evaluate trypsin-independent endogenous proteolytic events and to compare them be-

Table 2

Number and proportion of identified pro-tumorigenic and tumor suppressor proteins in acquired explorative pancreatic dataset. The acquired dataset described and analyzed in Fig. 3 was screened for commonly known (a) pro-tumorigenic and (b) tumor suppressor proteins in tissues of PDAC, CP and Normal Pancreas. Listed values represent the percentage of samples in which the respective Protein Of Interest (POI) has been identified as well as the corresponding absolute number of samples in brackets.

a)	Pro-tumorigenic Protein	Proportion of PDAC Samples identifying POI (total n=14)	Proportion of CP Samples identifying POI (total n=7)	Proportion of Normal Pancreas Samples identifying POI (total n=15)
	ADAM9	57.1 % (10)	14.3 % (1)	6.7 % (1)
	AGR2	100.0 % (14)	100.0 % (7)	100.0 % (15)
	AKT2	57.1 % (8)	28.6 % (2)	20.0 % (3)
	CDH1	100.0 % (14)	100.0 % (7)	100.0 % (15)
	ERBB2	28.6 % (4)	0.0 % (0)	13.3 % (2)
	KRAS	92.9 % (13)	71.4 % (5)	33.3 % (5)
	MSLN	35.7 % (5)	0.0 % (0)	0.0 % (0)
	S100A4	100.0 % (14)	100.0 % (7)	93.3 % (14)
	S100A6	85.7 % (12)	71.4 % (5)	60.0 % (9)
	S100P	100.0 % (14)	85.7 % (6)	46.7 % (7)
b)	Tumor Suppressor Protein	Proportion of PDAC Samples identifying POI (total n=14)	Proportion of CP Samples identifying POI (total n=7)	Proportion of Normal Pancreas Samples identifying POI (total n=15)
	EPB41L3	100.0 % (14)	100.0 % (7)	100.0 % (15)
	MAP2K4	92.9 % (13)	71.4 % (5)	93.3 % (14)
	NDRG2	92.9 % (13)	85.7 % (6)	86.7 % (13)
	PDCD4	100.0 % (14)	100.0 % (7)	100.0 % (15)
	SMAD4	71.4 % (10)	42.9 % (3)	60.0 % (9)

$$Intensity - Proportion_{Condition} = \frac{Sum\ (Intensity\ "SemiTryptic")_{Condition}}{Sum\ (Intensity\ "All\ peptides")_{Condition}}$$

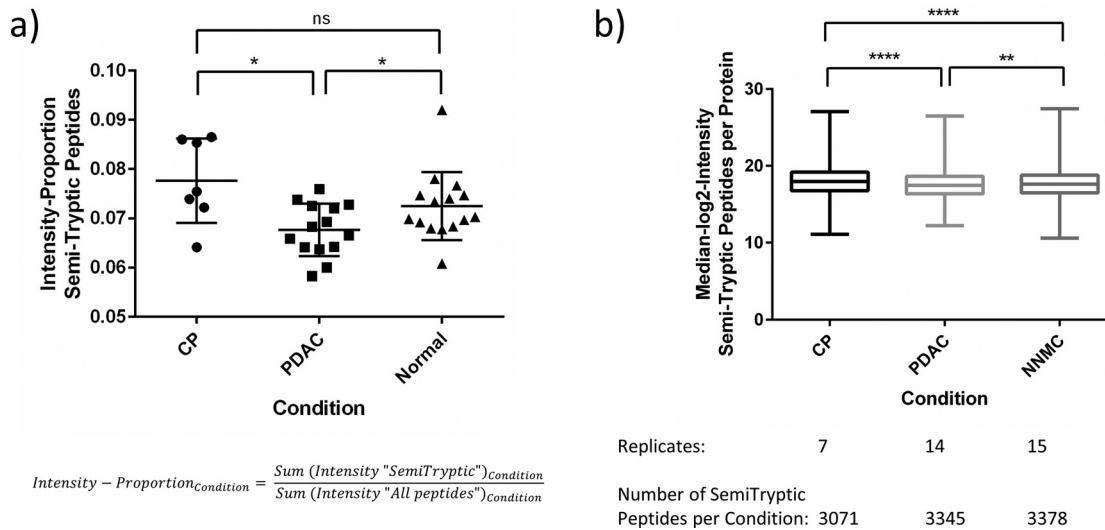| Replicates:	7	14	15
Number of SemiTryptic Peptides per Condition:	3071	3345	3378

Fig. 4. *Semi-tryptic peptide analysis revealed different proteolytic activities in the considered entities.* The acquired dataset described and analyzed in Fig. 3, was additionally analyzed with regard to semi-tryptic peptides. (a) Intensity proportion of semi-tryptic peptides in relation to all detected peptides is shown for each entity (CP, PDAC, Normal) as scatter plot with the horizontal line representing the mean and error bars corresponding to standard deviation. (b) Absolute median log2 intensities of all semi-tryptic peptides per protein were calculated for each entity (CP, PDAC, Normal) and are shown as Box-and-Whisker-Plot. For significance testing, a two-tailed Mann-Whitney U test was used because of unequal sample sizes between groups and not all groups passing normality tests (ns – non significant, *P ≤ 0.05, **P ≤ 0.01, ****P ≤ 0.0001).

tween the considered entities. Only semi-tryptic peptides were considered but not fully unspecific peptides. This revealed 3747 semi-tryptic peptides and corresponds to approximately 10% of all peptides (35,167 peptides) (data not shown). Comparing the intensity proportion of semi-tryptic peptides in relation to the overall intensity of all identified peptides revealed a significantly lower intensity-proportion in PDAC compared to the benign controls (Fig. 4 a). Chronic pancreatitis showed the highest intensity proportion for semi-tryptic peptides, although its difference to normal pancreas was not significant. Likewise, when considering the absolute median log2 intensities of all semi-tryptic peptides per protein, we observed a slightly but significantly higher intensity for chronic pancreatitis although revealing less semi-tryptic peptides than PDAC and normal pancreas (Fig. 4 b). This is expected and thereby validates previous literature as inflammatory tissue was previously described to reveal a higher proteolytic activity [64,65].

Proteogenomic analysis and identification of single amino acid variants

Proteogenomic analysis uses a combination of proteomics, genomics and/or transcriptomics data and enables the additional identification of potential sequence variants or copy number variations (CNVs) that could have not been detected when only considering a default human reference proteome [66]. This has recently been published for PDAC, where several correlations between CNVs and RNA or CNVs and protein level have been described [41]. Although we do not have genomic data for our cohort and thereby cannot validate them as CNVs, we noticed that four differentially expressed proteins enriched in PDAC and two proteins depleted in PDAC were annotated as CNV in the aforementioned publication.

In order to identify PDAC- and CP-related single amino acid variants (SAAVs) in our data, proteogenomic analysis was performed on the ac-
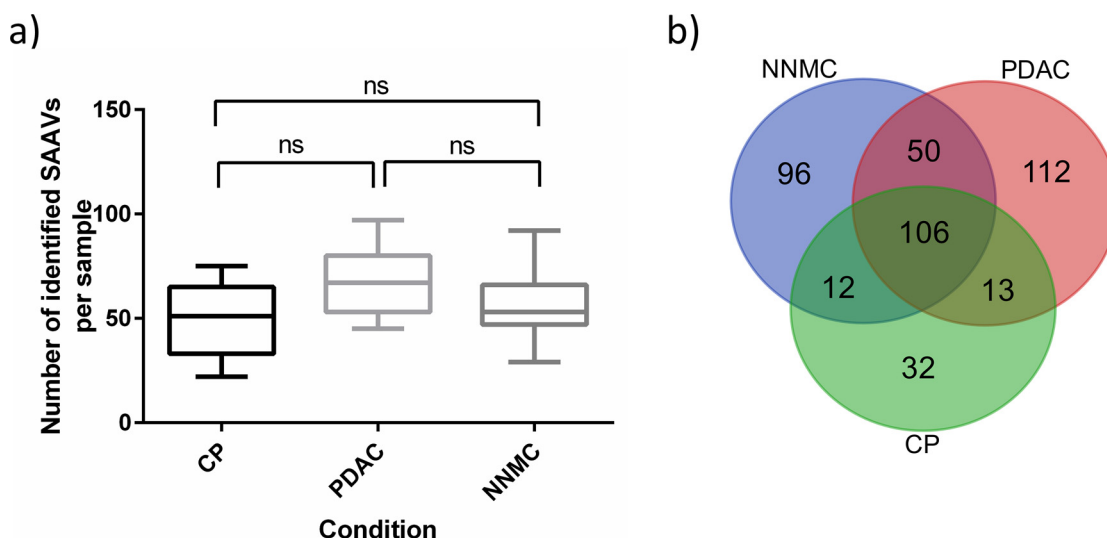
a)

b)

Fig. 5. *Proteogenomic analysis revealed single amino acid variants in PDAC and chronic pancreatitis.* The acquired dataset described and analyzed in Fig. 3 was additionally analyzed with regard to single amino acid variants (SAAVs). (a) Number of identified SAAVs per sample is shown for each entity (CP, PDAC, Normal) as Box-and-Whisker plot. For statistical testing, one-way ANOVA with Tukey´s multiple comparisons test was used (ns – non significant). (b) The Venn diagram shows overlapping and specific SAAVs for each condition.

quired proteomic data together with publicly accessible transcriptomic data for PDAC (Table S4 – a for used SRA-identifier) [67] as a proof-of-concept study. This resulted in the median identification of 51, 67 and 53 SAAVs per CP, PDAC and NNMC sample, respectively (Fig. 5 a). Although PDAC reveals a tendency for higher mutational burden/load, we do not see a significant difference by orders of magnitude in PDAC or CP in SAAV numbers, nor when comparing the intensity of the respective SAAV peptides (Fig. S6). This lack of differentiation may be due to the fact that we are not only considering mutational burden but also, to a large extent, naturally occurring allele variants which are also not part of the default proteome sequence database.

Comparing the identified SAAVs between the conditions, revealed 32 and 112 SAAVs, which specifically occur in CP and PDAC, respectively (Fig. 5 b, Sup.File2). From the 112 PDAC-specific SAAVs, 16 were annotated as "pathogenic" (FATHMM-prediction) in the publicly available COSMIC database [68]. Among these, the GTPase KRAS was found to be mutated in PDAC by the G12V substitution, which has been described to be one of the most frequently accumulated hotspots point mutations in PDAC [41,69] and thereby validates our findings. Of note, for CP, we found a S106L substitution of HRAS, which is another member of the RAS family comprising the most frequently mutated oncogene family in cancer. In addition, we found an E2410* nonsense mutation (prematurely terminated protein) of the Ankyrin repeat and KH domain-containing protein 1 (ANKHD1, UniprotID: Q8IWZ3) in PDAC, which native form has been suggested to be involved in cell proliferation of various cancers [70,71]. Furthermore, ANKHD1 has been described to inhibit Cyclin-dependent kinase inhibitor 1 (CDKN1A) expression in pancreatic cancer and is even suggested as potential therapeutic target [72]. However, to evaluate the influence of the here observed SAAV, further investigations are required which is beyond the scope of this project.

Conclusion

Our study suggests that KLK6 and 10 contribute to the tumor biology of pancreatic cancer due to their cell line specific secretion and significant upregulation in malignant entities such as PDAC and ampullary cancer compared to chronic pancreatitis and normal non-malignant pancreatic tissue. Explorative mass spectrometry clearly showed demarcated proteomes of malignant pancreatic tissue from protein expres-

sion patterns of the benign pancreas, among others illustrating the Warburg effect and a strong ECM fingerprint in pancreatic cancer. Furthermore, chronic pancreatitis was shown to reveal a higher endogenous proteolytic activity and proteogenomic analysis identified KRAS and ANKHD1 sequence variants. Our study provides novel insights into the proteomic characterization of PDAC and chronic pancreatitis and its association with kallikrein proteases, although larger patient cohorts will be required to validate our findings.

Availability of data

All mass spectrometry proteomics datasets used and/or analyzed during this study are available online at the MassIVE repository and are available from the corresponding author upon reasonable request. (http://massive.ucsd.edu/; dataset identifier: MSV000090255).

Ethics

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Heidelberg University (02.01.2017 / 2012-293N-MA). Informed consent was obtained from all individual participants included in the study.

Declaration of Competing Interest

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

CRediT authorship contribution statement

Janina Werner: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Patrick Bernhard:** Conceptualization, Investigation, Visualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Miguel Cosenza-Contreras:** Conceptualization, Formal analysis, Methodology, Software, Writing – review & editing. **Niko Pinter:** Conceptualization, Formal analysis, Methodology, Software, Writing – review & editing. **Matthias Fahrner:** Conceptualization, Methodology, Writing – review & editing. **Prama Pallavi:** Data curation, Writing – review & editing. **Johannes Eberhard:** Data curation, Writing – review & editing. **Peter Bronsert:** Data curation, Writing – review & editing. **Felix Rückert:** Conceptualization, Resources, Writing – review & editing. **Oliver Schilling:** Conceptualization, Resources, Funding acquisition, Writing – review & editing.

Acknowledgment

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neo.2022.100871.

References

[1] P. Sarantis, E. Koustas, A. Papadimitropoulou, A.G. Papavassiliou, M.V. Karamouzis, Pancreatic ductal adenocarcinoma: treatment hurdles, tumor microenvironment and immunotherapy, World J. Gastrointest. Oncol. 12 (2020) 173–181, doi:10.4251/wjgo.v12.i2.173.

[2] S.P. Pereira, L. Oldfield, A. Ney, P.A. Hart, M.G. Keane, S.J. Pandol, D. Li, W. Greenhalf, C.Y. Jeon, E.J. Koay, C.V. Almario, C. Halloran, A.M. Lennon, E. Costello, Early detection of pancreatic cancer, Lancet Gastroenterol. Hepatol. 5 (2020) 698–710, doi:10.1016/S2468-1253(19)30416-9.

[3] R. Casolino, C. Braconi, G. Malleo, S. Paiella, C. Bassi, M. Milella, S.B. Dreyer, F.E.M. Froeling, D.K. Chang, A.V. Biankin, T. Golan, Reshaping preoperative treatment of pancreatic cancer in the era of precision medicine, Ann. Oncol. 32 (2021) 183–196, doi:10.1016/j.annonc.2020.11.013.

[4] A. McGuigan, P. Kelly, R.C. Turkington, C. Jones, H.G. Coleman, R.S. McCain, Pancreatic cancer: a review of clinical diagnosis, epidemiology, treatment and outcomes, World J. Gastroenterol. 24 (2018) 4846–4861, doi:10.3748/wjg.v24.i43.4846.

[5] G. Luo, K. Jin, S. Deng, H. Cheng, Z. Fan, Y. Gong, Y. Qian, Q. Huang, Q. Ni, C. Liu, X. Yu, Roles of CA19-9 in pancreatic cancer: biomarker, predictor and promoter, Biochim. Biophys. Acta Rev. Cancer 1875 (2021) 188409, doi:10.1016/j.bbcan.2020.188409.

[6] J.E. Kim, K.T. Lee, J.K. Lee, S.W. Paik, J.C. Rhee, K.W. Choi, Clinical usefulness of carbohydrate antigen 19-9 as a screening test for pancreatic cancer in an asymptomatic population, J. Gastroenterol. Hepatol. 19 (2004) 182–186, doi:10.1111/j.1440-1746.2004.03219.x.

[7] A.D. Singhi, E.J. Koay, S.T. Chari, A. Maitra, Early detection of pancreatic cancer: opportunities and challenges, Gastroenterology 156 (2019) 2024–2040, doi:10.1053/j.gastro.2019.01.259.

[8] M. Vujasinovic, A. Dugic, P. Maisonneuve, A. Aljic, R. Berggren, N. Panic, R. Valente, R. Pozzi Mucelli, A. Waldthaler, P. Ghorbani, M. Kordes, H. Hagstrom, J.M. Lohr, Risk of developing pancreatic cancer in patients with chronic pancreatitis, J. Clin. Med. 9 (2020), doi:10.3390/jcm9113720.

[9] E. Del Poggetto, I.L. Ho, C. Balestrieri, E.Y. Yen, S. Zhang, F. Citron, R. Shah, D. Corti, G.R. Diaferia, C.Y. Li, S. Loponte, F. Carbone, Y. Hayakawa, G. Valenti, S. Jiang, L. Sapio, H. Jiang, P. Dey, S. Gao, A.K. Deem, S. Rose-John, W. Yao, H. Ying, A.D. Rhim, G. Genovese, T.P. Heffernan, A. Maitra, T.C. Wang, L. Wang, G.F. Draetta, A. Carugo, G. Natoli, A. Viale, Epithelial memory of inflammation limits tissue damage while promoting pancreatic tumorigenesis, Science 373 (2021) eabj0486, doi:10.1126/science.abj0486.

[10] T. Alhobayb, R. Peravali, M. Ashkar, The relationship between acute and chronic pancreatitis with pancreatic adenocarcinoma: review, Diseases 9 (2021), doi:10.3390/diseases9040093.

[11] X.Y. Cao, X.X. Zhang, M.W. Yang, L.P. Hu, S.H. Jiang, G.A. Tian, L.L. Zhu, Q. Li, Y.W. Sun, Z.G. Zhang, Aberrant upregulation of KLK10 promotes metastasis *via* enhancement of EMT and FAK/SRC/ERK axis in PDAC, Biochem. Biophys. Res. Commun. 499 (2018) 584–593, doi:10.1016/j.bbrc.2018.03.194.

[12] S.R. Hustinx, D. Cao, A. Maitra, N. Sato, S.T. Martin, D. Sudhir, C. Iacobuzio-Donahue, J.L. Cameron, C.J. Yeo, S.E. Kern, M. Goggins, J. Mollenhauer, A. Pandey, R.H. Hruban, Differentially expressed genes in pancreatic ductal adenocarcinomas identified through serial analysis of gene expression, Cancer Biol. Ther. 3 (2004) 1254–1261, doi:10.4161/cbt.3.12.1238.

[13] C.A. Iacobuzio-Donahue, R. Ashfaq, A. Maitra, N.V. Adsay, G.L. Shen-Ong, K. Berg, M.A. Hollingsworth, J.L. Cameron, C.J. Yeo, S.E. Kern, M. Goggins, R.H. Hruban, Highly expressed genes in pancreatic ductal adenocarcinomas: a comprehensive characterization and comparison of the transcription profiles obtained from three major technologies, Cancer Res. 63 (2003) 8614–8622.

[14] F. Ruckert, M. Hennig, C.D. Petraki, D. Wehrum, M. Distler, A. Denz, M. Schroder, G. Dawelbait, H. Kalthoff, H.D. Saeger, E.P. Diamandis, C. Pilarsky, R. Grutzmann, Co-expression of KLK6 and KLK10 as prognostic factors for survival in pancreatic ductal adenocarcinoma, Br. J. Cancer 99 (2008) 1484–1492, doi:10.1038/sj.bjc.6604717.

[15] G.M. Yousef, C.A. Borgono, N.M. White, J.D. Robb, I.P. Michael, K. Oikonomopoulou, S. Khan, E.P. Diamandis, In silico analysis of the human kallikrein gene 6, Tumour Biol. 25 (2004) 282–289, doi:10.1159/000081393.

[16] S.K. Hong, Kallikreins as biomarkers for prostate cancer, Biomed. Res. Int. 2014 (2014) 526341, doi:10.1155/2014/526341.

[17] J.B. Candido, O. Maiques, M. Boxberg, V. Kast, E. Peerani, E. Tomas-Bort, W. Weichert, A. Sananes, N. Papo, V. Magdolen, V. Sanz-Moreno, D. Loessner, Kallikrein-related peptidase 6 is associated with the tumour microenvironment of pancreatic ductal adenocarcinoma, Cancers (Basel) 13 (2021), doi:10.3390/cancers13163969.

[18] V. Iakovlev, E.R. Siegel, M.S. Tsao, R.S. Haun, Expression of kallikrein-related peptidase 7 predicts poor prognosis in patients with unresectable pancreatic ductal adenocarcinoma, Cancer Epidemiol. Biomarkers Prev. 21 (2012) 1135–1142, doi:10.1158/1055-9965.EPI-11-1079.

[19] B. Klucky, R. Mueller, I. Vogt, S. Teurich, B. Hartenstein, K. Breuhahn, C. Flechtenmacher, P. Angel, J. Hess, Kallikrein 6 induces E-cadherin shedding and promotes cell proliferation, migration, and invasion, Cancer Res. 67 (2007) 8198–8206, doi:10.1158/0008-5472.CAN-07-0607.

[20] S.K. Johnson, V.C. Ramani, L. Hennings, R.S. Haun, Kallikrein 7 enhances pancreatic cancer cell invasion by shedding E-cadherin, Cancer 109 (2007) 1811–1820, doi:10.1002/cncr.22606.

[21] A. Magklara, A.A. Mellati, G.A. Wasney, S.P. Little, G. Sotiropoulou, G.W. Becker, E.P. Diamandis, Characterization of the enzymatic activity of human kallikrein 6: autoactivation, substrate specificity, and regulation by inhibitors, Biochem. Biophys. Res. Commun. 307 (2003) 948–955, doi:10.1016/s0006-291x(03)01271-3.

[22] S. Srinivasan, T. Kryza, J. Batra, J. Clements, Remodelling of the tumour microenvironment by the kallikrein-related peptidases, Nat. Rev. Cancer 22 (2022) 223–238, doi:10.1038/s41568-021-00436-z.

[23] J.L. Shaw, E.P. Diamandis, Regulation of human tissue kallikrein-related peptidase expression by steroid hormones in 32 cell lines, Biol. Chem. 389 (2008) 1409–1419, doi:10.1515/bc.2008.158.

[24] C.S. Hughes, S. Moggridge, T. Muller, P.H. Sorensen, G.B. Morin, J. Krijgsveld, Single-pot, solid-phase-enhanced sample preparation for proteomics experiments, Nat. Protoc. 14 (2019) 68–85, doi:10.1038/s41596-018-0082-x.

[25] B.C. Searle, L.K. Pino, J.D. Egertson, Y.S. Ting, R.T. Lawrence, B.X. MacLean, J. Villen, M.J. MacCoss, Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry, Nat. Commun. 9 (2018) 5128, doi:10.1038/s41467-018-07454-w.

[26] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics 26 (2010) 966–968, doi:10.1093/bioinformatics/btq054.

[27] V. Demichev, C.B. Messner, S.I. Vernardis, K.S. Lilley, M. Ralser, DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput, Nat. Methods 17 (2020) 41–44, doi:10.1038/s41592-019-0638-x.

[28] W.E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods, Biostatistics 8 (2007) 118–127, doi:10.1093/biostatistics/kxj037.

[29] E. Afgan, D. Baker, B. Batut, M. van den Beek, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, B.A. Grüning, A. Guerler, J. Hillman-Jackson, S. Hiltemann, V. Jalili, H. Rasche, N. Soranzo, J. Goecks, J. Taylor, A. Nekrutenko, D. Blankenberg, The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update, Nucleic Acids Res. 46 (2018) W537–W544, doi:10.1093/nar/gky379.

[30] F. Rückert, K. Werner, D. Aust, S. Hering, H. Saeger, R. Grützmann, C. Pilarsky, Establishment and characterization of six primary pancreatic cancer cell lines, Austin J. Cancer Clin. Res. 2 (2015) 1055.

[31] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, Quantitative mass spectrometry in proteomics: a critical review, Anal. Bioanal. Chem. 389 (2007) 1017–1031, doi:10.1007/s00216-007-1486-6.

[32] G.M. Yousef, C.A. Borgono, C. Popalis, G.M. Yacoub, M.E. Polymeris, A. Soosaipillai, E.P. Diamandis, In-silico analysis of kallikrein gene expression in pancreatic and colon cancers, Anticancer Res. 24 (2004) 43–51.

[33] N.T. Van Heek, A. Maitra, J. Koopmann, N. Fedarko, A. Jain, A. Rahman, C.A. Iacobuzio-Donahue, V. Adsay, R. Ashfaq, C.J. Yeo, J.L. Cameron, J.A. Offerhaus, R.H. Hruban, K.D. Berg, M. Goggins, Gene expression profiling identifies markers of ampullary adenocarcinoma, Cancer Biol. Ther. 3 (2004) 651–656, doi:10.4161/cbt.3.7.919.

[34] J.P. Du, L. Li, J. Zheng, D. Zhang, W. Liu, W.H. Zheng, X.S. Li, R.C. Yao, F. Wang, S. Liu, X. Tan, Kallikrein-related peptidase 7 is a potential target for the treatment of pancreatic cancer, Oncotarget 9 (2018) 12894–12906, doi:10.18632/oncotarget.24132.

[35] C. Rubie, K. Kempf, J. Hans, T. Su, B. Tilton, T. Georg, B. Brittner, B. Ludwig, M. Schilling, Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues, Mol. Cell Probes 19 (2005) 101–109, doi:10.1016/j.mcp.2004.10.001.

[36] S. Bose, C. Zhang, A. Le, Glucose metabolism in cancer: the warburg effect and beyond, Adv. Exp. Med. Biol. 1311 (2021) 3–15, doi:10.1007/978-3-030-65768-0_1.

[37] J. Yang, B. Ren, G. Yang, H. Wang, G. Chen, L. You, T. Zhang, Y. Zhao, The enhancement of glycolysis regulates pancreatic cancer metastasis, Cell Mol. Life Sci. 77 (2020) 305–321, doi:10.1007/s00018-019-03278-z.

[38] O. Warburg, Über den Stoffwechsel der Carcinomzelle, Naturwissenschaften 12 (1924) 1131–1137, doi:10.1007/BF01504608.

[39] K. Fröhlich, E. Brombacher, M. Fahrner, D. Vogele, L. Kook, N. Pinter, P. Bronsert, S. Timme-Bronsert, A. Schmidt, K. Bärenfaller, C. Kreutz, O. Schilling, Benchmarking of analysis strategies for data-independent acquisition proteomics using a large-scale dataset comprising inter-patient heterogeneity, Nat. Commun. 13 (2022) 2622, doi:10.1038/s41467-022-30094-0.

[40] B.C. Searle, K.E. Swearingen, C.A. Barnes, T. Schmidt, S. Gessulat, B. Küster, M. Wilhelm, Generating high quality libraries for DIA MS with empirically corrected peptide predictions, Nat. Commun. 11 (2020) 1548, doi:10.1038/s41467-020-15346-1.

[41] L. Cao, C. Huang, D. Cui Zhou, Y. Hu, T.M. Lih, S.R. Savage, K. Krug, D.J. Clark, M. Schnaubelt, L. Chen, F. da Veiga Leprevost, R.V. Eguez, W. Yang, J. Pan, B. Wen, Y. Dou, W. Jiang, Y. Liao, Z. Shi, N.V. Terekhanova, S. Cao, R.J. Lu, Y. Li, R. Liu, H. Zhu, P. Ronning, Y. Wu, M.A. Wyczalkowski, H. Easwaran, L. Danilova, A.S. Mer, S. Yoo, J.M. Wang, W. Liu, B. Haibe-Kains, M. Thiagarajan, S.D. Jewell, G. Hostetter, C.J. Newton, Q.K. Li, M.H. Roehrl, D. Fenyo, P. Wang, A.I. Nesvizhskii, D.R. Mani, G.S. Omenn, E.S. Boja, M. Mesri, A.I. Robles, H. Rodriguez, O.F. Bathe, D.W. Chan, R.H. Hruban, L. Ding, B. Zhang, H. Zhang, Clinical Proteomic Tumor Analysis, C: Proteogenomic characterization of pancreatic ductal adenocarcinoma, Cell 184 (2021) 5031–5052 e5026, doi:10.1016/j.cell.2021.08.023.

[42] V.O. Oria, P. Bronsert, A.R. Thomsen, M.C. Foll, C. Zamboglou, L. Hannibal, S. Behringer, M.L. Biniossek, C. Schreiber, A.L. Grosu, L. Bolm, D. Rades, T. Keck, M. Werner, U.F. Wellner, O. Schilling, Proteome profiling of primary pancreatic ductal adenocarcinomas undergoing additive chemoradiation link ALDH1A1 to early local recurrence and chemoradiation resistance, Transl. Oncol. 11 (2018) 1307–1322, doi:10.1016/j.tranon.2018.08.001.

[43] Y. Cui, M. Tian, M. Zong, M. Teng, Y. Chen, J. Lu, J. Jiang, X. Liu, J. Han, Proteomic analysis of pancreatic ductal adenocarcinoma compared with normal adjacent pancreatic tissue and pancreatic benign cystadenoma, Pancreatology 9 (2009) 89–98, doi:10.1159/000178879.

[44] B. Abu-Jamous, S. Kelly, Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data, Genome Biol. 19 (2018) 172, doi:10.1186/s13059-018-1536-8.

[45] J. Kleeff, X. Shi, H.P. Bode, K. Hoover, S. Shrikhande, P.J. Bryant, M. Korc, M.W. Buchler, H. Friess, Altered expression and localization of the tight junction protein ZO-1 in primary and metastatic pancreatic cancer, Pancreas 23 (2001) 259–265, doi:10.1097/00006676-200110000-00006.

[46] R. Pandey, M. Zhou, S. Islam, B. Chen, N.K. Barker, P. Langlais, A. Srivastava, M. Luo, L.S. Cooke, E. Weterings, D. Mahadevan, Carcinoembryonic antigen cell adhesion molecule 6 (CEACAM6) in Pancreatic Ductal Adenocarcinoma (PDA): an integrative analysis of a novel therapeutic target, Sci. Rep. 9 (2019) 18347, doi:10.1038/s41598-019-54545-9.

[47] J.D. Humphries, J. Zha, J. Burns, J.A. Askari, C.R. Below, M.R. Chastney, M.C. Jones, A. Mironov, D. Knight, D.A. O'Reilly, M.J. Dunne, D.R. Garrod, C. Jorgensen, M.J. Humphries, Pancreatic ductal adenocarcinoma cells employ integrin alpha6beta4 to form hemidesmosomes and regulate cell proliferation, Matrix Biol. 110 (2022) 16–39, doi:10.1016/j.matbio.2022.03.010.

[48] V.M. Perez, J.F. Kearney, J.J. Yeh, The PDAC extracellular matrix: a review of the ECM protein composition, tumor cell interaction, and therapeutic strategies, Front Oncol. 11 (2021) 751311, doi:10.3389/fonc.2021.751311.

[49] L. Bettac, S. Denk, T. Seufferlein, M. Huber-Lang, Complement in pancreatic disease-perpetrator or savior? Front Immunol. 8 (2017) 15, doi:10.3389/fimmu.2017.00015.

[50] M.C. Whittle, S.R. Hingorani, Fibroblasts in pancreatic ductal adenocarcinoma: biological mechanisms and therapeutic targets, Gastroenterology 156 (2019) 2085–2096, doi:10.1053/j.gastro.2018.12.044.

[51] M.M. Markiewski, B. Nilsson, K.N. Ekdahl, T.E. Mollnes, J.D. Lambris, Complement and coagulation: strangers or partners in crime? Trends Immunol. 28 (2007) 184–192, doi:10.1016/j.it.2007.02.006.

[52] t2T.F.G. Shouji Shimoyama, T.O. Susanne Gansauge, G.B. Hans, Altered expression of extracellular matrix molecules and their receptors in chronic pancreatitis and pancreatic adenocarcinoma in comparison with normal pancreas, Int. J. Pancreatol. 18 (1995) 227–234.

[53] C. Tian, K.R. Clauser, D. Ohlund, S. Rickelt, Y. Huang, M. Gupta, D.R. Mani, S.A. Carr, D.A. Tuveson, R.O. Hynes, Proteomic analyses of ECM during pancreatic ductal adenocarcinoma progression reveal different contributions by tmor and stromal cells, Proc Natl Acad Sci U S A 116 (2019) 19609–19618, doi:10.1073/pnas.1908626116.

[54] M. Bauer, E. Ahrne, A.P. Baron, T. Glatter, L.L. Fava, A. Santamaria, E.A. Nigg, A. Schmidt, Assessment of current mass spectrometric workflows for the quantification of low abundant proteins and phosphorylation sites, Data Brief 5 (2015) 297–304, doi:10.1016/j.dib.2015.08.015.

[55] G. Hart-Smith, R.S. Reis, P.M. Waterhouse, M.R. Wilkins, Improved quantitative plant proteomics *via* the combination of targeted and untargeted data acquisition, Front Plant Sci. 8 (2017) 1669, doi:10.3389/fpls.2017.01669.

[56] D. Hanahan, Hallmarks of cancer: new dimensions, Cancer Discov. 12 (2022) 31–46, doi:10.1158/2159-8290.CD-21-1059.

[57] F.T. BosmanWorld Health Organization., International Agency for Research on Cancer, WHO Classification of Tumours of the Digestive System, International Agency for Research on Cancer, Lyon, 2010.

[58] K. Le, J. Wang, T. Zhang, Y. Guo, H. Chang, S. Wang, B. Zhu, Overexpression of mesothelin in Pancreatic Ductal Adenocarcinoma (PDAC), Int. J. Med. Sci. 17 (2020) 422–427, doi:10.7150/ijms.39012.

[59] R. Grutzmann, H. Boriss, O. Ammerpohl, J. Luttges, H. Kalthoff, H.K. Schackert, G. Kloppel, H.D. Saeger, C. Pilarsky, Meta-analysis of microarray data on pancreatic cancer defines a set of commonly dysregulated genes, Oncogene 24 (2005) 5079–5088, doi:10.1038/sj.onc.1208696.

[60] V.O. Oria, P. Lopatta, T. Schmitz, B.T. Preca, A. Nystrom, C. Conrad, J.W. Bartsch, B. Kulemann, J. Hoeppner, J. Maurer, P. Bronsert, O. Schilling, ADAM9 contributes to vascular invasion in pancreatic ductal adenocarcinoma, Mol. Oncol. 13 (2019) 456–479, doi:10.1002/1878-0261.12426.

[61] M. Fahrner, L. Kook, K. Frohlich, M.L. Biniossek, O. Schilling, A systematic evaluation of semispecific peptide search parameter enables identification of previously undescribed N-terminal peptides and conserved proteolytic processing in cancer cell lines, Proteomes 9 (2021), doi:10.3390/proteomes9020026.

[62] J.H. Shahinian, B. Mayer, S. Tholen, K. Brehm, M.L. Biniossek, H. Fullgraf, S. Kiefer, U. Heizmann, C. Heilmann, F. Ruter, M. Grapow, O.T. Reuthebuch, F. Eckstein, F. Beyersdorf, O. Schilling, M. Siepe, Proteomics highlights decrease of matricellular proteins in left ventricular assist device therapydagger, Eur. J. Cardiothorac. Surg 51 (2017) 1063–1071, doi:10.1093/ejcts/ezx023.

[63] T. Fretwurst, I. Tritschler, R. Rothweiler, S. Nahles, B. Altmann, O. Schilling, K. Nelson, Proteomic profiling of human bone from different anatomical sites - a pilot study, Proteomics Clin. Appl. (2022) e2100049, doi:10.1002/prca.202100049.

[64] J.P. Motta, C. Rolland, A. Edir, A.-.C. Florence, D. Sagnat, C. Bonnart, P. Rousset, L. Guiraud, M. Quaranta-Nicaise, E. Mas, D. Bonnet, E.F. Verdu, D.M. McKay, E. Buscail, L. Alric, N. Vergnolle, C. Deraison, Epithelial production of elastase is increased in inflammatory bowel disease and causes mucosal inflammation, Mucosal. Immunol. 14 (2021) 667–678, doi:10.1038/s41385-021-00375-w.

[65] edited by K. Chakraborty, A. Bhattacharyya, Role of proteases in inflammatory lung diseases, in: S. CHAKRABORTI, N.S. DHALLA (Eds.), Proteases in Health and Disease, Springer New York, New York, NY, 2013, pp. 361–385. edited by.

[66] A.I. Nesvizhskii, Proteogenomics: concepts, applications and computational strategies, Nat. Methods 11 (2014) 1114–1125, doi:10.1038/nmeth.3144.

[67] B.N. Mills, H. Qiu, M.G. Drage, C. Chen, J.S. Mathew, J. Garrett-Larsen, J. Ye, T.P. Uccello, J.D. Murphy, B.A. Belt, E.M. Lord, A.W. Katz, D.C. Linehan, S.A. Gerber, Modulation of the human pancreatic ductal adenocarcinoma immune microenvironment by stereotactic body radiotherapy, Clin. Cancer Res. 28 (2022) 150–162, doi:10.1158/1078-0432.CCR-21-2495.

[68] J.G. Tate, S. Bamford, H.C. Jubb, Z. Sondka, D.M. Beare, N. Bindal, H. Boutselakis, C.G. Cole, C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S.C. Jupe, C.Y. Kok, K. Noble, L. Ponting, C.C. Ramshaw, C.E. Rye, H.E. Speedy, R. Stefancsik, S.L. Thompson, S. Wang, S. Ward, P.J. Campbell, J.A. Forbes, COSMIC: the catalogue of somatic mutations in cancer, Nucleic Acids Res. 47 (2018) D941–D947, doi:10.1093/nar/gky1015.

[69] J.T. Topham, E.S. Tsang, J.M. Karasinska, A. Metcalfe, H. Ali, S.E. Kalloger, V. Csizmok, L.M. Williamson, E. Titmuss, K. Nielsen, G.L. Negri, S.E. Spencer Miko, G.H. Jang, R.E. Denroche, H.L. Wong, G.M. O'Kane, R.A. Moore, A.J. Mungall, J.M. Loree, F. Notta, J.M. Wilson, O.F. Bathe, P.A. Tang, R. Goodwin, G.B. Morin, J.J. Knox, S. Gallinger, J. Laskin, M.A. Marra, S.J.M. Jones, D.F. Schaeffer, D.J. Renouf, Integrative analysis of KRAS wildtype metastatic pancreatic ductal adenocarcinoma reveals mutation and expression-based similarities to cholangiocarcinoma, Nat. Commun. 13 (2022) 5941, doi:10.1038/s41467-022-33718-7.

[70] M. Fragiadaki, M.P. Zeidler, Ankyrin repeat and single KH domain 1 (ANKHD1) drives renal cancer cell proliferation *via* binding to and altering a subset of miRNAs, J. Biol. Chem. 293 (2018) 9570–9579, doi:10.1074/jbc.RA117.000975.

[71] B.O. Almeida, J.A. Machado-Neto, Emerging functions for ANKHD1 in cancer-related signaling pathways and cellular processes, BMB Rep. 53 (2020) 413–418.

[72] D. Ren, Y. Sun, D. Li, H. Wu, X. Jin, USP22-mediated deubiquitination of PTEN inhibits pancreatic cancer progression by inducing p21 expression, Mol. Oncol. 16 (2022) 1200–1217, doi:10.1002/1878-0261.13137.