BJC

Pancreatic preneoplastic lesions plasma signatures and biomarkers based on proteome profiling of mouse models

Laetitia Ligat[1],[8], Nathalie Saint-Laurent[1],[8], Aïcha El-Mrani[1], Véronique Gigoux[2], Talal Al Saati[3], Richard Tomasini[4], Jérémy Nigri[4], Sébastien Dejean[5], Frédéric Pont[1], Romain Baer[6], Julie Guillermet-Guibert[6], Pierre Cordelier[7], Frédéric Lopez[1] and Marlène Dufresne*,[7]

[1]Proteomics Group, Centre de Recherches en Cancérologie de Toulouse (CRCT), INSERM UMR1037, Université Toulouse III-Paul Sabatier, Toulouse 31037, France; [2]EA 4552, Université Toulouse III-Paul Sabatier, Toulouse 31432, France; [3]INSERM-US006 ANEXPLO/CREFRE-Histology facility, Toulouse 31024, France; [4]Centre de Recherches en Cancérologie de Marseille (CRCM), Inserm U1068, Institut Paoli-Calmettes, Université Aix-Marseille, CNRS, UMR7258, Marseille 13288, France; [5]Institut de Mathématiques de Toulouse, UMR 5219 CNRS, Université Toulouse III-Paul Sabatier, Toulouse 31062, France; [6]Centre de Recherches en Cancérologie de Toulouse (CRCT), INSERM UMR1037, Université Toulouse III-Paul Sabatier, Equipe Labellisée Ligue Contre le Cancer, Toulouse 31037, France and [7]Centre de Recherches en Cancérologie de Toulouse (CRCT), INSERM UMR1037, Université Toulouse III-Paul Sabatier, Toulouse 31037, France

Background: Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal malignancies with a mortality that is almost identical to incidence. Because early detected PDAC is potentially curable, blood-based biomarkers that could detect currently developing neoplasia would improve patient survival and management. PDAC develops from pancreatic intraepithelial neoplasia (PanIN) lesions, graded from low grade (PanIN1) to high grade (PanIN3). We made the hypothesis that specific proteomic signatures from each precancerous stage exist and are detectable in plasma.

Methods: We explored the peptide profiles of microdissected PanIN cells and of plasma samples corresponding to the different PanIN grade from genetically engineered mouse models of PDAC using capillary electrophoresis coupled to mass spectrometry (CE-MS) and Chip-MS/MS.

Results: We successfully characterised differential peptides profiles from PanIN microdissected cells. We found that plasma from tumor-bearing mice and age-matched controls exhibit discriminative peptide signatures. We also determined plasma peptide signatures corresponding to low- and high-grade precancerous step present in the mice pancreas using the two mass spectrometry technologies. Importantly, we identified biomarkers specific of PanIN3.

Conclusions: We demonstrate that benign and advanced PanIN lesions display distinct plasma peptide patterns. This strongly supports the perspectives of developing a non-invasive screening test for prediction and early detection of PDAC.

Pancreatic ductal adenocarcinoma (PDAC) is a rare disease but this diagnosis sounds like a death sentence for most cases. Incidence equals mortality because of a late diagnosis and the lack of curative treatment for advanced disease. Therefore, one vital question to answer is: who is likely to develop this cancer and who already has it?

*Correspondence: Professor M Dufresne; E-mail: marlene.dufresne@inserm.fr
[8]These authors contributed equally to this work.

This is an essential point because up to now surgical resection remains the only potentially curative treatment for pancreatic cancer patients. Patients with surgically resected localised disease have a 5-year relative survival around 20%. The success rate of this approach could be significantly improved if early forms of pancreatic tumours could be routinely and reliably detected as recently demonstrated (Canto et al, 2012). However, high-risk individuals have significantly more prevalent pancreatic cysts lesions and current imaging devices are still not sensitive enough to detect pancreatic intraepithelial neoplasia (PanIN) that are microscopic premalignant epithelial lesions.

Many studies using different kinds of –omics have been performed in order to propose biomarkers for PDAC. Differentially expressed genes, miRNA or proteins were included in a catalog of potential biomarkers of pancreatic cancer based on their report in published articles and in public repositories (Harsha et al, 2009; Cutts et al, 2011). However, they represent candidate markers rather than validated markers and for most of them lack sensitivity, specificity or reproducibility (Jamieson et al, 2011). The main limitation is that the vast majority of previous published studies used samples from patients that underwent surgical resection and were therefore already diagnosed. The advanced tumor stage restricts the value of these studies for screening purposes. Their results may also reflect consequences rather than predict the disease. Therefore the challenge of discovering markers applicable to early detection of PDAC is still substantial.

Proteomic profiling of biological samples is increasingly used in basic research to obtain more insight into the pathophysiology of diseases, as well as in translational clinical studies to better diagnose and monitor diseases, predict or follow response to therapy. There has been a great interest in applying this technology to pancreatic cancer blood biomarker, as detection of peptides as plasma signatures predictive of PDAC would represent a great advance in this field (Koomen et al, 2005; Tonack et al, 2009a; 2009b; Cecconi et al, 2011; Pan et al, 2011). Several previous findings indicate that plasma proteomic analysis of genetically engineered mouse (GEM) models of cancer provide a useful strategy to identify candidate markers applicable to human cancer (Kuick et al, 2007; Hung et al, 2009; Taguchi et al, 2011). In the case of PDAC, GEM models expressing the driving oncogene KrasG12D in the pancreas develop PanIN lesions and invasive tumours histologically identical to those in humans (Hingorani et al, 2003). Their potential utility for detection of early stage PDAC has been recently published (Faca et al, 2008; Hocker et al, 2013).

Capillary electrophoresis coupled to mass spectrometry (CE-MS) is increasingly employed in proteome analysis with focus on biomarkers discovery to successfully improve diagnosis and treatments of severe pathologies (Kolch et al, 2005; Mischak and Schanstra, 2011). Chip-MS system is a microfluidic chip-based technology designed for nanospray that allows for highly efficient HPLC separation and superior sensitive MS detection of complex proteomic mixtures (Vollmer and van de Goor, 2009). We conducted proteomic studies using several GEM models expressing KrasG12D in the pancreas using both technologies. We characterised differential peptides profiles from PanIN microdissected cells, determined plasma peptide profiles corresponding to benign and advanced PanIN and identified biomarkers specific of each precancerous stage.

MATERIALS AND METHODS

Mouse strains. The LSL-Kras$^{G12D/+}$ and the LSL-p53$^{R172H/+}$ knock-in mouse strains were obtained from the Mouse Models of Human Cancers Consortium Repository (NCI, Frederick, USA).

The Pdx1-Cre mouse strain was from the Douglas A. Melton's laboratory (Cambridge, MA, USA; Gu et al, 2002). The conditional Ink4a/Arf$^{lox/lox}$ mouse strain was from Ronald A. DePinho's laboratory (Aguirre et al, 2003). LSL-Kras$^{G12D/+}$ and Pdx1-Cre strains were bred to generate Pdx1-Cre;LSL-Kras$^{G12D/+}$ mice (named KC) and control littermates as previously described (Hingorani et al, 2003). They were also bred to Ink4a/Arf$^{lox/lox}$ mice and to LSL-p53^{R172H} to produce the Pdx1-Cre;LSL-Kras$^{G12D/+}$; Ink4a/Arf$^{lox/lox}$ and Pdx1-Cre;LSL-Kras$^{G12D/+}$;LSL-p53$^{R172H/+}$ (named KPC), respectively, as well as corresponding controls. All experiments were approved by the animal care committee of animal facility of INSERM-US006 and INSERM-U1068.

Sample collection. For the training plasma sample set, KC and control mice were killed at different time points during 19 months and plasma samples were collected at the time of killing from the orbital sinus and frozen. Plasmas were collected from Pdx1-Cre;LSL-Kras$^{G12D/+}$; Ink4a/Arf$^{lox/lox}$ mice exhibiting pancreatic tumours at 8 weeks of age. For the validation set, plasma were from KC and KPC killed at 6 months. Half of each pancreas was fixed in 10% neutral buffered formalin, embedded in paraffin, serially sectioned (5 μm) and every 10 sections stained with hematoxylin and eosin for histo-pathological analysis. Scoring of PanIN lesions was performed using consensus criteria established at the Penn Workshop (Hruban et al, 2006). Pancreas were graded by a pathologist (TAS) according to the highest grade of PanIN found in half of each pancreas serially sectioned. The other half of each pancreas was frozen in cooled liquid isopentane for laser microdissection.

Laser microdissection. Laser microdissection was performed with an Arcturus ARC 2000 microscope using 10-μm thick pancreas frozen sections. Sections were stained with cresyl violet for the identification and grading of PanIN by a pathologist (TAS) before microdissection. We microdissected 500–4000 cells from PanIN lesions or acini in a maximum of 20–30 min. Cells from several caps were pooled until 4000 cells were collected and processed for CE-MS analysis. In this way we collected 52 samples: 12 PanIN1, 20 PanIN2, 10 PanIN3 from 42 KC and 10 normal pancreatic exocrine cells samples from 10 control mice.

Protein extraction from microdissected cells. Isolated cells attached to the caps were suspended in 40 μl of Rapigest 0.1% buffer (Waters). After 2 min of sonication, solubilized proteins were subjected to trypsin digestion (0.5ng μl^{-1} sample, porcine trypsin; Promega), 16 h at 37 °C. Reaction was stopped with 0.5% trifluoroacetic acid before centrifugation at 13 000 r.p.m. for 5 min. Digested samples were subjected to a 20 kDa cut-off by centrifugation at 2000 g, 45 min, 4 °C in urea 2 M, 100 mM NaCl, 0.0125% NH$_4$OH, 0.01% SDS buffer on Vivaspin 2 column (Sartorius) and final desalting on NAP5 column (GE Healthcare).

Plasma sample preparation. To decrease the amount of the high-abundance proteins and concentrate low-abundance proteins, we used the ProteoMiner protein enrichment kit (Sennels et al, 2007), according to Bio-Rad's instructions. Briefly, 20 μl of plasma were loaded on 20 μl of beads. For CE-MS analysis, they were eluted in 40 μl of buffer with 8M urea, 2% CHAPS and 5% acetic acid. Adjustment of pH to 8 was performed before digestion of proteins with trypsin (0.5 ng μl^{-1} sample). Digestion was stopped and peptides were desalted as described for microdissected cells. For Chip-MS, proteins were eluted from beads with 2× Laemmli Buffer (Tris 80 mM, pH 6.8, SDS 3%, glycerol 10%, DTT 80 mM). The samples were then reduced, alkylated before concentration and desalting in a 7.5% SDS-PAGE. Proteins were stained by Coomassie Blue. Each lane was cut and each gel piece was washed several times in acetonitrile 100%, ammonium bicarbonate 100 mM and dried in vacuo. Gel pieces were rehydrated with 20 ng μl^{-1} trypsin prepared in ammonium bicarbonate 100 mM,

and submitted to in gel-digestion overnight at 37 °C. Peptides were then extracted and purified from gel then subjected to mass spectrometry analysis.

CE-MS technology. CE was performed on a P/ACE ProteomLAb PA800 (Beckman Coulter, Fullerton, CA, USA). Fused silica capillaries with a total length of 90 cm with 50 μm inner diameter were used for separation. Twenty percent acetonitrile, 250 mM formic acid was used as background electrolyte. Sample injection was performed for 99 s under pressure (1–6 psi); the sample plug injected corresponding to 50–300 nl volume and separation voltage was set at 25 kV for 60 min. The temperature of capillary was set to 35 °C during the entire run. A washing step with successive $H_2O/NH_4OH/H_2O$/background electrolyte injections was included between runs.

MS was performed using a microTOF II (Bruker Daltonics, Bremen, Germany) apparatus in a positive electrospray mode with an ESI-TOF sprayer-kit from Agilent Technologies (Palo Alto, CA, USA). The ESI sprayer was grounded and the ionspray interface potential was set to −4500 V. CE-MS coupling was realized by a coaxial sheath liquid interface (Agilent Technologies) with 30% v/v isopropanol and 1% v/v formic acid in HPLC-grade water as sheath liquid and a rate of 2 μl min⁻¹. Data acquisition and MS methods were automatically controlled by CE apparatus via contact-closure relays. Spectra were accumulated for 3 s each, over a mass range from 350 to 3000 m/z (mass/charge ratio).

Chip-MS/MS technology. The peptides mixtures were analysed by nanoHPLC-chip-MS/MS with a system consisting of a nano-pump, a capillary-pump (G1376A and G2226, Agilent) with two four-channel micro-vacuum degassers (G1379B, Agilent), a microfluidic chip cube (G4240-64000, Agilent) interfaced to an Amazon ETD mass spectrometer (Bruker Daltonics, Germany). A microfluidic reversed-phase HPLC chip (Zorbax 300SB-C18, 5 μm particle size, 75 μm internal diameter and 150 mm length) was used for peptide separation. Peptides were eluted using the following gradient of solvent A (0.1% formic acid) and B (10% acetonitrile, 0.1% formic acid) at 300 nl min⁻¹ flow rate: 0–1 min, 3% B; 1–3 min, 3–7% B; 3–44 min, 7–45% B; 44–46 min, 45–95% B; 46–50 min, 95% B; 50–51 min, 95–3% B.

For the profiling experiment, the mass spectrometer was operated in full scan MS. Scans MS were acquired on the 300–1500 m/z range in the enhanced-resolution mode. The maximum accumulation time was set to 75 ms. The voltage applied to the chip was 1800–1900 V, the dry gas flow was 4.0 l min⁻¹ at a temperature of 180 °C. A set of target peptides was obtained after data analysis and an inclusion list was built to fragment preferentially relevant peptides.

For peptide fragmentation, the Amazon was operated in data-dependent acquisition mode with the trap control software. Scans MS were acquired on the 300–1500 m/z range in the enhanced-resolution mode. MS/MS spectra were acquired using the scheduled precursor list but the mass spectrometer was allowed to acquire spectra in empty time intervals. The most intense ions per scan were selected for collision induced fragmentation and dynamic exclusion was employed within 12 s to prevent repetitive selection of the same peptide. The fragmentation amplitude was set to 0.6 V.

CE-MS and Chip-MS data analysis. The Bruker data files (.d folder) generated with the CE-MS or Chip-MS technology were loaded to Progenesis LC-MS version 4.0 (Nonlinear Dynamics, UK). Automatic alignment was performed using an algorithm that corrects retention time shifts. After checking manually all runs, the peak picking was performed to detect features (i.e., ions detected on the mass spectrometer) using automatic parameters for sensitivity and retention time window. To lower noise detection, sensitivity was set to its lower level.

Only features with <6 charges were kept for further analysis. Normalisation to all features was also processed. An experimental design was set up to compare the different groups. In order to select features of interest, statistical filters were set and only features matching all filters were kept. Filters used were: P-value <5% (Student's t-test), max fold change >1.5, power >80%. A .csv file was then exported from Progenesis and loaded to R 2.13.2. Descriptive statistics and principal component analysis (PCA) were performed using mixOmics R Package (Le Cao et al, 2009).

Identification of proteins. As potential biomarkers, relevant features selected by Progenesis were exported in .mgf files and the corresponding proteins were identified using the MASCOT software (http://www.matrixscience.com/) and SwissProt database (http://web.expasy.org/docs/swiss-prot_guideline.html). Mascot files were then imported in Progenesis software to select the most relevant identified proteins. Protein hits were validated if they satisfied one of the following criteria: identification with at least one top ranking peptide with a Mascot score of >43 or at least two top ranking peptides each with a Mascot score of >23.

Statistical analysis. Descriptive statistics are presented as PCA that allows exploratory data analysis combining samples and proteins. A Tukey test was used to analyse coefficient of variation (CV) in order to compare variability between groups and variability within groups. Statistical analysis were done with FactoMineR package for R software (http://cran.r-project.org/web/packages/FactoMineR/).

Enzyme-linked immunosorbent assay. Plasma levels of ITIH3 were measured using commercially available mouse kit obtained from Mybiosource.Com. Plasma was diluted 1 : 2000. Enzyme-linked immunosorbent assay (ELISA) was performed according to the manufacturers' protocols and samples were assayed in duplicate. Statistical significance (P<0.05) was determined by Mann-Whitney's test.

RESULTS

Differential peptide profiles from microdissected PanIN. PDAC develops through increasing grades of non-invasive PanIN lesions. PanIN are divided into three grades based on cytological and architectural abnormalities: benign PanIN1 lesions show no nuclear abnormalities and develop cytoplasmic mucin, PanIN2 demonstrate moderate dysplasia with stratification and nuclear crowding, PanIN3 have lost their polarity and show marked atypia similar to in situ carcinoma, they are the immediate precursor of metastatic PDAC (Figure 1A). As phenotypical changes were observed during PanIN development in parallel to molecular alterations, we wanted to compare peptide signatures from cells from different PanIN grades to test the accuracy of the peptide profiling approach. For this purpose we performed laser-based microdissection to get cells free of contaminating and unwanted tissue components and determine differentially expressed proteins in the different PanINs using control and KC (Figures 1B and C). We chose acinar cells as matched normal cells because of the scarcity of normal ductal epithelial cells. They were microdissected from control littermates pancreas.

Proteins were extracted from 4000 control acinar cells or PanIN, separated and analysed with the CE-MS technology. CE-MS raw files were pre-processed and compared using Progenesis LC-MS Software. In this software package, all peaks in the raw files are aligned according to their retention time by a graphical detection algorithm. This algorithm detects the peptides peaks in a gel-view representation of the mass spectrometry data and matches corresponding peaks, termed as features, between samples.

Each feature corresponds to a peptide characterised by its migration time, m/z and intensity. Then PCA is performed in order to reveal whether main sources of variability in the data are owing to the groups of samples.

To validate the overall CE-MS workflow, we first compared the biological variability between groups and within groups. As shown in Figure 1D, biological variability between groups (70%) was higher than the variability within groups (49–57%). Moreover Tukey test on CV (defined as the ratio of the standard deviation to the mean) shows that the difference observed between each CV is

significant (adjusted P-value < 5%), thus validating that statistical analysis of differences between groups was strictly related to biological parameters. PCA was then used to analyse peptide profiles, visualise each data set and detect eventual outliers. After statistical filters (P-value < 5%, max fold change > 1.5, power > 80% and q-value < 0.0001), a total of 1556 selected features (over 15 381) were retained for further analysis. Individual factor map shows that controls, PanIN1 and PanIN2/PanIN3 are sequentially separated in the first dimension (Dim1 on the horizontal axis) from left to right (Figure 1E). This result is slightly blurred by the specific position of sample P3.52 on the left side of the plot.

These results argue in favour of differences in peptide profiles between cells from normal tissue and different PanIN grades in agreement with phenotype and molecular differences between these cells which could induce different pathophysiological responses detectable in plasma.

Differential plasma peptide profiles from mice with PDAC. We first subjected plasma proteins from tumor-bearing mice and age-matched littermate controls to CE-MS profiling. Plasma was collected from one well-characterised model of PDAC, the Pdx1-Cre;LSL-Kras$^{G12D/+}$;Ink4a/Arf$^{lox/lox}$ mice and corresponding control littermates at 8 weeks of age. Although we analysed plasma samples from several mice, all Pdx1-Cre;LSL-Kras$^{G12D/+}$;Ink4a/Arf$^{lox/lox}$ mice were housed in the same facility and had the same genetic background. Therefore, to better reflect the clinical situation where individuals with different genetic background and environmental conditions would be screened, we also analysed plasma samples from KC that developed tumours after 12 months of age. These KC and corresponding control mice were housed in a different facility.

We validated that statistical analysis of differences between groups was strictly related to biological parameters (Figure 2A) and representative score plots illustrate classification of samples in two separated groups (control and tumours) emphasising the hypothesis of peptide profiles specific of each group (Figure 2B). These results overcome differences in genetic background and clearly argue in favour of potential discriminant plasmatic signatures from PDAC-bearing mice compared with control mice.

Differential plasma peptide profiles from mice with PanIN. Next, we made the hypothesis of proteomic signatures that are specific of each precancerous stage and are detectable in plasma.
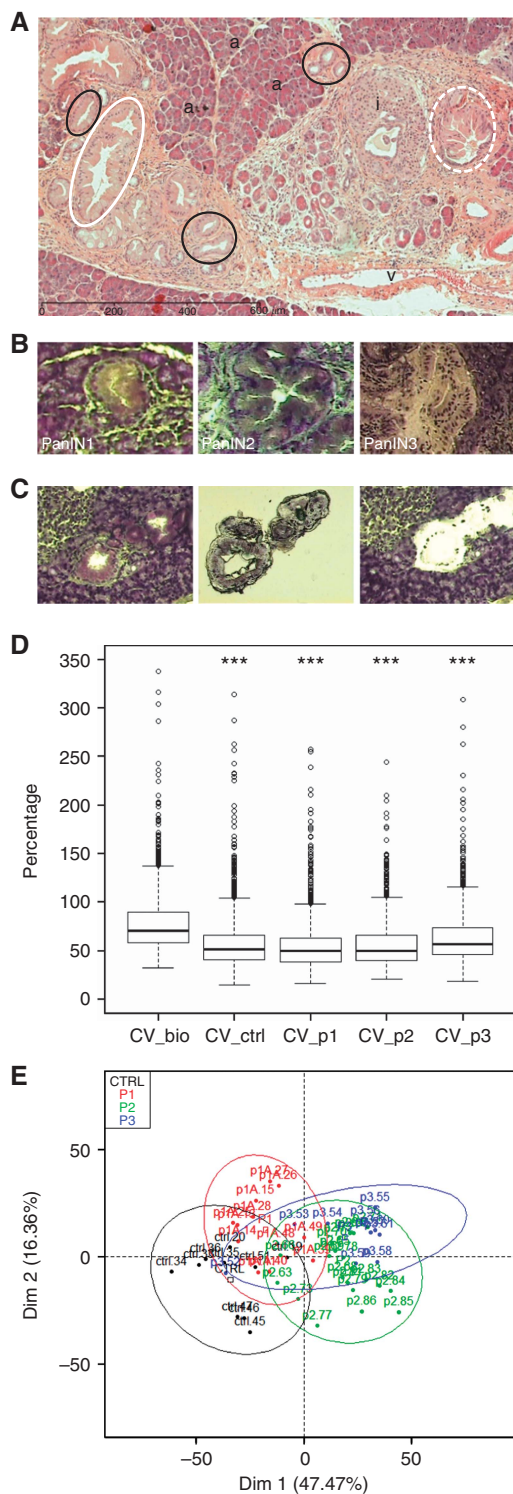

Figure 1. Histology, laser-microdissection and CE-MS of PanIN. (**A**) Representative PanIN1 (black line), PanIN2 (white line) and PanIN3 (white dashed line) on a hematoxylin and eosin stained KC mice pancreas section are shown; 'a' indicates acinar cells; 'i' indicates islets; 'v' indicates vessels. (**B**) Representative PanIN1 (left), PanIN2 (middle) and PanIN3 on a cresyl violet stained frozen section of murine pancreas. (**C**) Representative example of PanIN1 before and after laser microdissection. The area of interest is pulsed with laser (left) and captured cells are retrieved on a cap (middle), remnant tissue remains on the slide (right). (**D**) Validation of CE-MS workflow. Box plots representative of intra-groups and inter-groups CV (in percentage) from control cells (ctrl) and PanIN1 (p1), PanIN2 (p2), PanIN3 (p3) lesions in comparison with biological CV (CV_bio) of all samples taken together. ***P < 0.001 adjusted from Tukey's test, in comparison with CV_bio. (**E**) Score plots from results of PCA of peptide profiles from micro-dissected control cells (black, n = 10) and PanIN1 (red, n = 12), PanIN2 (green, n = 20), PanIN3 (blue, n = 10) lesions according to the first two principal components (Dim 1: 44.47%; Dim 2: 16.36%). Colour code is used to distinguish samples in accordance to their group memberships. Confidence ellipses at 95% around each samples from each group are also represented.
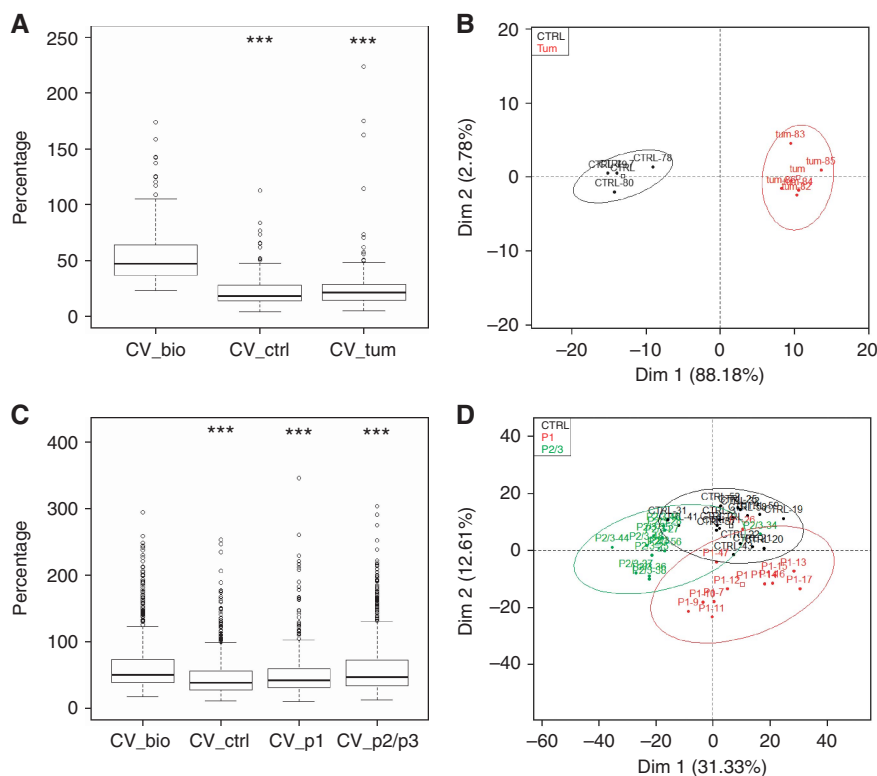
Figure 2. PCA analysis from plasma samples submitted to CE-MS analysis. (**A**) Box plots representative of intra-groups and inter-groups CV (in percentage) from control mice (ctrl) and PDAC-bearing mice plasmas in comparison with biological CV (CV_bio) of all samples taken together. ***$P<0.001$ adjusted from Tukey's test, in comparison with CV_bio. (**B**) Score plots from results of PCA of peptide profiles from control mice (black, $n=4$) and PDAC-bearing mice (red, $n=5$) plasmas according to the first two principal components (Dim 1: 88.18%; Dim 2: 2.78%). Confidence ellipses at 95% around each samples from each group are also represented. Colour code is used to distinguish samples in accordance to their group memberships. (**C**) Box plots representative of intra-groups and inter-groups CV (in percentage) from control mice (ctrl) and PanIN-bearing mice plasmas (p1, p2/3) in comparison with biological CV (CV_bio) of all samples taken together. ***$P<0.001$ adjusted from Tukey's test, in comparison with CV_bio. (**D**) Score plots from results of PCA of peptide profiles from control mice (black, $n=16$), PanIN1-bearing mice (red, $n=12$) and PanIN2/3-bearing mice (green, $n=12$) plasmas according to the first two principal components (Dim 1: 31.33%; Dim 2: 12.61%). Confidence ellipses at 95% around each samples from each group are also represented. Colour code is used to distinguish samples in accordance to their group memberships.

We tested this hypothesis by determining plasma peptide profiles of individuals for which we have performed the histo-pathological analysis and scoring of PanIN lesions of their pancreas. Proteins from 40 plasma samples were separated and analysed with CE-MS. Over these plasma samples, 16 were from control mice, 12 from KC with benign (PanIN1) and 12 from KC with advanced (PanIN2/3) precursor lesions as diagnosed by histopathology.

To validate the overall CE-MS workflow, we first compared the biological variability between groups and intra-groups. Biological variability between groups (51%) was higher than the intra-group variability (39–47%; Figure 2C). Moreover Tukey test on CV shows that the difference observed between each CV is significant (adjusted P-value < 5%), thus validating that statistical analysis of differences between groups was strictly related to biological parameters. After statistical filters (P-value < 5%, max fold change > 1.5, power > 80% and q-value < 0.05), a total of 902 selected features were retained among the 22 277 for further analysis. Individual factor map illustrates group separation and shows a horizontal axis (Dim1 on the horizontal axis) that separates distinctly advanced PanIN2/3 lesions from controls and benign PanIN1 lesions (Figure 2D). This result is slightly blurred by the specific position of sample P2/3.34 on the right side of the plot.

We then used a second proteomic approach to confirm differential plasma peptide profiling and identify plasma proteins specific of each pancreatic precancerous state PanIN1, PanIN2 and

PanIN3. Comparative protein expression differences were measured by Chip-MS/MS. The inter-groups biological variability was higher (53%) than the intra-group variability (40–44%) and Tukey test on CV shows that the difference observed between each CV was significant (adjusted P-value < 0.005%; Figure 3A). After statistical filters (P-value < 5%, max fold change > 1.5, power > 80% and q-value < 0.05), a total of 6401 selected features among the 290 075 were retained for classification. Individual factor map illustrates group separation and shows a horizontal axis (Dim1) that separates distinctly PanIN3 from controls, PanIN1, PanIN2 and PanIN3 and a vertical axis that separates distinctly controls and PanIN1 from PanIN2 (Figure 3B).

Identification of PanIN biomarkers. Fourteen proteins with a significant Mascot score were identified. Table 1 shows that proteins identified are differentially expressed in plasma from mice with PanIN2 or PanIN3 compared with plasma from control mice. Inter-alpha-trypsin inhibitor, heavy chain 4 (ITIH4), inter-alpha-trypsin inhibitor, heavy chain H3 (ITIH3), Ig alpha chain C region (IGHA) and complement C3 (CO3) discriminate mice with PanIN3 from control mice. Fibronectin and alpha-2-macroglobulin discriminate mice with PanIN2 from mice with PanIN3.

A separate mice cohort composed of KC, KPC and control mice aged of 6 months was used for validation. Individual factor map shows a horizontal axis (Dim1) that separates distinctly two groups (Figure 3C). Importantly, again we identified ITIH3 as
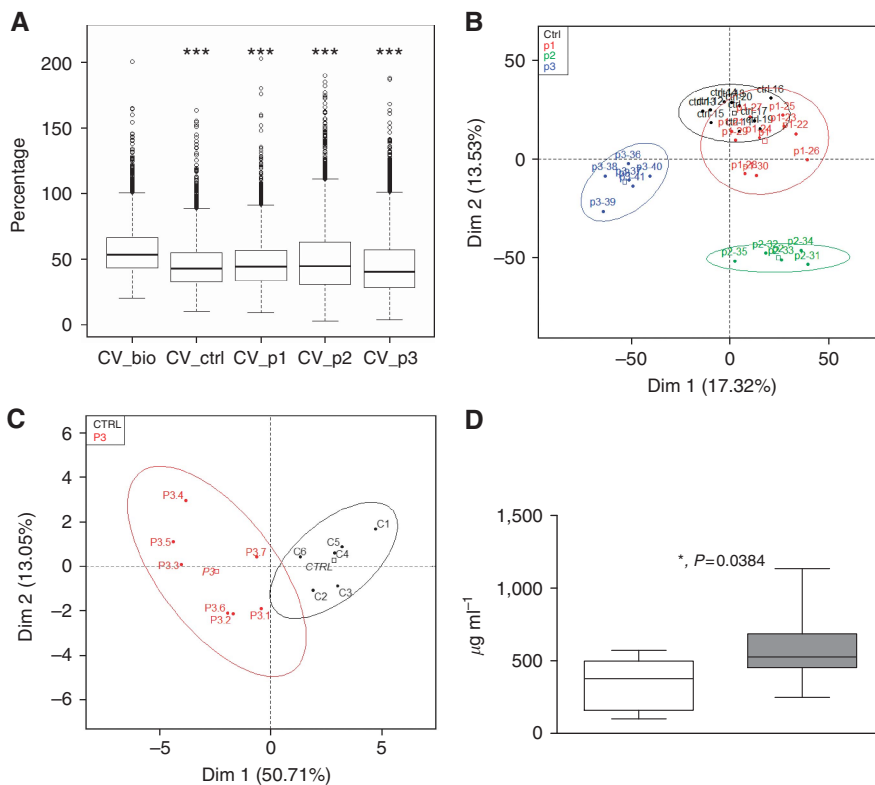
Figure 3. Analysis of plasma samples submitted to Chip-MS analysis. (**A**) Box plots representative of intra-groups and inter-groups CV (in percentage) from control mice (ctrl) and PanIN-bearing mice plasmas (p1, p2, p3) in comparison with biological CV (CV_bio) of all samples taken together.***P<0.001 adjusted from Tukey's test, in comparison with CV_bio. (**B**) Score plots from results of PCA of peptide profiles from control mice (black, n = 10), PanIN1-bearing mice (red, n = 10), PanIN2 (green, n = 5) and PanIN3-bearing mice (blue, n = 12) plasmas according to the first two principal components (Dim 1: 17.32%; Dim 2: 13.53%). Confidence ellipses at 95% around each samples from each group are also represented. Colour code is used to distinguish samples in accordance to their group memberships. (**C**) Score plots from results of PCA of peptide profiles from control mice (black, n = 6) and PanIN3-bearing mice (red, n = 7) plasmas according to the first two principal components (Dim 1: 50.71%; Dim 2: 13.05%). Confidence ellipses at 95% around each samples from each group are also represented. (**D**) Quantitative analysis by Elisa of ITIH3 in the plasma of control mice (white box, n = 5) and of PanIN3-bearing KC and KPC mice (grey box, n = 13), *P<0.05.

Table 1. List of proteins identified by Chip-MS/MS in the plasma of Ctrl and of mice having P1, P2 and P3 in their pancreas

Accession	Peptides	Score	Anova (P)	Fold	High	Low	Description
ITIH4_MOUSE	13	763	0.00040	1.7	P3	Ctrl	Inter-alpha-trypsin inhibitor, heavy chain
CO3_MOUSE	21	808	0.00055	1.7	Ctrl	P3	Complement C3
IGHA_MOUSE	10	478	0.00197	3.1	P3	Ctrl	Ig alpha chain C region
FINC_MOUSE	19	735	0.00267	1.6	P3	P2	Fibronectin
APOA4_MOUSE	4	167	0.00298	2.1	P1	P3	Apolipoprotein A-IV
ITIH1_MOUSE	2	75	0.00303	1.8	P3	P2	Inter-alpha-trypsin inhibitor heavy chain H1
HVM16_MOUSE	1	77	0.00337	1.6	P3	P1	Ig heavy chain V region MOPC 21 (fragment)
ITIH3_MOUSE	11	413	0.00515	1.9	P3	Ctrl	Inter-alpha-trypsin inhibitor heavy chain H3
IGG2B_MOUSE	4	98	0.00519	2.1	P2	Ctrl	Ig gamma-2B chain C region
FETUA_MOUSE	2	171	0.00544	3.4	P2	P3	Alpha-2-HS-glycoprotein
LUM_MOUSE	2	55	0.01094	1.6	Ctrl	P3	Lumican
IGHM_MOUSE	6	278	0.01632	1.9	Ctrl	P2	Ig mu chain C region secreted form
THRB_MOUSE	3	165	0.01663	1.6	P3	P2	Prothrombin
A2M_MOUSE	9	341	0.02465	1.9	P2	P3	Alpha-2-macroglobulin

Abbreviations: Chip-MS = Chip-mass spectrometry; Ctrl = control mice; P1 = PanIN1; P2 = PanIN2; P3 = PanIN3.

differentially expressed in plasma of these two groups (23 peptides, score: 739, P = 0.0378, 1.9 fold). Pathological analysis validated the presence of PanIN3 in the pancreas of KC and KPC (Table 3).

To verify the differential abundance of plasma proteins from PanIN3-bearing mice, we measured the levels of ITIH3 by ELISA in plasma from the same individual mice used for the Chip-MS/MS discovery and validation sets (Table 2 and Table 3). As shown in Figure 3D, plasma levels of these two proteins allowed to distinguish PanIN3-bearing mice from mice without PanIN lesions with statistical significance.

Table 2. Quantification of PanIN lesions in the pancreas of the KC training set mice (Chip-MS/MS plasma set 1)

			PanIN1		PanIN2		PanIN3	
Training set (n = 21)	Number	Age (months)	n	Area (% total)	n	Area (% total)	n	Area (% total)
PanIN1 (n = 10)	P1–21	6	3	0.044	0	—	—	—
	P1–22	1	5	0.075	0	—	—	—
	P1–23	1	6	0.044	0	—	—	—
	P1–24	15	9	0.234	0	—	—	—
	P1–25	2	13	0.322	0	—	—	—
	P1–26	2	33	0.499	0	—	—	—
	P1–27	4	13	0.208	0	—	—	—
	P1–28	9	80	1.737	0	—	—	—
	P1–29	6	18	0.144	0	—	—	—
	P1–30	6	11	0.283	0	—	—	—
PanIN2 (n = 5)	P2–31	9	24	0.466	4	0.603	0	—
	P2–32	9	2	0.072	2	0.138	0	—
	P2–33	9	10	0.370	2	0.230	0	—
	P2–34	9	30	0.273	3	0.181	0	—
	P2–35	11	29	0.776	3	0.356	0	—
PanIN3 (n = 6)	P3–36[a]	11	47	1.352	2	0.543	8	1.981
	P3–37[a]	11	31	0.651	3	0.155	5	0.625
	P3–38[a]	15	15	0.374	2	0.467	1	0.182
	P3–39[a]	15	15	0.492	2	0.613	2	0.417
	P3–40[a]	18	56	1.811	17	0.912	7	1.315
	P3–41[a]	18	144	4.964	14	1.130	13	2.724

Abbreviations: Chip-MS = Chip-mass spectrometry; Ctrl = control mice; KC = Pdx1-Cre;LSL-Kras$^{G12D/+}$ mice.
[a]Indicate plasma used for validation by ELISA.

Table 3. Grade of detected PanIN lesions in the pancreas of the mice used for Chip-MS/MS plasma set 2

Number	Strain	Age (months)	PanIN3
407[a]	KC	6	Yes
371[a]	KPC	6	Yes
384[a]	KPC	6	Yes
415[a]	KPC	6	Yes
359[a]	KC	6	Yes
363[a]	KC	6	Yes
B790[a]	KC	6	Yes

Abbreviations: Chip-MS = Chip-mass spectrometry; KC = Pdx1-Cre;LSL-Kras$^{G12D/+}$ mice; KPC = Pdx1-Cre;LSL-Kras$^{G12D/+}$;LSL-p53$^{R172H/+}$.
[a]Indicate plasma used for validation by ELISA.

DISCUSSION

Using a well-established GEM recapitulating the molecular and morphological stages of human PDAC development we first identified peptide signatures that can discriminate between morphologically normal pancreatic tissue and each of the PanIN: PanIN1, PanIN2 and PanIN3. Until now, most of tissue-based proteomics studies did not differentially discriminate PanIN according to their grade or identified biomarkers from pancreatic cancer (Cui *et al*, 2009; Pan *et al*, 2009; Turtoi *et al*, 2011; Wang *et al*, 2011). To our knowledge only one other study used microdissected cells of different PanIN grade to investigate the pancreatic proteome by 2D-DIGE and reported differentially regulated proteins involved in pancreatic tumor progression

(Sitek *et al*, 2009). Therefore, this first set of results establishes a solid basis for developing a non-invasive screening method for early diagnosis using proteomics.

Many studies reported the identification of circulating proteins as markers of pancreatic cancer. For example, 18 peaks were identified to be differentially expressed between pancreatic cancer and healthy volunteers (Xue *et al*, 2010). Several proteins were characterised and evaluated alone or in combination with CA19-9 for their discriminating power for pancreatic cancer diagnosis (Fiedler *et al*, 2009; Takano *et al*, 2010; Fakelman *et al*, 2010; Matsubara *et al*, 2011; Tonack *et al*, 2013). However none of them could detect early stages of cancer development as these markers have been investigated on patients with established PDAC. To circumvent this issue, we explored differences in plasmatic protein profiles in validated GEM models. Profiling of blood samples is usually quite challenging because of the presence of abundant proteins such as albumin. However, we overcame this major obstacle to blood biomarker discovery as we obtained differential peptide profiles in plasma from mice. A further noticeable result is that we were able to detect precancerous pancreatic lesions without sample pooling thus providing an effective methodology for the screening of individuals.

Importantly we discovered candidate blood biomarkers that discriminate high-grade PanIN lesions showing that the presence of microscopic and non-invasive precancerous lesions may be detected in blood circulation. Of note, plasma was obtained from mice before pancreatic cancer has developed. This gives to the proteins that we identified the value of true precancerous biomarkers whose expression is not influenced by the presence of a tumor. Four plasma proteins were differentially present in the plasma of mice with PanIN3 and of control mice suggesting potential early detection of PDAC by combining these proteins.

Among them, two proteins, ITIH3 and ITIH4, belong to the family of inter-alpha-trypsin inhibitor (ITIH). They were previously identified in plasma of mice with PanIN and PDAC, but not linked specifically to the presence of PanIN3 (Faca *et al*, 2008). ITIH4 has been associated with several disease conditions and proposed to be used as potential biomarker of breast cancer (Song *et al*, 2006). Importantly, we validated ITIH3 using both mass spectrometry and ELISA in mice with different ages, genotypes and strain backgrounds. It is also worthy of note that ITIH3 has been identified as a novel potential biomarker for PDAC detection in a recent proteomics determination (Tonack *et al*, 2013). Our data lead us to believe that this protein might be present at high amount in the blood of individuals with non-invasive cancer. ITIH3 screening could help to characterize the critical step before the fatal disease and the presence of PanIN3 that is currently impossible to do. This is strongly supported by the identification of ITIH3 as a biomarker of early stage of gastric cancer patients and of the formation of neoplastic polyps in the Apc Min/ + mouse (Chong *et al*, 2010; Ivancic *et al*, 2013). In contrast with our results, CO3 was found to be elevated in plasma of mice with PanIN (Faca *et al*, 2008). Reasons for this discrepancy may rely on methodological differences as our study detected candidate biomarkers in individuals and not in pooled samples. Indeed, CO3 is an acute-phase protein and its presence in the blood of pancreatic cancer patients was linked to the inflammatory component of PDAC (Hanas *et al*, 2008). Thus one cannot exclude that outliers of the pooled PanIN samples could mask other samples.

Taken together, our data could argue in favour of strategies for discriminating 'low-risk' or 'high-risk' patients for PDAC development as advanced PanIN3 lesions probably represent a decisive step to neoplastic lesion (Crnogorac-Jurcevic *et al*, 2013). Beyond PDAC prediction, such approaches could also be relevant for follow-up of 'high-risk' patients and monitoring of future tailored therapies of PDAC.

Because of biological heterogeneity between species, it is possible that biomarkers from mice will not necessarily be of clinical utility in humans even if strong concordance between mouse and human cancer has already been observed (Kuick *et al*, 2007; Faca *et al*, 2008; Hung *et al*, 2009; Taguchi *et al*, 2011). Nonetheless, one can expect that validation of pre-analytical sample management and technological performance (sensitivity, robustness, specificity and multiplex characteristics) could be extrapolated to studies in humans. For now, accessibility to human samples, especially at early stages of pancreatic cancer development remains the main drawback for prospective validation studies. However, future studies conducted on samples from patients associated with known pancreatic cancer risk and on pre-diagnosed samples from large cohorts should allow the development of early detection tools for this lethal pathology.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Aguirre AJ, Bardeesy N, Sinha M, Lopez L, Tuveson DA, Horner J, Redston MS, DePinho RA (2003) Activated Kras and Ink4a/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. *Genes Dev* **17**: 3112–3126.

Canto MI, Hruban RH, Fishman EK, Kamel IR, Schulick R, Zhang Z, Topazian M, Takahashi N, Fletcher J, Petersen G, Klein AP, Axilbund J, Griffin C, Syngal S, Saltzman JR, Mortele KJ, Lee J, Tamm E, Vikram R, Bhosale P, Margolis D, Farrell J, Goggins M (2012) Frequent detection of pancreatic lesions in asymptomatic high-risk individuals. *Gastroenterology* **142**: 796–804.

Cecconi D, Palmieri M, Donadelli M (2011) Proteomics in pancreatic cancer research. *Proteomics* **11**: 816–828.

Chong PK, Lee H, Zhou J, Liu SC, Loh MC, Wang TT, Chan SP, Smoot DT, Ashktorab H, So JB, Lim KH, Yeoh KG, Lim YP (2010) ITIH3 is a potential biomarker for early detection of gastric cancer. *J Proteome Res* **9**: 3671–3679.

Crnogorac-Jurcevic T, Chelala C, Barry S, Harada T, Bhakta V, Lattimore S, Jurcevic S, Bronner M, Lemoine NR, Brentnall TA (2013) Molecular analysis of precursor lesions in familial pancreatic cancer. *PLoS One* **8**: e54830.

Cui Y, Zhang D, Jia Q, Li T, Zhang W, Han J (2009) Proteomic and tissue array profiling identifies elevated hypoxia-regulated proteins in pancreatic ductal adenocarcinoma. *Cancer Invest* **27**: 747–755.

Cutts RJ, Gadaleta E, Hahn SA, Crnogorac-Jurcevic T, Lemoine NR, Chelala C (2011) The Pancreatic Expression Database: 2011 update. *Nucleic Acids Res* **39**: D1023–D1028.

Faca VM, Song KS, Wang H, Zhang Q, Krasnoselsky AL, Newcomb LF, Plentz RR, Gurumurthy S, Redston MS, Pitteri SJ, Pereira-Faca SR, Ireton RC, Katayama H, Glukhova V, Phanstiel D, Brenner DE, Anderson MA, Misek D, Scholler N, Urban ND, Barnett MJ, Edelstein C, Goodman GE, Thornquist MD, McIntosh MW, DePinho RA, Bardeesy N, Hanash SM (2008) A mouse to human search for plasma proteome changes associated with pancreatic tumor development. *PLoS Med* **5**: e123.

Fakelman F, Felix K, Buchler MW, Werner J (2010) New pre-analytical approach for the deep proteome analysis of sera from pancreatitis and pancreas cancer patients. *Arch Physiol Biochem* **116**: 208–217.

Fiedler GM, Leichtle AB, Kase J, Baumann S, Ceglarek U, Felix K, Conrad T, Witzigmann H, Weimann A, Schutte C, Hauss J, Buchler M, Thiery J (2009) Serum peptidome profiling revealed platelet factor 4 as a potential discriminating peptide associated with pancreatic cancer. *Clin Cancer Res* **15**: 3812–3819.

Gu G, Dubauskaite J, Melton DA (2002) Direct evidence for the pancreatic lineage: NGN3 + cells are islet progenitors and are distinct from duct progenitors. *Development* **129**: 2447–2457.

Hanas JS, Hocker JR, Cheung JY, Larabee JL, Lerner MR, Lightfoot SA, Morgan DL, Denson KD, Prejeant KC, Gusev Y, Smith BJ, Hanas RJ, Postier RG, Brackett DJ (2008) Biomarker identification in human pancreatic cancer sera. *Pancreas* **36**: 61–69.

Harsha HC, Kandasamy K, Ranganathan P, Rani S, Ramabadran S, Gollapudi S, Balakrishnan L, Dwivedi SB, Telikicherla D, Selvan LD, Goel R, Mathivanan S, Marimuthu A, Kashyap M, Vizza RF, Mayer RJ, Decaprio JA, Srivastava S, Hanash SM, Hruban RH, Pandey A (2009) A compendium of potential biomarkers of pancreatic cancer. *PLoS Med* **6**: e1000046.

Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, Jacobetz MA, Ross S, Conrads TP, Veenstra TD, Hitt BA, Kawaguchi Y, Johann D, Liotta LA, Crawford HC, Putt ME, Jacks T, Wright CV, Hruban RH, Lowy AM, Tuveson DA (2003) Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell* **4**: 437–450.

Hocker JR, Mohammed A, Aston CE, Brewer M, Lightfoot SA, Rao CV, Hanas JS (2013) Mass profiling of serum to distinguish mice with pancreatic cancer induced by a transgenic Kras mutation. *Int J Cancer* **133**: 2662–2671.

Hruban RH, Adsay NV, Albores-Saavedra J, Anver MR, Biankin AV, Boivin GP, Furth EE, Furukawa T, Klein A, Klimstra DS, Kloppel G, Lauwers GY, Longnecker DS, Luttges J, Maitra A, Offerhaus GJ, Perez-Gallego L, Redston M, Tuveson DA (2006) Pathology of genetically engineered mouse models of pancreatic exocrine cancer: consensus report and recommendations. *Cancer Res* **66**: 95–106.

Hung KE, Faca V, Song K, Sarracino DA, Richard LG, Krastins B, Forrester S, Porter A, Kunin A, Mahmood U, Haab BB, Hanash SM, Kucherlapati R (2009) Comprehensive proteome analysis of an Apc mouse model uncovers proteins associated with intestinal tumorigenesis. *Cancer Prev Res* **2**: 224–233.

Ivancic MM, Huttlin EL, Chen X, Pleiman JK, Irving AA, Hegeman AD, Dove WF, Sussman MR (2013) Candidate serum biomarkers for early intestinal cancer using 15 N metabolic labeling and quantitative proteomics in the ApcMin/ + mouse. *J Proteome Res* **12**: 4152–4166.

Jamieson NB, Carter CR, McKay CJ, Oien KA (2011) Tissue biomarkers for prognosis in pancreatic ductal adenocarcinoma: a systematic review and meta-analysis. *Clin Cancer Res* **17**: 3316–3331.

Kolch W, Neususs C, Pelzing M, Mischak H (2005) Capillary electrophoresis-mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. *Mass Spectrom Rev* **24**: 959–977.

Koomen JM, Shih LN, Coombes KR, Li D, Xiao LC, Fidler IJ, Abbruzzese JL, Kobayashi R (2005) Plasma protein profiling for diagnosis of pancreatic cancer reveals the presence of host response proteins. *Clin Cancer Res* **11**: 1110–1118.

Kuick R, Misek DE, Monsma DJ, Webb CP, Wang H, Peterson KJ, Pisano M, Omenn GS, Hanash SM (2007) Discovery of cancer biomarkers through the use of mouse models. *Cancer Lett* **249**: 40–48.

Le Cao KA, Gonzalez I, Dejean S (2009) integrOmics: an R package to unravel relationships between two omics datasets. *Bioinformatics* **25**: 2855–2856.

Matsubara J, Honda K, Ono M, Tanaka Y, Kobayashi M, Jung G, Yanagisawa K, Sakuma T, Nakamori S, Sata N, Nagai H, Ioka T, Okusaka T, Kosuge T, Tsuchida A, Shimahara M, Yasunami Y, Chiba T, Hirohashi S, Yamada T (2011) Reduced plasma level of CXC chemokine ligand 7 in patients with pancreatic cancer. *Cancer Epidemiol Biomarkers Prev* **20**: 160–171.

Mischak H, Schanstra JP (2011) CE-MS in biomarker discovery, validation, and clinical application. *Proteomics Clin Appl* **5**: 9–23.

Pan S, Chen R, Crispin DA, May D, Stevens T, McIntosh MW, Bronner MP, Ziogas A, Anton-Culver H, Brentnall TA (2011) Protein alterations associated with pancreatic cancer and chronic pancreatitis found in human plasma using global quantitative proteomics profiling. *J Proteome Res* **10**: 2359–2376.

Pan S, Chen R, Reimel BA, Crispin DA, Mirzaei H, Cooke K, Coleman JF, Lane Z, Bronner MP, Goodlett DR, McIntosh MW, Traverso W, Aebersold R, Brentnall TA (2009) Quantitative proteomics investigation of pancreatic intraepithelial neoplasia. *Electrophoresis* **30**: 1132–1144.

Sennels L, Salek M, Lomas L, Boschetti E, Righetti PG, Rappsilber J (2007) Proteomic analysis of human blood serum using peptide library beads. *J Proteome Res* **6**: 4055–4062.

Sitek B, Sipos B, Alkatout I, Poschmann G, Stephan C, Schulenborg T, Marcus K, Luttges J, Dittert DD, Baretton G, Schmiegel W, Hahn SA, Kloppel G, Meyer HE, Stuhler K (2009) Analysis of the pancreatic tumor progression by a quantitative proteomic approach and immunhistochemical validation. *J Proteome Res* **8**: 1647–1656.

Song J, Patel M, Rosenzweig CN, Chan-Li Y, Sokoll LJ, Fung ET, Choi-Miura NH, Goggins M, Chan DW, Zhang Z (2006) Quantification of fragments of human serum inter-alpha-trypsin inhibitor heavy chain 4 by a surface-enhanced laser desorption/ionization-based immunoassay. *Clin Chem* **52**: 1045–1053.

Taguchi A, Politi K, Pitteri SJ, Lockwood WW, Faca VM, Kelly-Spratt K, Wong CH, Zhang Q, Chin A, Park KS, Goodman G, Gazdar AF, Sage J, Dinulescu DM, Kucherlapati R, Depinho RA, Kemp CJ, Varmus HE, Hanash SM (2011) Lung cancer signatures in plasma based on proteome profiling of mouse tumor models. *Cancer Cell* **20**: 289–299.

Takano S, Sogawa K, Yoshitomi H, Shida T, Mogushi K, Kimura F, Shimizu H, Yoshidome H, Ohtsuka M, Kato A, Ishihara T, Tanaka H, Yokosuka O, Nomura F, Miyazaki M (2010) Increased circulating cell signalling phosphoproteins in sera are useful for the detection of pancreatic cancer. *Br J Cancer* **103**: 223–231.

Tonack S, Aspinall-O'Dea M, Jenkins RE, Elliot V, Murray S, Lane CS, Kitteringham NR, Neoptolemos JP, Costello E (2009a) A technically detailed and pragmatic protocol for quantitative serum proteomics using iTRAQ. *J Proteomics* **73**: 352–356.

Tonack S, Aspinall-O'Dea M, Neoptolemos JP, Costello E (2009b) Pancreatic cancer: proteomic approaches to a challenging disease. *Pancreatology* **9**: 567–576.

Tonack S, Jenkinson C, Cox T, Elliott V, Jenkins RE, Kitteringham NR, Greenhalf W, Shaw V, Michalski CW, Friess H, Neoptolemos JP, Costello E (2013) iTRAQ reveals candidate pancreatic cancer serum biomarkers: influence of obstructive jaundice on their performance. *Br J Cancer* **108**: 1846–1853.

Turtoi A, Musmeci D, Wang Y, Dumont B, Somja J, Bevilacqua G, De Pauw E, Delvenne P, Castronovo V (2011) Identification of novel accessible proteins bearing diagnostic and therapeutic potential in human pancreatic ductal adenocarcinoma. *J Proteome Res* **10**: 4302–4313.

Vollmer M, van de Goor T (2009) HPLC-Chip/MS technology in proteomic profiling. *Methods Mol Biol* **544**: 3–15.

Wang Y, Kuramitsu Y, Ueno T, Suzuki N, Yoshino S, Iizuka N, Zhang X, Oka M, Nakamura K (2011) Differential expression of up-regulated cofilin-1 and down-regulated cofilin-2 characteristic of pancreatic cancer tissues. *Oncol Rep* **26**: 1595–1599.

Xue A, Scarlett CJ, Chung L, Butturini G, Scarpa A, Gandy R, Wilson SR, Baxter RC, Smith RC (2010) Discovery of serum biomarkers for pancreatic adenocarcinoma using proteomic analysis. *Br J Cancer* **103**: 391–400.