RESEARCH ARTICLE

Mining the malignant ascites proteome for pancreatic cancer biomarkers

Hari Kosanam[1,2,3*], Shalini Makawita[1,3*], Bramford Judd[1], Alice Newman[4] and
Eleftherios P. Diamandis[1,3,5]

[1] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada
[2] Proteomic Methods Inc., Toronto, ON, Canada
[3] Samuel Lunenfeld Research Institute, Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, ON, Canada
[4] Princess Margaret Hospital, University Health Network, Toronto, ON, Canada
[5] Department of Clinical Biochemistry, University Health Network, Toronto, ON, Canada

Pancreatic cancer (PC) is one of the most lethal malignancies and disease-specific biomarkers are desperately needed for better diagnosis, prognosis, monitoring treatment efficacy and for accelerating the development of novel targeted therapeutics. Being an advanced stage manifestation and a proximal fluid in contact with cancer tissues, the ascitic fluid presents itself as a promising rich source of biomarkers. Herein, we present a comprehensive proteomic analysis of pancreatic ascitic fluid. To fractionate the complex ascites proteome, we adopted a multi-dimensional chromatographic approach that included size-exclusion, ion-exchange and lectin-affinity chromatographic techniques. Our detailed proteomic analysis with high-resolution Orbitrap® mass spectrometer resulted in the identification of 816 proteins. We followed rigorous filtering criteria that consisted of PC-specific information obtained from three publicly available databases (Oncomine, Protein Atlas and Unigene) to segregate 20 putative biomarker candidates for future validation. Since these proteins are of membranous and extra-cellular origin, most are glycosylated, and many of them are over-expressed in cancer tissues, the probability of these proteins entering the peripheral blood circulation is high. Their validation as serological PC biomarkers in the future is highly warranted.

1 Introduction

Accounting for over 43 000 estimated new cases and 37 000 estimated deaths per year in USA alone, pancreatic cancer (PC) is one of the most aggressive of all cancer types and the fourth leading cause of deaths due to cancer [1]. With its asymptomatic early stages, the dismal course of PC is due in large part to the advanced-stage at which the disease is detected in most of the patients. Currently, 5-year survival rates are less than 5%; however, detection of PC in its early stages can increase 5-year survival rates to 20–40% [1]. Until better therapeutic measures are developed, the earliest detection of PC is key to improving patient survival [2].

The most widely used serum biomarker for PC is CA19.9. However, due to its poor sensitivity and specificity for early-stage detection, and its elevation in benign diseases of the gastrointestinal system, it is used primarily for monitoring therapeutic response in patients with PC [3, 4]. Other

Correspondence: Dr. Eleftherios P. Diamandis, Head of Clinical Biochemistry, Mount Sinai Hospital and University Health Network, 60 Murray St., Box 32, Floor 6, Rm. L6-201, Toronto, ON, M5 T 3L9 Canada
E-mail: ediamandis@mtsinai.on.ca
Fax: +1-416-586-8628

Abbreviations: EST, expressed sequence tag; **FA,** formic acid; **FPR,** false-positive rate; **IHC,** immunohistochemical; **IXC,** ion-exchange chromatography; **MLAC,** multi-lectin affinity chromatography; **MUC,** mucin-related proteins; **PC,** pancreatic cancer; **SAX,** strong anion-exchange chromatography; **SCX,** strong cation-exchange chromatography

---

* These authors contributed equally to this work.

methods of detection are based on imaging techniques that are used for the detection of pancreatic masses or suspected cancerous lesions, in individuals who present with nonspecific abdominal complaints or symptoms suggestive of PC [5, 6]. Such methods are expensive and nonapplicable for efficient early detection [7]. In this regard, the need for new cancer biomarkers is immense and towards this goal, in the present study, we describe the proteomic analysis of ascites fluid collected from PC patients.

Ascites fluid can form in cancer patients when metastases occur in the peritoneal cavity due to processes such as increased fluid filtration, neovascularization, angiogenesis and lymphatic obstruction [8–10]. Such processes are initiated largely by the secretions of cancer cells, and as a result, ascites fluid can serve as a local microenvironment to mine for biomarkers. A recent analysis of ascites fluid from ovarian cancer patients has proven fruitful for the identification of new disease biomarkers [11]. Proteomic analysis yielded 445 proteins and identification of one novel serum biomarker, Nidogen-2 [12]. Previously, Gortzak-Uzan et al. identified over 2500 proteins in the cellular and soluble ascites fractions and approximately 80 candidate cancer biomarkers [13]. Many known biomarkers were identified in ascites fluid in these analyses, which provides further credibility to this discovery approach [14].

To our knowledge, proteomic analysis of ascitic fluid from PC patients has not been reported. In this regard, we performed MS analysis of three ascites fluid samples from patients with pancreatic ductal adenocarcinomas. Ascites is an albumin- and immunoglobulin (IG)-rich medium [11]. Depleting high-abundance proteins helps to access less abundant and probably clinically significant proteins in complex proximal fluids like ascites. Several albumin and IG affinity depletion methods were attempted previously [15, 16]. The limited success of these methods is largely due to the fact that albumin and IG may bind to certain low molecular weight proteins and peptides, a phenomenon known as sponge effect. Thus, depleting albumin and IG could result in removal of low-abundance and probably informative proteins [17, 18]. Therefore, to extend the depth of proteomic analysis, the combined use of different chromatographic and affinity fractionation procedures is preferable over a unilateral approach [11, 17]. In the current study, we fractionated the ascites proteome by size-exclusion chromatography (SEC), ion-exchange chromatography (IXC), one-dimensional gel electrophoresis (SDS-PAGE) and multi-lectin affinity chromatography (MLAC) methods. We trypsin-digested the fractions separately and then used strong cation-exchange chromatography (SCX) and reverse-phase high-performance chromatography (RP-HPLC) to resolve tryptic-peptides before nano-ESI-Orbitrap® MS analysis. Since it is time-consuming to process each patient's ascites by all fractionation methods, we chose one highly effective method, which allows detection of more proteins with the least number of samples (fractions)

submitted to MS analysis, to fractionate the other two ascites samples. From a relatively large data set of 816 ascites proteins, 20 highly plausible putative biomarker candidates were identified for future validation studies. We pursued stringent filtering criteria that relied on mRNA expression fold changes (Oncomine), immunohistochemical (IHC) staining levels (Protein Atlas) and expressed sequence tag (EST) counts (Unigene).

2 Materials and methods

2.1 Patients and specimens

Three ascites samples from patients with pancreatic ductal adenocarcinomas were obtained with informed consent and institutional ethics board approval at Princess Margaret Hospital, Toronto, Canada. The three patients (two males and one female; ages 76, 76 and 56 years, respectively) had advanced PC with presence of metastases. The patients were receiving chemotherapy or palliative treatment at the time of ascites collection. Ascites was collected through paracentesis. One liter of fluid was obtained from each patient, aliquoted and frozen immediately at −80°C until further use.

2.2 Sample preparation

Ascites fluids were centrifuged at $20\,000 \times g$ for 45 min at 4°C to separate lipids and cellular debris. The clarified supernatant was filtered through a 0.22 µM membrane (Millipore, Billerica, MA, USA) and buffer-exchanged to the appropriate chromatographic loading buffer using Amicon centrifugal filters (MWCO 3000 Da, Millipore). The ascitic fluids were assayed for total protein and albumin content at Mount Sinai Hospital clinical biochemistry laboratories. From our repository of ascites samples, we selected these three samples that were free from blood or lipid contamination. One of the samples was processed by all fractionation methods, as part of our optimization studies, and the remaining two samples were fractionated by one fractionation method. The same amount of total protein was used in all fractionation methods.

2.3 SEC

SEC was performed on a 7.5 mm × 60 cm TSK-Gel G3000SW column (Tosoh Bioscience) connected to an Agilent 1100 HPLC system equipped with a diode array detector. The chromatographic conditions were: sample load, 0.5 mL; mobile phase, 100 mM phosphate/150 mM sodium chloride; flow rate, 0.5 mL/min; time, 60 min. The eluted proteins were monitored at 280 nm. Thirty injections were performed to fractionate 15 mL of ascites fluid.

Proteins eluted at similar elution times in different chromatographic runs were pooled and desalted/concentrated by Amicon (MWCO 3000 Da) centrifugal filters. The fractions were trypsin-digested and the resulting peptides were further separated by SCX. Prior to sample analysis, the chromatographic conditions (flow rate and mobile phase composition) were optimized for enhanced resolution of SEC standards (BIO-RAD gel-filtration standards; thyroglobulin, 670 kDa; γ-globulin, 158 kDa; ovalbumin, 44 kDa; myoglobin, 17 kDa; and vitamin B12, 1.35 kDa). As a quality control procedure, the standards were analyzed before and after sample analysis to ensure the quality of the fractionation procedure.

2.4 IXC

IXC constitutes both strong anion-exchange chromatography (SAX) and SCX. SAX was carried out on a 10 mL HiTrap Q sepharose column connected to an AKTA fast protein LC system (GE Healthcare, NJ, USA). A 15 mL aliquot of ascites fluid was diluted fourfold with buffer A (50 mM Tris, pH 8.5) and loaded onto the column at 0.25 mL/min. The column was washed with five column volumes of buffer A (50 mM Tris, pH 8.5) and proteins were eluted with a 60 min linear gradient profile from 100% buffer A to 100% buffer B (50 mM Tris, 1 M sodium chloride, pH 8.5) at a 2 mL/min flow rate. Ten fractions collected from the SAX run were desalted and concentrated prior to trypsin digestion. The flow through (proteins not bound to anion-exchange column) after buffer exchange to an SCX loading buffer of pH 5.5 (see below) was loaded on to a 5 mL HiTrap SP-HP (SCX) column. The cation-exchange chromatography was performed using a 60 min gradient at 2 mL/min (buffer A: 50 mM MES, pH 5.5; and buffer B: 50 mM MES+1 M NaCl, pH 5.5). The fractions collected were trypsin-digested as described previously [11] and SCX performed on the tryptic peptides as described below.

2.5 MLAC

We prepared the MLAC column with 1:1 addition of agarose-bound concanavalin A, and wheat germ agglutinin to an empty 25 mL glass column (GE Healthcare). The column was equilibrated with ten column volumes of equilibration buffer containing 20 mM Tris, 150 mM NaCl, 1 mM CaCl$_2$ and 1 mM MnCl$_2$ (pH 7.4). A 15 mL aliquot of ascites fluid in equilibration buffer was applied to the column and incubated for 4 h with end-to-end mixing. The column was washed with ten column volumes of equilibration buffer. The lectin-bound proteins were eluted with three column volumes of lectin elution buffer (20 mM Tris, 150 mM NaCl, 1 mM CaCl$_2$, 1 mM MnCl$_2$, 0.5 M methyl-α-D-mannopyranoside, and 0.5 M *N*-acetylglucosamine). The lectin eluate was concentrated using Amicon 3000 Da centrifugal filters. The

lectin-bound proteins were further separated by SEC. A total of 12 fractions per SEC run were collected and trypsinized. The tryptic peptides were desalted and analyzed by LC-MS/MS. SCX was not performed on tryptic-peptides from lectin-bound proteins.

2.6 SCX

Tryptic-peptides were lyophilized and reconstituted in 500 μL of buffer A (0.26 M formic acid (FA) in 5% ACN). The samples were loaded onto a PolySULFOETHYL aspartamide SCX column (4.6 mm × 50 mm, 200A and 5 μm) (The Nest Group, MA, USA) and fractionation was performed using an Agilent 1100 HPLC system. A 60 min linear gradient method was operated buffer A → B (B: 0.26 M FA in 5% ACN and 1 M ammonium formate) at a flow rate of 250 μL/min. The eluate absorbance was monitored at 280 nm. Fractions were collected in 250 μL aliquots, which were later pooled. A total of 190 and 120 fractions collected from SCX analysis of SEC and IXC fractions, respectively, were subjected to RP-LC-MS/MS analysis. To evaluate column performance, a peptide cation exchange standard mixture was applied before and after chromatographic analysis.

2.7 One-dimensional gel electrophoresis

Ascites fluid containing 50 μg of total protein (less than the total protein used for other methods, but the maximum possible load for SDS-PAGE) was subjected to gel electrophoresis and stained with Coomassie blue. The gel was de-stained; 19 gel bands were excised and subjected to in-gel trypsin digestion as described previously [19].

2.8 LC-MS/MS analysis on LTQ-Orbitrap®

The peptides from SEC/SCX, IXC/SCX, lectin chromatography/SEC and SDS-PAGE/in-gel trypsin digestion were desalted using the Omix C18 tips (Varian, Palo Alto, CA, USA). Samples were diluted with Buffer A (0.1% FA in water) and injected into a C$_{18}$ trap column (150 μm packed in-house) using the EASY-nLC system (Proxeon Biosystems, Odense, Denmark). Peptides were eluted from the trap column with an increasing concentration of Buffer B (0.1% FA in ACN) onto a resolving 5 cm long PicoTip Emitter (75 μm inner diameter, 8 μm tip, New Objective) packed in-house with 3 μm Pursuit C$_{18}$ (Varian). Peptides were resolved using gradient RP chromatography at a flow rate of 400 nL/min for 90 min. The chromatography system was connected online to LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) via a nano-ESI source (Proxeon Biosystems). The capillary temperature was 160°C and spray voltage was 2 kV. The mass spectra

were acquired in data-dependent mode. The first scan event was a full MS scan from m/z 450 to 1450 and the next six scans were MS/MS scans on the six most intense parent ions observed in scan 1. These scan events were alternated between each other for 90 min MS acquisition time. Collision dissociation energy for MS/MS was set at 30%. Dynamic exclusion, monoisotopic precursor selection and charge state screening were enabled.

All the MS analyses were performed in duplicate.

2.9   Database searching and bioinformatics

The MS spectra were searched against the nonredundant IPI human database (version 3.71 containing both forward and reverse protein sequences) using two search engines separately; MASCOT, version 2.1.03 (Matrix Science) and the Global Proteome Machine manager version 2006.06.01 (GPM X! Tandem; Beavis Informatics, Canada). The following parameters were used: (i) enzyme: trypsin; (ii) one missed cleavage allowed; (iii) fixed modification: carbamidomethylation of cysteines; (iv) variable modifications: oxidation of methionines; (v) MS^1 tolerance, 7 ppm; and (vi) MS^2 tolerance, 0.4 Da. The resulting MASCOT DAT and X! Tandem XML files were merged using Scaffold® (version 2.06, Proteome Software, Portland, Oregon) with 'MudPIT' (multidimensional protein identification technology) option checked. Scaffold result data was filtered using the X! Tandem LogE (min 3.0) and MASCOT ion-score filters [ion score 15, 30(+2) and 40 (+3)] in order to obtain a protein false-positive rate (FPR) of ≤1%. FPR = 2 × (number of proteins identified by searching the reverse sequences)/(the total number of identified proteins). Scaffold® protXML reports were exported and uploaded into Protein Center (Proxeon Biosystems) to retrieve gene ontology annotations. Publicly available Oncomine, Protein Atlas and Unigene were searched for biomarker candidate selection.
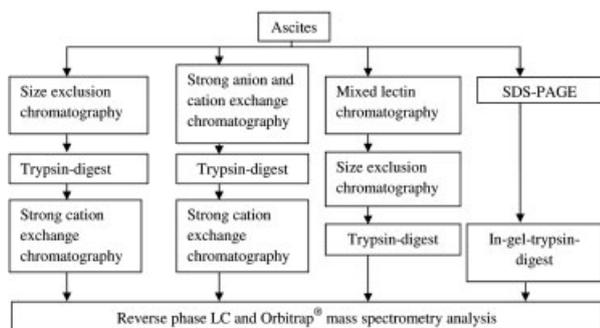

Figure 1. Experimental strategy used for pancreatic ascitic fluid fractionation; four methods (size-exclusion, ion-exchange, lectin chromatographies and one-dimensional electrophoresis) were used to fractionate ascites. For detailed description see Section 2 in the text.

3   Results and discussion

Although serum is a promising source for biomarkers, the detection of low-abundance disease-specific proteins is a challenging task. This is largely due to the masking effect rendered by a relatively small number of high-abundance proteins. Proximal fluids surround the cancer tissues and absorb shed and secreted proteins from them. In addition, these fluids are usually less complex than serum and provide an excellent alternative source for biomarker discovery. Our laboratory explored the proteomes of several proximal fluids such as ovarian ascites [11], amniotic fluid [17], seminal plasma [20], nipple aspirate fluid [21], pancreatic juice and bronchoalveolar lavage fluid and identified many candidate biomarkers by applying sets of hypothesis-driven, stringent filtering criteria. Such candidates are subsequently validated in serum by enzyme-linked immunosorbent assays or selected reaction monitoring assays. Following a similar path, we performed proteomic analysis of pancreatic ascites fluid in the pursuit of novel biomarker candidates. Ascites is an end-stage manifestation in 20% of PC patients. Death usually follows the accumulation of ascites. The cause of ascites formation is multifactorial, but primarily it is due to obstruction of the diaphragmatic lymphatic system and overproduction of exudates by tumor cells. Ascitic fluid is composed of tumor-originating cells and their secretome; therefore it is a gold mine for biomarkers and potential drug targets. To date, the proteome of pancreatic ascites has not been delineated. Like serum, ascites is also an albumin-rich fluid and approximately 40–50% of its proteome is albumin. To circumvent the 'sponge' effects of albumin, we avoided using albumin-depletion methods. Instead, we adopted a multi-dimensional chromatographic approach, which included SEC, IXC, SDS-PAGE and lectins at the protein level and SCX and RP-HPLC at the peptide level. Our experimental strategy is summarized in Fig. 1.

3.1   Evaluation of fractionation methods

We identified 813 proteins using four fractionation methods from one ascites sample; 773 were identified from SEC/SCX; 574 from MLAC/SEC; 463 from IXC/SCX and 303 from SDS-PAGE. Figure 2 shows the overlap of proteins between the three most efficient methods (SEC/SCX, MLAC/SEC and IXC/SCX). From SEC/SCX, a total of 160 fractions were analyzed by LC-MS/MS, requiring 240 instrument hours; with IXC/SCX, 120 fractions were analyzed by MS (180 LC-MS/MS hours); Only 12 fractions were analyzed with MLAC/SEC (18 h of LC-MS/MS time). The MS analysis of 19 SDS-PAGE gel bands identified 303 proteins (29 LC-MS/MS hours). The lower number of proteins identified from SDS-PAGE may be due to lower total protein load (50 µg). The MS spectral counts are usually correlated to the abundance

of the corresponding proteins. To assess the high-abundance protein depleting ability of the fractionation methods, we summed the spectral counts of the ten most-abundant proteins (including albumin) detected from MLAC/SEC, SEC/SCX and IXC/SCX methods. The cumulative spectral counts in MLAC/SEC were fivefold lower than the other two methods (Supporting Information Fig. 1). This reduction in high-abundance proteins enabled lectins to unveil the complexity of the ascites proteome more efficiently than any of the other tested methods. As a result, a larger number of proteins were detected with the least number of fractions submitted to MS analysis.

3.2 Overlap between the ascites samples from three patients

In total, 816 proteins were identified with 2% FPR in ascites sampled from three patients. In sample 1 we identified 813 proteins (fractionated by four methods); in samples 2 and 3, we identified 514 and 512 proteins (fractionated only by MLAC), respectively, and 493 proteins were identified in all the three samples. All 816 proteins are listed in Supporting Information Table 1.

3.3 MLAC identifies more N-linked glycoproteins

MLAC column contains two lectins, wheat germ agglutinin and concanavalin A, which selectively bind to glycoproteins containing *N*-acetyl glucosamine and α-D-mannose and

Figure 2. (A) Number of proteins identified with three fractionation methods: IXC, SEC and MLAC. More proteins (773) were detected with the SEC method, but SEC identified less extracellular and membrane proteins, while requiring longer sample analysis time (180 h); MLAC identified 574 proteins with higher number of extracellular and membranous proteins and with only 10% of the SEC analytical time (18 h). (B) Overlap of pancreatic ascites proteins with human plasma proteome (HIP-2) and ovarian ascites proteome; only 39 ovarian ascitic proteins were also found in pancreatic ascites; this provides some evidence that ovarian and pancreatic ascitic composition is dissimilar and that PC-specific markers can be found in pancreatic ascites.

α-D-glucose, respectively. It is believed that 60–70% of human proteins are glycosylated; many glycosylated proteins are typically of extracellular and membranous origin, and they are highly likely to be secreted into the circulation from cancer tissues. Moreover, many of the biomarkers in clinical use are glycosylated; therefore, mining glycosylated proteins in proximal fluids such as ascites is an attractive approach for biomarker discovery. We included lectins in our strategy to specifically capture the N-linked glyco-sub proteome. A total of 304 (Supporting Information Table 2) proteins in ascites were identified as known N-linked glycoproteins (Swiss-Prot annotation); among those 222 proteins were identified by MLAC, and the remaining 82 by all other fractionation methods. This data shows that as expected, lectins capture a larger proportion of the glycoproteome. Confirming the glycosylation status of these proteins would require a different experimental set-up; however this objective is beyond the scope of the current study. The inclusion of lectins in the fractionation methodology served three purposes; enrichment of N-linked glycoproteins, decreased levels of high-abundance proteins and savings in precious analysis time.

3.4 Gene ontology analysis

We used Protein Center (Thermo Fisher) to classify ascites proteins based on their sub-cellular localization (Supporting Information Fig. 2). The proteins are predominantly of extracellular (561 proteins), membranous (437) and cytoplasmic origin (396). Each protein may be localized in multiple cellular compartments. Therefore, the total number of proteins in Supporting Information Fig. 2 exceeds 816. We then categorized the cellular localization based on fractionation methods. MLAC/SEC identified a marginally higher number of extracellular and membranous proteins (60%) compared to SEC/SCX (52%) and IXC/SCX (45%).

3.5 Comparison of ascites proteome with previously published biomarkers

We compared our database of 816 proteins with the compendium of PC biomarkers documented by Harsha et al. [22]. One hundred sixty-six ascites proteins were previously reported as elevated in PC (see Supporting Information Table 3). Bunger et al. [23] recently reviewed and compiled a list of 35 serum (protein) biomarkers, which were previously validated in serum for PC diagnostics; among those, ten proteins were identified in our study. Most of the proteins identified in ascitic fluid may be associated with the pathophysiology of PC. Our study detected a number of such proteins that have been previously implicated in PC pathogenesis. Such proteins include: matrix
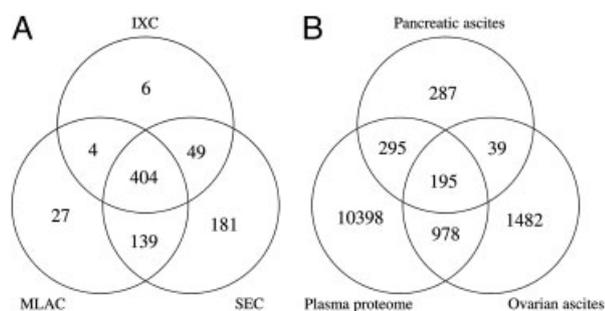
metalloproteinase-2, which is associated with pancreatic tumor progression and development of desmoplasia, was detected in the current study with 13 unique peptide hits. Matrix metalloproteinase-2 is also a potential therapeutic target to control desmoplasia [24]. Neutrophil gelatinase-associated lipocalin (four peptide hits) was reported to be 27-fold up-regulated in PC cells compared to normal human pancreatic ductal epithelial cells [25]. Mucin-related proteins (MUC1, MUC5B, MUC5AC and MUC6) were detected in ascitic fluid with two or more peptide hits. Mucin-1 carries the PC-specific CA19-9 epitope [4]. Mannose-binding lectin-2 (MBL2) [26] and carcinoembryonic antigen-related cell adhesion molecule (CEACAM1) [27], recently studied as PC biomarkers in serum, were also identified.

3.6 Comparison with serum and ovarian ascites protein databases

Ascites is the accumulation of serous fluid in the peritoneal cavity. Other than PC, the incidence of ascites has also been reported in ovarian, colon, liver and endometrial cancers [13]. Although the cancer pathways could be divergent, the composition of ascites may be similar between cancers. Comparing the pancreatic ascites proteome (this work) with other ascites proteomes may allow elimination of common proteins and aid in the selection of PC-specific biomarkers. With this in mind, we compared our data with the ovarian ascites proteome, consisting of more than 2500 proteins (no other cancer ascites proteomic data is currently available). Figure 2B shows the overlap of pancreatic ascites with the ovarian ascites proteome [13] and normal human plasma proteome [28]. Pancreatic and ovarian ascites proteomes are predominantly derived from serum. Approximately 60% (490) of pancreatic ascites proteins were also found in healthy serum. However, it is interesting to note that only 39 of 1521 ovarian ascites proteins not found in serum were identified in the current study; this underlines that both pancreatic and ovarian ascites proteomes have important compositional differences. Alternatively, the differences may be due to the fact that in the previous study [13] the proteome was delineated for both soluble and cellular components while in our study only soluble fraction was analyzed.

3.7 Selection of biomarker candidates

Major challenges in biomarker research arise in the verification and validation phases. The identification of 816 proteins generated a relatively large biomarker database; so the verification and validation of all of these proteins is an obvious herculean task. Therefore, it is necessary to reduce the number of biomarker candidates that may be eventually verified and validated. From our catalogue of 816 proteins,

we short-listed 20 proteins for future validation studies based on the following criteria (Fig. 3 and Table 1):

(i) Single peptide hits were included in proteome data as these proteins may be critical in understanding tumor biology. However, we excluded them from our mining strategy because their predicted low abundance in ascites makes their secretion and possible detection in serum less likely. Hence, we retained only those proteins that were identified with two or more peptide hits; this resulted in 346 proteins.

(ii) We then removed high-abundance serum proteins with concentrations >5 µg/mL; this shortened the list to 247. The database of 153 high-abundance proteins and their respective concentrations were reported by Hortin et al. [29] based on a literature survey.

(iii) We selected 180 proteins that were identified in at least one of the six PC cell line conditioned media (BxPc3, CAPAN1, CFPAC1, PANC1, MIA-PaCa2 and SU.86.86) and malignant pancreatic juice as described elsewhere (Makawita, S., Diamandis, E. P. et al. manuscript accepted for publication, *Mol. Cell. Proteomics*, June 7, 2011).

3.8 Bioinformatic filtering criteria

Finally, we searched the remaining 180 proteins in three publicly available databases: (i) Oncomine for mRNA expression changes (up-regulation) in PC cell-lines and tissues (a consistent >2-fold up-regulation reported in two or more publications). (ii) Protein Atlas for IHC staining levels; we followed a point system to grade the level of staining in PC for each of 180 proteins; 0-negative staining; 1-weak; 2-less moderate; 3-moderate; 4-less strong and 5-strong. (iii) Unigene for EST counts in ppm; we calculated the ratio of EST-PC over EST-healthy pancreatic tissue for all 180 proteins.
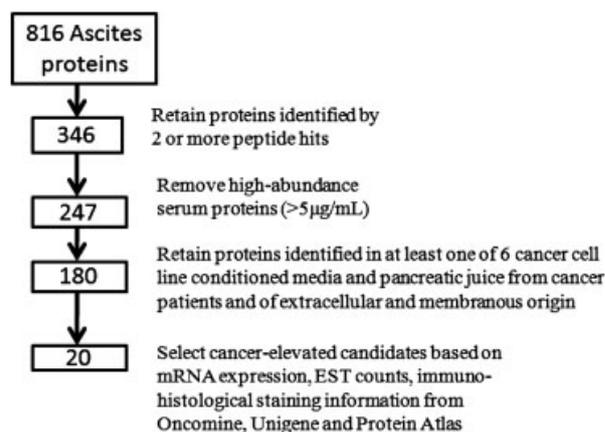

Figure 3. Selection criteria followed to segregate the 20 candidate pancreatic biomarkers.

Table 1. Panel of 20 candidate biomarkers suitable for future validation studies

IPI	Gene	Protein description	MW	Ascites-1[a]	Ascites-2[a]	Ascites-3[a]	Oncomine[b]	Protein Atlas[c]	Unigene[d]
IPI00410241	POSTN	Periostin, osteoblast specific factor	90	5	6	19	8.3	2	1.74
IPI00186903	APOL1	Isoform 2 of apolipoprotein L1	46	1	3	5	5.14	3	2.15
IPI00020986	LUM	Lumican	38	9	9	17	13.12	4	1.88
IPI00013933	DSP	Isoform DPI of desmoplakin	332	1	1	13	3.69	3	0.33
IPI00165438	NRP1	Muscle type neuropilin 1	72	1	5	11	6.53	3	1.02
IPI00024284	HSPG2	Basement membrane-specific heparan sulfate proteoglycan core protein	469	3	1	16	14.93	2	1.6
IPI00298281	LAMC1	Laminin subunit γ-1	178	1	3	5	4.043	2	1.91
IPI00299729	TCN1	Transcobalamin-1	48	1	1	4	24.48	x	2.06
IPI00006146	SAA2	Serum amyloid A2 isoform a	14	3	1	1	6.35	x	2.25
IPI00027780	MMP2	72 kDa type IV collagenase	74	3	3	13	2.12	2	1.98
IPI00011695	PRSS2	Protease serine 2 isoform B	46	2	2	2	33.73	3	0.18
IPI00555812	GC	Vitamin D-binding protein precursor	53	30	26	35	6.33	4	0.12
IPI00021000	SPP1	Isoform A of Osteopontin					5.32	4	0.6
IPI00220737	NCAM1	Isoform 3 of neural cell adhesion molecule 1	84	1	1	2	2.41	2	0.78
IPI00011218	CSF1R	Macrophage colony-stimulating factor 1 receptor	105	1	1	7	2.98	5	1.04
IPI00554711	JUP	Junction plakoglobin	82	1	1	8	3.76	3	2
IPI00465028	TPI1	Triosephosphate isomerase isoform 2	27	1	1	2	2.093	2	1.88
IPI00003351	ECM1	Isoform 1 of extracellular matrix protein 1	61	8	9	21	3.19	3	1.19
IPI00011588	PLA2G7	Platelet-activating factor acetylhydrolase	50	1	1	2	3.29	5	0
IPI00479997	STMN1	Stathmin	17	1	2	2	2.18	4	0.75

a) Number of unique peptides detected with ascites samples from three patients (Ascites1, 2 and 3).
b) Previously reported mRNA expression fold-change increase (derived from Oncomine) in pancreatic cancer.
c) Levels of immuno-histochemical staining in pancreatic cancer tissues (obtained from Protein Atlas; 0 for negative; 1 weak; 2 less moderate; 3 moderate; 4 less strong; 5 strong; X information not available.
d) Ratio of expressed sequence tags (ppm values) in cancer over normal tissues (Unigene).

To arrive at the final list of 20 candidates, we considered only those proteins that fulfilled at least two of the three following selection criteria: a minimum of (i) twofold increase on Oncomine scale; (ii) IHC staining level 3 from Protein Atlas; (iii) EST ratio ≥1.5 on Unigene scale. Table 1 lists the panel of 20 biomarker candidates.

3.9 Possible detection of candidate biomarkers in serum

Out of 20 candidates, 12 were identified by MS [28] (HIP-2 database for plasma proteome) in serum with one or more peptide hits. In general, the MS methods for protein detection and quantification are less sensitive than traditional ELISA assays. Therefore, these 12 proteins should be easily measurable in serum by ELISA. All 20 listed biomarker candidates have commercially available ELISA kits for their detection in serum. We have listed ELISA limits-of-detection and number of peptides detected in MS analysis for each protein in healthy serum samples (Supporting Information Table 4). To the best of our knowledge, none of these candidates were tested in PC sera. Hence, the verification and validation of these promising candidates in serum samples is worthwhile.

4 Concluding remarks

If the purpose is early detection, then, hypothetically, it may be best to study body fluids from patients in their earliest stages of cancer. But PC is diagnosed and ascites is developed during the final stages of the disease. Hence, many of the proteins found in ascites may not be present or detectable in sera of early stage patients. However, it is logical to believe that some of the moderately abundant ascites and cancer-specific proteins of extracellular or secretory origin may be detectable in the serum. Therefore, one of our future directions is to validate these 20 candidates in serum, in hopes that some or a panel could represent promising PC biomarkers.

The authors have declared no conflict of interest.

5   References

[1] Jemal, A., Bray, F., Center, M. M., Ferlay, J. et al., Global cancer statistics. *CA Cancer J. Clin.* 2011, *61*, 69–90.

[2] Etzioni, R., Urban, N., Ramsey, S., McIntosh, M. et al., The case for early detection. *Nat. Rev. Cancer* 2003, *3*, 243–252.

[3] Goonetilleke, K. S., Siriwardena, A. K., Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. *Eur. J. Surg. Oncol.* 2007, *33*, 266–270.

[4] Marrelli, D., Caruso, S., Pedrazzani, C., Neri, A. et al., CA19-9 serum levels in obstructive jaundice: clinical value in benign and malignant conditions. *Am. J. Surg.* 2009, *198*, 333–339.

[5] Soriano, A., Castells, A., Ayuso, C., Ayuso, J. R. et al., Preoperative staging and tumor resectability assessment of pancreatic cancer: prospective study comparing endoscopic ultrasonography, helical computed tomography, magnetic resonance imaging, and angiography. *Am. J. Gastroenterol.* 2004, *99*, 492–501.

[6] Irisawa, A., Sato, A., Sato, M., Ikeda, T. et al., Early diagnosis of small pancreatic cancer: role of endoscopic ultra-sonography. *Dig. Endosc.* 2009, *21*, S92–S96.

[7] Mendieta Zeron, H., Garcia Flores, J. R., Romero Prieto, M. L., Limitations in improving detection of pancreatic adenocarcinoma. *Future Oncol.* 2009, *5*, 657–668.

[8] Adam, R. A., Adam, Y. G., Malignant ascites: past, present, and future. *J. Am. Coll. Surg.* 2004, *198*, 999–1011.

[9] Tamsma, J. T., Keizer, H. J., Meinders, A. E., Pathogenesis of malignant ascites: Starling's law of capillary hemody-namics revisited. *Ann. Oncol.* 2001, *12*, 1353–1357.

[10] Zervos, E. E., Osborne, D., Boe, B. A., Luzardo, G. et al., Prognostic significance of new onset ascites in patients with pancreatic cancer. *World J. Surg. Oncol.* 2006, *4*, 16–22.

[11] Kuk, C., Kulasingam, V., Gunawardana, C. G., Smith, C. R. et al., Mining the ovarian cancer ascites proteome for potential ovarian cancer biomarkers. *Mol. Cell. Proteomics* 2009, *8*, 661–669.

[12] Kuk, C., Gunawardana, C. G., Soosaipillai, A., Kobayashi, H. et al., Nidogen-2: a new serum biomarker for ovarian cancer. *Clin. Biochem.* 2010, *43*, 355–361.

[13] Gortzak-Uzan, L., Ignatchenko, A., Evangelou, A. I., Agochiya, M. et al., A proteome resource of ovarian cancer ascites: integrated proteomic and bioinformatic analyses to identify putative biomarkers. *J. Proteome Res.* 2008, *7*, 339–351.

[14] Metzgar, R. S., Rodriguez, N., Finn, O. J., Lan, M. S. et al., Detection of a pancreatic cancer-associated antigen (DU-PAN-2 antigen) in serum and ascites of patients with adenocarcinoma. *Proc. Natl. Acad. Sci. USA* 1984, *81*, 5242–5246.

[15] Siegmund, R., Kiehntopf, M., Deufel, T., Evaluation of two different albumin depletion strategies for improved analysis

of human CSF by SELDI-TOF-MS. *Clin. Biochem.* 2009, *42*, 1136–1143.

[16] Bellei, E., Bergamini, S., Monari, E., Fantoni, L. I. et al., High-abundance proteins depletion for serum proteomic analy-sis: concomitant removal of non-targeted proteins. *Amino Acids* 2011, *40*, 145–156.

[17] Cho, C. K., Shan, S. J., Winsor, E. J., Diamandis, E. P., Proteomics analysis of human amniotic fluid. *Mol. Cell. Proteomics* 2007, *6*, 1406–1415.

[18] Magagnotti, C., Fermo, I., Carletti, R. M., Ferrari, M., Bachi, A., Comparison of different depletion strategies for improving resolution of the human urine proteome. *Clin. Chem. Lab Med.* 2010, *48*, 531–535.

[19] Granvogl, B., Gruber, P., Eichacker, L. A., Standardisation of rapid in-gel digestion by mass spectrometry. *Proteomics* 2007, *7*, 642–654.

[20] Batruch, I., Lecker, I., Kagedan, D., Smith, C. R. et al., Proteomic analysis of seminal plasma from normal volun-teers and post-vasectomy patients identifies over 2000 proteins and candidate biomarkers of the urogenital system. *J. Proteome Res.* 2011, *10*, 941–953.

[21] Pavlou, M. P., Kulasingam, V., Sauter, E. R., Kliethermes, B., Diamandis, E. P., Nipple aspirate fluid proteome of healthy females and patients with breast cancer. *Clin. Chem.* 2010, *56*, 848–855.

[22] Harsha, H. C., Kandasamy, K., Ranganathan, P., Rani, S. et al., A compendium of potential biomarkers of pancreatic cancer. *PLoS Med.* 2009, *6*, e1000046.

[23] Bunger, S., Laubert, T., Roblick, U. J., Habermann, J. K., Serum biomarkers for improved diagnostic of pancreatic cancer: a current overview. *J. Cancer Res. Clin. Oncol.* 2011, *137*, 375–389.

[24] Schultz, J. D., Rotunno, S., Erben, P., Sommer, J. U. et al., Down-regulation of MMP-2 expression due to inhibition of receptor tyrosine kinases by imatinib and carboplatin in HNSCC. *Oncol. Rep.* 2011, *25*, 1145–1151.

[25] Moniaux, N., Chakraborty, S., Yalniz, M., Gonzalez, J. et al., Early diagnosis of pancreatic cancer: neutrophil gelatinase-associated lipocalin as a marker of pancreatic intraepithelial neoplasia. *Br. J. Cancer* 2008, *98*, 1540–1547.

[26] Rong, Y., Jin, D., Hou, C., Hu, J. et al., Proteomics analysis of serum protein profiling in pancreatic cancer patients by DIGE: up-regulation of mannose-binding lectin 2 and myosin light chain kinase 2. *BMC Gastroenterol.* 2010, *10*, 68–77.

[27] Serra, S., Asa, S. L., Bamberger, A. M., Wagener, C., Chetty, R., CEACAM1 expression in pancreatic endocrine tumors. *Appl. Immunohistochem. Mol. Morphol.* 2009, *17*, 286–293.

[28] Saha, S., Harrison, S. H., Shen, C., Tang, H. et al., HIP2: an online database of human plasma proteins from healthy individuals. *BMC Med. Genomics* 2008, *1*, 12–21.

[29] Hortin, G. L., Sviridov, D., Anderson, N. L., High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. *Clin. Chem.* 2008, *54*, 1608–1616.