

Identification of Novel Accessible Proteins Bearing Diagnostic and Therapeutic Potential in Human Pancreatic Ductal Adenocarcinoma

Andrei Turtoi,^{†,‡} Davide Musmeci,^{†,§} Yinghong Wang,[†] Bruno Dumont,[†] Joan Somja,^{||} Generoso Bevilacqua,[§] Edwin De Pauw,[‡] Philippe Delvenne,^{||} and Vincent Castronovo^{*,†}

[†]Metastasis Research Laboratory, GIGA-Cancer, University Hospital CHU, University of Liege, Bat. B23, 4000 Liege, Belgium

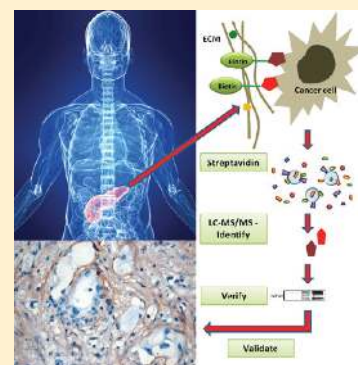
[‡]Laboratory of Mass Spectrometry, University of Liege, Bat. B6C, 4000 Liege, Belgium

[§]Division of Surgical, Molecular and Ultrastructural Pathology, University Hospital of Pisa, Italy

^{||}Faculty of Medicine, Department of Anatomy and Pathology, University of Liege, B-4000 Liege, Belgium

S Supporting Information

ABSTRACT: Pancreas ductal adenocarcinoma (PDAC) remains a deadly malignancy with poor early diagnostic and no effective therapy. Although several proteomic studies have performed comparative analysis between normal and malignant tissues, there is a lack of clear characterization of proteins that could be of clinical value. Systemically reachable (“potentially accessible”) proteins, suitable for imaging technologies and targeted therapies represent a major group of interest. The current study explores potentially accessible proteins overexpressed in PDAC, employing innovative proteomics technologies. In the discovery phase, potentially accessible proteins from fresh human normal and PDAC tissues were *ex vivo* biotinylated, isolated and identified using 2D-nano-HPLC–MS/MS method. The analysis revealed 422 up-regulated proteins in the tumor, of which 83 (including protein isoforms) were evaluated as potentially accessible. Eleven selected candidates were further confirmed as up-regulated using Western blot and multiple reaction monitoring protein quantification. Of these, transforming growth factor beta-induced (TGFBI), latent transforming growth factor beta binding 2 (LTBP2), and asporin (ASPN) were further investigated by employing large scale immunohistochemistry-based validations. They were found to be significantly expressed in a large group of clinical PDAC samples compared to corresponding normal and inflammatory tissues. In conclusion, TGFBI, LTBP2, and ASPN are novel, overexpressed, and potentially accessible proteins in human PDAC. They bear the potential to be of clinical value for diagnostic and therapeutic applications and merit further studies using *in vivo* models.



KEYWORDS: pancreas cancer, accessible proteins, biomarkers, proteomics

INTRODUCTION

Pancreas cancer is one of the deadliest malignancies with highly unmet needs.¹ The high death burden is explained by the fact that the disease has a nonsymptomatic course of development, resulting in patients being diagnosed at a very late stage. Unfortunately, even though during the past decades several biomarkers have been put forward, none has led to a major progress in early diagnosis and more effective therapies.² Practically, to date, only surgical resection and systemically toxic radio- and chemotherapy are building the first line of treatment. Current efforts are concentrated toward the assessment of novel targeted therapies focused around tumor cell and stem cell signaling pathways as well as tumor microenvironment.³ Unfortunately, selective inhibitors against specific molecular targets (alternatively combined with chemo- or radio-therapy) did not bring significant additional benefits to pancreas cancer patients.³ Therefore, the major challenge is to identify new strategies and targets enabling to interfere more efficiently against pancreas cancer development and progression.

To this end, the transition from genomic to proteomic era has prompted several studies to better characterize the proteins modulated in pancreatic cancer.⁴ Nevertheless, the enormous complexity of the human proteome, high dynamic range of concentrations, as well as the selection and exploitation of precious clinical material still remain a major challenge. In contemporary clinical proteomics there is a move from entire proteome studies towards a more dedicated approach where the focus lies on a smaller but relevant group of proteins. Accordingly, cancer specific membrane and extracellular proteins, which are theoretically accessible to a systemically delivered ligands, represent such a group deserving an augmented attention. These accessible cancer proteins are the first line of physical interaction of the cancer with the surrounding environment. Furthermore such proteins bear the potential to serve as targets in attractive therapeutic and diagnostic approaches, where vehicles such as antibodies can be used to selectively bring to the tumor high loads of cytotoxic compounds

Received: June 2, 2011

Published: July 14, 2011

(or imaging agents).⁵ The present study attempts to fill in exactly this void in the current knowledge concerning the potentially accessible proteome in pancreas cancer. To accomplish this, a specific strategy of immersing fresh tissue biopsies in chemically modified biotin solution is employed. In intact tissues, this chemical compound labels proteins external to the cell membrane. Subsequently, it is possible to isolate this group of labeled proteins and, using mass spectrometry, to reveal their identity. Even though there is an apparent dearth of proteomic information relating to pancreas cancer, the current work compares the most relevant studies to the newly generated data. The study clearly demonstrates the additional benefit of limiting the investigation to the specific group of potentially accessible proteins. Owing to this unique approach, a set of novel proteins previously not or poorly described in the context of pancreas cancer are highlighted. Importantly, the current study goes beyond the comparison of normal and tumoral tissues but also involves a significant number of inflammatory (pancreatitis) tissue samples. This is important especially because a significant number of cancer related proteins are not specific to cancer but are also found in inflammation.

MATERIAL AND METHODS

The current study uses a discovery phase that investigates a limited number of cases originating from precious, fresh, clinical material. The second phase represents a verification study, where a set of modulated proteins are further quantified in a targeted mode. Finally, using a large cohort of individual samples ($N = 62$), the current work validates several novel and potentially accessible proteins for their significant presence in a large series of pancreas cancer. The expression levels of these proteins are additionally assessed in a panel of ($N = 38$) normal human tissues. The tissue collection is described in more detail in the Supporting Information. This study was approved by the ethical committee of the University Pisa (Italy) and University of Liege (Belgium). An informed consent was obtained from all the patients.

DISCOVERY PHASE

Isolation of Accessible Proteins

The detailed procedure was described previously.⁶ Briefly, fresh tissue samples were soaked into EZ-link Sulfo NHS-SS-biotin (Pierce, Rockford, IL) solution. When sufficient material was available, a part of biotinylated tissue was fixed in formalin and embedded in paraffin. This material served to verify the biotin labeling (Supporting Information, Figure S3). Remaining tissue samples were snap-frozen and pulverized. Subsequently, 100 mg of tissue powder was lysed and subjected to protein extraction followed by affinity purification of biotinylated proteins using streptavidin. The purified proteins were precipitated with 20% trichloroacetic acid; the pellets were suspended in 100 μ L of 100 mM NH_4HCO_3 (Sigma Aldrich) and digested with trypsin (Promega, Madison, WI) using 1:50 protease to protein ratio. Following this, 5 μ g of peptides were desiccated and resuspended in 0.1% formic acid (Sigma Aldrich) water solution. Prior to the MS analysis, the sample was purified using ZipTip (Millipore, Billerica, MA).

MS-based Identification of Potentially Accessible Proteins

The peptide samples were analyzed using the 2D-nano-HPLC system Ultimate 3000 (Dionex, Sunnyvale, CA), which was

connected online with the electro spray ion-trap mass spectrometer Esquire HCT ultra (Bruker Daltonics, Bremen, Germany). A total of 5 μ g of sample was injected on the Bio-X-SCX column (500 μ m i.d. \times 15 mm; Dionex, p/n: 161395). Peptides were fractionated using four different concentrations of salt (45, 75, 150, and 500 mM ammonium acetate, 2% acetonitrile and 2% formic acid) and analyzed using the C18 analytical column (Acclaim 75 μ m \times 150 mm, p/n: 162224; Dionex, Sunnyvale, CA) with a gradient of 0–40% phase B (80% acetonitrile, 19.9% water and 0.1% formic acid) for 120 min at the flow rate of 0.3 μ L/min. Mass spectral data were acquired over a mass range from 200 to 1600 m/z . One full MS scan was automatically followed by four MS/MS scans of the four most intensive peptides found in this mass range (bearing +2 or +3 charges). The acquired spectra were processed for peak list generation using the Data Analysis software version 3.4 (Bruker Daltonics, Bremen, Germany). The MS/MS database search was conducted using the Mascot search engine version 2.2.2 (Matrix Sciences, Boston, MA). The human nonredundant and nonidentical protein database Swiss-Prot (Swiss Institute for Bioinformatics, Basel, Switzerland) version 57.7 was used and a total of 20 405 entries were searched. False positive rate of detection (FPR) was calculated using a scrambled version of the same database and reperforming the search against it. The number of false positive hits was expressed as a percentage compared to the original search with nonscrambled database.

Semiquantitative analysis (according to the previously published method⁷) was performed for the proteins that were identified in both tumoral and normal samples. Accordingly, the peak-list files from all the fractions were combined into one file for each individual and disease state. These combined files were submitted to the Swiss-Prot database using the Mascot search engine. The results yielded protein identifications along with the exponentially modified protein abundance index (emPAI). The emPAI is calculated using the equation

$$\text{emPAI} = 10^{(\text{Nobsd}/\text{Nobsbl})} - 1$$

where Nobsd is the number of observed and Nobsbl the number of observable peptides per protein. Ishihama et al.⁷ showed that the emPAI value is directly proportional to the protein quantity. Abundance ratios (protein emPAI tumoral vs normal) were calculated for each individual patient separately. Only proteins displaying scores above 100 were considered for the emPAI quantification. The normalization was conducted assuming the Gaussian distribution of the data, where the maximum number of the proteins has an emPAI abundance ratio of 1.0. Specifically, a median value was calculated and all the ratios in a given sample were divided with this value. The proteins with the increased emPAI abundance ratio (tumoral vs normal ≥ 2) were regarded as differentially expressed and included in further analysis.

Proteins that were not eligible for emPAI quantification (scores lower than 100 and/or not present in each disease condition) were screened for the frequency of presence or absence in all of the samples. Those proteins which were identified more frequently in the tumoral conditions than in the normal counterpart (in at least 2 individuals more) were also considered as potentially modulated. In the present study, no outlier correction was performed.

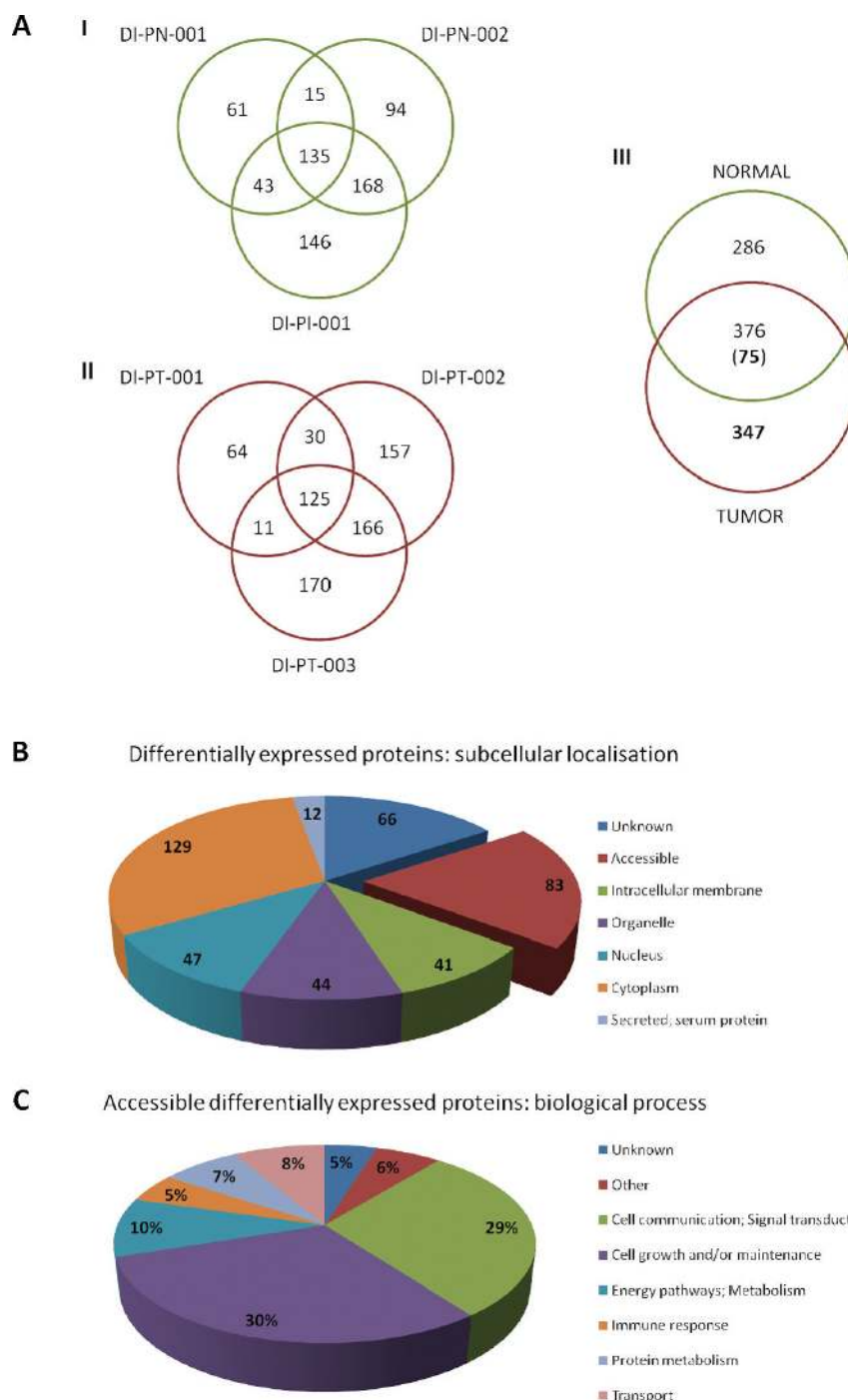


Figure 1. (A) [I–II] The assessment of the number of proteins found in common between the biological replicates of normal (N) and tumoral (T) pancreas samples. [III] Protein distribution between the normal and tumoral conditions. The number of proteins found as overexpressed is indicated in bold. A protein is considered modulated when (i) detected only in the tumoral condition, (ii) found in both conditions but more often identified in tumoral one or (iii) had a fold change ratio (emPAI tumor/normal) > 2. (B) Subcellular localization of the differentially expressed proteins. A protein is potentially accessible when it has been described as present at the cell surface or as deposited in the extracellular matrix. (C) Distribution of the accessible and differentially expressed proteins according to their biological process indicated in the Gene Ontology.

The subcellular localization of the modulated proteins was obtained using the Uniprot database, Gene Ontology information. Proteins localized in the plasma membrane and/or extracellular matrix were considered as potentially accessible. Known serum proteins and proteins located at the inner side of plasma membrane were not further considered.

■ VERIFICATION PHASE

Selected, potentially accessible proteins identified as modulated were further verified using alternative quantification strategies (including patients as indicated in the Table S2 of Supporting Information). Modulation of the following proteins was examined: asporin (ASPN), galectin-3 (LGALS3), latent-transforming growth

Table 1. Potentially Accessible Differentially Expressed Proteins Identified as Overexpressed in Pancreas Tumor Compared to Normal Tissue ($n = 3$)^a

accession	gene	protein name	N	T	emPAI T/N	subcellular location ^b	biological process ^b	biological function
P02751	FN1	Fibronectin	1	3	1.19	E	GM	angiogenesis; cell migration; collagen binding
P20774	OGN	Mimecan	2	3	2.19	E	Cc; St	growth factor activity
Q15063	POSTN	Periostin	2	3	5.63	E	Cc; St	cell adhesion; heparin binding; protein binding
P50395	GDI2	Rab GDP dissociation inhibitor beta	2	3	2.06	Cy; C	T	protein binding; protein transport; signal transduction
P12111	COL6A3	Collagen alpha-3(VI) chain	3	3	2.33	E	GM	cell adhesion
Q9BXN1	ASPN	Asporin	0	2		E	GM	bone mineralization; protein binding
P07384	CAPN1	Calpain-1 ^c	0	2		Cy; C	PM	cell proliferation; proteolysis
P23946	CMA1	Chymase	0	2		S	PM	proteolysis; regulation of inflammatory response
P08572	COL4A2	Collagen ^c	0	2		E	GM	angiogenesis; collagen type IV; protein binding
P16444	DPEP1	Dipeptidase 1	0	2		C	PM	protein binding; proteolysis; zinc ion binding
P08294	SOD3	Extracellular superoxide dismutase	0	2		E	M; Ep	heparin binding; oxidation reduction
P17931	LGALS3	Galectin3	0	2		N; CS; Cy, C	R	sugar binding
P22352	GPX3	Glutathione peroxidase 3	0	2		S	M; Ep	glutathione peroxidase activity; transcription factor binding
P04899	GNAI2	Guanine nucleotide-binding protein G ^c	0	2		C; Cy; ER; G	Cc; St	protein binding; signal transducer activity
Q00839	HNRNPU	Heterogeneous nuclear ribonucleoprotein U	0	2		N; CS	R	protein binding
Q16363	LAMA4	Laminin subunit alpha-4	0	2		E	GM	cell adhesion; regulation of cell migration
Q14767	LTBP2	Latent-transforming growth factor beta-binding protein 2	0	2		S	GM	calcium ion binding; growth factor binding
Q6UWY5	OLFML1	Olfactomedin-like protein 1	0	2		S	U	Unknown
P60903	S100A10	Protein S100-A10	0	2		C; N; Cy	Cc; St	calcium ion binding; receptor binding; signal transduction
Q13733	ATP1A4	Sodium/potassium-transporting ATPase subunit alpha-4	0	2		C	T	regulation of cellular pH
P35442	THBS2	Thrombospondin ^c	0	2		E	GM	calcium ion binding; cell adhesion; heparin binding; protein binding
Q15582	TGFBI	Transforming growth factor-beta-induced protein ig-h3	0	2		E	Cc; St	cell adhesion; integrin binding
P38606	ATP6 V1A	V-type proton ATPase catalytic subunit A	0	2		C	M; Ep	proton-transporting V-type ATPase
Q9NZN4	EHD2	EH domain-containing protein 2	0	2		C	Cc	ATP binding; calcium ion binding
Q01518	CAP1	Adenylyl cyclase-associated protein 1	1	2	3.08	C	GM	signal transduction
P46940	IQGAP1	Ras GTPase-activating-like protein IQGAP1	1	2	2.75	C	Cc; St	protein phosphatase binding
P07585	DCN	Decorin	2	2	2.17	E	GM	unknown
P51888	PRELP	Prolargin	2	2	10.36	E	GM	protein binding
P21980	TGM2	Protein-glutamine gamma-glutamyltransferase 2	2	2	8.50	E; C; Cy; N	M; Ep	calcium ion binding; cell adhesion
Q05682	CALD1	Caldesmon	3	2	2.05	Cy; C	GM	cytoskeleton binding
P08253	MMP2	72 kDa type IV collagenase	0	1		E	PM	calcium ion binding; collagen catabolic; proteolysis

Table 1. Continued

accession	gene	protein name	N	T	emPAI T/N	subcellular location ^b	biological process ^b	biological function
Q10588	BST1	ADP-ribosyl cyclase 2	0	1		C	I	unknown
O00468	AGRN	Aggrin	0	1		E	GM	laminin binding; signal transduction
P05186	ALPL	Alkaline phosphatase, tissue-nonspecific isozyme	0	1		C	M; Ep	alkaline phosphatase activity
P04745	AMY1A	Alpha-amylase 1	0	1		S	M	protein binding
P01019	AGT	Angiotensinogen	0	1		S	Cc; St	blood vessel remodeling
P20073	ANXA7	Annexin A7	0	1		N; C; ER; Cy	T	calcium ion binding
P80723	BASPN1	Brain acid soluble protein 1	0	1		C	R	promoter binding
P16870	CPE	Carboxypeptidase E	0	1		S	PM	protein modification process
P49961	ENTPD1	Ectonucleoside triphosphate diphosphohydrolase 1	0	1		C	M; Ep	calcium ion binding; cell adhesion
Q12805	EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1	0	1		S	GM	calcium ion binding; protein binding
P23142	FBLN1	Fibulin ^c	0	1		E	GM	calcium ion binding; protein binding
P01906	HLA-DQA2	HLA class II histocompatibility antigen ^c	0	1		C	I	immune response
P56199	ITGA1	Integrin alpha-1	0	1		C	Cc; St	calcium ion binding; cell-matrix adhesion
P24043	LAMA2	Laminin ^c	0	1		E	GM	cell adhesion
Q13449	LSAMP	Limbic system-associated membrane protein	0	1		C	Cc; St	cell adhesion; protein binding
Q93052	LPP	Lipoma-preferred partner	0	1		N; Cy; C	GM	cell adhesion
P01033	TIMP1	Metalloproteinase inhibitor 1	0	1		S	GM	protein binding
P55083	MFAP4	Microfibril-associated glycoprotein 4	0	1		E	GM	calcium ion binding; cell adhesion; receptor binding; signal transduction
Q9NZM1	MYOF	Myoferlin	0	1		C	GM	protein binding
P14543	NID1	Nidogen-1	0	1		E	GM	protein binding
P54315	PNLIPRP1	Pancreatic lipase-related protein 1	0	1		S	M; Ep	lipid catabolic process
Q13492	PICALM	Phosphatidylinositol-binding clathrin assembly protein	0	1		C	U	unknown
Q7ZSL7	PODN	Podocan	0	1		E	U	protein binding
P54707	ATP12A	Potassium-transporting ATPase alpha chain 2	0	1		C	T	proton transport
Q07954	LRP1	Prolow-density lipoprotein receptor-related protein 1	0	1		C	Cc; St	apolipoprotein bindin
P61224	RAP1B	Ras-related protein Rap-1b	0	1		C	Cc; St	protein binding
P54920	NAPA	Alpha-soluble NSF attachment protein	0	1		C	T	cellular membrane fusion
P35613	BSG	Basigin	0	1		C	Cc	protein binding
Q9UHD8	SEPT9	Septin-9	0	1		C	P	protein binding
P24821	TNC	Tenascin ^c	0	1		E	Cc; St	cell adhesion; receptor binding
P63010	AP2B1	AP-2 complex ^c	0	1		C	U	clathrin binding; protein transporter activity
Q0P6H9	TMEM62	Transmembrane protein 62	0	1		C	U	unknown
Q92597	NDRG1	Protein NDRG1	1	1	2.00	Cy; C; N	Cc; St	protein binding; response to metal ion

^a A protein is considered overexpressed when (i) it was detected only in the tumoral condition, (ii) found in both conditions but more often identified in tumoral with respect to normal samples, and (iii) found with a fold change ratio of the emPAI tumor/normal > 2. ^b Subcellular location: C, cell membrane; CS, cell surface; E, extracellular; N, nucleus; Cy, cytoplasm; S, secreted; ER, endoplasmic reticulum; G, golgi. Biological process: U, unknown; R, regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism; Cc, cell communication; St, signal transduction; GM, cell growth and/or maintenance; P, cell proliferation; I, immune response; M, metabolism; Ep, energy pathways; Pm, protein metabolism; T, transport. ^c Identified as multiple isoforms/subunits.

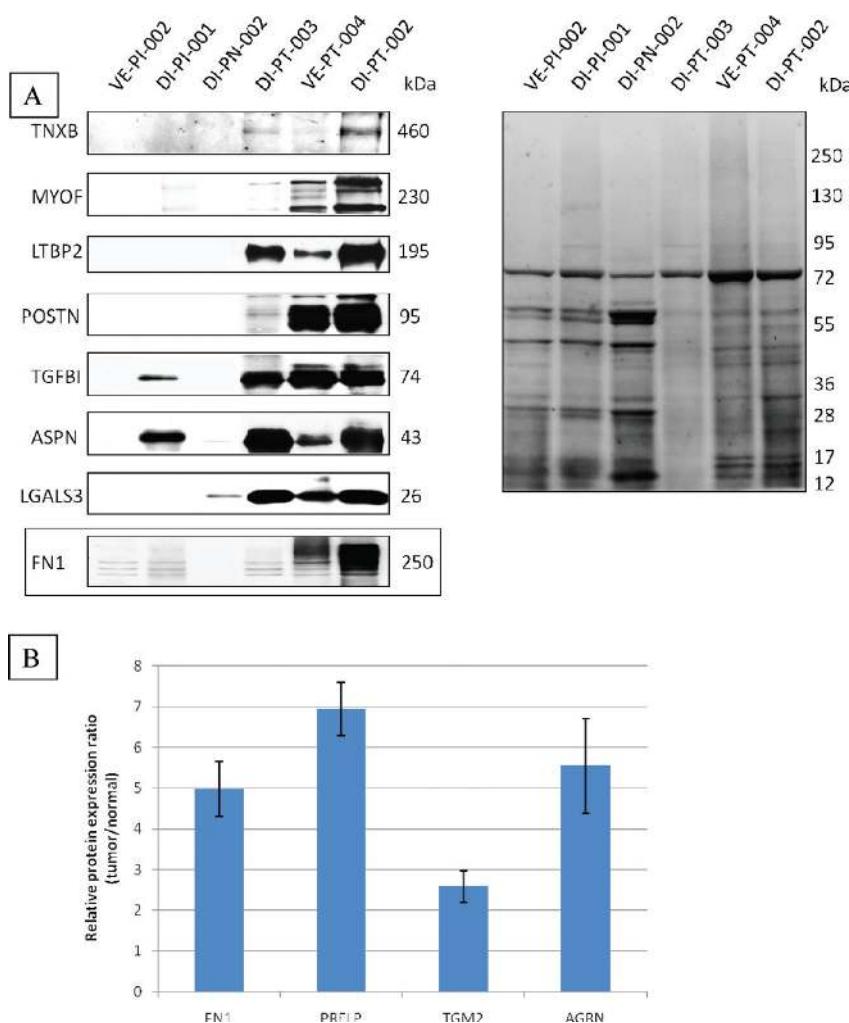


Figure 2. (A) Validation of ASPN, LGALS3, LTPB2, MYOF, POSTN, TNXB and TGFBI proteins using Western blot analysis. In addition, FN1 was quantified using both WB and MRM methods (see B). The sample identifier refers to patients outlined in the Supporting Information Table S2. The right pane shows the Ponceau red staining of the WB membrane following the transfer. (B) MRM-based validation of FN1, PRELP, TGM2 and AGRN modulated proteins. The relative quantification is calculated as a mean value with a corresponding standard deviation.

factor beta-binding protein-2 (LTPB2), myoferlin (MYOF), periostin (POSTN), tenascin-X (TNXB), transforming growth factor beta-induced (TGFBI), fibronectin (FN1), proline/arginine-rich end leucine-rich repeat protein (PRELP), transglutaminase 2 (TGM2) and agrin (AGRN). When suitable commercial antibodies were available, Western blotting was performed. Otherwise, the proteins were quantified using a targeted MS method, commonly known as MRM (multiple reaction monitoring). For detailed description of the WB and MRM methods refer to the Supporting Information.

■ VALIDATION PHASE

From the pool of proteins identified by proteomic analysis and confirmed to be largely overexpressed in the pancreas cancer tissues (using either WB or MRM), three were chosen (ASPN, TGFBI and LTPB2) to be further studied using IHC (immunohistochemistry) in human pancreas adenocarcinoma, pancreatitis, and normal pancreas tissue samples. The expression levels of selected antigens were also evaluated in a series of further normal tissues using IHC (Supporting Information). The detailed description of the IHC method is outlined in the Supporting Information.

■ RESULTS

Identification of Novel Modulated Proteins in Human Pancreatic Ductal Adenocarcinoma

The particular interest within the present work lied in the identification of differentially expressed proteins in human pancreas cancer, focused on a restricted area of the proteome: the potentially accessible proteins. In this study, three non-tumoral pancreas tissues and three tumor lesions were analyzed in the proteomic discovery phase using the ex vivo biotinylation procedure, which led to the identification of 662 and 723 proteins in the two respective conditions. Of these, 376 proteins were identified in both conditions. The average FPR was ~6% (normal and tumoral conditions). As shown in the Figure 1A I/II (both in normal and tumoral biological replicates), ~1/5 of identified proteins were observed in all three replicates, ~1/3 in two out of three replicates, and the remaining proteins were uniquely found in the respective replicate. The modulation of the identified proteins was assessed semiquantitatively primarily by their “presence” or “absence” in tumoral with respect to normal condition. Moreover, emPAI values were used to further semiquantify proteins found in both conditions.⁷ The emPAI value is

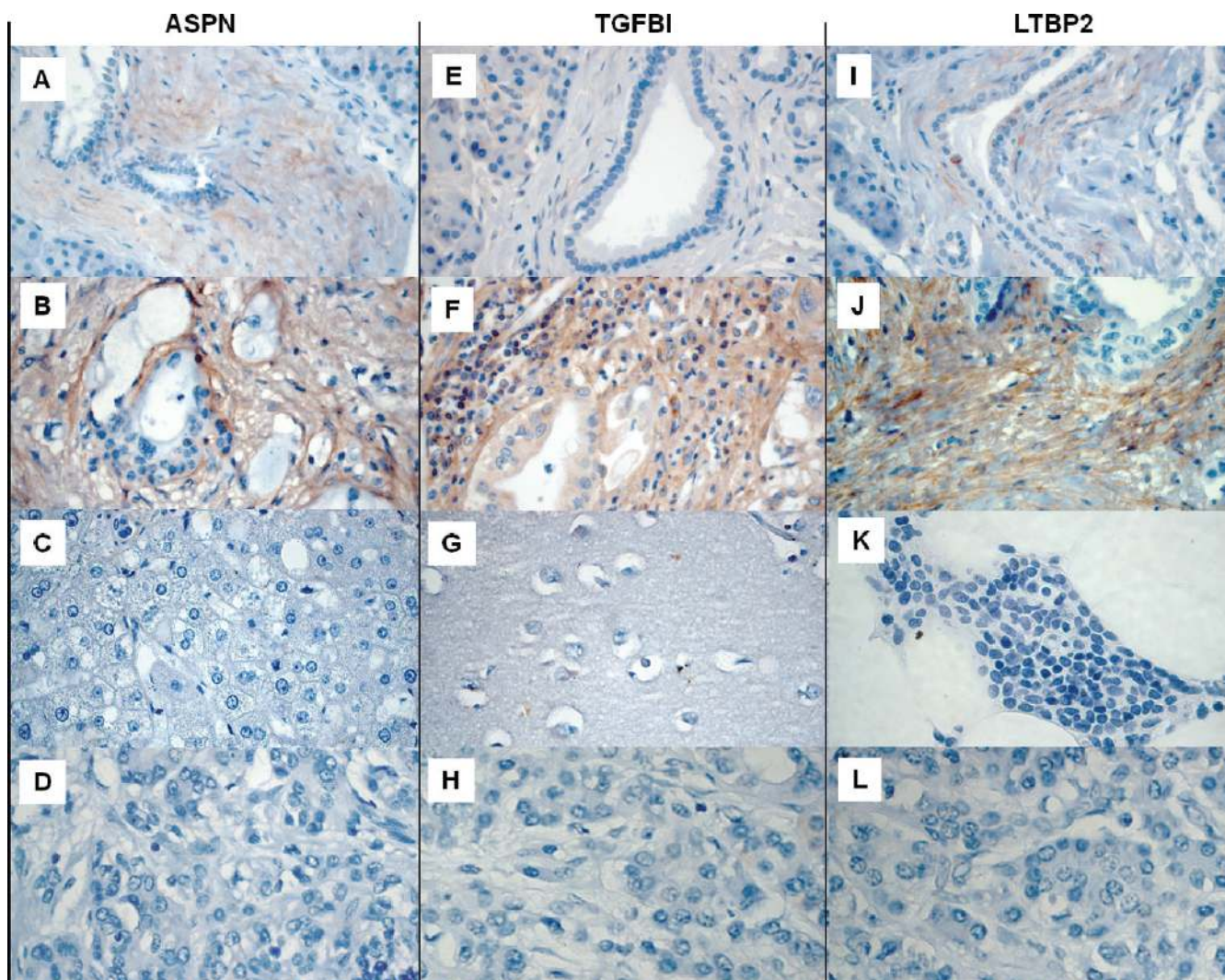


Figure 3. Immunohistochemical staining of the three selected modulated proteins in pancreas noncancerous tissue (normal pancreas and pancreatitis) and adenocarcinoma. The image shows ASPN expression in (A) normal peri-ductal tissue, (B) pancreatic adenocarcinoma and (C) normal liver (negative control 1). (D) Second negative control, where the primary antibody was omitted, was performed on pancreatic adenocarcinoma. TGFBI was (E) predominantly negative in the normal and (F) strongly positive in tumoral tissue. Corresponding negative controls for TGFBI were the (G) normal brain and (H) pancreatic adenocarcinoma with omitted primary AB. Immunostaining against LTBP2 indicated (J) high protein expression levels in tumoral compared to (I) normal tissue. (K) Normal thymus represents the negative control. (L) When primary AB was omitted, no positivity was detectable in the pancreatic adenocarcinoma.

an estimation of the protein abundance in complex mixtures. In the current study, 422 proteins were found as overexpressed in the pancreas tumor. Among these, 347 were exclusive to the tumoral samples while 75 were present in both conditions, however up-regulated in the tumoral one (Figure 1A III). Of the 422 proteins overexpressed in the tumoral condition, 83 proteins were eligible as potentially accessible (Figure 1B and Table 1). These included 64 distinct proteins, whereas the remaining part consisted of diverse protein isoforms, subunits, and family members. The score, the number of unique peptides, and the sequence coverage of each of the modulated proteins are outlined in the Supporting Information, Table S6. The biological profiles for the 83 differentially expressed and potentially accessible proteins were determined according to their Uniprot database Gene Ontology annotation (Figure 1C). The majority of the proteins were involved in cell communication and signal transduction (29%) as well

as cell growth and/or maintenance (30%). The other biological processes were less represented: protein metabolism (7%), general metabolism/energy pathways (10%) and transport (8%).

Western Blot and MRM Verification of Selected, Modulated, Proteins in Pancreas Cancer

Out of 83 modulated and potentially accessible proteins, 11 were selected for further verification regarding their differential expression in pancreas cancer vs normal tissues. The selection of the proteins was based on the MS results (overexpressed or uniquely detected in pancreas tumor), and the fact that they were not known as implicated in pancreas cancer. In particular the differential expression of asporin and myoferlin proteins has never been reported in any malignancies. The verification step for the 11 selected proteins was performed on three normal (including two chronic pancreatitis) and four (three for WB and

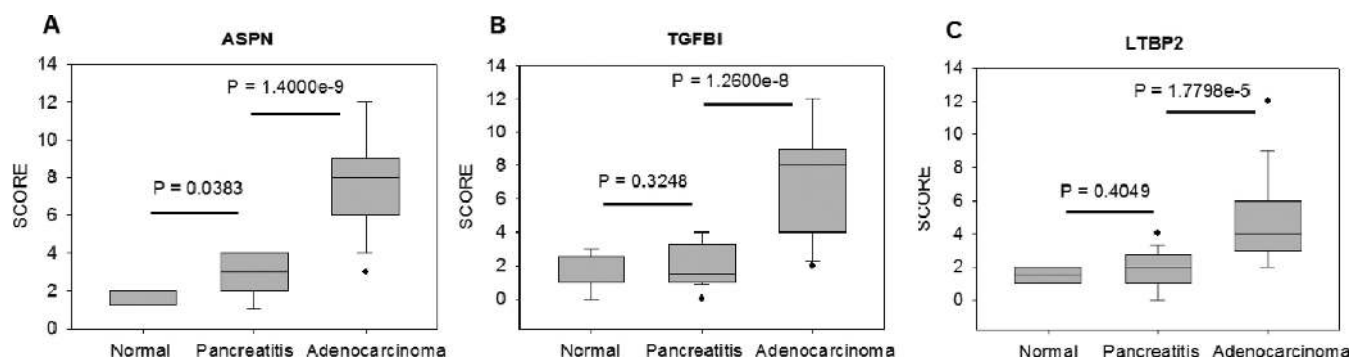


Figure 4. Immunohistochemical evaluation (Box-Plot) of the identified modulated proteins in pancreas noncancerous tissue (normal pancreas and pancreatitis) and adenocarcinoma. (A) ASPN: significant positivity was observed in the tumoral with respect to pancreatitis tissue ($P \leq 0.0001$); pancreatitis was only somewhat positive in comparison to the normal pancreas, however the difference was statistically significant ($P \leq 0.05$). (B) TGFBI: compared to pancreatitis, the positivity in the tumor was significant ($P \leq 0.0001$); the expression level of TGFBI in normal and pancreatitis tissues was not significantly different. (C) LTBP2: statistical testing indicated significant difference between these two groups ($P \leq 0.0001$); LTBP2 expression level in the normal tissue was not significantly different compared to pancreatitis.

four for MRM) pancreas adenocarcinoma tissue samples (detailed in the Table S2, Supplemental Data). WB was used for seven of these proteins because suitable antibodies were available while the four other were quantified using MRM. Additionally, FN1 was quantified using both verification strategies. As shown in Figure 2A, ASPN, LGALS3, LTBP2 and TGFBI were expressed in all cancer samples. MYOF, POSTN and TNXB were detectable in all the tumoral samples, however the proteins were found less expressed in the patient DI-PT-003 when compared to other individuals. Nontumoral specimens were largely negative for all the examined proteins, however TGFBI and ASPN tested positive in one of the two pancreatitis patients (DI-PI-001). FN1 was strongly positive for 2/3 tumoral samples, negative for one normal patient and weakly positive in the two inflammatory and one tumoral patient.

MRM was used for FN1, PRELP, TGM2 and AGRN and confirmed that all proteins were overexpressed in human pancreas cancers compared to the normal counterpart tissue (Figure 2B). Overall, the verification step enabled to confirm that 11 selected potentially accessible proteins are overexpressed in pancreas adenocarcinoma as found by the proteomic investigation in the Discovery Phase.

Immunohistochemistry Validation in Large Collection of Pancreas Human Tissue Samples

Following the verification phase, several proteins have prompted our interest to be further examined in a larger collection of normal, pancreatitis and adenocarcinoma samples. To our knowledge and out of the 11 proteins confirmed to be overexpressed in pancreas cancer, ASPN, TGFBI and LTBP2 are novel in the context of this malignancy. The expression of these three proteins were assessed using IHC on a series of tumoral ($n = 34$) and nontumoral ($n = 28$) pancreas tissues. The non-tumoral tissues were further subdivided in normal pancreas and chronic pancreatitis (Table S2, Supporting Information). This served to better characterize the tumor specificity of the selected proteins. As shown in Figure 3A–B and Figure 4A, ASPN was exclusively detected in the extracellular matrix (ECM) and no staining was observed in either normal or tumoral epithelial cells. ASPN expression was significantly ($p \leq 0.0001$) stronger in cancer specimens (score ≈ 8) in comparison with the inflammatory (score ≈ 3) and normal (score ≈ 2) tissues, where no or very low staining was noted. Although the inflammatory tissues showed only low positivity, the protein level of ASPN was significantly

elevated ($p \leq 0.05$) with respect to the normal specimen. The staining of tissue microarray encompassing a wide collection of normal human tissues showed no ASPN staining for adrenal gland (cortex and medulla), brain, breast, bronchus (mucous and cartilage), colon, endometrium, intestine (smooth muscle and villus), liver, myenteric plexus, pancreas, placenta, prostate (both gland and ducts), lung vein and artery, stomach, tongue, thymus, thyroid and umbilical cord. Some weak positivity (score ≈ 4) was found in gallbladder, kidney cortex, myometrium, esophagus and tongue. Strong staining was only detectable in the kidney pelvis (score ≈ 6) (detailed in Supporting Information, Table S1). As far as TGFBI is concerned, the staining was generally found in the ECM of the ductal adenocarcinoma tissues (Figure 3E–F and Figure 4B). Tumoral epithelial cells showed moderate cytoplasmic staining. This was not observed in the normal tissue. A strong ECM expression was found in 50% of all of the adenocarcinoma analyzed, while the remaining cases showed lower staining. Some ducts localized in the areas of chronic pancreatitis were moderately positive. Normal ducts showed no expression of the protein. Altogether, the average expression of TGFBI in tumoral tissue was significantly stronger ($p \leq 0.0001$) in the tumor (score ≈ 8) when compared to the inflammatory (score ≈ 2) and normal (score ≈ 2) pancreas tissue. In contrast to ASPN, TGFBI levels in chronic pancreatitis and normal tissues were not significantly different. Regarding the expression of TGFBI in other normal human tissues, the tissue microarray analysis indicated that brain, breast, colon, endometrium (inactive), intestine (smooth muscle and villus), myenteric plexus, prostate (gland), lung vein, thymus and thyroid were negative. Low positivity (score ≈ 3) was detectable in several tissues including adrenal gland (cortex and medulla), gallbladder, liver, esophagus, tongue and prostate (ducts) (detailed in the Supporting Information, Table S1). LTBP2 immunoreactivity was mainly found in tumoral ductal and peri-ductal stromal tissue (Figure 3I–J and Figure 4C). More than 50% of the adenocarcinoma showed a strong expression of the protein, which was significantly higher ($p \leq 0.0001$) in the tumoral (score ≈ 5) comparing to the inflammatory (score ≈ 2) and normal (score ≈ 2) tissue. LTBP2 protein expression in normal and pancreatitis tissues was not significantly different. Concerning the expression of LTBP2 in further normal tissue, the IHC experiments demonstrated that this protein was not

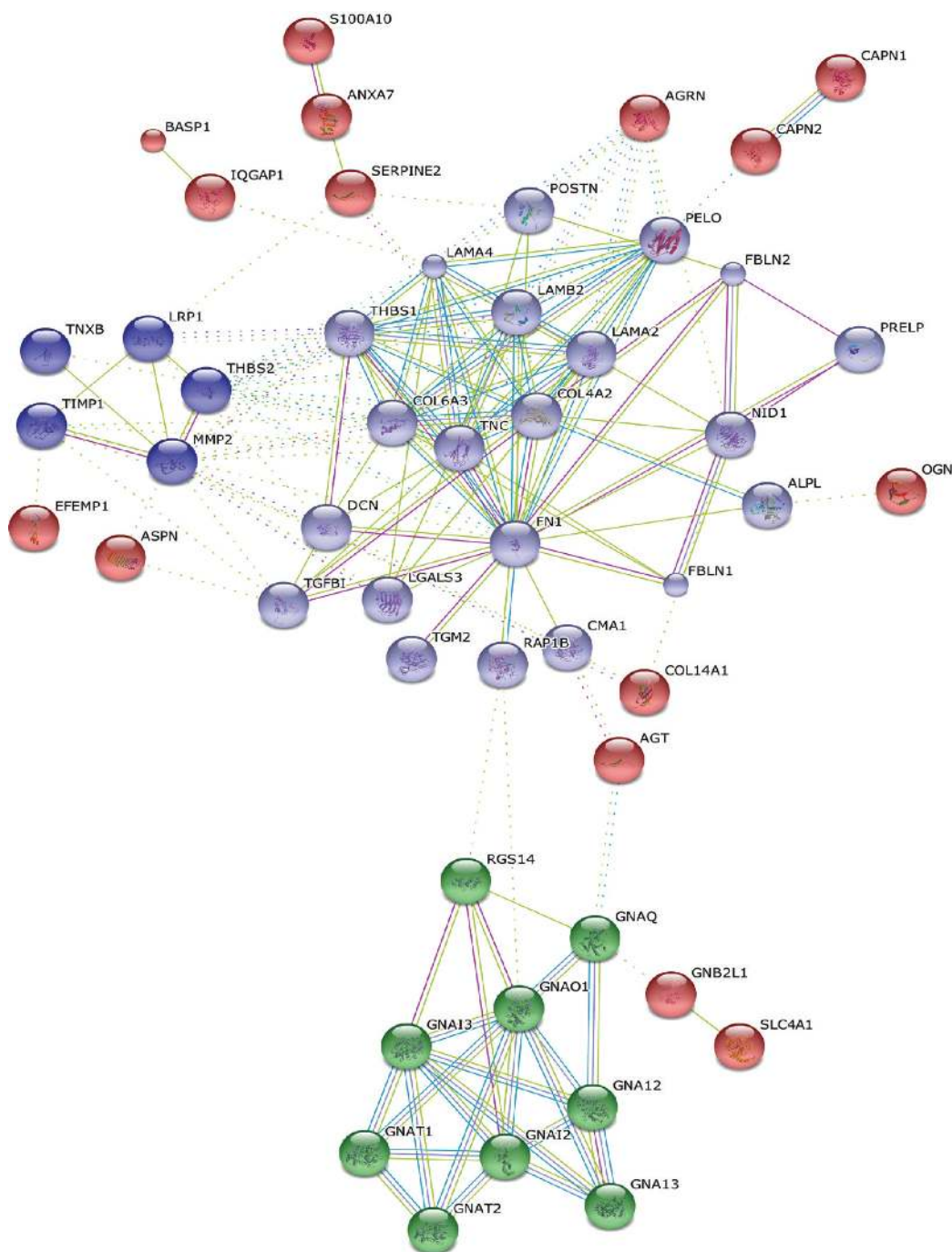


Figure 5. K-means clustering and functional network of modulated proteins in pancreas cancer. The individual proteins are listed in the Table 1. Networking and clustering was performed using the STRING software (for details: <http://string-db.org>). In the k-means clustering analysis, the k value was set at 4. The four individual colors denote the 4 highlighted clusters.

present in most of them: adrenal gland (cortex), brain, breast, colon, endometrium, intestine (smooth muscle and villus), liver, myenteric plexus, esophagus, prostate (gland), lung vein, tongue (muscle), thymus, thyroid and umbilical cord. Weak positivity (score ≈ 2) was detectable in few tissues including the medulla of the adrenal gland, epithelium of the tongue, mucosa of the bronchus and lung artery. Complete evaluation of all normal tissues is provided in the Supporting Information Table S1. Some examples of normal tissues tested are shown in the Figure S2 (Supporting Information).

Bioinformatic Analysis and Potential Networking of Modulated Proteins Found in Pancreas Cancer

The pool of proteins found to be potentially accessible and modulated in human pancreas adenocarcinoma were further examined using the STRING program⁸ (Figure 5). The rationale behind this analysis was to tease out and recognize the most relevant biological clusters concerning the accessible proteome in pancreas cancer. Two separate cycles of analysis were performed. First, the modulated proteins were connected according to their biological coexpression, coregulation and function they

share. Second analysis comprised of k-means clustering, where *k* variable was set at the value of 4. Using this, three distinct clusters were observable, within which proteins shared particular similarity regarding their biological role. The fourth cluster (proteins in red) contained all proteins which were either not linked enough to form a separate cluster or are linked to one of the three main clusters but were functionally not similar enough to be included in that respective cluster. The first cluster (marked dark blue) consists of 5 proteins concentrated around matrix metalloproteinase 2 (MMP2). The second group (light blue) is composed primarily of proteins forming the extracellular matrix. The most prominent proteins within this cluster are the members of collagen and laminin family. The last group (indicated in green) is almost exclusively composed of guanine nucleotide-binding protein members.

DISCUSSION

The MS analysis performed in this study led to the identification of 1009 proteins among which 422 were overexpressed in pancreas ductal adenocarcinoma. Of these proteins and owing to the technology of *ex vivo* biotinylation, 83 (including protein isoforms) were identified as potentially accessible and targetable because they are reported to be on the outer side of the membrane and/or in the extracellular matrix (Table 1). Noteworthy, a large proportion of the identified proteins had unknown subcellular localization (*n* = 66). A considerable number of these modulated proteins could be accessible, but due to the current uncertainty, they were not further considered here. Several proteins that were grouped in other cellular localizations were also reported (or predicted) to be membrane-bound or extracellular. As far as the literature research allowed, these proteins were further investigated and appropriately included in the list of potentially accessible proteins.

While up to now there has not been any similar proteomic investigation focused on the systemically reachable proteins, several studies have searched pancreatic cancer biomarkers at the global proteome level. Chen et al. used isotope-coded affinity tag (ICAT) to profile pancreatic cancer tissue.⁹ Other studies employed the two-dimensional polyacrylamide gel-electrophoresis (2D-PAGE) method.^{10–15} Here it is important to note that the 2D-PAGE is problematic in resolving hydrophobic proteins while it is this group of proteins that bears the highest interest because of their cell membrane location and possibility to be accessible. Although no enrichment of the membrane proteome was performed, these studies led to the identification of a certain number of potentially accessible proteins (Supporting Information, Table S3). Noteworthy is especially the last column in the Table S3, which displays the number of potentially accessible proteins (evaluated according to the criteria used in the present study). Most of these were also identified in the present study (Supporting Information, Table S4). In summary, 27 proteins have been reported. Eleven of them were also identified modulated in the present study. Interestingly, our study has shed light on several novel proteins previously not reported to be related to the process of pancreas ductal adenocarcinoma. One of these is myoferlin (MYOF), whose function seems to be necessary for the proliferation of endothelial cells in response to VEGFR-2 signaling.¹⁶ This finding could suggest that cancer cells use the overexpression of this protein to accelerate angiogenesis allowing faster tumor growth.¹⁶ Similarly, Tenascin-X has never been described in the context of pancreatic cancer, while its

overexpression has been recently described in malignant mesothelioma, both at mRNA and protein level.¹⁷ Further investigations are needed to clarify if MYOF and TNXB are consistently overexpressed in human pancreas cancer as well as to examine their role in pancreas cancer progression.

In the present study, the expression of LTBP2, TGFBI and ASPN are further investigated employing a large collection of human pancreas cancer tissues. These candidates have previously not been associated with this malignancy at the level of protein expression. LTBP2 colocalizes with fibronectin and collagen type I in extracellular matrix and is involved in extracellular matrix regulation and remodeling.^{18,19} Although there is apparent dearth of data describing the function of LTBP2, two studies have suggested that TGF-beta 1 increases the transcription of LTBP2 gene, which participates in cell adhesion processes.^{20,21} The expression of LTBP2 in normal and pancreatitis patients was not statistically different. This observation encourages the idea that the up-regulation of LTBP2 protein is probably tumor specific.

TGFBI is another protein found overexpressed in pancreas ductal adenocarcinoma. The expression of TGFBI was reported to be increased at the mRNA level in pancreatic cancers and the overexpression of the protein was found in other human cancers such as colorectal, glial, and clear cell renal tumors.^{22–25} The immunohistochemistry results confirmed on a larger series of specimens the increased expression of TGFBI in the ECM of the pancreas carcinoma. The analysis also showed that there was no difference between normal and pancreatitis patients with respect to the expression of TGFBI. This result indicates the possibility that the overexpression of this protein is more specific to cancer. One protein that appeared to be completely novel in the context of pancreas adenocarcinoma is asporin. This protein is a new member of the class I small leucine-rich repeat proteoglycan (SLRP) family. The ASPN mRNA is expressed in some adult and fetal tissues with the highest concentrations in aorta and uterus and very weak levels in pancreas.²⁶ The expression of ASPN protein was however predominantly identified in human dentin and periodontal ligament.²⁷ Regarding its function, ASPN and TGF-beta1 form a regulatory feedback loop in particular cartilage cells, where ASPN inhibits TGF-beta/Smad signaling by colocalizing with TGF-beta1 on the cell surface.²⁸ Up to now, overexpression of ASPN protein in cancer is unknown and only one report using DNA microarrays has revealed ASPN mRNA as up-regulated in breast cancer.²⁹ In the current study, the proteomic investigation revealed the overexpression of ASPN in the pancreas carcinoma and has provided the information that this protein might be an accessible biomarker. The validation experiments based upon IHC have demonstrated the significant overexpression of ASPN in a large group of pancreas carcinoma patients. However, the comparison of the normal and pancreatitis IHC data suggest that ASPN might be overexpressed in inflammation conditions in addition to the cancerous state. The data appear to strengthen the available information, demonstrating that the increased levels of this protein most probably represent the stromal reaction to an inflammatory or inflammatory like condition. Taken together and considering the IHC data on the normal tissues from different organs, all three proteins may prove to be of value for subsequent investigations aiming at diagnostic and therapeutic applications.

Given the large overlap between the current study and previously published data, it was of interest to examine all the modulated proteins using a suitable bioinformatic tool. For this

purpose, the publicly available STRING software was used. The main cluster of modulated proteins found in the current study includes collagen and laminin family members as well as other predominantly ECM proteins. Additionally, the data revealed a cancer relevant group of proteins consisting of matrix metalloproteinase 2 (MMP2) and related proteins. MMPs are the main actors of cancer progression.³⁰ They are secreted by cancer cells as pro-enzymes and are consequently activated in the extracellular milieu. Finally, a third group of proteins consists of several different guanine-binding protein isoforms. These are very important in relationship with growth factor, hormone or neurotransmitter signal transduction cascades, regulating cell motility and proliferation. Current data leave no additional space for further clarification regarding the implication of these proteins in the context of pancreatic carcinoma. Certainly, further work is required to clarify whether these protein groups are in fact participating in one of the above-mentioned signaling cascades and what is their relevance with respect to pancreas adenocarcinoma.

■ ASSOCIATED CONTENT

Supporting Information

Supplemental data, tables, and figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Vincent Castronovo, M.D., Ph.D., Head of Metastasis Research Laboratory, GIGA-Cancer, University of Liege. Fax: 003243662975. E-mail: vcastronovo@ulg.ac.be.

■ ACKNOWLEDGMENT

This work was supported by a grant from the Research Concerted Action (IDEA project) of the University of Liège (ULG), Belgium, from the CEE (FP7 network: ADAMANT-Antibody Derivatives As Molecular Agents for Neoplastic Targeting (HEALTH-F2-2007-201342)), from the National Fund for Scientific Research (NFSR, Belgium) and TELEVIE as well as from the Centre Anti-Cancéreux of the ULG. We acknowledge the support of GIGA-Proteomics Platform of the ULG for experimental support and are particularly thankful to Pascale Heneaux for IHC experiments.

■ REFERENCES

- (1) Jemal, A.; Siegel, R.; Ward, E.; Hao, Y.; Xu, J.; Thun, M. J. Cancer statistics, 2009. *CA Cancer J. Clin.* **2009**, *59* (4), 225–49.
- (2) Duffy, M. J.; Sturgeon, C.; Lamerz, R.; Haglund, C.; Holubec, V. L.; Klapdor, R.; Nicolini, A.; Topolcan, O.; Heinemann, V. Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report. *Ann. Oncol.* **2010**, *21* (3), 441–7.
- (3) Philip, P. A.; Mooney, M.; Jaffe, D.; Eckhardt, G.; Moore, M.; Meropol, N.; Emens, L.; O'Reilly, E.; Korc, M.; Ellis, L.; Benedetti, J.; Rothenberg, M.; Willett, C.; Tempero, M.; Lowy, A.; Abbruzzese, J.; Simeone, D.; Hingorani, S.; Berlin, J.; Tepper, J. Consensus report of the national cancer institute clinical trials planning meeting on pancreas cancer treatment. *J. Clin. Oncol.* **2009**, *27* (33), S660–9.
- (4) Tonack, S.; ASPNinall-O'Dea, M.; Neoptolemos, J. P.; Costello, E. Pancreatic cancer: proteomic approaches to a challenging disease. *Pancreatology* **2009**, *9* (5), S67–76.

- (5) Turtoi, A.; De Pauw, E.; Castronovo, V. Innovative proteomics for the discovery of systemically accessible cancer biomarkers suitable for imaging and targeted therapies. *Am. J. Pathol.* **2011**, *178* (1), 12–8.

- (6) Castronovo, V.; Kischel, P.; Guillonueau, F.; de Leval, L.; Defécherieux, T.; De Pauw, E.; Neri, D.; Waltregny, D. Identification of specific reachable molecular targets in human breast cancer using a versatile ex vivo proteomic method. *Proteomics* **2007**, *7*, 1188–1196.

- (7) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **2005**, *4* (9), 1265–72.

- (8) Jensen, L. J.; Kuhn, M.; Stark, M.; Chaffron, S.; Creevey, C.; Muller, J.; Doerks, T.; Julien, P.; Roth, A.; Simonovic, M.; Bork, P.; von Mering, C. STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **2009**, No. Database, D412–6.

- (9) Chen, R.; Yi, E. C.; Donohoe, S.; Pan, S.; Eng, J.; Cooke, K.; Crispin, D. A.; Lane, Z.; Goodlett, D. R.; Bronner, M. P.; Aebersold, R.; Brentnall, T. A. Pancreatic cancer proteome: the proteins that underlie invasion, metastasis, and immunologic escape. *Gastroenterology* **2005**, *129*, 1187–97.

- (10) Chung, J. C.; Oh, M. J.; Choi, S. H.; Bae, C. D. Proteomic analysis to identify biomarker proteins in pancreatic ductal adenocarcinoma. *ANZ J. Surg.* **2008**, *78*, 245–51.

- (11) Qi, T.; Han, J.; Cui, Y.; Zong, M.; Liu, X.; Zhu, B. Comparative proteomic analysis for the detection of biomarkers in pancreatic ductal adenocarcinomas. *J. Clin. Pathol.* **2008**, *61*, 49–58.

- (12) Lu, Z.; Hu, L.; Evers, S.; Chen, J.; Shen, Y. Differential expression profiling of human pancreatic adenocarcinoma and healthy pancreatic tissue. *Proteomics* **2004**, *4*, 3975–88.

- (13) Shen, J.; Person, M. D.; Zhu, J.; Abbruzzese, J. L.; Li, D. Protein expression profiles in pancreatic adenocarcinoma compared with normal pancreatic tissue and tissue affected by pancreatitis as detected by two-dimensional gel electrophoresis and mass spectrometry. *Cancer Res.* **2004**, *64*, 9018–26.

- (14) Shekouh, A. R.; Thompson, C. C.; Prime, W.; Campbell, F.; Hamlett, J.; Herrington, C. S.; Lemoine, N. R.; Crnogorac-Jurcevic, T.; Buechler, M. W.; Friess, H.; Neoptolemos, J. P.; Pennington, S. R.; Costello, E. Application of laser capture microdissection combined with two-dimensional electrophoresis for the discovery of differentially regulated proteins in pancreatic ductal adenocarcinoma. *Proteomics* **2003**, *3*, 1988–2001.

- (15) Mikuriya, K.; Kuramitsu, Y.; Ryozaawa, S.; Fujimoto, M.; Mori, S.; Oka, M.; Hamano, K.; Okita, K.; Sakaida, I.; Nakamura, K. Expression of glycolytic enzymes is increased in pancreatic cancerous tissues as evidenced by proteomic profiling by two-dimensional electrophoresis and liquid chromatography-mass spectrometry/mass spectrometry. *Int. J. Oncol.* **2007**, *30*, 849–55.

- (16) Bernatchez, P. N.; Acevedo, L.; Fernandez-Hernando, C.; Murata, T.; Chalouni, C.; Kim, J.; Erdjument-Bromage, H.; Shah, V.; Gratton, J. P.; McNally, E. M.; Tempst, P.; Sessa, W. C. Myoferlin regulates vascular endothelial growth factor receptor-2 stability and function. *J. Biol. Chem.* **2007**, *282* (42), 30745–53.

- (17) Yuan, Y.; Nymoan, D. A.; Stavnes, H. T.; Rosnes, A. K.; Bjørang, O.; Wu, C.; Nesland, J. M.; Davidson, B. Tenascin-X is a novel diagnostic marker of malignant mesothelioma. *Am. J. Surg. Pathol.* **2009**, *33* (11), 1673–82.

- (18) Mangasser-Stephan, K.; Gartung, C.; Lahme, B.; Gressner, A. M. Expression of isoforms and splice variants of the latent transforming growth factor beta binding protein (LTBP) in cultured human liver myofibroblasts. *Liver* **2001**, *21*, 105–113.

- (19) Gabrielsen, A.; Lawler, P. R.; Yongzhong, W.; Steinbruchel, D.; Blagoja, D.; Paulsson-Berne, G.; Kastrup, J.; Hansson, G. K. Gene expression signals involved in ischemic injury, extracellular matrix composition and fibrosis defined by global mRNA profiling of the human left ventricular myocardium. *J. Mol. Cell. Cardiol.* **2007**, *42*, 870–83.

- (20) Vehvilainen, P.; Hyytiainen, M.; Keski-Oja, J. Latent transforming growth factor-beta-binding protein 2 is an adhesion protein for melanoma cells. *J. Biol. Chem.* **2003**, *278*, 24705–13.

(21) Ahmed, W; Kucich, U; Abrams, W; Bashir, M; Rosenbloom, J; Segade, F; Mecham, R. Signaling pathway by which TGF-beta1 increases expression of latent TGF-beta binding protein-2 at the transcriptional level. *Connect. Tissue Res.* **1998**, *37*, 263–76.

(22) Ivanov, S. V.; Ivanova, A. V.; Salnikow, K; Timofeeva, O; Subramaniam, M; Lerman, M. I. Two novel VHL targets, TGFBI (BIGH3) and its transactivator KLF10, are up-regulated in renal clear cell carcinoma and other tumors. *Biochem. Biophys. Res. Commun.* **2008**, *370*, 536–40.

(23) Ma, C; Rong, Y; Radloff, D. R.; Datto, M. B.; Centeno, B; Bao, S; Cheng, A. W.; Lin, F; Jiang, S; Yeatman, T. J.; Wang, X. F. Extracellular matrix protein betaig-h3/TGFBI promotes metastasis of colon cancer by enhancing cell extravasation. *Genes Dev.* **2008**, *22*, 308–21.

(24) Golembieski, W. A.; Rempel, S. A. cDNA array analysis of SPARC-modulated changes in glioma gene expression. *J. Neurooncol.* **2002**, *60*, 213–26.

(25) Yamanaka, M; Kimura, F; Kagata, Y; Kondoh, N; Asano, T; Yamamoto, M; Hayakawa, M. BIGH3 is overexpressed in clear cell renal cell carcinoma. *Oncol. Rep.* **2008**, *19*, 865–874.

(26) Lorenzo, P; Aspberg, A; Onnerfjord, P; Bayliss, M. T.; Neame, P. J.; Heinegard, D. Identification and characterization of asporin, a novel member of the leucine-rich repeat protein family closely related to decorin and biglycan. *J. Biol. Chem.* **2001**, *276*, 12201–11.

(27) Park, E. S.; Cho, H. S.; Kwon, T. G.; Jang, S. N.; Lee, S. H.; An, C. H.; Shin, H. I.; Kim, J. Y.; Cho, J. Y. Proteomics Analysis of Human Dentin Reveals Distinct Protein Expression Profiles. *J. Proteome Res.* **2009**, *8* (3), 1338–46.

(28) Nakajima, M; Kizawa, H; Saitoh, M; Kou, I; Miyazono, K; Ikegawa, S. Mechanisms for asporin function and regulation in articular cartilage. *J. Biol. Chem.* **2007**, *282* (44), 32185–92.

(29) Turashvili, G; Bouchal, J; Baumforth, K; Wei, W; Dziechciarkova, M; Ehrmann, J; Klein, J; Fridman, E; Skarda, J; Srovnal, J; Hajdich, M; Murray, P; Kolar, Z. Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. *BMC Cancer* **2007**, *7*, 55.

(30) Martin, M. D.; Matrisian, L. M. The other side of MMPs: protective roles in tumor progression. *Cancer Metastasis Rev.* **2007**, *26* (3–4), 717–24.