Application of a Global Proteomic Approach to Archival Precursor Lesions: Deleted in Malignant Brain Tumors 1 and Tissue Transglutaminase 2 Are Upregulated in Pancreatic Cancer Precursors

Wang Cheung a  Marlene M. Darfler d  Hector Alvarez a  Brian L. Hood e  Thomas P. Conrads e  Nils Habbe a  David B. Krizman d  Jan Mollenhauer f, g  Georg Feldmann a  Anirban Maitra a–c

Departments of a Pathology, b Oncology and c Genetic Medicine, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins University School of Medicine, Baltimore, Md., d Expression Pathology Incorporated, Gaithersburg, Md., and e Department of Pharmacology, University of Pittsburgh Cancer Institute, University of Pittsburgh School of Medicine, Pittsburgh, Pa., USA; f Department of Molecular Genome Analysis, Deutsches Krebsforschungszentrum, Heidelberg, Germany; g Department of Molecular Oncology, Institute of Medical Biology, University of Southern Denmark, Odense, Denmark

Key Words
Pancreatic cancer · Intraductal papillary mucinous neoplasms · Proteomics · Mass spectrometry

Abstract
Background: Pancreatic cancer is an almost uniformly fatal disease, and early detection is a critical determinant of improved survival. A variety of noninvasive precursor lesions of pancreatic adenocarcinoma have been identified, which provide a unique opportunity for intervention prior to onset of invasive cancer. Biomarker discovery in precursor lesions has been hampered by the ready availability of fresh specimens, and limited yields of proteins suitable for large scale screening. Methods: We utilized Liquid Tissue®, a novel technique for protein extraction from archival formalin-fixed material, and mass spectrometry to conduct a global proteomic analysis of an intraductal papillary mucinous neoplasm (IPMN). Tissue microarrays comprised of 38 IPMNs were used for validation of candidate proteins. Results: The proteomic analysis of the IPMN Liquid Tissue lysate resulted in identification of 1,534 peptides corresponding to 523 unique proteins. A subset of 25 proteins was identified that had previously been reported as upregulated in pancreatic cancer. Immunohistochemical analysis for two of these, deleted in malignant brain tumors 1 (DMBT1) and tissue transglutaminase 2 (TGM2), confirmed their overexpression in IPMNs. Conclusion: Global proteomics analysis using the Liquid Tissue workflow is a feasible approach for unbiased biomarker discovery in limited archival material, particularly applicable to precursor lesions of cancer.

Introduction
Pancreatic cancer is the fourth most common cause of cancer-related mortality in the United States, accounting for 32,000 deaths annually [1]. The overwhelming majority of patients present with advanced metastatic disease, rendering their tumors inoperable. Diagnosis of pancreatic cancer at an early stage provides the best option for successfully treating and improving the dire prognosis of this malignancy. It is now well established that pancrea-
Pancreatic adenocarcinomas do not arise de novo, but rather progress through a series of histologically distinct non-invasive precursor lesions [2, 3]. The three most common subtypes of pancreatic cancer precursors include pancreatic intraepithelial neoplasia (PanIN), intraductal papillary mucinous neoplasm (IPMN), and mucinous cystic neoplasm (MCN). Surgical resection of noninvasive cystic precursors (IPMN and MCNs) is typically curative, while the presence of any invasive component confers an adverse prognosis on outcome [4, 5]. The identification of biomarkers of pancreatic cancer precursors has the potential to yield diagnostic candidates that will facilitate the early detection of these lesions in at-risk individuals [6].

Global proteomic analysis has emerged as a promising strategy to elucidate potential biomarkers in various cancer subtypes [7–10]. Unfortunately, these analyses typically mandate the use of snap-frozen or fresh cancer tissues, or the use of cancer cell lines, precluding the direct analysis of smaller, precursor lesions. The ability to perform global proteomic analyses from microscopic precursor lesions in archival tissues would greatly expand the repertoire of specimens for cancer biomarker discovery. Standard formalin-fixed, paraffin-embedded (FFPE) samples are abundant in surgical pathology archives and permit the accurate histological designation of precursor lesions in these cases, without the presence of contaminating cancer cells. Several approaches have recently been described for mass spectrometric analysis of peptides from FFPE specimens, including cancer [11–13], as well as non-cancerous tissues [14, 15]. The Liquid Tissue® platform is a technology that permits facile global proteomic analysis of archival material by mass spectrometry [16] and has been utilized for identifying candidate biomarkers of prostate cancer from microdissected, formalin-fixed radical prostatectomy specimens [17], and most recently, from archival sections of head and neck squamous cell carcinomas (HNSCCs) [18]. Notably, these reports confirm that formalin fixation does not significantly alter the ability to generate tryptic peptides for subsequent global proteomic analysis by mass spectrometry.

The objectives of this pilot study were twofold. The first goal was to establish that global proteomic approaches are not limited to cancer cells, but can also be extended to the archival precursor lesions, using the Liquid Tissue platform. The second goal was toward elucidation of specific proteins whose dysregulation is associated with an early stage in the multistep progression of pancreatic cancer, by performing this analysis on a noninvasive IPMN. Two significant candidate marker proteins for IPMN were identified by mass spectrometry, deleted in malignant brain tumors 1 (DMBT1) and tissue transglutaminase 2 (TGM2). The expression levels of these proteins were validated in a larger series of IPMNs by immunohistochemistry on tissue microarrays (TMAs). These results confirm that global proteomic analysis of archival precursor lesions is feasible, and can facilitate the identification of candidate biomarkers applicable in subsequent translational research.

**Materials and Methods**

*Tissue Processing and Sample Preparation*

An FFPE block of a noninvasive IPMN with carcinoma in situ was selected for obtaining epithelial cells. A representative hematoxylin and eosin stained section was utilized to confirm the pathological diagnosis of IPMN by established criteria [19] (fig. 1). For tissue collection and protein preparation, a single 10-μm-thick tissue section was cut from the tissue block, placed on a standard microscope slide, and heated for 1 h at 60°C. Paraffin was removed with SubX (Surgipath Medical Industries, Richmond, Ill., USA) followed by tissue rehydration through a series of graded ethanol solutions and distilled water. Approximately 30,000 cells from the relevant region were procured by manual (needle) microdissection for processing, as described previously [16–18]. The collected tissue was processed with the Liquid Tissue MS Protein Prep Kit according to manufacturer’s recommendations (Expression Pathology, Gaithersburg, Md., USA). Briefly, the tissue was suspended in 20 μl of Liquid Tissue buffer, incubated at 95°C for 90 min, then cooled on ice for 2 min, at which time 1 μg of sequencing-grade porcine trypsin (Promega, Madison, WI) was added and the solution was incubated for an additional 90 min. The solution was then desalted on spin columns (140 μl; Millipore, Bedford, Mass., USA) and the eluate was dried using a SpeedVac (Thermo Fisher Scientific, Waltham, Mass., USA) before being reconstituted in Liquid Tissue buffer. 

![Fig. 1. Photomicrograph of HE-stained, noninvasive IPMN used for microdissection and Liquid Tissue lysate preparation.](https://example.com/fig1)
Antigen retrieval was performed by incubating with H9262 surgically resected specimens at The Johns Hopkins Hospital; all chival IPMN samples for TMA construction were obtained from minimum delta correlation score ([20, 21]. Gene ontology (GO) annotation was performed using a ‘Batch Gene Finder’ (http://cgap.nci.nih.gov/Genes/BatchGeneFinders) and nucleic acid, DNA and protein binding.

Proteomic analysis of the IPMN Liquid Tissue lysate using nanoLC-MS/MS yielded a total of 1,534 peptides, corresponding to 869 total proteins and 523 unique proteins, which are presented in their entirety in online suppl. table 1 (www.karger.com/doi/10.1159/000161012). Further, 95 proteins (~18%) were identified by two or more unique tryptic peptides. In online supplement table 2, we have provided the SEQUEST parameters (Xcorr, ΔCn, Sp, Rsp, and the coverage of γ- and β-ions), which were determined in order to reduce the numbers of false positives, and increase the proportion of true positives. The calculated false-positive rate was 3.26%. Representative tandem mass spectra, including peptides corresponding to three of the proteins identified as expressed in IPMN (cathepsin D, DMBT1, and clusterin), are shown in figure 2. Additional annotation for each unique protein is provided in online supplement table 1, including NCBI Entrez ID, unigene cluster for the corresponding transcript, cytogenetic location, and gene ontology (GO) classification by molecular function, process, and component. Summarized graphically in online supplement figure 1 is the GO molecular function, demonstrating that the top categories of functional proteins identified include those involved in structural and catalytic activity, and nucleic acid, DNA and protein binding.

It is likely that the vast majority of proteins whose peptides were identified by LC-MS/MS in the IPMN Liquid Tissue lysate are constitutively expressed ‘native’ proteins in the pancreas (including those expressed in the contaminating stromal component). Upon manual curation, however, the list of expressed proteins in the noninvasive IPMN included several whose mRNA transcripts are overexpressed in invasive pancreatic cancer by global mRNA expression profiling technologies (SAGE, cDNA and oligonucleotide microarrays) [24–31]. Comparison with quantitative proteomic studies of invasive pancreatic cancer [32, 33] also revealed additional examples of upregulated proteins that were detected in the IPMN. This ‘enriched’ subset of 25 previously identified cancer-

Liquid Chromatography-Tandem Mass Spectrometry and Bioinformatic Analysis

Liquid chromatography (LC) was performed using a Dionex Ultimate 3000 nanoflow LC system (Dionex Corporation, Sunnyvale, Calif., USA) coupled on-line to a linear ion trap (LIT) mass spectrometer (MS) (LTQ; ThermoFisher Scientific Inc., San Jose, Calif., USA). Separation of the sample was performed using a 75-μm inner diameter ×360 outer diameter ×10 cm-long fused silica capillary column (Polymeric Technologies, Phoenix, Ariz., USA) packed in house with 5 μm, 300 Å pore size Jupiter C-18 stationary phase (Phenomenex, Torrance, Calif., USA). After injecting 250 ng of the global protein digest, the column was washed with 98% mobile phase A (0.1% formic acid in water) for 30 min and peptides were eluted by development of a linear gradient of 2% mobile phase B (0.1% formic acid in acetonitrile) to 42% mobile phase B in 140 min, then to 98% B in an additional 20 min, all at a constant flow rate of 250 nl/min. The LTQ-MS was operated in a data-dependent MS/MS mode in which each full MS scan (precursor ion selection scan range of m/z 350–1,800) was followed by seven MS/MS scans where the seven most abundant peptide molecular ions dynamically determined from the MS scan were selected for tandem MS using a relative collision-induced dissociation (CID) energy of 35%. Dynamic exclusion was utilized to minimize redundant selection of peptides for CID. Tandem mass spectra were searched against the UniProt Homo sapiens proteome database (http://www.expasy.org) using SEQUEST (ThermoFisher Scientific, Inc.). Peptides were considered legitimately identified if they achieved specific charge state and proteolytic cleavage-dependent cross-correlation (Xcorr) scores of 1.9 for [M+H]+, 2.2 for [M+2H]2+, and 3.1 for [M+3H]3+, and a minimum delta correlation score (ΔCn) of 0.08. Further manual annotation of the identified proteins was performed using publicly available internet-accessible tools, including the NCBI CGAP ‘Batch Gene Finder’ (http://cgap.nci.nih.gov/Genes/BatchGeneFinder). Gene ontology (GO) annotation was performed using a full-featured version of the Expression Analysis Systematic Explorer (EASE, version 2.0 for Windows operating systems) available at http://abcc.ncbi.nlm.nih.gov [20, 21].

Immunohistochemistry

TMAs comprising tissue cores from 38 IPMNs were used for immunohistochemistry as previously described [22, 23]. The archival IPMN samples for TMA construction were obtained from surgically resected specimens at The Johns Hopkins Hospital; all samples were de-linked from direct patient identifiers. Briefly, 3-μm sections of the IPMN TMAs were deparaffinized by routine techniques. Antigen retrieval was performed by incubating with DAKO low pH antigen retrieval reagent (1:10 dilution) for 20 min at 95°C, followed by 20 min of cooling to ambient temperature. Primary and secondary incubation steps for anti-deleted in malignant breast tumor 1 (anti-DMBT1h12, 1/100 dilution, mouse monoclonal) and anti-tissue transglutaminase 2 (anti-TGM2, 1/100 dilution of rabbit polyclonal; Neomarkers; LabVision Corp., Fremont, Calif., USA) were performed as described previously [24, 25]. Scoring of TMAs utilized a four tier classification (0, 1, 2 and 3) where less than 5% of labeled epithelial cells was scored as ‘0’, 5–10% of epithelial labeling was scored as ‘1’, 10–30% of epithelium labeling was scored as ‘2’, and >30% of cells labeling was scored as ‘3’.
associated proteins in IPMN (table 1) represents a source of markers with aberrant expression at the earliest stages of pancreatic neoplasia. For further validation in a larger series of IPMN TMAs we selected two proteins, DMBT1 and TGM2, which have – to the best of our knowledge – so far never been reported to be overexpressed in IPMNs. However, we and others have previously described their upregulation in invasive pancreatic cancers [24, 25, 34–36].

Immunohistochemical analysis confirmed the overexpression of both DMBT1 and TGM2 in IPMN tissues (fig. 3, 4). Specifically, 2–3+ staining was observed in 15 of 38 (39%) IPMNs labeled with anti-DMBT1 antibody, and in 12 of 33 (36%) cases labeled with anti-TGM2 (five IPMN cores stained with anti-TGM2 were not able to be evaluated) (table 2). While DMBT1 was primarily expressed in the IPMN epithelium (fig. 3), TGM2 expression was also seen in the juxta-tumoral stroma, including in occasional cases where the epithelium itself did not express TGM2 (fig. 4). In contrast, minimal or no DMBT1 or TGM2 expression was observed in the non-neoplastic pancreas, including sections of normal intralobular ducts when these were available for evaluation. Although the numbers of cases are limited, we did not observe any sig-
significant correlation between aberrant protein expression and histologic grade of the IPMN epithelium (adenoma, moderate dysplasia, or carcinoma in situ) [19, 37].

### Discussion

Few public health measures have had as much success in ameliorating cancer mortality as early detection at an operable, and hence potentially curable, stage [38, 39]. The National Cancer Institute recognizes this and has therefore invested considerable efforts in early detection strategies [40, 41]. In order to implement effective early diagnosis of cancer, two prerequisite tenets need to be met. First, there must be a recognizable precursor stage before invasive cancer develops which provides a window of opportunity for preventative intervention. This criterion has now been met for most common cancers [42–46], including pancreatic cancer, where tangible noninvasive precursor lesions are recognized, and for which standardized diagnostic criteria have been established [2, 37]. The second requirement is the identification of sensi-

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene symbol</th>
<th>Unigene ID</th>
<th>Number of peptides</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annexin A1</td>
<td>ANXA1</td>
<td>Hs.494173</td>
<td>2</td>
<td>[25–27, 31, 33]</td>
</tr>
<tr>
<td>Annexin A2</td>
<td>ANXA2</td>
<td>Hs.511605</td>
<td>1</td>
<td>[32, 33]</td>
</tr>
<tr>
<td>Caldesmon</td>
<td>CALD1</td>
<td>Hs.490203</td>
<td>4</td>
<td>[33]</td>
</tr>
<tr>
<td>Cathepsin D</td>
<td>CTSD</td>
<td>Hs.654447</td>
<td>2</td>
<td>[32]</td>
</tr>
<tr>
<td>Cytokeratin-7</td>
<td>KRT7</td>
<td>Hs.670221</td>
<td>4</td>
<td>[25, 26, 28]</td>
</tr>
<tr>
<td>Cytokeratin-8</td>
<td>KRT8</td>
<td>Hs.533782</td>
<td>8</td>
<td>[27]</td>
</tr>
<tr>
<td>Cytokeratin-19</td>
<td>KRT19</td>
<td>Hs.654568</td>
<td>4</td>
<td>[25, 27, 28]</td>
</tr>
<tr>
<td>Cofilin</td>
<td>CFL1</td>
<td>Hs.170622</td>
<td>2</td>
<td>[32]</td>
</tr>
<tr>
<td>Collagen type I, alpha 1</td>
<td>COL1A1</td>
<td>Hs.172928</td>
<td>32</td>
<td>[26, 30]</td>
</tr>
<tr>
<td>Collagen type I, alpha 2</td>
<td>COL1A2</td>
<td>Hs.489142</td>
<td>19</td>
<td>[26, 28]</td>
</tr>
<tr>
<td>Collagen type IV, alpha 3</td>
<td>COL4A3</td>
<td>Hs.570065</td>
<td>3</td>
<td>[32]</td>
</tr>
<tr>
<td>Deleted in malignant brain tumors 1</td>
<td>DMBT1</td>
<td>Hs.279611</td>
<td>1</td>
<td>[24]</td>
</tr>
<tr>
<td>Fibrilllin 1</td>
<td>FBN1</td>
<td>Hs.591133</td>
<td>4</td>
<td>[32]</td>
</tr>
<tr>
<td>Fibrinogen gamma chain</td>
<td>FGG</td>
<td>Hs.546255</td>
<td>1</td>
<td>[32]</td>
</tr>
<tr>
<td>Gelsolin</td>
<td>GSN</td>
<td>Hs.522373</td>
<td>5</td>
<td>[32]</td>
</tr>
<tr>
<td>Lumican</td>
<td>LUM</td>
<td>Hs.406475</td>
<td>5</td>
<td>[32]</td>
</tr>
<tr>
<td>Peroxiredoxin 1</td>
<td>PRDX1</td>
<td>Hs.180909</td>
<td>1</td>
<td>[27]</td>
</tr>
<tr>
<td>Phosphoglycerate kinase 1</td>
<td>PGK1</td>
<td>Hs.654578</td>
<td>5</td>
<td>[27]</td>
</tr>
<tr>
<td>Plectin 1</td>
<td>PLEC1</td>
<td>Hs.434248</td>
<td>3</td>
<td>[25, 32]</td>
</tr>
<tr>
<td>Rho GDP Dissociation Inhibitor alpha</td>
<td>ARHGDIA</td>
<td>Hs.159161</td>
<td>2</td>
<td>[31]</td>
</tr>
<tr>
<td>Transglutaminase 2</td>
<td>TGM2</td>
<td>Hs.517033</td>
<td>4</td>
<td>[25–27, 29, 32]</td>
</tr>
<tr>
<td>Transgelin 2</td>
<td>TAGLN2</td>
<td>Hs.517168</td>
<td>1</td>
<td>[32]</td>
</tr>
<tr>
<td>Tropomysin 5</td>
<td>TPM4</td>
<td>Hs.631618</td>
<td>1</td>
<td>[27]</td>
</tr>
<tr>
<td>Tubulin alpha-1 chain</td>
<td>TUBA1A</td>
<td>Hs.654422</td>
<td>4</td>
<td>[27]</td>
</tr>
<tr>
<td>Vimentin</td>
<td>VIM</td>
<td>Hs.642813</td>
<td>21</td>
<td>[29]</td>
</tr>
</tbody>
</table>

Five TMA cores dropped out and were therefore not evaluable in the TGM2 series. DMBT1 labeling is restricted essentially to the IPMN epithelium, while TGM2 labeling is observed in both the neoplastic epithelium and the juxta-tumoral stroma.
Reactive and specific surrogate biomarkers that can provide ancillary information to routine clinical assays for diagnosing the presence of high-risk precursor lesions [41, 47]. The latter remains a work in evolution. For example, in the case of pancreatic cancer, the most commonly utilized serological marker – CA19-9 – has suboptimal performance in the diagnosis of early pancreatic neoplasia, and is probably best utilized in monitoring disease recurrence in patients who have already received therapy [48, 49].

Large-scale proteomic approaches have provided a major impetus to the search for new biomarkers in cancers and in their precursor lesions [8, 50]. One of the major limitations of proteomic applications in clinical samples has been the requirement for high-quality snap-frozen or live tissue specimens [51], essentially excluding the repositories of fixed, archival paraffin-embedded tissues available in most institutions. Further, the current approach to biomarker discovery in precursor lesions involves the retrospective analysis of tumor antigens which are first identified in invasive cancers (‘reverse proteomics’) [52]. An unbiased ‘forward’ proteomics analysis of precursor lesions might directly elucidate biomarkers that are aberrantly expressed at the earliest stages of multistep cancer development.

Fig. 3. DMBT1 labeling in IPMN tissue microarrays. a, b Two examples of noninvasive IPMNs with intense DMBT1 expression in the neoplastic epithelium. c A representative example of a DMBT1-expressing invasive adenocarcinoma arising in an IPMN. Notice the absence of labeling in the stroma. d Example of IPMN with no expression of DMBT1 protein.
progression. The Liquid Tissue platform has previously been applied to the discovery of differentially expressed proteins in formalin-fixed paraffin embedded tissues [16–18]. In this study, we have applied the Liquid Tissue platform to elucidate the proteome of a noninvasive precursor lesion of pancreatic cancer (IPMN). We have confirmed the upregulation of two of the expressed proteins (DMBT1 and TGM2) in IPMNs versus normal pancreatic tissues by immunohistochemistry, underscoring the validity of our approach for tumor marker discovery in precursor lesions. Although our proof-of-principle analysis in IPMNs was not quantitative in nature, the Liquid Tissue platform has been shown to be amenable to quantitative proteomic analyses (e.g. stable isotope labeling) for biomarker discovery in prostate cancer [17]. Nevertheless, by simple manual curation, we were able to focus on those proteins most likely to be aberrantly expressed in IPMNs compared to normal pancreas based on their known association with the invasive pancreatic cancer proteome (table 1). One of the validated proteins, DMBT1, belongs to the superfamily of scavenger receptor cysteine-rich proteins and is a secreted glycoprotein [53]. Loss of DMBT1 expression was first recognized in brain tumors, but overexpression of this protein has now been documented in

---

**Fig. 4.** TGM2 labeling in IPMN TMAs. **a** Absence of TGM2 labeling in non-neoplastic pancreatic acini and in small intra-lobular ducts (arrow); labeling is observed only in endothelial cells. **b, c** Two examples of noninvasive IPMNs with intense TGM2 expression in the neoplastic epithelium. Unlike DMBT1, TGM2 expression is often observed in the juxta-tumoral stroma (d), even in the absence of demonstrable labeling in the IPMN epithelium (arrowhead).
pancreatic cancers [24, 34]. Hence, our findings presented here, demonstrating for the first time overexpression of DMBT1 in IPMNs, are in line with these previous reports and suggest that upregulation of DMBT1 is a comparatively early event in pancreatic carcinogenesis. TGM2 plays a critical role in catalyzing the calcium-dependent transamidation of polypeptides, but is unique in that it also has protein isomerase and kinase activities [54]. Several lines of evidence support a role for TGM2 in cancer biology [55], with distinct and potentially contrasting effects of expression in neoplastic cells [56] and in the juxta-tumoral stroma [57]. We and others have established TGM2 overexpression as a phenotype of invasive pancreatic cancer [25, 36], and this study extends this finding to IPMNs as well.

In conclusion: In the present study, we show that our recently described technology for protein extraction from archival tissue sample material is amenable to mass spectrometric analysis of tryptic peptides for analyzing the proteome of noninvasive pancreatic cancer precursor lesions. We hope that application of the Liquid Tissue platform to proteomic analysis of archival formalin-fixed tissues will prove to be a valuable additional tool for biomarker discovery in the future.

Acknowledgements

The Sol Goldman Pancreatic Cancer Research Center, the Michael Rolfe Foundation, and the NIH SPORE (Specialized Programs of Research Excellence) in Gastrointestinal Cancer P50CA62924. H.A. received generous support from the D’Amato Family. G.F. was supported by a fellowship grant within the post-doctoral program of the German Academic Exchange Service (DAAD). J.M. received support from the Future Award in Health Sciences 2005 of the HGF.

Marlene M. Darfler and David B. Krizman are full-time employees of Expression Pathology Incorporated, Gaithersburg, Md. Other than that, the authors declare no conflict of interest.

References

Pancreatology 2008;8:608–616


Cheung et al.