Cholangiocellular carcinoma (CCC) and pancreatic ductal adenocarcinoma (PDAC) are two highly aggressive cancer types that arise from epithelial cells of the pancreaticobiliary system. Owing to their histological and morphological similarity, differential diagnosis between CCC and metastasis of PDAC located in the liver frequently proves an unsolvable issue for pathologists. The detection of biomarkers with high specificity and sensitivity for the differentiation of these tumor types would therefore be a valuable tool. Here, we address this problem by comparing microdissected CCC and PDAC tumor cells from nine and eleven cancer patients, respectively, in a label-free proteomics approach. The novel biomarker candidates were subsequently verified by immunohistochemical staining of 73 CCC, 78 primary, and 18 metastatic PDAC tissue sections. In the proteome analysis, we found 180 proteins with a significantly differential expression between CCC and PDAC cells (p value < 0.05, absolute fold change > 2). Nine candidate proteins were chosen for an immunohistochemical verification out of which three showed very promising results. These were the annexins ANXA1, ANXA10, and ANXA13. For the correct classification of PDAC, ANXA1 showed a sensitivity of 84% and a specificity of 85% and ANXA10 a sensitivity of 90% at a specificity of 66%. ANXA13 was higher abundant in CCC.

This is an open access article under the CC BY license.

From the ‡Medizinisches Proteom-Center, Ruhr-Universität Bochum, Germany; §Institut für Pathologie, Universitätsklinikum Essen, Universität-Duisburg-Essen, Germany; ¶Institut für Allgemeinchirurgie, Viszeral- und Transplantationschirurgie, Universitätsklinikum Essen, Universität-Duisburg-Essen, Germany; ‖Klinik für Pathologie und Neuropathologie, Abteilung Allgemeine Pathologie, Universitätsklinikum Tübingen, Germany

Received August 17, 2015, and in revised form, November 9, 2015
Published, MCP Papers in Press, DOI 10.1074/mcp.M115.054585


It presented a sensitivity of 84% at a specificity of 55%. In metastatic PDAC tissue ANXA1 and ANXA10 showed similar staining behavior as in the primary PDAC tumors (13/18 and 17/18 positive, respectively). ANXA13, however, presented positive staining in eight out of eighteen secondary PDAC tumors and was therefore not suitable for the differentiation of these from CCC. We conclude that ANXA1 and ANXA10 are promising biomarker candidates with high diagnostic values for the differential diagnosis of intrahepatic CCC and metastatic liver tumors deriving from PDAC. Molecular & Cellular Proteomics 15: 10.1074/mcp.M115.054585, 1072–1082, 2016.

The majority of malignant neoplasms located in the liver are metastases originating from primary tumor sites in other organs, most commonly the colon or the pancreas (1). In many cases, a histological or immunohistological examination by an experienced pathologist can specify the type and origin of the underlying cancer. Hepatocellular carcinoma or hepatic metastasis from primaries such as pulmonary adenocarcinoma, colorectal adenocarcinoma, and breast carcinoma are usually easily distinguishable by morphology and means of known immunohistochemical markers. For primary cholangiocellular carcinoma (CCC) and metastases of pancreatic ductal adenocarcinoma (PDAC), however, the distinction in a liver biopsy is basically an unsolvable task because of their high similarity.

1 The abbreviations used are: CCC, cholangiocellular carcinoma; ANXA, annexin; AUC, area under the curve; BH, Benjamini-Hochberg; CKMT1A, creatine kinase U-type mitochondrial; CV, coefficient of variation; EGFR, epidermal growth factor receptor; FDR, false discovery rate; FLNB, filamin-B; HPC2, human pancreatic cancer fusion protein #2; IHC, immunohistochemistry; IRS, immuno-reactive score; LCM, laser capture microdissection; MAOB, Amine oxidase [flavin-containing] B; MMP, matrix metalloproteinases; (m)PDAC, (metastases of) pancreatic ductal adenocarcinoma; NTLT, nontumorous liver tissue; PODXL-1, podocalyxin-like protein 1; ROC, receiver operating characteristics; TGF, transforming growth factor.
This is an important and frequently asked clinical question, because treatment options differ significantly for the two cancer types. In the case of CCC, a surgical approach can be beneficial if the diagnosis is made at an early stage of tumor progression. In contrast, palliation is often the only option if the cancer’s origin is the pancreas (2).

In most cases, pathologists rely on supporting information gained from radiologic or sonographic examinations to enable a differential diagnosis. High-quality imaging may detect the majority of pancreatic masses. However, these techniques, especially MRI, which would be the most suitable, are not available in every clinic. Furthermore, radiologic examinations are higher in cost than e.g. an immunohistochemical analysis and if a metastatic PDAC in the liver is suspected, a biopsy is mandatory before palliative chemotherapy in any case (3). Hence, immunohistochemical biomarkers supporting pathologists in the differential diagnosis of CCC and PDAC would be of vital importance.

Although there are genomic differences in CCC and PDAC, the value of these in daily practice is uncertain. For example, mutations in isocitrate dehydrogenase IDH1 and IDH2 were found only in cholangiocarcinomas of intrahepatic origin (about one third). Extrahepatic cholangiocarcinomas do not show these mutations (4). Furthermore, 20–54% of intrahepatic CCC tumors harbor k-ras-mutations in contrast to 90% of PDAC cases (5). Nevertheless, neither IDH nor KRAS mutations appear to be suitable for distinction between the two tumor entities.

Although several immunohistochemical markers have been tested in regard to this challenge, so far, none have presented results sufficient for a clinical implementation. In 2007, Ney et al. described the use of podocalyxin-like protein 1 (PODXL-1) for differentiating PDAC from adenocarcinomas of the biliary and gastrointestinal tracts. The immunohistochemical study revealed the expression of PODXL-1 in 44% of the PDAC cases (71/160), whereas none of the intrahepatic (0/18) and only one of the extrahepatic CCC (1/13) were stained (6). The protein agrin, on the other hand, was proposed to aid in differential diagnosis of primary and metastatic cancers to the liver because of different expression patterns. Although in PDAC it showed a faint staining over broad areas, CCC tissue was stained stronger and more extensively (2). A combination of two markers, N-cadherin and the antibody human pancreatic cancer fusion protein #2 (HPC2), was suggested by Hooper et al. who showed that N-cadherin is expressed predominantly in CCC tissue whereas HPC2 stained a higher number of PDAC cases. Combining both markers significantly increased specificity, but at the expense of sensitivity (7).

Likewise, a biomarker panel consisting of four proteins (S100P, pVHL, MUC5AC, and CK17) was recently described to be helpful in distinguishing intrahepatic CCC from metastatic PDAC. There, the staining pattern S100P+/pVHL+/MUC5AC−/CK17− was observed in 59% of the 41 CCC and none of the 60 PDAC cases. However, PDAC metastases were not tested (8). It becomes clear, that despite the effort, there is to date no biomarker or marker panel that can distinguish reliably between these two entities.

While in all of the above mentioned studies the selection of the candidates was hypothesis-driven, we here conducted a biomarker discovery study comparing CCC and PDAC cells by highly sensitive and accurate proteomics techniques to identify novel biomarker candidates. To overcome possible bias because of the high heterogeneity of CCC and PDAC tumor tissue, the malignant cells were isolated specifically by laser capture microdissection. Promising candidate proteins resulting from the quantitative proteomics analysis were subsequently verified by immunohistochemistry (IHC) in a cohort of 73 CCC and 78 PDAC sections as well as thirteen metastatic liver tumors from PDAC.

**EXPERIMENTAL PROCEDURES**

**Clinical Data**—In total, tumor tissue from 189 cancer patients was collected (supplemental Table S1). Intrahepatic CCC tissue and corresponding nonmalignant liver tissue from 84 patients was excised during partial hepatectomy at the University Hospital of Essen, Department of General, Visceral, and Transplantation Surgery, Germany. All tumors were diagnosed according to current WHO-criteria (9). All cases were intrahepatic cholangiocarcinoma, grossly the mass-forming type (MF-type) with adenocarcinoma morphology. Rare histologic variants and combined hepatocellular-cholangiocarcinoma were not included. Tissue from 87 primary PDAC tumors and eighteen metastatic liver tumors from PDAC was collected during surgery at University Hospital of Essen and University Hospital of Tübingen, Germany. Informed consent was obtained from each patient and the study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki. The local ethics committee approved the study (11–4839-BO).

**Experimental Design**—For this study, three sample cohorts were assembled. The set for the proteome analysis comprised 11 CCC and nine PDAC samples (cohort 1). Experience from similar studies conducted previously has proven sample sizes in this range sufficient to detect reproducible proteome patterns and therefore reveal potential biomarkers that were successfully verified in subsequent experiments (10–12). From each sample, three technical replicates were analyzed by LC-MS/MS to improve overall protein identification and quantification accuracy.

A set of 10 CCC and 10 PDAC tumors was composed for a first immunohistochemical screening of candidates in a time and material saving way (cohort 2). This set included four of the CCC samples used for the proteome analysis and the other 16 from cohort 3, assembled for the verification (supplemental Table S1). They were chosen according to availability of material. Cohort 3, which was independent from cohort 1, was used for the verification of candidate proteins by immunohistochemistry. It comprised 73 CCC, 78 PDAC, and 18 PDAC metastases from the liver (mPDAC). The aim of the immunohistochemical study was the first-time estimation and assessment of the diagnostic accuracy of the selected markers in a large cohort. Therefore, the highest number of samples available to us was included in order to improve the estimates and decrease the size of the confidence intervals. The number of PDAC metastases available was heavily limited because only cases with a definite diagnosis were included.

More detailed information concerning the sample cohort composition is summarized in Table I. Fig. 1 visualizes the study design.

**Proteome Analysis**—

**Tissue Preparation**—Directly after surgery tissue samples were frozen in liquid nitrogen and stored at −80 °C. For laser capture...
microdissection the tissue was cryosectioned at $-20^\circ$C. The 10 $\mu$m sections were placed on 1.0 polyethylene naphthalate (PEN) membrane-covered slides (Carl Zeiss Microscopy, Thornwood, NY) and dehydrated by dipping into 70% ethanol for 5 s. Cell nuclei were stained for 30 s with 1% (w/v) cresyl violet in 50% ethanol. To remove excess cresyl violet and dehydrate the tissue sections, slides were dipped five times in 70% ethanol, once in 100% ethanol and then air dried.

**Laser Capture Microdissection**—Selective isolation of malignant cells from the cryo-sectioned tumor tissue was achieved by laser capture microdissection (LCM) on a Palm MicroBeam (Carl Zeiss Microscopy). A total area of 1 mm$^2$ was dissected from each sample and collected in 15 $\mu$L 50 mM ammonium bicarbonate with 0.1% RapiGest SF Surfactant (Waters, Eschborn, Germany). Until further processing, samples were stored at $-80^\circ$C.

**In-solution Digestion**—Cell lysis was performed by sonicating twice for 1 min 20 s on ice in an ultrasonic bath and subsequent incubation at 95°C for 5 min. Proteins’ disulfide bonds were reduced with 5 mM DTT at 60°C for 30 min and alkylated with 15 mM iodoacetamide for 30 min at ambient temperature in the dark. For enzymatic protein digestion, 50 ng trypsin (SERVA Electrophoresis, Heidelberg, Germany) was added and samples were incubated at 37°C for 16 h.

**TABLE I**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Analysis method</th>
<th>n</th>
<th>Sex</th>
<th>Age</th>
<th></th>
<th>n</th>
<th>Sex</th>
<th>Age</th>
<th></th>
<th>n</th>
<th>Sex</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LC-MS</td>
<td>11</td>
<td>9</td>
<td>2</td>
<td>31</td>
<td>77</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>IHC</td>
<td>10</td>
<td>9</td>
<td>1</td>
<td>47</td>
<td>74</td>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>IHC</td>
<td>73</td>
<td>47</td>
<td>26</td>
<td>28</td>
<td>81</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIG. 1.** The discovery and verification of biomarker candidates for the differential diagnosis of CCC and PDAC. The experimental workflow used for this study can be divided into two phases (A). During the discovery phase, the protein list resulting from the LC-MS experiment was filtered in several steps (B). This led to the selection of proteins for the verification by immunohistochemistry (IHC). Abbreviations: m/z, mass to charge ratio; RT, retention time; ROC, receiver operating characteristic; $p_{FDR}$, t test p value after adjustment (Benjamini, Hochberg); FC, fold change.
Acidification with 0.5% TFA for 30 min at 37 °C ensured quenching of enzyme activity and precipitation of hydrolyzed RapiGest. The latter was removed by centrifugation. The supernatant was dried in a centrifugal evaporator and resolved in 0.1% TFA. From each sample containing 1 mm² tumor cell area 1/10 (corresponding to 100,000 μm²) was used for a sample pool (used as a reference sample later). The remaining material of every sample was split into three equal amounts (corresponding to about 300,000 μm² each) for triplicate measurement of each biological replicate.

**LC-MS/MS Analysis**—For the LC-MS/MS study the 60 samples were randomized within three blocks corresponding to the three technical replicates, with CCC and PDAC samples alternating. In addition, six pools of all samples were analyzed, one after every tenth sample. For each measurement, tryptic peptides originating from −300,000 μm² of tumor cells dissolved in 15 μl 0.1% TFA were injected into an Ultimate 3000 RSLCnano system (Dionex, Idstein, Germany) online coupled to an LTQ Orbitrap Elite mass spectrometer (Thermo Scientific, Bremen, Germany). The peptides were precenetrated for 7 min on a trap column (Acclaim® PepMap 100, 75 μm × 2 cm, C18, 5 μm, 100 Å) using 30 μl/min 0.1% TFA and subsequently separated on an analytical column (Acclaim® PepMap RSLC, 75 μm × 50 cm, C18, 5 μm, 100 Å) by applying a gradient from 5–40% solvent B over 98 min (solvent A, 0.1% formic acid; solvent B, 0.1% formic acid, 84% acetonitrile; 400 nl/min; column oven temperature 60 °C). Peptide spectra were acquired in the Orbitrap analyzer at a resolution of 60,000. Using a data-dependent acquisition mode, the 20 most abundant peptides were then fragmented by collision-induced dissociation and measured in the linear ion trap. Further details concerning MS operating parameters have been described previously (13).

**Data Analysis**—Protein identification was performed using Progenesis QI for proteomics (ver. 1.4) (Thermo Scientific, Bremen, Germany) searching the UniProt database (Release 2014_05; 545,388 entries; 20,339 human entries) via Mascot (ver. 2.3.0.2) (Matrix Sciences Ltd., London, UK). Search parameters were as follows: enzyme, trypsin; maximum missed cleavage sites, 1; taxonomy, Homo sapiens; mass tolerance, 5 ppm for precursor and 0.4 Da for fragment ions; dynamic modification, oxidation (M); static modification, carbamidomethyl (C). On peptide level, search results were filtered for a false discovery rate (FDR) < 1% using the PSM Validator tool implemented in Progenesis Discoverer. Protein group was not applied.

Ion intensity-based label-free quantification was performed using Progenesis QI for proteomics (ver. 2.0.5387.52102, Nonlinear Dynamics Ltd., Newcastle upon Tyne, UK). LC-MS runs were aligned to account for retention time shifts. As a reference, one of the sample pool’s measurements was used. Detected ions were filtered for those eluting between 18 and 112 min of the LC run, charged positively two, three, or fourfold and exhibiting more than two isotopes. For ion intensity quantification, an aggregate peak map containing the m/z and retention time information of all sample files was created and applied to each sample. This ensures no missing values. Peptide and protein identifications from Proteome Discoverer were imported into Progenesis QI for proteomics and matched to peptide spectra. Protein abundances were calculated by averaging normalized ion intensities of all unique peptides of one protein – including those measured at different charge states. Proteins quantified with only one unique peptide were removed from the study. Further details about peptide quantification have been described previously (10).

The mass spectrometric data has been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (14) with the dataset identifier PXD001833 and 10.6019/PXD001833. In addition, the peptide and protein identification and quantification data is listed in supplemental Tables S2, S3, and S4.

**Statistical Rationale for Proteome Analysis**—Normalized protein abundances were obtained from the software Progenesis and transformed by arcsinh to reduce the skewness of the data and meet the assumptions of the t test applied later on. Resulting values were averaged for each triple of technical replicates, and used to calculate the p value of Student’s t test (independent samples, two-sided, assuming equal variances). Furthermore, the fold changes between groups were determined on the scale of normalized abundances. Coefficients of variation (CV) of technical replicates were calculated, using the transformed normalized protein abundances, to show reproducibility of the LC-MS system. A significantly differential expression was defined as an absolute fold change > 2 and an FDR-corrected p value (pFDR) < 0.05. The FDR was controlled by adjusting the raw p values of the t test using the method of Benjamini and Hochberg (15). For all applications of statistical inference, the significance level equals 0.05. The described analysis was conducted with the software R (R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria).

**Immunohistochemistry**—Preparation and Staining of Tissue Samples—Tissue microarrays were constructed using three cores per case with a diameter of 1 mm. In addition, tissue sections from metastatic PDAC tumors excised from the liver were prepared. Tissue was formalin fixed, paraffin embedded and sectioned at 4 μm thickness before dewaxing and pretreatment in EDTA buffer (pH 9) for 20 min at 95 °C. All staining reactions were carried out with a Dako Autostainer (Dako, Glostrup, Denmark). Details regarding the applied primary antibodies and the respective reaction conditions are listed in supplemental Table S5. Negative controls were prepared by incubation with nonimmune immunoglobulin instead of the primary antibody at the same concentration and included in every run. Tumor cases that presented with a specific staining during antibody establishment were included as positive controls in every subsequent run. Exemplary control stainings are depicted in supplemental Fig. S1.

**Evaluation of Immunohistochemical Stains**—The immunohistochemical staining of tissue sections was evaluated using an immuno-reactive score (IRS) as described previously (11). Briefly, the proportion of positively stained cells was graded as 0 (0%), 1 (1–5%), 2 (6–10%), 3 (11–50%), or 4 (51–100%) and the staining intensity was expressed as 0 (none), 1 (weak), 2 (intermediate), or 3 (strong). Both values were multiplied to obtain an IRS between 0 and 12. The evaluation was carried out independently by two observers. Cases with slight discrepancies concerning staining intensity or proportion of stained cells were additionally evaluated by a third independent observer. The values given by two out of three observers were used for further calculations. There were no cases with both values differing by more than one level. The inter-rater variability was determined by calculating Cohen’s kappa with squared weights.

**Statistical Rationale for Immunohistochemical Analysis**—Previous to the analysis of the distributions of the IR scores, the interobserver variability was assessed by means of the simple percentage of exact agreement as well as by weighted Kappa, where disagreements are weighted according to their squared distance from perfect agreement. For these computations, the R package irr was used (16). In contrast to the (transformed) abundancies in the proteome study, the IRS data set is analyzed using a nonparametric approach owing to the underlying distribution. The significance of the difference in immunohistochemical staining between all three groups (CCC, PDAC, and mPDAC) was assessed for each candidate, using Kruskal-Wallis’ test based on the IRS. After adjustment of resulting p values for multiple testing according to Benjamini-Hochberg, the maximum p value was below 0.001. Subsequently, the post-hoc pairwise comparisons using Mann-Whitney’s U test (again adjusted using Benja-
mini-Hochberg) were performed to determine which groups differ significantly in their staining patterns. Diagnostic characteristics of the univariate marker candidates were assessed by means of ROC (receiver operating characteristics) analyses performed with the R package pROC (17). First, the overall discriminative power of the candidates was compared based on the area under the ROC curve (AUC). Along with the AUC values, corresponding 95% confidence intervals were calculated as defined by DeLong (18). For each antibody, the optimal IRS cut-off value was determined in order to best discriminate the two groups according to Youden’s criterion. This is equivalent to the maximization of the sum of sensitivity and specificity. It is noted that the obtained diagnostic values are overoptimistic because of the optimization process. However, they are considered to be suitable for the comparison of different candidates within this study. A remark regarding the notation: In the direct comparison of CCC and PDAC, we considered PDAC as “cases” and CCC as “controls,” thus, in the following, sensitivity is the (true) probability of classifying a PDAC sample correctly whereas specificity is the (true) probability of a correct result for CCC.

When comparing diagnostic accuracy of two markers based on the same sample cohort, the McNemar test is used to account for the dependence of observations. Given $p$ values correspond to the two-sided case.

RESULTS

In order to identify potential biomarkers for the differentiation of CCC and PDAC, the proteins extracted from the respective tumor cells were analyzed by label-free LC-MS/MS. The malignant cells were previously isolated from tissue sections by laser capture microdissection to account for the heterogeneity of these desmoplastic neoplasms. After analysis of the proteomics data set, suitable candidate proteins were evaluated by IHC—first in a small test set, then in a large cohort of primary CCC and PDAC as well as metastatic PDAC tumors from the liver. An overview of the described study design is depicted in Fig. 1.

**Discovery Phase**—In the label-free LC-MS/MS data, a total of 61,775 peptide ions were quantified in an ion intensity–based manner. This led to the identification and quantification of 2701 proteins of which 1814 were quantified with more than one unique peptide. This analysis was performed using three technical replicates of each sample. In comparison, using only one replicate each, an average of 2395 proteins could be quantified in total, and 1519 with more than one unique peptide. Technical replicates presented a very high repeatability, which was assessed by calculating the coefficient of variation (CV) of each triplicate measurement of every protein. For 92% of all protein abundances a CV of less than 10% was observed (39% < 1% CV; 85% < 5% CV), and only for 5% it was above 20%. A significantly differential expression between the two experimental groups (absolute fold change $> 2$, FDR-corrected $p$ value of $t$ test $< 0.05$) was observed for 180 proteins in which 85 were more abundant in CCC and 95 in PDAC cells (supplemental Table S6). Analyzing only one technical of each biological replicate, an average of 148 significantly differential proteins were identified. This means that the triplicate measurements increased the number of potentially relevant proteins by almost 22%. In order to select the most suitable candidates from the 180 differentially regulated proteins for the subsequent verification, the Euclidian distance was calculated for each protein. This considers both fold change and $p$ value and is visualized in Fig. 2.

In addition, the expression profiles—the relative peptide and protein abundances in all samples—were observed. These should show little variation between samples from one group and of different peptides of one protein as well as a clear separation of the two groups. Especially promising
results were achieved for nine proteins that were, therefore, selected for immunohistochemical verification (Table II).

**Verification Phase**—The nine candidates were first tested in a tissue microarray including ten CCC and ten PDAC tumor samples to assess cell type specificity of the proteins and the performance of the antibodies in a time and material saving fashion. Here, the annexins ANXA1, ANXA10, and ANXA13 showed the most promising results (Table II) and were therefore transferred to the next verification step in the form of a tissue microarray consisting of tumor sections from 73 CCC and 78 PDAC patients. Corresponding nontumorous liver tissue (NTLT) from the CCC patients was also analyzed to determine the expression of the candidate proteins in normal bile duct epithelium. Representative staining patterns are shown in Fig. 3 (and in supplemental Fig. S2 for the remaining six proteins).

For all of the following statistical analyses the immunoreactive scores (IRS) that take the percentage of stained cells and the staining intensity into account were considered. The scoring was performed by two independent observers with a good to very good agreement as reflected by high values of weighted kappa: 0.836 for ANXA1, 0.782 for ANXA10, and 0.852 for ANXA13. Concordance of IRS was achieved in 77% of all cases for ANXA1 staining, 64% for ANXA10, and 71% for ANXA13. The remaining samples, which showed only slight discrepancies, were scored by a third observer resulting in final IRS for all samples. These confirmed the significantly differential expression of ANXA1, ANXA10, and ANXA13 in primary PDAC cells as compared with CCC cells ($p < 0.001$, Mann-Whitney, adjusted) observed in the proteomics experiment (Fig. 4, upper panel).

To evaluate optimal cut-off values for the IRS for each antibody, ROC analysis was performed comparing primary PDAC and CCC samples. Here, sensitivity was defined as the percentage of correctly classified PDAC samples that, in this case, means a positive staining for ANXA1 and ANXA10 and negative immuno-reactivity for ANXA13. Specificity is therefore to be understood as the proportion of correctly classified CCC samples. ROC curves including AUC values, optimized IRS cut-offs as well as the sensitivity and specificity at this threshold are depicted in Fig. 4 (lower panel). The highest AUC was thereby obtained for ANXA1 (0.907). It exhibits a sensitivity of 84% at a specificity of 85%. The latter is lower for ANXA10 with 66% specificity, however, sensitivity reaches almost 90%. Analysis of ANXA13 resulted in an AUC of 0.697 with a sensitivity of 84% and a specificity of 55%. In addition, NTLT was evaluated regarding staining of nonmalignant cholangiocytes. Here, 7% of all samples showed positive staining for ANXA1, whereas none were positive for ANXA10. On the other hand, in 70% of the samples positive ANXA13 staining in cholangiocytes was visible (Fig. 4, upper panel). However, for none of these marker candidates, significant correlation between the expression in nonmalignant cholan-

<table>
<thead>
<tr>
<th>Accession</th>
<th>Protein</th>
<th>Gene</th>
<th>LC-MS (cohort 1)</th>
<th>IHC (cohort 2)</th>
<th>IHC (cohort 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>94ULC5</td>
<td>Long-chain-fatty-acid-CoA ligase 5</td>
<td>ACSL5</td>
<td>1.04E-07</td>
<td>11.7</td>
<td>PDAC</td>
</tr>
<tr>
<td>P04083</td>
<td>Annexin A1</td>
<td>ANXA1</td>
<td>3.96E-04</td>
<td>8</td>
<td>PDAC</td>
</tr>
<tr>
<td>Q9UJ72</td>
<td>Annexin A10</td>
<td>ANXA10</td>
<td>4.83E-05</td>
<td>169.2</td>
<td>PDAC</td>
</tr>
<tr>
<td>P27216</td>
<td>Annexin A13</td>
<td>ANXA13</td>
<td>1.45E-04</td>
<td>6.4</td>
<td>CCC</td>
</tr>
<tr>
<td>P12532</td>
<td>Creatine kinase U-type, mitochondrial</td>
<td>CKMT1A</td>
<td>4.47E-04</td>
<td>8</td>
<td>PDAC</td>
</tr>
<tr>
<td>O75369</td>
<td>Filamin-B</td>
<td>FLNB</td>
<td>1.45E-04</td>
<td>4.4</td>
<td>PDAC</td>
</tr>
<tr>
<td>P27338</td>
<td>Amine oxidase [flavin-containing] D</td>
<td>MAOB</td>
<td>4.78E-04</td>
<td>4</td>
<td>CCC</td>
</tr>
<tr>
<td>P36952</td>
<td>Serpin B5</td>
<td>SERPINB5</td>
<td>5.37E-05</td>
<td>7.8</td>
<td>PDAC</td>
</tr>
<tr>
<td>O15195</td>
<td>Villin-like protein</td>
<td>VILL</td>
<td>7.30E-07</td>
<td>8.1</td>
<td>PDAC</td>
</tr>
</tbody>
</table>

**Table II**

Expression of candidate proteins in CCC and PDAC samples determined by label-free proteomics and IHC. For sample set 2, immunohistochemical staining was considered to be positive if at least 10% of the tumour cells were stained, regardless of the staining intensity. For sample cohort 3, individual cut-off values for each antibody were established by ROC analysis (Fig. 4).
giocytes and CCC cells was verifiable (for more detailed data see supplemental Table S7).

**Analysis of Metastatic PDAC Tumor Tissue**—Because all three potential markers performed well in the comparison of primary PDAC and CCC tumors, they were also tested in eighteen cases of metastatic PDAC. Here, ANXA1 and ANXA10 showed similar staining as in the primary PDAC samples. Applying the previously established IRS cut-off values, 13 out of 18 and 17 out of 18 tissue sections were positive for ANXA1 and ANXA10, respectively (Table II). These markers, therefore, show a significantly differential expression between mPDAC and CCC tissue ($p < 0.001$) (Fig. 4, upper panel) with sensitivities and specificities of 72 and 85% for ANXA1 and 94 and 66% for ANXA10 (specificity refers to the cohort of CCC samples previously described). ANXA13, on the other hand, was expressed in 8 out of 18 (56% Sensitivity) mPDAC cases (Table II). This means that, in terms of ANXA13 expression, mPDAC cells behave differently than PDAC cells that hardly express this protein ($p < 0.001$). Thus, ANXA13 is in this case not suitable to differentiate between CCC and mPDAC ($p = 0.61$; Fig. 4, upper panel).

**Combination of ANXA1 and ANXA10 in a Biomarker Panel**—In many cases the combination of single biomarkers in form of a biomarker panel can be beneficial. Because ANXA1 and ANXA10 showed the most promising results so far, a multivariate analysis including these two proteins was performed. Table III summarizes the congruency and discrepancies of CCC and PDAC sample classification by ANXA1 and ANXA10 scoring. Although there is no significant difference in the classification accuracy regarding the PDAC samples, ANXA1 outperforms ANXA10 in the CCC cohort. Overall, the combination of ANXA1 and ANXA10 did not improve the diagnostic values of the single marker ANXA1. Considering the panel to be positive for PDAC if at least one marker classifies the samples as PDAC (applying above mentioned cut-off values) resulted in sensitivity of 96% for PDAC and 100% for mPDAC but specificity of only 61%. If the panel is considered to be positive for PDAC only if both antibodies show positive staining, the specificity is improved to 90%, whereas the sensitivity for PDAC drops to 78% and for mPDAC even to 69%.

**DISCUSSION**

The issue of classifying a liver tumor correctly regarding a possible pancreatic or cholangiocellular origin is equally challenging because of morphologic similarities as well as crucial for therapeutic decisions. It is a question pathologists are regularly confronted with. Since, up to now, no reliable diagnostic method such as the application of tissue markers is available, this study aimed at identifying novel biomarkers for identification of PDAC and CCC.
differentiating CCC from PDAC. We therefore chose a global unbiased approach in form of LC-MS-based shotgun proteomics for the discovery phase. This has previously proven to be a suitable method for biomarker discovery studies (10, 11) and incidentally holds the advantage of enabling subsequent re- or meta-analysis of the data for, e.g. functional studies.

Here, each sample was analyzed in technical triplicates. This increased the total number of quantifiable proteins by 13% and those quantified by at least two unique peptides by 19% highlighting the benefit of technical replicate measurements. Tabb et al. showed a repeatability of protein identifications by LC-MS/MS on LTQ Orbitrap instruments of 70–80%. Supposedly, especially peptides with an abundance at the threshold of triggering an MS/MS can be lost if measuring only once (19). This also leads to a lower number of peptides per protein resulting in the exclusion of some proteins and a less exact quantification of others. Repeated measurement of an ion intensity also improves the confidence of the quantification for each peptide although our results revealed a very high repeatability between multiple LC-MS/MS measurements of the same sample. The latter has also been described

**TABLE III**

<table>
<thead>
<tr>
<th>PDAC-Samples</th>
<th>ANXA10</th>
<th>CCC-Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANXA1</td>
<td>Classified as CCC</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Classified as PDAC</td>
<td>5</td>
</tr>
<tr>
<td>ANXA10</td>
<td>Classified as CCC</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Classified as PDAC</td>
<td>4</td>
</tr>
</tbody>
</table>

**Fig. 4.** Statistical evaluation of the immunohistochemical analysis of annexin A1, A10, and A13 in sample cohort 3. Box plots (upper panel) visualize immuno-reactive scores (IRS) for CCC (n = 73), PDAC (n = 78), and mPDAC (n = 13) tissue as well as nontumorous bile duct cells (BD; n = 67). Boxes represent the 25th and 75th percentiles. White dashed lines indicate the median, squares within the boxes the mean value and whiskers the 1.5-fold standard deviation. p values reflecting the significance of differences between two groups were calculated using the Mann-Whitney test after adjustment. Receiver operating characteristic (ROC) curves (lower panel) were calculated to determine the optimal IRS cut-off value for each potential biomarker. These are indicated on the curve including the sensitivity and specificity values at this point in brackets. Abbreviations: AUC, area under the curve; CI: 95% confidence interval.
in a previous study published by our group (20). The number of significantly differential proteins was increased by 22%. These additional proteins of interest can be of high importance for future studies such as pathway analyses. For example, they include several proteins known to be associated with cancer and metastasis formation such as protein kinase C delta (21), CD44 (22), cathepsin E (23), or Ras-related protein Rab-34 (24). Also serpin B5, which was chosen as a candidate protein in this study, would not have been identified as differentially expressed using only one technical replicate.

For this project, the malignant cells were isolated from the heterogeneous tumor tissue by LCM. We were able to achieve solid and reproducible LC-MS/MS data while applying as little as 300,000 μm² of tissue that corresponds to only few thousand tumor cells. Our results underline the necessity of this step because other cell types within the tumor show different protein expression patterns than the actual tumor cells. ANXA1 e.g. was often strongly detectable in stromal or inflammatory cells, also in CCC tissue, but not in tumorous cholangiocytes. Analysis of the whole tumor tissue would have therefore falsified quantitative results.

The verification of selected biomarker candidates was performed via IHC because this is a well-established method used in pathological routine diagnostic. Out of the nine proteins tested in the first verification step, three could be transferred to the second stage involving a larger sample cohort. The reasons for discrepancies between proteomics and IHC results as seen for the remaining six candidates can be diverse. On the one hand, different patient cohorts were used for the identification and the subsequent verification. Therefore, differing protein expressions can be due to biological variances. On the other hand, the reason might be related to the difference in methodologies. While LC-MS allows continuous protein quantification, IHC — although characterized by low limits of detection — is merely a semiquantitative method. Especially for proteins of rather high abundance it is often not possible to assess subtler quantitative variations. This might be the case for CKMT1A, FLNB, MAOB, and Serpin B5 for which positive reactivity was observed in all twenty tissue samples from cohort 2. Also, some antibodies used for IHC show unspecific binding, or bind only to certain protein isoforms that often cannot be distinguished in shotgun LC-MS experiments, in most cases. Nevertheless, these proteins might still be of relevance in a different context, e.g. in relation with tumor biology or as noninvasive biomarkers detectable in body fluids. This was, however, not within the scope of this study.

Here, we identified and verified the three annexin proteins ANXA1, ANXA10, and ANXA13 as being differentially expressed between primary CCC and PDAC cells. They belong to a highly conserved protein family comprising twelve members in vertebrates. They are characterized by the ability of binding negatively charged phospholipids in a calcium-dependent manner. This can lead to a rapid translocation from the cytosol to plasma or intracellular membranes and is important for their interaction with binding partners (25). They are involved in several cellular processes such as calcium signaling, vesicle trafficking, cell division, growth regulation, and apoptosis, which is why their differential expression in tumor cells that has been reported for many of the annexins in several cancer types is plausible (26).

ANXA1 is the best studied annexin, also in association with cancer. For example, it has been shown to be decreased in breast carcinoma, prostatic adenocarcinoma, and esophageal squamous carcinoma, but increased in esophageal adenocarcinoma and hairy cell leukemia (26). In PDAC, a study comparing tumors to matched normal pancreas tissue showed a higher ANXA1 expression in the malignant tissue that was correlated with poor tumor differentiation (27). Our study confirmed the abundance of ANXA1 in PDAC cells while it furthermore provided the best diagnostic values for distinguishing PDAC from CCC (84% sensitivity, 85% specificity). In contrast to our findings, Hongsrichan et al. discovered an overexpression of ANXA1 in 64 out of 68 CCC cases. This was in comparison to hepatocellular carcinoma and healthy liver tissue (28). However, all of these CCC tumors were associated to an *Opisthorchis viverrini* infection that, consequently, seems to have an effect on the protein expression profile of the cancer cells. ANXA1 has also been linked to tumorigenesis and metastasis formation on the molecular level as it is a substrate for several serine/threonine and tyrosine kinases (25). Upon phosphorylation, for instance, it migrates to the cell membrane where it interacts with epidermal growth factor receptor (EGFR) (29). It has been shown that upon nontransient EGFR activation with consequential tumor growth, both ANXA1 expression and phosphorylation were increased (30). Furthermore, silencing ANXA1 in human CCC cell lines by siRNA induced down-regulation of transforming growth factor (TGF)-β, matrix metalloproteinases (MMP) 2 and MMP 9 as well as up-regulation of NF-κB (28), all of which are proteins involved in cancer development and metastasis.

Next to ANXA1, in the present study, ANXA10 received promising diagnostic values with a sensitivity of 90% and a specificity of 66% for the identification of PDAC versus CCC. Recently, an immunohistochemical study examining ANXA10 expression in several types of tumorous and nontumorous tissue was published. Here, normal expression was observed in the gastric mucosa, the duodenum, the kidney, and the urinary bladder with a mainly nuclear and weak cytoplasmic staining, and none in normal bile ducts and pancreatic ducts. Within their PDAC cohort, 78% (57/73) of the primary and 83% (19/23) of the metastatic tumors showed positive ANXA10 staining, while it was only 17% (10/59) of the intrahepatic CCC cases. Taking the general biological variance into consideration, these results are very similar to ours. Concerning the pancreatobiliary system, this study also included adenocarcinomas of the ampulla of Vater, the extrahepatic bile ducts and the gallbladder. Interestingly, this indicated
that the closer the primary tumor site is located to the pancreas, the more likely a positive ANXA10 staining is and, conversely, the closer it is to the liver, the less primary tumors express ANXA10 (31). Thus, although ANXA10 expression is not visible in healthy cells of the pancreatobiliary system, their localization has an effect on the ANXA10 abundance in tumor cells deriving from them. This knowledge is also important for the potential establishment of ANXA10 as a biomarker.

For ANXA13 our observations were different. The abundance of this protein did not seem to correlate with the primary tumor site but with the actual localization of the malignancy because it was detected in 55% of the CCC and only 16% of the PDAC sections, but it was also expressed by 44% of the mPDAC cases. ANXA13 is thought to be the probable common ancestor of all vertebrate annexins (32). It is supposed to have a high tissue specificity in healthy tissue, restricted only to intestinal and kidney epithelial cells (33), although it was also detected in nontumorous cholangiocytes in the present study. In malignant cells, the decreased ANXA13 expression has been described to cause protection against Rapamycin, an anticancer drug (34). Because of our findings, it is not suitable for distinguishing mPDAC in the liver from CCC. However, it might be of relevance in the context of the molecular mechanisms of PDAC metastasis formation because of the significantly differential expression between the primary and the secondary tumors.

The combination of biomarkers in the form of a panel can often improve diagnostic values. For ANXA1 and ANXA10 this was not the case. Nevertheless, in comparison to marker panels described in previous publications, ANXA1 delivers better diagnostic values even on its own. The panel S100P-/pVHL+/MUC5AC-/CK17-, for example, has a specificity of 100% for CCC when compared with PDAC but only a sensitivity of 59% (8) and combining MUC1 and CK17 led to positive predictive values of 76% for PDAC and 58% for CCC (35). It could therefore be advantageous to implement ANXA1, or also ANXA10, into these panels or construct a new panel with the most suitable proteins. In conclusion, ANXA1 has a high potential of becoming a valuable marker for the differential diagnosis of CCC and metastatic PDAC in the liver.

Acknowledgments—We thank Laura Malkus, Birgit Korte, Stephanie Tautges, Kristin Rosowski, Dorothe Möllmann, and Don Marvin Voss for their excellent technical assistance.

REFERENCES

3. German S3-Guideline “S3-Leitlinie zum exokrinen Pankreaskarzinom” (2013) Version 1.0, AWMF registration number: 032/010OL.


