Urine proteomic analysis differentiates cholangiocarcinoma from primary sclerosing cholangitis and other benign biliary disorders

Jochen Metzger,1 Ahmed A Negm,2 Ruben R Plentz,3 Tobias J Weismüller,2,4 Jochen Wedemeyer,5 Tom H Karlsen,6 Mohammed Dakna,1 William Mullen,7 Harald Mischak,1,7 Michael P Manns,2,4 Tim O Lankisch2,4

ABSTRACT

Background Diagnosis and curative treatment of cholangiocarcinoma (CC) often comes too late due to the lack of reliable tumour markers especially in patients with primary sclerosing cholangitis (PSC). The authors recently introduced bile proteomic analysis for CC diagnosis. Nevertheless, bile collection depends on invasive endoscopic retrograde cholangiography. The authors therefore evaluated urine proteomic analysis for non-invasive CC diagnosis.

Methods Using capillary electrophoresis mass spectrometry the authors established a CC-specific peptide marker model based on the distribution of 42 peptides in 14 CC, 13 PSC and 14 benign biliary disorder (BBD) patients.

Results In cross-sectional validation of 123 patients, the urine peptide marker model correctly classified 35 of 42 CC patients and 64 of 81 PSC and BBD patients with an area under the curve value of 0.87 (95% CI 0.80 to 0.92, p = 0.0001, 83% sensitivity, 79% specificity). Evaluation of 101 normal controls resulted in 86% specificity. All 10 patients with CC on top of PSC were correctly classified. The majority of sequence-identified peptides are fragments of interstitial collagens with some of them also detected in blood indicating their extra-renal origin. Immunostaining of liver sections for matrix metalloproteinase 1 indicated increased activity of the interstitial collagenase in liver epithelial cells of CC patients.

Conclusion The urine test differentiates CC from PSC and other BBD and may provide a new diagnostic non-invasive tool for PSC surveillance and CC detection.

INTRODUCTION

Cholangiocarcinoma (CC) originates from cholangiocytes of the intrahepatic and extrahepatic biliary tract. Although still rare, its incidence is increasing over the years. The prognosis of CC remains poor as curative treatment options such as surgery or orthotopic liver transplantation (OLT) can only be performed at an early stage of CC. Unfortunately, CC is often detected in an unresectable stage as specific tumour markers or other reliable diagnostic methods are still lacking.2

In clinical practice, the diagnosis of CC is based on a combination of imaging techniques and tissue sampling. Tumour markers, like serum carbohy-
eventually end in biliary cirrhosis and hepatic decompensation. Furthermore, PSC usually starts at a young age, affects twice as many men as women and is accompanied by an inflammatory bowel disease in 60%–80% of patients. In patients with PSC, the differentiation between benign and malignant strictures is particularly difficult because CC as well as chronic or acute inflammation may result in similar cholangiographic findings. In a pilot study we used capillary electrophoresis mass spectrometry (CE-MS) to identify disease specific peptide distribution patterns in the bile of patients with choledocholithiasis, PSC and CC.³ Nine bile proteomic analysis (BPA) discriminated CC from benign conditions accurately with a sensitivity and specificity of 84% and 78%, respectively. However, the collection of bile is dependent on invasive, time consuming and expensive ERCP.

In contrast to bile, spontaneous urine appears to be an attractive diagnostic medium, because it harbours lower numbers of proteins and lipids, remains relatively stable due to minimal proteolysis after sampling and is non-invasively and frequently accessible.¹⁰ ¹¹ A non-invasive test for CC would be of great interest for PSC surveillance, because these young patients are at a special risk for CC development and the diagnosis or exclusion of such malignant transformation at an optimum time point is crucial to define resection or their priority on the waiting list for liver transplantation. Therefore, our aim was to identify urinary peptide markers with altered distribution in CC compared with PSC and other benign biliary disorder (BBD) patients and to establish a support vector machine (SVM)-based peptide classifier model on the basis of the peptide discriminate properties that distinguishes CC from PSC, other BBD and normal controls.

PATIENTS AND METHODS

Patients

Urine samples from 164 consecutive patients with cholestasis who presented for diagnostic and therapeutic ERC were included in the study at the Hannover Medical School, Germany and at the Norwegian PSC Research Center, Oslo, Norway (for patient characteristics see table 1). Variation between the centres was kept to a minimum by urine sampling according to a standard protocol (available at http://www.eurokup.org) and by consistent preparation, measurement and data analysis at Mosaïques (Hannover, Germany). Urine from 101 age- and sex-matched volunteers served as control.

The diagnosis of PSC was based on the typical cholangiographic findings such as strictures or irregularity of intrahepatic or extrahepatic bile ducts and exclusion of secondary causes for sclerosing cholangitis. CC was proven by histology in 53 out of 56 patients with CC. In three patients a definite histology could not be obtained, but clinical, laboratory, radiological and ERCP findings enabled the diagnosis of CC. Forty-one patients with CC were therapy-naive. Thirteen patients received chemotherapy and two patients underwent photodynamic therapy.

Table 1  Clinical and demographic data of the patients included in the discovery and validation phase of the study

<table>
<thead>
<tr>
<th>Type of biliary disease (n)</th>
<th>Training cohort</th>
<th>Validation cohort</th>
<th>Normal controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>14</td>
<td>27</td>
<td>101</td>
</tr>
<tr>
<td>CC on top of PSC</td>
<td>11</td>
<td>32</td>
<td>NA</td>
</tr>
<tr>
<td>PSC</td>
<td>13</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Choledocholithiasis</td>
<td>7</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Chronic pancreatitis</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SSC</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholecystectomy injury</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBD stones</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBD dilatation</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biliary casts</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other benign strictures, that is, leakage</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Normal finding</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BBD, benign biliary disorder; CBD, common bile duct; CC, cholangiocarcinoma; NA, not analysed; PSC, primary sclerosing cholangitis; SSC, secondary sclerosing cholangitis.
Urine sample preparation and CE-MS analysis

Midstream spot urine samples were collected from patients according to the EuroKUP standard protocol. The samples were stored at −20°C (at a maximum of 2 years) and thawed again once, just before preparation. Urine samples were analysed by dipstick to exclude patients with haematuria or bacterial infection from CE-MS analysis. Sample preparation was performed as described before. Briefly, 0.7 μl of the urine sample was diluted 1:2 in an alkaline buffer containing 2 M urea, 10 mM NH4OH and 0.02% SDS (pH 10.5). This was followed by size-exclusion over 20 kDa Centrisart ultracentrifugation filters (Sartorius, Göttingen, Germany) for the removal of proteins >20 kDa and by desalting over PD-10 columns (GE Healthcare, Munich, Germany). After lyophilisation, samples were resuspended in HPLC-grade water to a final protein concentration of 0.8 μg/μl as verified with the bicinchoninic acid assay (Interchim, Montlucon, France). A 100 nL aliquot was injected into the CE-MS under constant flow and pressure conditions at a pH of 2.2 to ensure that all peptides are positively charged. CE-MS analysis was carried out on a P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton, USA) on-line coupled to a Micro-TOF MS (Bruker Daltonic, Bremen, Germany).

Proteomic data processing and cluster analysis

Mass spectral ion peaks with a charge >1, a signal to noise ratio >4 and minimum detection in three consecutive spectra assignable to the same peptide entity were deconvoluted into single masses using the MosaicsVisu software. CE migration time and ion signal intensity were normalised using internal peptides as calibrants. Sample-specific peptide lists were deposited in a Microsoft SQL database for subsequent statistical analyses and for generation of a CC-specific peptide model using the SVM-based MosaCluster software.

Statistical analysis

Values for sensitivity, specificity and 95% CIs were calculated using receiver operating characteristic (ROC) plots (MedCalc Software 8.1.1.0, Belgium). p Values were calculated on the basis of natural logarithm transformed intensities and the Wilcoxon rank sum test.

Peptide sequencing

Urine samples were separated on a Dionex Ultimate 3000 RSLS nano flow system (Dionex, Camberly UK). A 5 μl sample was loaded in 0.1% formic acid and acetonitrile (98:2) onto a Dionex 100 μm×2 cm, 5 μm C18 nano trap column at a flowrate of 5 μl/min. Elution was performed on an Acclaim PepMap C18 nano column 75 μm×15 cm, 2 μm, 100 Å with a linear gradient of 0.1% formic acid against 100% acetonitrile starting at 5%–50% over 100 min. The sample was ionised in positive ion mode using a Proxeon nano spray ESI source (Thermo Fisher Hemel UK) and analysed in an Orbitrap Velos FTMS (Thermo Finnigan, Bremen, Germany). The MS was operated in data-dependent mode to switch between MS and MS/MS acquisition and parent ions were fragmented by high-energy collision-induced dissociation. Data files were searched against the IPI human non-redundant database using SEQUEST without enzyme specificity. No fixed modification and oxidation of methionine and proline as variable modifications were selected. Mass error windows of 10 ppm and 0.05 Da were allowed for MS and MS/MS, respectively. In SEQUEST, the peptide data were extracted using high peptide confidence and top one peptide rank filters. The correlation between peptide charge at the working pH of 2.2 and CE-migration time was utilised to minimise false-positive identification rates: Calculated CE-migration time based on the number of basic amino acids was compared with the experimental migration time.

Immunohistochemistry

Staining was performed on histology samples from explanted livers of patients with PSC undergoing OLT. CC histology was performed using sections obtained during surgery. Slides were deparaffinised with Roti-Histol (Roti, Karlsruhe, Germany) and rehydrated in ethanol. After PBS washing steps, antigen retrieval was performed by heating the slides in a pressure cooker with an antigen unmasking solution (Vector, California, USA). Slides were then washed in PBS and incubated for 10 min in 10% H2O2, rinsed with PBS, and incubated 1 h in blocking solution (5% normal serum, 0.5% Triton X-100). Tissue sections were incubated overnight at 4°C with anti-matrix metalloproteinase 1 (MMP-1) antibody (1:100, Abnova, Heidelberg, Germany). The manufacturer’s protocols were used for ABC and DAB substrates (Vector), slides were counterstained with haematox- ylin. Finally, slides were dehydrated in ethanol, cleared with xylenes and mounted with Permount (Fisher, Wiesbaden, Germany). Staining intensity was subdivided into absent, weak, moderate and strong according to pathology recommendations.

RESULTS

Establishment of a urine peptide marker model for differentiation of CC from PSC and other BBD

CE-MS measurements were performed in urine samples of 164 patients admitted for ERCP. All patients were sampled only at one single time point and urine was tested negative for signs of haematuria or bacterial infection in each case. The low molecular weight proteome of urine in the range of 0.8–20 kDa was resolved down to the femtomolar range. This resulted in the detection of approximately 5600 peptides with a distribution of >20% over all urine samples. The mean number of peptides and small proteins detected in the urine samples after deconvolution of CE-MS ion signals (a process where signals of the same analyte at different charge states are merged to one single ion peak) was 1765 (range: 592–3800). From the 164 urine samples analysed in this study, 41 were randomly selected for statistical marker discovery and model establishment (discovery phase), whereas 123 were used for subsequent model validation (validation phase). In the discovery phase, urine samples from 14 CC patients were compared with urine samples from patients with PSC (n=13), choleodocholithiasis (n=7), chronic pancreatitis (n=5), secondary sclerosing cholangitis (n=1) and bile leakage following cholecystectomy (n=1) in order to identify differentially excreted peptides between CC cases and PSC and BBD controls (see table 1 for composition of the training cohort and for demographic and clinical data of patients). When evaluating the peptide lists from the CE-MS analyses, 313 urinary peptides showed differences between patients with CC compared with those with PSC and BBD (p<0.05). From the 313 peptides, 106 remained statistically significant after false discovery rate-adjustment by the method of Benjamini and Hochberg. Since the number of peptides showing statistical differences between the groups substantially exceeds the number of samples, we aimed at reducing the number of peptides included in the peptide marker model. This was achieved by take-one-out optimisation on the original data set of 313 peptides and evaluation in AdaBoost, an adaptive boosting algorithm sensitive to outliers. Twenty-nine peptides significant.
in false discovery rate-adjusted p value statistics and 13 additional peptides were identified to provide optimum differentiation of CC from PSC and BBD in the training cohort. The selected 42 peptides are presented in table 2. A SVM-based classifier model based on the 42 peptides was established using the MosaCluster software. Total cross-validation of classification scores (defining membership probabilities in arbitrary units) for the CC case and non-CC reference groups of the training cohort resulted in an AUC of 0.98 and a 95% CI from 0.89 to 0.99 in ROC analysis proving the models ability to differentiate all except one sample of the training cohort correctly. The compiled CE-MS marker profiles for the CC cases and the PSC and BBD controls of the training cohort are presented in figure 1. As indicated by the figure, the peptide marker model represents a combination of discriminatory peptides for CC against PSC and CC against BBD.

**Performance evaluation of the urine peptide marker model**

To avoid errors caused by overfitting of the data, the accuracy of the urine peptide marker model was subsequently determined in a large set of patient samples not used for marker discovery and model establishment which consisted of 42 CC including 10 with CC on top of PSC, 45 PSC and 36 BBD patients (for a more detailed list of patient characteristics see table 1). Test characteristics such as AUC, sensitivity and specificity were evaluated solely on this set of patients. As indicated by the ROC curve in figure 2A the peptide marker model differentiates CC from all other biliary diseases with an AUC of 0.87 (95% CI 0.80 to 0.92; p=0.0001). At a cut-off of −0.89 this resulted in a correct classification of 35 from 42 CC (83% sensitivity (95% CI 69% to 95%)) and 64 from 81 controls (79% specificity (95% CI 69% to 87%)). Most notably, all 10 patients with CC on top of PSC scored positive for CC.

Although classification by urine proteomic analysis (UPA) is only moderately correlated with the demographic and clinical parameters presented in table 1 (for correlation coefficients see table 2), we tested our marker model on a selection of CC (n=16) and benign stricture (n=34) patients from the validation cohort who are matched for demographic and clinical parameters. Statistical evaluation on this stratified subsample of patients revealed that UPA outperforms all classifiers based on these classical parameters. All relevant statistical information on this stratification approach is presented in table 2.

We evaluated the distribution of misclassified patients among the different control groups of the validation cohort. As shown in the box and whisker representation of group-specific classification in figure 2B the specificity values for PSC and BBD patients were equally high at 78% and 81%, respectively, and similar to the 86% specificity obtained for 101 age- and gender-matched normal controls.

**Combined bile proteomic analysis and UPA for CC diagnosis**

In a substudy, we investigated combined proteomic analysis of bile and urine samples (BPA–UPA) of patients for whom both fluids were available at the same day of ERCP. For classification of bile we used our previously established CC from PSC differentiation model.9 Urine and bile samples were available from nine CC, nine PSC and nine BBD patients not used in the discovery phase of either the urine or bile model. As presented in figure 3, combined BPA–UPA analysis resulted in complete discrimination when CC diagnosis was based on the positive results in both the bile and urine tests. This is indicated in the Cartesian graph of figure 3: all samples of CC patients are located in the upper left quadrant, whereas all samples of patients with PSC and BBD are exclusively distributed in the double negative and single positive quadrants.

**Peptide sequencing for pathology analysis**

We characterised the urinary peptide markers in respect of their amino acid sequence by MS/MS peptide sequencing.13 The majority of sequence-identified peptides are fragments of type I and associated collagens from restricted areas in the primary sequence. Amino acid sequence of the selected markers is given in table 3. Four of the differentially excreted collagen fragments in CC patients were also identified in blood plasma (see the supplementary figure for a comparison of the respective urine and blood plasma MS/MS sequence spectra). It appears plausible to assume an extra-renal origin of their proteolytic release.

**Table 2** Correlation of UPA’s numerical values to the clinical parameter and statistical evaluation of matched CC cases and benign stricture controls in a stratified subsample of the validation cohort

| Parameter | CC (n=16) | Benign stricture (n=34) | Two-tailed probability p* AUC in ROC 95% CI in ROC p of AUC difference to UPA† |
|-----------|----------|------------------------|--------------------------------|----------------------------------|-------------------------------|
| UPA scores | -0.01 to 0.32 | 0.61 -1.46 to -1.22 | 0.0001 0.92 0.80 to 0.98 | 0.92 0.80 to 0.98 |
| Age, y | 0.19 | 52.7 to 58.5 | 118.9 | 48.2 to 52.2 | 132.8 | 0.19 0.63 0.48 to 0.76 0.003 |
| Gender | 0.05 | n.d. n.d. | n.d. | n.d. | 0.75 | n.d. n.d. |
| Fever | 0.13 | n.d. n.d. | n.d. | n.d. | 0.65 | n.d. n.d. |
| Alkaline phosphatase, U/L | 0.49 | 30 to 58.5 | 118.9 | 48.2 to 52.2 | 132.8 | 0.19 0.63 0.48 to 0.76 0.003 |
| γ-Glutamyltransferase, U/L | 0.31 | 192 to 394 | 118.9 | 48.2 to 52.2 | 132.8 | 0.19 0.63 0.48 to 0.76 0.003 |
| Bilirubin, μmol/L | 0.43 | 29 to 79 | 2284 | 0.75 | n.d. n.d. |
| Leucocyte count, ×10⁹/L | 0.23 | 7 to 9 | 17 | 8 | 0.56 | 0.62 0.47 to 0.75 0.001 |
| C reactive protein, mg/L | 0.43 | 1 to 5 | 12 | 2 | 1 to 3 | 0.48 to 0.77 0.001 |
| Alanine aminotransferase, U/L | 0.28 | 70 to 93 | 1837 | 69 | 0.51 0.36 to 0.65 0.001 |
| Aspartate aminotransferase, U/L | 0.42 | 85 to 112 | 1806 | 68 | 0.52 0.50 to 0.80 0.005 |
| Carbohydrate antigen 19-9, U/L | 0.40 | 140 to 207 | 15917 | 108 | 0.40 0.75 0.003 |
| Lactate dehydrogenase, U/L | 0.08 | 222 to 263 | 4304 | 199 | 0.29 0.62 0.46 to 0.75 0.003 |

For this table, CC and benign stricture patients of the validation cohort were matched in their clinical parameters to obtain subgroups with no statistical significant differences in mean, 95% CI and variance values. Uniqueness of the two groups was verified by a two-tailed probability p >0.05 in the Fisher exact t test for categorical data or in the Student t test for equally distributed data. For all parameters with equal distribution the values for AUC and 95% CI were determined and used in a painwise comparison of ROC curves to determine a p value for the difference of the AUC to that of the UPA model.18

*Fisher exact t test for categorical data and Student t test for equally distributed data.
†Pairwise comparison of ROC curves.
AUC, area under the curve; CC, cholangiocarcinoma; n.d., not determined; ROC, receiver operating characteristics; UPA, urine proteomic analysis.
MMP-1 immunostaining of liver histology sections
We performed immunostaining of liver histology sections from CC and PSC patients to confirm our hypothesis that collagen type I fragments derive from the liver through enhanced MMP activity, especially by MMP-1.\(^{20,21}\) We analysed differential MMP-1 expression in liver biopsy sections of six CC and six

---

**Figure 1** Compiled capillary electrophoresis mass spectrometry (CE-MS) spectra of peptides included in the urine peptide marker model for the cholangiocarcinoma (CC) case and benign biliary disorders (BBD) and primary sclerosing cholangitis (PSC) control groups of the training cohort. On the x-axis: CE-migration time in the range of 15–50 min; on the y-axis: log molecular mass in the range of 0.8–20 kDa; on the z-axis: mean signal intensity expressed as peak height with a maximum relative intensity of 0.1 (colour scale, in arbitrary units).

**Figure 2** Performance characteristics of the urine peptide marker model for cholangiocarcinoma (CC) detection. (A) Receiver operating characteristics (ROC) curve and ROC characteristics such as area under the curve (AUC), 95% CIs, and p value for the differentiation of CC from primary sclerosing cholangitis (PSC) and benign biliary disorders (BBD) in patients of the validation cohort. Of note: the \(p\) value considers the AUC as a random normal variable. For testing the performance of a classifier the departure of its AUC from 0.5 (a random classifier) is assessed by using a standard t test. (B) Specificity of CC detection in normal controls and in the PSC and BBD patient groups included in the validation cohort. A post hoc rank-test was performed for average rank differences between the non-CC reference groups and the CC case group (each with \(p<0.05\)) after a significant result in the global Kruskal–Wallis test (\(p<0.0001\)).

<table>
<thead>
<tr>
<th>Group comparison</th>
<th>PSC/BBD versus CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size control case</td>
<td>81/42</td>
</tr>
<tr>
<td>AUC*</td>
<td>0.87</td>
</tr>
<tr>
<td>Standard error</td>
<td>0.039</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.80 to 0.92</td>
</tr>
<tr>
<td>Significance level P</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

\* At \(-0.89\) as cut-off values for sensitivity and specificity are 83 and 79%, respectively.

---

Hepatology

---
Figure 3 Combined bile proteomic analysis and urine proteomic analysis (BPA—UPA test) of patients at the same endoscopic retrograde cholangiopancreatography date. Patients were included if both bile and urine were available from the same sampling date and patients were not included in one of the models training groups. Of note: in BPA there is a second model for exclusion of choledocholithiasis before primary sclerosing cholangitis (PSC) from cholangiocarcinoma (CC) differentiation, which was not applied in this context leading to a higher frequency of false positive classified benign biliary disorders (BBD) samples by BPA testing (lower right quadrant).

PSC patients. Immunostaining from resected CC biopsies exhibited different staining intensities from weak, moderate to strong. One of the tissues had a weak, four a moderate and one a strong MMP-1 staining (figure 4). The expression of MMP-1 was located in epithelial cells, mostly in the cell membrane. In contrast, PSC tissues showed in two of the six cases no MMP-1 expression, in two low and in two moderate staining intensities.

DISCUSSION
The detection of CC remains a diagnostic challenge particularly in patients suffering from PSC who are known to develop CC in approximately 15% of cases which means a 160-fold increased risk to develop CC. In the last decade, invasive methods such as ERC with or without cholangioscopy, intraductal ultrasound, brush cytology or forceps biopsies have found their way into clinical practice without having demonstrated a significant impact on the differentiation between malignant and benign biliary lesions. Furthermore, tumour markers like CA 19-9 demonstrated a low sensitivity and specificity in different studies and are therefore not useful to detect CC at an early stage or for surveillance of patients with risk factors for CC. Appropriate surveillance is especially required for young patients with PSC since OLT performed before CC development is the only life-saving therapy.

A straight forward approach for the early detection of CC is the identification of markers in bile, as the development of carcinoma takes place at the biliary epithelium and tumour-related proteins are secreted or shed into the bile. BPA was successfully performed by our group leading to 84% sensitivity and 78% specificity in discriminating CC from PSC in a validation cohort of 25 CC and 18 PSC patients. Bile analysis is limited to some extent due to its dependence on ERC/P, an invasive and potentially life-threatening procedure. Furthermore, access to bile becomes difficult if patients received a bile-digestive anastomosis or other surgical procedures. Thus, CC diagnosis by use of a non-invasive source of body fluid will be of benefit in order to allow better monitoring in patients with risk factors for cancer.

In this study, we hypothesised that CC progression contributes to changes in bile and is reflected in other body fluids. Therefore, we analysed blood and urine as a source for CC-specific small protein and peptide markers.

At first sight, blood appears to be an attractive diagnostic source. However, we were unable to establish a diagnostic model for CC in blood (data not shown), mainly due to the high dynamic range between low-level and high-abundance proteins. In contrast to blood, urine as a sample matrix provides several advantages. First, it is non-invasively accessible and can be obtained in large quantities whenever needed. Second, urine is stable in its composition if handled properly mainly due to the completion of endogenous proteolysis at the time of urination. Third, since urine originates as the ultrafiltrate of plasma, the urinary proteome is highly sensitive towards changes of renal origin and of a wide range of non-renal diseases including cardiovascular, autoimmune and infectious diseases as well as certain types of cancer. Finally, urine is enriched in low molecular weight proteins and peptides, which can be transferred without an initial protease digestion step directly to MS.

As demonstrated in this study, we were successful in establishing a peptide marker model based on urinary peptides that mirrors systemic effects of CC tumour progression and is of equal diagnostic precision as in bile. Characterisation of peptides by amino acid sequencing provides insights into the mechanisms responsible for systemic changes during CC. The prevalence of fragments of different types of collagens in the peptide model suggests remodelling of the extracellular matrix (ECM) and release of collagen peptides into the secretome, that includes all molecules released by cells and tissues, during cholangiocarcinogenesis. The effect of CC on the ECM is further evidenced by increased levels of an osteopontin-derived peptide in CC patients, since this oxidant stress-sensitive ECM cell adhesion molecule induces upregulation of type I collagens. In contrast, decreased in CC compared with PSC and BBD are peptide fragments from the cell surface molecules CD99, membrane-associated progesterone receptor component 1 and Na/K-ATPase ion channel subunit γ.

In our CC-specific urine peptide marker model some of the most discriminatory sequenced peptide markers arise from collagen α-2(I), which is considered a mesenchymal marker. All collagen α-2(I)-derived peptides were increased in frequency...
### Table 3: Characteristics of peptides in the urine peptide marker model for differentiation of CC from PSC and other BBDs

<table>
<thead>
<tr>
<th>Peptide-ID</th>
<th>Mass (Da)</th>
<th>CE-time (min)</th>
<th>Sequence information</th>
<th>Protein</th>
<th>AA</th>
<th>Peptide distribution in the training cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>8503</td>
<td>947.47</td>
<td>24.74</td>
<td>SGVDDQSR</td>
<td>Uromodulin</td>
<td>589–597</td>
<td>Controls (n = 27)</td>
</tr>
<tr>
<td>13746</td>
<td>1025.47</td>
<td>25.00</td>
<td>AKTGVGQFDT</td>
<td>Kininogen-1</td>
<td>62–71</td>
<td>27 (62)</td>
</tr>
<tr>
<td>19773</td>
<td>1128.39</td>
<td>37.03</td>
<td>DFFDFNLED</td>
<td>CD99 antigen-like protein 2</td>
<td>26–34</td>
<td>1612 (1901)</td>
</tr>
<tr>
<td>20334</td>
<td>1138.47</td>
<td>26.20</td>
<td>DGEAQGQPSyGPA</td>
<td>Collagen α-1(III) chain</td>
<td>613–625</td>
<td>34 (62)</td>
</tr>
<tr>
<td>23628</td>
<td>1184.56</td>
<td>25.95</td>
<td>KgQGEGQVGCQG</td>
<td>Collagen α-1(III) chain</td>
<td>657–666</td>
<td>25 (48)</td>
</tr>
<tr>
<td>24933</td>
<td>1198.54</td>
<td>39.57</td>
<td>GPpGEQGREGPE</td>
<td>Collagen α-1(III) chain</td>
<td>1007–1019</td>
<td>50 (144)</td>
</tr>
<tr>
<td>26431</td>
<td>1231.49</td>
<td>39.57</td>
<td>GpGEQGREGPE</td>
<td>Collagen α-1(III) chain</td>
<td>1701–1713</td>
<td>39 (69)</td>
</tr>
<tr>
<td>28306</td>
<td>1260.61</td>
<td>27.43</td>
<td>RpGEVpGpGpPE</td>
<td>Collagen α-1(III) chain</td>
<td>918–930</td>
<td>1 (4)</td>
</tr>
<tr>
<td>33732</td>
<td>1350.63</td>
<td>27.09</td>
<td>pGEpGQpGpGVP</td>
<td>Collagen α-1(III) chain</td>
<td>651–664</td>
<td>325 (420)</td>
</tr>
<tr>
<td>42404</td>
<td>1487.65</td>
<td>29.62</td>
<td>GLSMDGGSQKGQVP</td>
<td>Sodium/potassium-transporting ATPase subunit γ</td>
<td>3–18</td>
<td>285 (501)</td>
</tr>
</tbody>
</table>

Continued
and amplitude in CC compared with benign strictures. We also
detected higher abundance of collagen α-1(XVII) peptides in CC
patients, a transmembrane molecule on keratinocytes shed
from the cell surface by metallopeptidases.\textsuperscript{27} The high abun-
dance of specific collagen fragments in urine of CC patients
together with the detection of several of these peptides also in
blood led us to analyse expression of MMP-1 in liver sections of
CC patients at the sites of tumour. We focused on MMP-1 since
this is the only protease able to initiate breakdown of inter-
stitial collagens.\textsuperscript{20, 21} We observed increased expression of MMP-
1 mainly on the surface of epithelial cells at the CC tumour
sites compared with PSC controls. We suggest that the high
expression of MMP-1 in CC is attributed to keratinocyte
responses since these cells release factors upon activation, that
is, 14-3-3\textsubscript{z}, which induce expression of MMP-1 and other
MMPs in fibroblasts.\textsuperscript{28} In agreement with this hypothesis,
we recently identified a 14-3-3\textsubscript{z/d} derived peptide (28-SVTEQ-
GAELSNEER-41) increased in bile of CC patients.\textsuperscript{9} The same
peptide was previously described by Schilling and Overall to
originate from trypsin 1 cleavage.\textsuperscript{31} Based on UPA we were able to detect CC in 35 of 42 patients
in cross-sectional validation. In fact, the urine peptide marker
model identified CC in all 10 patients with CC on top of PSC,
indicating its potential for PSC surveillance. Instead of
performing expensive MRCP or invasive diagnostic ERCP,
a simple urine test may be used to diagnose or exclude CC.
Moreover, as this model accurately discriminates CC from
normal controls and other BBDs it may be of relevance for
population at risk-based CC screening. This is of special interest
as so far no marker or test is useful for CC screening and only
the early detection can extend life expectancy for patients with
CC.

One limitation is that we have not investigated the influence of
demographic and lifestyle differences between the Norwegian
and German patients on the classification results. Nevertheless,
we found no apparent differences in CE-MS peptide profiles
between Norwegian and German cohorts. Another limitation is
that the effect of long-term storage of urine at −20°C on the
proteomic pattern was not explicitly assessed. However, no
significant alterations in the low molecular weight urinary

\begin{table}
\begin{center}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
Peptide & Mass (Da) & CE-time (min) & Protein & Amino acid positions & FDR adj. p* & AUC \n\hline
150909 & 4082.94 & 21.09 & Osteopontin & ESEELNGAYKAIPVAQDLNAPSDWDSRGKDSYETSQL & 2.1E-02 & 0.65 NS \\
168079 & 4787.06 & 22.45 & n.i. & n.i. & 4.5E-02 & 0.69 4.6E-02 \\
181591 & 8110.83 & 19.82 & n.i. & n.i. & 5.2E-03 & 0.74 5.2E-02 NS \\
189663 & 12716.80 & 25.90 & n.i. & n.i. & 5.2E-03 & 0.74 5.2E-02 NS \\
\hline
\end{tabular}
\end{center}
\textbf{Table 3 Continued}
\end{table}

---

**Figure 4** Immunohistochemistry staining for matrix metallopeptidase 1 (MMP-1) in liver tissues from patients with cholangiocarcinoma (CC) and primary sclerosing cholangitis (PSC). Note the strong to moderate staining intensity of MMP-1 in CC compared with the absent expression of MMP-1 in PSC. Magnification 20x and 40x.

---
proteome were observed in a retrospective study in which urine samples were stored up to 2 years at −20°C before analysis.32

The 95% CI of the AUC for patients in the validation cohort was 0.80 to 0.92 with the upper limit being just 1.15 times higher than the lower limit. Hence, while the data indicate high confidence for the point estimate of the AUC of 0.87, it is possible that lower performance may appear in an independent study. Therefore, an attractive approach is the combination of bile and UPA (BPA–UPA) to achieve high specificity. Our data demonstrate that combining both tests from fluid obtained at the same day resulted in complete discrimination of benign strictures from CC. The potential of this combined bile and urine test for PSC surveillance and CC screening must be further assessed in longitudinal observational studies. It will also be considered in future studies whether a combination of UPA with other available non-invasive diagnostic tools, that is, MRI, may be an alternate strategy to improve accuracy of CC diagnosis in patients with unclear biliary strictures.

In summary, the urine peptide marker model differentiates CC from PSC and other BBD and may become, in combination with or as replacement of bile fluid analysis, a new diagnostic non-invasive tool for PSC surveillance and early CC detection, hopefully at a curable stage.

Contributors JM and TOL designed the experiments and wrote the manuscript; AAN, T.J.W, THK and TOL performed patient diagnosis, patient enrolment and collection of clinical data; JM, RRP and WM performed the experimental work; JW, HM, MD and MPM supervised the experiments and their analysis.

Funding This work was supported by the Niedersächsischen Krebsgesellschaft and a grant from the German Federal Ministry of Education and Research (reference number: O1E00802). JM and HM are supported by grant 260844 from the HEPACUTE EU-project.

Competing interests HM is founder and co-owner of Mosaikes diagnostics, which developed the CE–MS technology and the MosaikesViu and MosaCluster software. JM and MD are employees of Mosaikes diagnostics.

Patient consent Obtained.

Ethics approval Ethics approval was provided by the local ethical committees. Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


Urine proteomic analysis differentiates cholangiocarcinoma from primary sclerosing cholangitis and other benign biliary disorders


Gut 2013 62: 122-130 originally published online May 12, 2012
doi: 10.1136/gutjnl-2012-302047

Updated information and services can be found at:
http://gut.bmj.com/content/62/1/122.full.html

These include:

Data Supplement
"Supplementary Data"
http://gut.bmj.com/content/suppl/2012/05/12/gutjnl-2012-302047.DC1.html

References
This article cites 32 articles, 5 of which can be accessed free at:
http://gut.bmj.com/content/62/1/122.full.html#ref-list-1

Article cited in:
http://gut.bmj.com/content/62/1/122.full.html#related-urls

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Hepatic cancer (399 articles)
Pancreas and biliary tract (1826 articles)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/