BILIARY TRACT CANCER SERUM PROFILING USING MAGNETIC BEAD-BASED PEPTIDE EXTRACTION AND MALDI-TOF MASS SPECTROMETRY

Stéphane Camuzeaux

Cancer Proteomics Laboratory
Institute for Women’s Health
University College London

Joint BSPR/EBI Conference 2011
Background

- Cholangiocarcinoma and gall bladder cancer are referred to as Biliary Tract Cancer (BTC). Around 1,600 death each year in the UK

- Early diagnosis is key to reducing mortality

- BTC is uniformly fatal unless detected early with the potential for surgical resection and has a dismal prognosis

- The standard for diagnosis is cytological or histological confirmation of malignancy within the biliary stricture; involving invasive procedures

- Need non-invasive alternatives
Background

- Most commonly used blood marker is CA19-9 (carbohydrate antigen 19-9/sialylated Lewis (a) antigen)

- However, CA19-9 lacks adequate sensitivity and specificity; it is often elevated in benign conditions (cholangitis and pancreatitis) and undetectable in 7% of the population who are Lewis (a) negative

- Need to find better diagnostic markers for early BTC detection, preferably from human serum

- It has been postulated that the serum peptidome may be a valuable source of diagnostic cancer biomarkers, specifically in relation to the activity of tumour-related exopeptidases
Generation of surrogate tumour markers in the blood

Villanueva et al., 2005 JCI

Tumour Cell → Protease → Serum protein → Tissue microenvironment → Circulation → Serum protein → Clotting proteases → Surrogate tumour marker → Clotting

Villanueva et al., 2005 JCI
Serum peptidome profiling

- Peptide signatures diagnostic and specific for prostate, breast and bladder cancer were found by MALDI-TOF profiling using C8-coated magnetic beads for peptide extraction

- Ladders of peptides identified as products of abundant serum proteins hypothesised to be generated ex vivo (at clotting) by tumour-specific exopeptidase activities – surrogate markers

From Villanueva et al., 2006 JCI
High-throughput semi-automated serum peptide profiling by MALDI-TOF MS

Serum collection from patient

Robotic Handling

C18 magnetic bead extraction

Automated mass spectrometry acquisition

Data processing and statistical analysis

Two independent sets: - discovery set - validation set
Biomarker discovery workflow

- Platform reproducibility assessed (intra/inter assay precision) using QC Sigma serum
- Analyse case and control serum samples
- Data analysis to find discriminatory peaks between groups
- Generation of a model
  - Discovery set: training set and test set
- Validation of the model (validation set)
- Identification of peaks of interest
Sample handling

- Blood collected in gel tubes (gel plug)
- Tubes inverted five times
- 60 min clotting at room temperature
- Centrifuged, aliquoted
- Storage at -80°C

- **Discovery set:** 95 case control serum samples collected from patients diagnosed with BTC, benign biliary strictures and healthy volunteers attending University College London Hospital between 2006 and 2008

- **Validation set:** 14 BTC and 16 healthy volunteer samples collected in 2009 and 2010

(Healthy control volunteers had no active illnesses and were not on medication)
Platform reproducibility

-Three replicate runs, four spotting replicates per sample

-Intra- and inter-assay variation for quality control using Sigma serum:

**Discovery set:**
-27 Sigma serum across 12 MALDI targets:
  -Average intra-assay variation (all peaks; S/N>3): 10.6% +/- 7.2
  -Inter-assay variation (all peaks; S/N>3): 12.8% +/- 6.7

**Validation set:**
-24 Sigma serum across 4 MALDI targets:
  -Average intra-assay variation (all peaks; S/N>3): 12.8% +/- 10.0
  -Inter-assay variation (all peaks; S/N>3): 14.5% +/- 10.8
Platform reproducibility

![Graph showing platform reproducibility](image)
Spectral filtering

- 1000 shots acquired for each spotting (up to 12 positions, 100 shots per position)
- At least 3 spotting replicates with 1000 shots per sample
- At least 2 run replicates per sample

<table>
<thead>
<tr>
<th>Number of samples included after filtering:</th>
<th>Discovery set</th>
<th>Validation set</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of samples run</td>
<td>95</td>
<td>30</td>
</tr>
<tr>
<td>number of spectra acquired</td>
<td>1140</td>
<td>360</td>
</tr>
<tr>
<td>number of spectra included after filtering</td>
<td>1079</td>
<td>355</td>
</tr>
<tr>
<td>% of spectra included</td>
<td>94.6</td>
<td>98.6</td>
</tr>
<tr>
<td>number of spectra included after visual inspection</td>
<td>1069</td>
<td>347</td>
</tr>
<tr>
<td>% of spectra included after visual inspection</td>
<td>93.8</td>
<td>96.4</td>
</tr>
<tr>
<td>total number of samples included</td>
<td>92</td>
<td>30</td>
</tr>
<tr>
<td>% of sample included</td>
<td>96.8</td>
<td>100</td>
</tr>
</tbody>
</table>

3 samples removed from discovery set
## Sample sets after filtering

**- Discovery set: 92 samples**

<table>
<thead>
<tr>
<th>Group</th>
<th>Patients</th>
<th>Gender</th>
<th>Median age (yrs)</th>
<th>Age range (yrs)</th>
<th>Median bilirubin (g/L)</th>
<th>Median CA19-9 (IU/mL)</th>
<th>CA19-9 &gt;37 IU/mL</th>
<th>BTC stage &lt;T3</th>
<th>BTC stage ≥T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>22</td>
<td>7F:15M</td>
<td>60</td>
<td>39-78</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>PSC</td>
<td>10</td>
<td>3F:7M</td>
<td>48</td>
<td>22-76</td>
<td>17</td>
<td>17</td>
<td>3/10</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>AIP/IAC</td>
<td>7</td>
<td>7M</td>
<td>63</td>
<td>43-71</td>
<td>12</td>
<td>15</td>
<td>1/4</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Benign other</td>
<td>14</td>
<td>9F:5M</td>
<td>53</td>
<td>35-74</td>
<td>8</td>
<td>--</td>
<td>--</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**- Validation set: 30 samples**

<table>
<thead>
<tr>
<th>Group</th>
<th>Patients</th>
<th>Gender</th>
<th>Median age (yrs)</th>
<th>Age range (yrs)</th>
<th>Median bilirubin (g/L)</th>
<th>Median CA19-9 (IU/mL)</th>
<th>CA19-9 &gt;37 IU/mL</th>
<th>BTC stage &lt;T3</th>
<th>BTC stage ≥T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTC</td>
<td>14</td>
<td>7F:7M</td>
<td>73</td>
<td>44-90</td>
<td>19</td>
<td>404</td>
<td>11/14</td>
<td>6/14</td>
<td>8/14</td>
</tr>
<tr>
<td>Healthy</td>
<td>16</td>
<td>8F:8M</td>
<td>34</td>
<td>23-80</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Average MALDI-TOF spectra

Discovery set: BTC (red) and healthy (green)
Discriminatory peaks in discovery set

- 8 peaks found to significantly discriminate BTC from healthy (p<0.001, average fold ≥ 2)
- 5 peaks discriminate BTC from benign group (p<0.05, average fold ≥ 1.5); 4 common
ROC Curve Analysis

- ROC curve BTC vs. Healthy AUC:

  - m/z 2606.6 : 0.82
  - m/z 2084.1 : 0.86
  - m/z 2556.8 : 0.91
  - m/z 1265.3 : 0.91
  - m/z 2212.3 : 0.92
  - m/z 1352.6 : 0.92
  - m/z 5812.6 : 0.96
  - m/z 2906.0 : 0.97

  ![BTC vs. H m/z 5812.6](image1)
  ![BTC vs. H m/z 2212.3](image2)

  **BTC vs. H m/z 5812.6**
  **AUC = 0.96**

  **BTC vs. H m/z 2212.3**
  **AUC = 0.92**

- ROC curves BTC vs. Benign AUC:

  - m/z 5812.6 : 0.70
  - m/z 2935.7 : 0.71
  - m/z 2556.8 : 0.73
  - m/z 2906.0 : 0.76
  - m/z 2212.3 : 0.77

  ![BTC vs. Benign CA19-9](image3)

  **BTC vs. Benign CA19-9**
  **AUC = 0.82**
Model generation and validation

- 5 permutations of: 75% BTC/Healthy training set; 25% BTC/Healthy test set

- Models generated per permutation:
  - Genetic Algorithms (GA) or Support Vector Machine (SVM)
  - Number of k-Nearest Neighbours (k-NN)
  - Number of peaks

- 20% Leave Out Cross Validation (20% LOCV)

- Best performing model SVM 8 peaks, 3 k-NN, 20% LOCV 95.2% :

  \[ m/z: \] 1021.7, **1265.3**, 1352.6, 1364.9, **2556.8**, **2906.0**, 5070.7, 8779.6

- Classification of independent validation set: BTC (n=14) / Healthy (n=16)

  **Sensitivity** = 85.7%; **Specificity** = 100%; **PPV** = 100%; **NPV** = 88.9%
Peak identification

- BTC samples pool and healthy samples pool prepared
- Extraction using C18 magnetic bead
- Extracts split in two for parallel top-down analysis: GeLC-MS/MS / Zip-Tip LC-MS/MS

<table>
<thead>
<tr>
<th>Av Mass (m/z)</th>
<th>Name</th>
<th>Fragment Sequence</th>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1021.7</td>
<td>Fibrinopeptide A</td>
<td>DFLAEGGGVR</td>
<td>yes; Villanueva et al; Tiss et al</td>
</tr>
<tr>
<td>1265.3</td>
<td>Fibrinopeptide A</td>
<td>GEGDFLAEGGGVR</td>
<td>yes; Villanueva et al; Tiss et al</td>
</tr>
<tr>
<td>1352.6</td>
<td>Fibrinopeptide A</td>
<td>SGEGDFLAEGGGVR</td>
<td>yes; Villanueva et al; Tiss et al</td>
</tr>
<tr>
<td>2084.1</td>
<td>Fibrinogen alpha</td>
<td>GGSTSYGTGSETESPRNPSSAG</td>
<td>Koomen et al;</td>
</tr>
<tr>
<td>2212.3</td>
<td>HMW kininogen</td>
<td>KHNLSGGHHERQGQGHGQ</td>
<td>Villanueva et al; Tiss et al</td>
</tr>
<tr>
<td>2556.8</td>
<td>Fibrinogen alpha</td>
<td>SSSYSKQFTSTSYNRGDESTFES</td>
<td>Villanueva et al; Tiss et al</td>
</tr>
<tr>
<td>2935.7</td>
<td>Fibrinogen alpha</td>
<td>SSSYSKQFTSTSYNRGDESTFESKSY</td>
<td>Villanueva et al; Tiss et al</td>
</tr>
<tr>
<td>5812.6</td>
<td>Fibrinogen alpha</td>
<td>SSSYSKQFTSTSYNRGDESTFESKSYKMADEAGSEADHEGTHSTKRGHAHKS</td>
<td>yes</td>
</tr>
<tr>
<td>1364.9</td>
<td></td>
<td></td>
<td>no</td>
</tr>
<tr>
<td>2606.6</td>
<td></td>
<td></td>
<td>no</td>
</tr>
<tr>
<td>2906.0*</td>
<td></td>
<td></td>
<td>no</td>
</tr>
<tr>
<td>5070.7</td>
<td></td>
<td></td>
<td>no</td>
</tr>
<tr>
<td>8779.6</td>
<td></td>
<td></td>
<td>no</td>
</tr>
</tbody>
</table>

**bold**=discriminatory peak  ■■ = used in model  BTC vs healthy
Conclusions

- Applied a semi-automated MALDI-TOF MS serum peptidome profiling strategy to a set of BTC case control samples

- Identified discriminatory peaks for BTC vs. healthy. Less robust discrimination of BTC vs. benign group (inflammatory response?)

- Tested a model on an independent validation set that accurately classified BTC cases from healthy control

- Identified peaks used in the model; mostly fragments of abundant serum proteins, suggesting tumour-specific exopeptidase activities

- Need to further define and test models for discrimination of BTC vs. benign group

- Further work: specific assays to develop a clinical test (e.g. using SRM)
Acknowledgements

Cancer Proteomics Laboratory
Institute for Women’s Health UCL

Dr John F. Timms
Dr John Sinclair

Institute of Hepatology UCL & UCLH

Dr Stephen P. Pereira
Dr Neomal S. Sandanayake
Dr George Webster
BILIARY TRACT CANCER SERUM PROFILING USING MAGNETIC BEAD-BASED PEPTIDE EXTRACTION AND MALDI-TOF MASS SPECTROMETRY

Stéphane Camuzeaux

Cancer Proteomics Laboratory
Institute for Women’s Health
University College London