Coronary heart disease related proteomics

Jaakko Niemi
Contents

- Background of coronary heart disease
- Shotgun proteomics
- SEQUEST algorithm
- Application: HDL proteomics
- Conclusions
Coronary heart disease

- the most common cause of death in the Western World
- annual costs 360 billion US dollars (in USA alone)
- lipoprotein particles have a major role in the development and progression of the disease
Lipoprotein structure and main classes

<table>
<thead>
<tr>
<th>Lipoprotein Fraction</th>
<th>Density range, g/ml</th>
<th>Diameter range, nm</th>
<th>Average C, %</th>
<th>Average TG, %</th>
<th>Average Lipid, %</th>
<th>Average Protein, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM</td>
<td>&lt;0.960</td>
<td>80-1200</td>
<td>4</td>
<td>88</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>VLDL</td>
<td>0.960-1.006</td>
<td>30-80</td>
<td>23</td>
<td>56</td>
<td>92</td>
<td>8</td>
</tr>
<tr>
<td>IDL</td>
<td>1.006-1.019</td>
<td>25-35</td>
<td>43</td>
<td>29</td>
<td>89</td>
<td>11</td>
</tr>
<tr>
<td>LDL</td>
<td>1.019-1.063</td>
<td>16-25</td>
<td>58</td>
<td>13</td>
<td>79</td>
<td>21</td>
</tr>
<tr>
<td>HDL2</td>
<td>1.063-1.125</td>
<td>8.5-12</td>
<td>41</td>
<td>16</td>
<td>67</td>
<td>33</td>
</tr>
<tr>
<td>HDL3</td>
<td>1.125-1.210</td>
<td>4-8.5</td>
<td>35</td>
<td>13</td>
<td>48</td>
<td>52</td>
</tr>
</tbody>
</table>
Lipoprotein metabolism
Development of a foam cell
Progression of an atherosclerotic lesion
Plaque rupture and thrombosis
HDL particles

- "the good cholesterol"
- inversely related to CHD
- maintain reverse cholesterol transport
- may have anti-inflammatory and anti-oxidative properties
Shotgun proteomics

- direct analysis of large protein complexes (DALPC)
- capable of identifying individual proteins in large protein complexes without first purifying each protein component to homogeneity
Shotgun proteomics

- 2D Chromatography (charge and hydrophobicity)
  - Strong cation exchange (SCX)
  - Reversed-phase (RP)
- SEQUEST algorithm for identifying proteins in MS spectra
SEQUEST

- tandem mass spectrometry data analysis method used for protein identification
- identifies collections of tandem mass spectras comparing these spectras to the protein sequences found in protein databases
- identifies each tandem mass spectrum individually.
SEQUEST

- Step 1: Tandem mass spectrometry data reduction
- Step 2: Searching amino acid sequences from protein database
- Step 3: Scoring method and theoretical spectra construction
- Step 4: Cross-correlation analysis
Step 1: MS data reduction

- Aim: to eliminate noise and to reduce the number of ions to be considered
- all but the 200 most abundant ions are removed
- the remaining ions are renormalised

processed MS data for the peptide DLRSWTAADTAAQISQ
Step 2: Search method

- possible amino acid sequences are identified in a protein database
- protein sequences are retrieved and scanned to find linear combinations of amino acids that match the mass of the peptide
Step 3: Scoring method

- The fragment ions are predicted for the sequences derived from database
- They are compared to the MS spectra using a scoring function
- Theoretical spectra is created for the top 500 best fit sequences

Theoretical spectra for the peptide DLRSWTAADTAAQISQ
Step 4: Cross-correlation analysis

- Theoretical spectras are compared to the experimental spectra using cross-correlation.
- The final score attributed to each candidate peptide sequence is the value of the function at $\tau=0$ minus the mean of the cross-correlation function over the range $-75<\tau<75$.
- The scores are normalized to 1.0 and the peptide sequence with the biggest score is selected.

Cross-correlation function for the peptide DLRSWTAADTAQISQ
Application: HDL proteomics

- Quantify the protein composition of HDL → provide insights into the properties of HDL
- 33 subjects (7 CAD patients)
- Shotgun proteomics used for protein identification
Results (1/2)

- 48 proteins identified in HDL
  - 22 linked to cholesterol and lipoprotein metabolism
  - 23 acute-phase-response proteins
- Fundamental role for HDL in innate immunity
Results (2/2)

- HDL₃ from CAD patients is selectively enriched in some proteins related to inflammatory response
- Individual protein levels as markers of CAD?
Discussion

- HDL proteomics provides new information on HDL functions
- The degree of variation in the number and identity of HDL proteins large
- Future direction: find simple HDL biomarkers to predict CHD
Conclusions

- Lipoproteins are closely related to coronary heart disease
- Shotgun proteomics is a useful method to identify individual proteins in large protein complexes
- SEQUEST method can be used to identify proteins from experimental MS spectras
- Proteomics may provide new information on coronary heart disease