Urinary Proteomics in Kidney and Cardiovascular Disease

Harald Mischak

A complex organism, experiencing a multitude of complex environmental impacts, can not be described in detail by single features

Even microscopic structures reveal high complexity, e.g., vascular architecture
Urinary Proteomics in Kidney and Cardiovascular Disease
Harald Mischak

Single biomarkers are of limited specificity

- C-reactive protein

  Biomarker for e.g.,
  - Viral infection
  - Coronary artery disease
  - Sepsis
  - Diabetic nephropathy

Limited precision of single marker X

Increased precision of two markers, X and Y
Higher precision of 3 markers, X, Y and Z

Omics technologies or: why proteomics?

Genomics
Proteomics
Metabolics

Recommended steps for clinical proteomics:

1. Define a clear clinical question and how the outcome of the study would improve the diagnosis and/or treatment of the disease.
2. Define the patient and control populations, clinical data to be collected, as well as protocols for sampling and sample preparation.
3. Define the type of samples needed for the discovery and validation phases.
4. Define and validate the analytical platforms for discovery phases for validated assay set.
5. Obtain IRB approval and written informed consent from the participant.
6. Perform a pilot study on a validated discovery platform.
7. Statistically evaluate data from the pilot study to calculate the number of cases and controls for the training set.
8. Perform study of the training set on the validated platform based on the calculated number of cases and controls.
9. Evaluate findings on blinded samples.
10. Deposit datasets in a public database.
11. Using these results, transfer the assay to the application platform and test using a training (if applicable) and subsequently a blinded set.
12. Apply towards clinical use to show whether the findings improve the current clinical situation.
Clinical diagnosis of diabetic nephropathy

- Urine albumin: Urine albumin is commonly expressed as albumin/creatinine ratio; An increase $\geq 30$mg/g may be a sign of kidney damage
- eGFR (estimated glomerular filtration rate): The calculation of eGFR is based on the amount of serum creatinine

Markers of pathophysiological relevance are needed that are more likely to indicate early onset of chronic kidney diseases

Why urine?

- Easily accessible
- Obtained non invasive
- Available in large quantities
- Urinary polypeptides are stable, yielding comparable datasets
- Urinary polypeptides display the “status” of the kidney
Why not blood?

- Obtained invasive, in moderate quantities
- Complex mixture of particles (cells), lipids, and water-soluble compounds; First two have to be removed
- High variability due to intrinsic protease activity
- Highly abundant proteins that currently cannot be effectively removed obscure low abundance biomarkers
- While blood contains biomarkers for disease, these are generally not accessible with current proteomics technologies, due to:
  1. High variability due to pre-analytical parameters
  2. High abundant proteins obscuring biomarkers
  3. Requirement to analyze thousands of samples to obtain statistical significance

Technology platforms

2DE-MS

Technology platforms (2)

Multiple LC-MS-MS datasets
Urinary Proteomics in Kidney and Cardiovascular Disease
Harald Mischak

Technology platforms (3)
Proteomics Technology platform: CE/MS Technology
Capillary Electrophoresis coupled to Mass Spectrometry

- Separation and analysis of proteins and peptides (>1,000)
- Run time ~60 min
- CE
  - Fast
  - Robust
  - Inexpensive
  - Reproducible
- MS
  - Resolution
  - Scan speed

CE-MS analysis
Data deconvolution using MosaiquesVisu
Calibration using ProteinCluster on SQL
ID allocation on SQL Server
MS/MS sequencing of potential biomarkers
Sample
Sample preparation
Raw data
Data deconvolution using MosaiquesVisu
ID list
Predefined biomarkers
Biomarker model for Diagnosis, Prognosis

Video

Summary of data flow

The screen versions of these slides have full details of copyright and acknowledgements
Urinary Proteomics in Kidney and Cardiovascular Disease
Harald Mischak

Human urinary LMW proteome database

Clinical data
- Age
- Gender
- Urinary albumin/creatinine
- Cholesterol (mmol/l)
- Creatinine (micromol/l)
- Patients history

Database

Sequence Information

CE-MS peptidome profile

Biomarker selection

Statistics

100-Specificity

Sensitivity

Compiled pattern

Individual analyses

The screen versions of these slides have full details of copyright and acknowledgements
Classification using n-dimensional models

How to identify valid biomarkers and biomarker models?

- How many samples are required to define useful biomarkers?
- Relevance of Statistics?
- Which classifiers perform best?
- Is assessment in a blinded set necessary/helpful?

How to identify valid proteomic biomarkers?

CE-MS analysis

Healthy female 19-40 years
- Evaluation of all individual biomarkers
- Blinded testset
- 134 samples

Healthy male 19-40 years
- Evaluation of biomarker models
- 

The screen versions of these slides have full details of copyright and acknowledgements.
Adjustment for multiple is mandatory in statistical assessment of multidimensional proteomic data

<table>
<thead>
<tr>
<th>No. of samples</th>
<th>No. of biomarkers @ p-value &lt; 0.05</th>
<th>No. of biomarkers @ p-value &lt; 0.05 after FDR adjustment</th>
<th>No. of biomarkers @ p-value &lt; 0.05 after Bonferroni</th>
<th>Biomarkers found valid (p-value &lt; 0.05) in testset</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>41</td>
<td>0</td>
<td>0</td>
<td>1/0/0</td>
</tr>
<tr>
<td>40</td>
<td>151</td>
<td>0</td>
<td>0</td>
<td>25/0/0</td>
</tr>
<tr>
<td>65</td>
<td>232</td>
<td>48</td>
<td>1</td>
<td>67/18/1</td>
</tr>
<tr>
<td>134</td>
<td>323</td>
<td>136</td>
<td>23</td>
<td>124/71/19</td>
</tr>
</tbody>
</table>

Correct estimation of effect size is dependent on number of samples

Inappropriately low numbers of independent samples result in „erroneous biomarkers“
Urinary Proteomics in Kidney and Cardiovascular Disease
Harald Mischak

Increasing sample size results in increasing number of biomarkers

Classification performance using different size training sets

Classification error depends on sample size
Urinary Proteomics in Kidney and Cardiovascular Disease
Harald Mischak

Case
• Diseased
• Treated (drug)

Control
• Healthy
• Untreated (placebo)

Kolch et al., RCM 2004, 18: 2635-2636
Neuhoff et al., RCM 2004, 18: 149-156
Mischak et al., PROTEOMICS - Clinical Applications 2007, 1: 148-156

Proteomic CKD biomarker discovery

CKD pattern (n=279 biomarker):
- Different collagen
- Plasma proteins (serum albumin, transferrin, alpha-1-antitrypsin, alpha-1B-glycoprotein, alpha-2-HS-glycoprotein, antimicrotubulin, ApoA-I, beta-2-microglobulin, fibrinogen alpha)
- Quinine
- Urobilinogen
- Nephrectomy
- Complement component 4b
- Collagenase 2
- Prostaglandin D-synthase
- Proprotein convertase subtilisin/kexin type 1 inhibitor
- Polymeric immunoglobulin receptor
- Osteopontin
- Neurosecretory protein VGF
- Membrane associated progesterone receptor component 1
- CD99 antigen
- Ig lambda chain C regions

Structural (ECM) Proteins

Collagens:
- Specific fragments decreased in urine of CKD patients
- Function: main proteins of connective tissue (~30% of whole-body protein content)
- PTM: Proline residues at the second or third position of the tripeptide repeating unit (G-X-Y) are hydroxylated
- Cellular component: extracellular matrix, secreted
- Expressed in glomeruli and renal tubulointerstitial regions
- Products of myofibroblasts, degraded by MMPs

Most common collagens in the biomarker pattern:
- COL1A1 (n=126)
- COL3A1 (n=54)
- COL2A1 (n=15)

© 2010 www.proteinatlas.org

The screen versions of these slides have full details of copyright and acknowledgements
Urinary Proteomics in Kidney and Cardiovascular Disease
Harald Mischak

CKD-biomarkers and their regulation (2)

Apoptosis-related proteins

Osteopontin:
- Subcellular location: secreted
- Expressed in renal tubules and in mineralized tissues
- Remodeling of extracellular matrix
- Inhibition of apoptosis

Clusterin/Apolipoprotein J:
- Decreased in urine of CKD patients
- Subcellular location: secreted
- Expressed in different tissues
- Remodeling of extracellular matrix
- Correlated inversely with proteinuria in CKD patients
- Associated with apoptosis

Biomarker validation

- Diabetes scoring: 98% Sensitivity
- Chronic kidney disease: AUC < 0.95
- Diabetic Nephropathy (vs. Normo) AUC ~ 0.95

Pathophysiological suggestions

- Acute phase response: u-ATAT↑
- Inhibition of plasminogen: u-FIBA↓
- Glomerular basement deposition: u-α2M↓
- Apoptosis/chronic inflammation (ROS, AGE): u-ALB↑, u-B2M↑, u-FIBA↓
- Reduced renal function/proteinuria: Decrease of GFR
- Inhibition of matrix metalloproteinases (e.g., MMP2 and MMP9):
  - Increased collagen accumulation
  - Reduced degradation of collagens
- Glomerular fibrin deposition
- Renal damage
- Decrease of GFR

The screen versions of these slides have full details of copyright and acknowledgements
Correlation of selected urinary peptides with CKD staging

Longitudinal analysis of diabetic type 2 patients for progression of DN

Facts:
- Longitudinal study of type 2 diabetic patients (progressors and non-progressors), normoalbuminuric at beginning
- Urine samples nearly once a year of each patients over approximately 10 years
- Definition of diabetic nephropathy: AER>200 µg/min and retinopathy
- Progression: positive predictive value: 91%; negative predictive value: 87%

Prediction of DN

S-crea: serum creatinine [µmol/L]
UPA: urinary proteome analysis
AER: albumin excretion [µg/min]

The screen versions of these slides have full details of copyright and acknowledgements
Early detection of DN in normoalbuminuric diabetics

CKD biomarker profile vs. AER

Multicentric European PRIORITY trial

Diabetes mellitus Typ 2
normoalbuminuric N = 3280

Coronary artery disease
Urinary Proteomics in Kidney and Cardiovascular Disease
Harald Mischak

Proteomics of coronary artery disease

Assessment of therapy success

Assessment of therapy success
Summary

- (Urinary) proteomic biomarker panels enable (early) detection, prognosis, and assessment of therapy of a variety of diseases, with 80 - 98% sensitivity and specificity.
- Biomarkers and biomarker panels were validated in multiple independent multi-centric blinded studies.
- Variability in single biomarkers is counteracted by diagnostic patterns that tolerate instability and inconsistency of individual polypeptides/biomarkers.
- Strict statistical testing is key to clinical application.
- Large numbers (>1000) of comparable datasets is key to success.
- Urinary proteome indicates disturbed collagen turnover as key molecular change in diabetes-associated vascular (CKD and CAD) disease.
- Significant overlap, yet also difference between CAD and CKD/DN can be detected on a molecular level.
- Urinary biomarkers have the potential to effectively and substantially improve treatment as molecular indicators of disease.