Biomarker Discovery in the Time Dependant Progression of Lung Fibrosis

The 4800 Plus MALDI TOF/TOF™ Analyzer Combined with iTRAQ® Reagents-8plex and ProteinPilot™ Software Provides an Integrated Solution for Protein Quantification in Large Datasets

Katy Williams¹, Kate Blease², Lily Wong² and Christian Lombardo².
¹Applied Biosystems, Foster City, USA, ²Celgene Corporation, San Diego, USA

Introduction

Idiopathic Pulmonary Fibrosis (IPF) is characterized by excessive deposition of extracellular matrix in the interstitium resulting in respiratory failure that results in organ damage and eventually death. The mechanisms underlying the development and progression of the disease are not fully understood and no pharmacological therapy has been proven unequivocally to alter or reverse the inflammatory process of IPF¹. Drug development programs are assisted by identification of both markers of disease and mechanism, and also drug efficacy biomarkers.

The goal of this study is to identify biomarkers with clear linkage to the biology associated with the progression of fibrosis in an animal model that are translatable to humans. Intratracheal injection of bleomycin in rodents serves as a model for progression of lung fibrosis in humans.

We used an unbiased proteomic strategy to identify and quantify changes in the mouse lung proteome using iTRAQ® Reagents - 8plex labeling. The combined labeled samples were separated by 2D LC and analyzed on the 4800 Plus MALDI TOF/TOF™ Analyzer. Protein identification and quantitation of relative protein ratios were performed using ProteinPilot™ software 2.0. The expression profiles of several candidate biomarkers were verified by orthogonal measurements of protein abundance.

Key Features for Biomarker Discovery in Large Datasets

• Highest order multiplexing with iTRAQ® reagents - 8plex enables relative quantitation of up to eight samples in a single LC MALDI run. This results in increased statistical control with sample replicates as well as more confidence in protein quantitation from multiple peptide measurements per protein.

• High sensitivity MS/MS data enabled by the OptiBeam™ on-axis laser and true MS/MS in CID cell with tunable collision energy. The 4800 Plus MALDI TOF/TOF Analyzer provides excellent peptide fragmentation spectra plus high quality low mass data for reliable reporter ion signal.

• ProteinPilot™ software 2.0 with the Paragon™ Algorithm identifies more peptides with its unique ability to search for many peptide modifications and non-conformant enzyme cleavages simultaneously. In addition, the ProGroup™ algorithm enables isoform specific iTRAQ® reagent quantification, critical for correctly quantifying isoforms.
Experimental Design

**Animals:** C57Bl6 mice were used for this study. Bleomycin sulphate (Sigma; USA) was dissolved in 0.9% saline and administered as a single intratracheal dose of 0.075U in a volume of 30µL per animal (six animals for each time point). All animals received intratracheal instillations of either bleomycin or saline on day 0 as previously described. Animals were sacrificed on day 0 (control) or days 1, 7 and 14 post bleomycin administration by CO2 asphyxiation, and approximately 500µL of blood was removed from each mouse into heparinized tubes to yield plasma. The trachea was isolated and cannulated and a bronchoalveolar-lavage (BAL) was performed using 1mL filter-sterilized normal saline. Lungs were perfused with saline via the right ventricle to remove blood from the vascular bed of the lung. Finally, whole lungs were dissected from each mouse and snap frozen in liquid N2.

**iTRAQ® Reagent - 8plex labeling:** Proteins were isolated from lung tissue, digested with trypsin, and labeled with iTRAQ reagents according to the Applied Biosystems protocol.

**Chromatography:** The combined sample was separated by strong cation exchange chromatography (some SCX fractions were pooled) prior to separation by reverse-phase HPLC on a Pepmap C18 column 100 um x 150mm (LC Packings) using a 60 minute gradient (2-35% acetonitrile, 0.1% TFA, 500 nL/min). HPLC eluent was mixed 1:2 with matrix (alpha-cyano-hydroxy cinnamic acid) and 20 sec fractions spotted onto MALDI sample plates.

**Mass Spectrometry:** LC MALDI MS/MS was performed on using a 4800 Plus MALDI TOF/TOF™ Analyzer. MSMS data were acquired using 2 kV collision energy with CID pressure (air) of 5x10⁻⁶ Torr. Automatic stop conditions were employed that vary the total number of laser shots to maximize data quality while minimizing sample consumption.

**Data Processing:** Protein identification and quantitation were performed using the Paragon™ Algorithm in Thorough search mode in ProteinPilot™ Software 2.0. Proteins were identified by searching against the UniProt SwissProt database 20061012. The database was appended with a complete reverse sequence database to enable reverse database searching for the assessment of false positive rates. The protein ratios for all mice (n=6) at a given time point were averaged to obtain a final average ratio for each protein. Protein ratios at days 7 and 14 that were greater than 1.5 fold relative to day 0 and also showed less than 25% standard deviation between animals at each time point were considered putative biomarkers. The data were also analyzed using ANOVA, p=0.01.

**Western blots:** Samples were separated on an Invitrogen 4-12% Nupage Novex Bis-Tris Mini Gel. Gels were transferred to nitrocellulose membrane in an Invitrogen X Cell II Blot Module. Membranes were blocked in Odyssey Blocking Buffer for 1 hour. Membranes were incubated at 1:1000 dilution in Odyssey Blocking Buffer with 0.1% Tween-20 antibody over night at 4 deg C, washed three times with PBS/0.1% Tween-20, then incubated for 1 hour at room temp with Alexa-Fluor secondary antibody (goat-anti-mouse or goat anti-rabbit; 1:10,000 dilution). After washing three times with PBS/0.1% Tween-20 as before, the membrane was scanned on a LiCor instrument. Quantitation was performed by integrating intensities of the bands. Data were expressed as a ratio of putative biomarker/beta-tubulin and then normalized to one for control treatment.
Protein Identification and Quantitation Results using ProteinPilot™ Software

Over 850 proteins were identified with high degree of confidence (>95%) in each experiment (Table 1). The data were searched against a concatenated forward and reverse database to allow an estimation of false discovery rate. This reverse database searching confirmed a 5% instantaneous false discovery rate for each dataset at the 95% protein confidence level.

Data Analysis Criteria for Biomarker Identification

In this study of time-dependent disease progression, biomarkers of interest will show initial changes in protein abundances that are stable over time, i.e. stasis biomarkers that increase or decrease during disease progression and remain altered (Figure 3). Additionally, inter-animal variability exists in disease model studies and is often a confounding factor. Therefore, averaged protein quantitation data for replicate mice at each time point should show low variability in order to be most useful as biomarkers. A 1.5-fold change and standard deviation of less than 25% between animals were considered as potential biomarkers. Eleven proteins were identified as down-regulated and 10 proteins were identified as up-regulated on days 7 and 14. ANOVA analysis of the average iTRAQ® reagent ratios was also used to measure significant changes (p=0.01) in protein abundance on days 7 and 14 and a further 9 putative biomarkers were identified.

Table 1. Protein Identification Results. Summary of proteins identified at the 95% confidence level from mouse lung tissue using ProteinPilot™ Software.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Distinct Peptides</th>
<th>Distinct Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>6042</td>
<td>944</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>8432</td>
<td>907</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>5956</td>
<td>864</td>
</tr>
</tbody>
</table>

Figure 3. Identification and Quantitation Evidence for Peptides from Biomarker #26 and Biomarker #12. The top panel shows a spectrum from a protein whose abundance increases at the beginning of disease progression and remains high. The lower panel shows a spectrum from a protein whose abundance decreases and remains low. The insets show the normalized iTRAQ® reagent ratios for each peptide.
Confirmation of Expression Changes using Western blots.

To assess the correspondence of the iTRAQ® reagent data with an independent assay of protein abundance, we performed Western blot analysis. Of the 30 putative biomarkers, eight proteins have commercially available antibodies. Quantitation was performed by integrating fluorescent intensity of each band, the data are expressed as a ratio of putative biomarker/β-tubulin (control) and then normalized to one for control treatment, day 0 (Figure 5). The relative protein expression ratios of seven proteins over the time course measured by immunoblot analysis was generally in good agreement to the ratios derived by iTRAQ reagent analysis, and in all cases the trend of up- or down-regulation was confirmed (Table 2).

The immunostaining intensity of one protein, Marker 5, indicated a much higher abundance relative to the iTRAQ reagent ratios determined for this protein. Studies are underway both to confirm the specificity of this antibody and to verify the expression of this protein via orthogonal strategies.

Table 2. Comparison of Protein Expression Changes, MS and Western Blot Ratios. The protein expression changes for the iTRAQ reagent and Western Blot analysis are similar for eight putative markers. *The specificity of this antibody is being confirmed.

<table>
<thead>
<tr>
<th>Protein</th>
<th>iTRAQ</th>
<th>Western</th>
<th>Verified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker 1</td>
<td>2.3</td>
<td>4.8</td>
<td>✓</td>
</tr>
<tr>
<td>Marker 2</td>
<td>2.2</td>
<td>2.9</td>
<td>✓</td>
</tr>
<tr>
<td>Marker 3</td>
<td>1.8</td>
<td>2.3</td>
<td>✓</td>
</tr>
<tr>
<td>Marker 4</td>
<td>1.7</td>
<td>2.4</td>
<td>✓</td>
</tr>
<tr>
<td>Marker 5</td>
<td>1.6</td>
<td>37.5*</td>
<td>✓*</td>
</tr>
<tr>
<td>Marker 6</td>
<td>1.6</td>
<td>1.7</td>
<td>✓</td>
</tr>
<tr>
<td>Marker 7</td>
<td>1.6</td>
<td>1.6</td>
<td>✓</td>
</tr>
<tr>
<td>Marker 8</td>
<td>1.5</td>
<td>1.8</td>
<td>✓</td>
</tr>
</tbody>
</table>
Biologically Relevant Results

Differences in protein expression during the time-dependent progression of lung fibrosis were measured using iTRAQ reagents - 8pex. Since multiple animals must be analyzed for each time point, we used a “multiplexed multiplex” strategy to identify putative biomarkers from a total of 24 samples. Averaged protein ratios from 6 mice at each time point were compared over the 4 time points to identify proteins with increased or decreased abundance on days 7 and 14.

The eight putative biomarkers that have been verified by immunoblotting fall into several functional classes consistent with the pathology of fibrosis, which is characterized by protein deposition in lung tissue. Upregulation of several proteinase inhibitors suggests an imbalance of protein degradation contributes to fibrosis. Cell growth is regulated by cell-cell contact and the down-regulation of several cell-cell contact proteins suggests that the loss of this critical regulatory function is also a contributor to the fibrosis pathology. In concordance with this, the appearance of specific markers for the endothelial-mesenchymal transition have been also been observed on day 14. Down-regulation of specific ion channels in this fibrosis model is analogous to loss-of-function mutations observed of conductance channels in cystic fibrosis.

The functions of the differentially expressed proteins can be rationalized in the context of fibrosis and lung disease and suggest a protective response to injury or disease. This implies that measurement of these markers may be very useful for monitoring drug efficacy.
Conclusions

- The multiplexed use of iTRAQ® Reagents - 8plex allowed for the analysis of 24 lung tissue samples in this experiment.

- A 4 point time course was coupled with 6 mice at each time point in order to measure both protein abundance trends over a time course as well as inter-animal variability.

- High sensitivity and broad dynamic range of LC-MALDI on the 4800 Plus TOF/TOF™ Analyzer enabled confident identification of >850 proteins per 8-plex sample.

- Although inter-animal variability exists in disease model studies and is often a confounding factor, many proteins were observed that exhibited >1.5 fold changes with a S.D. of <25% on days 7 and 14 post-bleomycin administration.

- Good correlation between iTRAQ reagent ratios and antibody measurements via immunoblots was observed.

- The biomarkers fall into several functional classes consistent with the pathology of fibrosis.

Future Directions

Currently, only the expression changes for eight proteins have been verified using Western blots due to the lack of available antibodies. The next steps in this work include the verification of the expression changes of the other 22 putative markers from the prepared lung tissue proteins using quantitative Multiple Reaction Monitoring (MRM) on the 4000 QTRAP® system. In addition, the MIDAS™ Workflow will be used to determine which of these proteins are detectable and changing in the plasma of these 24 mice.

References