

Introduction

Biomarker discovery based on differential comparison of identified proteins relies on intensity-driven data dependent MS/MS spectral acquisition, which is widely acknowledged to result in confident identification of only a subset of the actual proteins present in a sample. Alternatively, combining retention time with the high mass accuracy of current TOF-based mass spectrometric systems is a powerful approach for profiling compounds present in complex biological matrices. In this work, we explore an approach that combines profile analysis of samples using accurate mass and retention time with targeted MS/MS of differentially expressed potential markers from the MS step. Informatics approaches to facilitate rapid differential analysis of samples for profiling applications are also described.

Experimental

Model Study

Initial studies were done using a tryptic digest of *E. coli* lysate (BioRad) spiked with both bovine serum albumin and transferrin digests. Samples were prepared at various levels of up- and down-regulation. Samples were prepared three times to model variability. The samples were prepared such that a 2 μ L injection resulted in the amounts on-column shown below in Table 1.

| Sample | <i>E. coli</i> lysate (ng total protein) | BSA (fmol) | Serotriferin (fmol) |
|--------|--|------------|---------------------|
| 1 | 400 | 0 | 0 |
| 2 | 400 | 100 | 200 |
| 3 | 400 | 200 | 100 |
| 4 | 400 | 400 | 50 |
| 5 | 400 | 50 | 400 |

Table 1: Composition of model study samples.

Tissue Culture Samples

In addition to the synthetic samples, tissue culture extracts that model human cardiac disease were analyzed using this approach. Smooth muscle cells (SMC) from human cardiac artery (CASM) were cultured and the cultured cells were exposed to different factors that are known to play a role in many cardiac disease processes. Gene expression studies have already been done on these samples¹.

Experimental

| CASM Samples | Treatment | Comment |
|------------------------|---------------------------------------|--|
| 1 (CASM_VEH) | Control | No treatment except vehicle |
| 2 (CASM_OXDL) | Oxidized low density lipoproteins | Main atherogenic factor for cardiovascular disease |
| 3 (CASM_TNF α) | TNF alpha | |
| 4 (CASMIL1 β) | Interleukin 1 beta | Pro-inflammatory cytokines |
| 5 (CASM_PDFG) | Platelet-derived growth factor (PDGF) | Very strong pro-proliferative cytokines |
| 6 (CASM_TGF β) | TGF beta | Can inhibit cell proliferation in some stages of cells |

Table 2: Treatment of CASM cells

CASM Sample Preparation

CASM cultures were obtained, grown and treated as previously described¹. Table 2 (above) shows the various treatments done. After treatment, cells were lysed in RIPA buffer which contained 0.5M Tris-HCl, 1.5 M NaCl, 2.5% deoxycholic acid, 10 mM EDTA and 10% NP-40. NP-40 is a non-ionic detergent which aids in lysing and solubilizing the proteins. As detergent can interfere with MS analysis, an attempt was made to remove the detergent prior to digestion by diluting and filter in a 5kDa MWCO filter. The samples were then digested using 2,2,2-trifluoroethanol to solubilize and denature proteins as previously described. However, LC/MS analysis showed both signal suppression and a dominant detergent peak at m/z 618. The low critical micelle concentration (CMC; 0.29 mM) of NP-40 resulted in the micelles reforming during concentration in the MWCO filters. However, using a 5 kDa MWCO after digestion and retaining the peptide containing flow-through was effective in removing the detergent.

Nanospray LC/MS/MS

The Agilent 1200 HPLC-Chip/MS system was used with the Agilent 6510 Q-TOF (Fig. 1) for the LC/MS/MS analyses. Conditions were as follows: HPLC-Chip: Protein ID chip with 150 x 0.075 mm analytical column and 40 nL enrichment column. Sample load: 1 μ g of total protein per injection. Flow: 300 nL/min analytical pump, 4 μ L/min loading pump. Mobile phases A: 0.1% FA, B: 90% Acetonitrile, 0.1% FA. Gradients: varied depending on experiment. MS Conditions: Drying gas: 4 L/min, 325°C; Vcap: 1800V; MS Scan range: 300-1800.

Figure 1: HPLC-Chip interfaced to an Agilent 6510 Q-TOF MS.



Results and Discussion

Molecular features were identified using an algorithm that finds the mass peaks in all mass spectra, removes chemical background, clusters by RT and m/z , centroids and determines a peak volume. Related peaks (isotopes, adducts, dimers, trimers, multiple charge states) are combined and assigned a neutral mass and total volume. Extracted features were then evaluated using software that identifies common features and calculates cross-sample response values. Preliminary data from the model study shows that the bovine proteins can be identified with confidence at both the 2X and 4X up/down regulated levels. Examination of mass spectra at the elution time for the peptides showed that the peptide ions were of low relative abundance and it is unlikely that these ions would be automatically selected for data-dependent MS/MS.

Model Study Results

The chromatogram below (Figure 2) shows the total ion chromatogram (TIC) for one of the Sample 2 replicates. The chromatogram shows the complexity of the sample. Figure 3 shows two contour plots; the raw data (left) and the results after processing by the feature extraction algorithm (right). Persistent chemical background evident in the left hand plot (horizontal streaks) has been removed in the right hand plot. Table 3 presents typical mass and retention time precision of extracted features (n=20).



Figure 2: TIC of Sample 2 (#2) from the model study of spiked *E. coli* lysate.

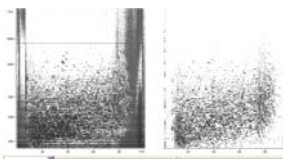


Figure 3: Contour plots of raw (left) and processed (right) data.

| RT (min) | % RSD of RT | Mass | S.D. of Mass (mDa) |
|----------|-------------|-----------|--------------------|
| 9.13 | 1.02 | 1509.7273 | 1.6 |
| 20.10 | 0.32 | 2016.9039 | 2.0 |
| 20.52 | 0.33 | 1310.6421 | 1.3 |
| 27.74 | 0.29 | 1388.6966 | 1.6 |
| 28.35 | 0.31 | 918.5474 | 1.3 |

Table 3: Reproducibility of mass and retention from an earlier model study performed by HPLC-Chip/TOF MS.

GeneSpring MS and Spectrum Mill Analysis

Extracted feature lists were clustered using GeneSpring MS. Clustering results are shown in Figure 4. Specific features were selected for targeted MS/MS analysis and protein database search. Figure 5 show the mass spectrum at the elution apex for one of the target peptides. In this case, the target peptide was the 25th most abundant species and would therefore most likely be missed by a classical data-dependent approach. The Spectrum Mill database search result for this spectrum is shown in Fig. 6.

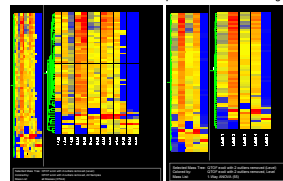


Figure 4: Clustering of features that were significantly different by ANOVA by sample (left) and by level (right).

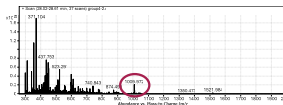


Figure 5: Apex mass spectrum for targeted species.

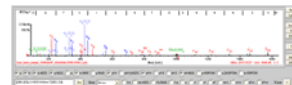


Figure 6: Fragment ion assignments for transferrin peptide

Results and Discussion

Tissue Culture Samples

The base peak chromatograms (BPC) below (Figure 7) show visible differences between three of the tissue culture samples. Differences are also evident in the statistical evaluation results from GeneSpring MS. ANOVA was performed on features passing a minimum cross-sample occurrence frequency filter and significant features were subjected to principal component analysis (PCA). The results (Figure 8-left) showed two clear outliers among the samples. After removal of the outliers, the analysis was repeated (Figure 8-right) with tight clustering of the replicate samples and clear differentiation between the treatments (Figure 9).

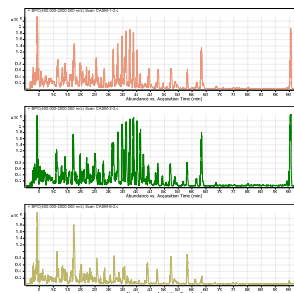


Figure 7: BPCs for control (top), oxidized LDL treated (middle) and TGF beta treated (bottom) cultures.

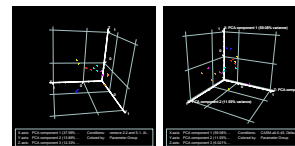


Figure 8: PCA analysis of CASM samples before (left) and after (right) outlier removal.

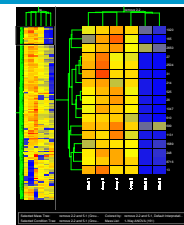


Figure 9: Clustering of the treatment groups. The figure on the right shows a zoom view from the top left tree.

Features for targeted MS/MS were selected using fold-change filters and were further filtered by abundance and/or charge state. The exported target list was directly imported into the Q-TOF acquisition method.

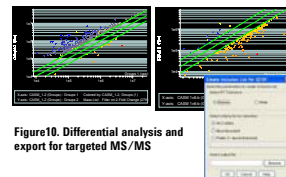


Figure 10: Differential analysis and export for targeted MS/MS

Conclusions

- The model study demonstrated the feasibility and power of a profiling/targeted approach to differential expression analysis
- Significant differences between treatment groups were found in the CASM pilot study
- Current work is focused on correlating target protein identifications with recently published gene expression profiling results

Reference

¹Deng, David et al. *Arterioscler Thromb Vasc Biol.* 2006, 26:1058.