Efficient Biomarker Discovery Coupled with SRM Assay Development in Maternal Serum from Normal and Trisomy 21 First Trimester Pregnancies

David Sarracino[1], Bryan Krastins[1], Michael Athanas[2], Amol Prakash[1], Taha Rezai[1], Scott Peterman[1], Kypros Nicolaides[3], Ramesh Kuppasamy[3] , Mary F Lopez[1]
[1]VAST Scientific, Cambridge, MA; [2]Thermo Fisher Scientific, Cambridge, MA; [3] Fetal Medicine Foundation , London, UK

Overview

Purpose: Mass spectrometric and bioinformatic analysis of maternal serum samples from Trisomy 21 and normal pregnancies in order to identify differentially expressed proteins. These discovery data were used to develop SRM assays.

Methods: High resolution mass spectrometry coupled with label-free differential analysis using SIEVE software. MS spectra from discovery experiments were mined to facilitate development of SRM assays using novel SRM workflow software.

Results: SIEVE analysis facilitated the discovery of a panel of differentially expressed proteins. SRM assays for a panel of the differentially expressed proteins were developed.

Introduction

Women over 35 years old constitute about 15% of pregnancies in most developed countries worldwide. The risk for many chromosomal defects increases with maternal age. Down's Syndrome (DS) occurs at a rate of 1 in 600-800 pregnancies and prompts most prenatal diagnoses that involve invasive procedures such as amniocentesis or chorionic villous sampling. These procedures carry a risk of miscarriage and therefore are only applied to women in a high risk groups (1). Over the past several years, there has been an effort to develop less invasive screening methods that can identify aneuploidy early in the pregnancy. Maternal blood tests that measure serum protein markers associated with DS such as free b-hCG, PAPP-A, alpha fetoprotein and inhibin A, combined with the measurement of nuchal translucency have garnered increased acceptance for first trimester screening (2) but the sensitivity and specificity of these markers could still be improved. Previous proteomic studies have identified other putative serum markers (3), however, the need remains for further discovery and confirmation in this area. In this study, we coupled high-resolution LC-MS/MS (LTQ Orbitrap hybrid mass spectrometer) with novel, label-free differential analysis software to analyze a cohort of maternal blood samples from first trimester DS and normal pregnancies. Bioinformatic analysis was carried out using SIEVE software. The SIEVE iterative workflow includes chromatographic alignment, global intensity-based feature extraction and aggregate protein identification assignment using SEQUEST®. Subsequent to the discovery experiments, candidate differentially expressed target proteins were selected and quantitative selective reaction monitoring (SRM) assays were developed on a TSQ Vantage triple quadrupole mass spectrometer. Novel SRM development software facilitated the efficient selection of signature peptides from the targeted proteins and the resulting SRM analyses monitored predicted precursor and fragment transition ion pairs. Heavy-isotope internal standards allowed for simultaneous quantification. The quantitative SRM assays were applied to the original cohort samples to verify putative biomarker abundance ratios. Future experiments will test the SRM marker panel on other sample cohorts. The described workflow demonstrates an approach that may in general be applied to link biomarker discovery studies with subsequent candidate verification assays.

Methods

Clinical samples

Maternal serum samples from Trisomy 21 and normal first trimester pregnancies were provided by the Fetal Medicine Foundation and collected from study participants with full consent and approval. The majority of these women were found to be high risk based on the FMF risk calculation for Trisomy 21. Blood samples were collected into red top tubes (BD Vacutainer® REF 367694 Z) which do not contain anti-coagulant. Samples were centrifuged at 1,500 rpm for 15 minutes at 4 °C and frozen at -70 °C. Samples were transported frozen and were thawed immediately before processing. Samples from Trisomy 21 patients and healthy control patients were processed in a random order to prevent systematic errors and variations from experiment to experiment.

Sample preparation

Serum samples without enrichment were prepared by diluting serum 1:4 v/v with 8 M GuHCl/150 mM Tris/10 mM DTT pH 8.5. Samples were incubated at 37 °C for 60 minutes and then cooled to room temperature. Next, 500 mM iodoacetic acid/1 M Tris pH 8.5 was added to each sample to a final concentration of 40 mM. Samples were alkylated in the dark at room temperature for 60 minutes. The reaction was quenched with the addition of 2M DTT to a final concentration of 5mM. Post quench, samples were diluted to 1 ml with the addition of 50 mM Tris/5 mM CaCl2 pH 8.0 and 10 ug of a sequencing grade enzyme (Promega) was added to each sample and samples were incubated at 37 °C for 24 hours. The digestion reaction was quenched with the addition of TFA to 1%. Subsequent to digestion, the samples were processed by solid phase extraction for desalting and clean-up using Thermo Scientific Hypersep C18 50 mg 96-well plates. Plates were equilibrated once with 100% ACN and then rinsed twice with 0.25% TFA. Samples were then loaded and wells were rinsed 5 times with 0.25% TFA. The peptides were eluted with 400 ml 75% ACN/0.1% formic acid at 300 rcf in a centrifuge equipped with a swinging bucket rotor. After elution, the samples were lyophilized. Just before loading on the mass spectrometer, peptides were resuspended in 97% H2O/3%ACN/0.25% formic acid.

High-resolution LC-MS/MS

High-resolution LC-MS/MS analysis was carried out on a Thermo Scientific LTQ Orbitrap XL mass spectrometer. Samples in 5% (v/v) acetonitrile 0.1% (v/v) formic acid were injected onto a 75 µm x 25 cm fused silica capillary column packed with Thermo Scientific Hypersil Gold C18AQ 5 µm media, in a 250 µL/min gradient of 5% (v/v) acetonitrile, 0.1% (v/v) formic acid to 30% (v/v) acetonitrile, 0.1% (v/v) formic acid over the course of 180 minutes with a total run length of 240 minutes. The LTQ Orbitrap XL™ was run in a top 5 configuration at 60K resolution for a full scan, with monoisotopic precursor selection enabled, and +1, and unassigned charge state rejected. The analysis was carried out with CID fragmentation modes.

Label-free, differential analysis and protein identification

Bioinformatic analysis was carried out using Thermo Scientific SIEVE software and SEQUEST.

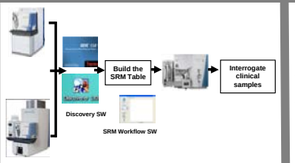
SRM assays

SRM assays were developed on a Thermo Scientific TSQ Vantage and TSQ Quantum Ultra triple quadrupole mass spectrometers, Surveyor MS pump, Micro autosampler and an IonMax Source equipped with a low flow metal needle. Reverse phase separations were carried out on a 1 mm X 50 mm Thermo Scientific Hypersil Gold 1.9µm C18 particle. Solvent A was LC-MS grade water with 0.2% (v/v) formic acid, and solvent B was LC-MS grade 30% (v/v) acetonitrile with 0.2% (v/v) formic acid (Optima grade reagents, Thermo Fisher Scientific). Thermo Scientific SRM workflow prototype software (available upon request) was used for targeted protein quantification. This software algorithm is used to predict candidate peptides and for choosing multiple fragment ions for SRM assay design, building an instrument method and a sequence file, and also for automatic peptide identity confirmation and quantitative data processing. For the workflow described herein, the differentially expressed peptides that were identified in the LC-MS/MS discovery MS data were imported directly into SRM builder software. Transitions were chosen based on the predominant fragments observed in the discovery data (>5 transitions per peptide). Peptides were identified by co-eluting light and heavy transitions derived from synthetic peptide standards.

Results

FIGURE 1. Discovery to targeted workflow.

A comprehensive mass spectrometry-based workflow linking biomarker discovery with targeted SRM assay development takes advantage of the similarity between peptide fragmentation behavior in trap and triple quadrupole instruments. This allows data generated in discovery experiments with SIEVE software to be mined for the optimization of SRM assay development using SRM Workflow software.



Maternal serum samples from Trisomy 21 and normal first trimester pregnancies were provided by the Fetal Medicine Foundation and collected from study participants with full consent and approval. The majority were found to be high risk based on the FMF risk calculation for Trisomy 21.

Samples	Diagnosis
21	Normal pregnancy
24	Trisomy 21 pregnancy

FIGURE 2. Comparative product ion spectra from trap and triple quadrupole instruments.

Algorithms can predict unique peptides and SRM transitions. However, this hypothetical approach provides no assurance that the peptides will actually be detected in the sample. Fortunately, MS/MS spectra from Discovery experiments can be mined to enhance the design of effective SRM assays. Synergy between linear ion trap MS/MS data and SRM assays depends on the similarity of parent ion fragmentation behavior in the linear trap and triple quadrupole mass spectrometers.

In order to confirm that MS2 data acquired during discovery experiments in ion traps could be used to predict behavior in the triple quadrupole mass spectrometers, we compared product ion spectra from both instruments .

The data below from two human apolipoprotein E peptides in +2 charge state demonstrate that fragmentation patterns and intensities are consistent between the two instruments.

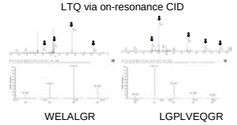
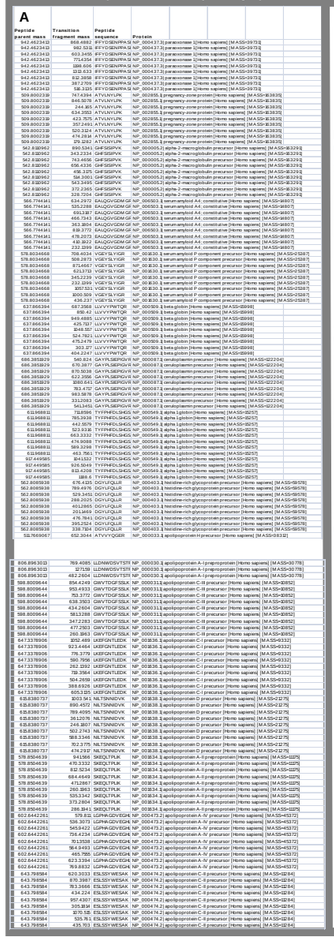


FIGURE 3. Label-free quantitative analysis using SIEVE software.

All individual samples were processed with LC-MS/MS as described in the methods and the resulting high resolution raw mass spectral files were analyzed with SIEVE software. The SIEVE iterative workflow included chromatographic alignment, global intensity-based feature extraction and aggregate protein identification assignment. Chromatographic alignment was based upon the pair-wise MS full scan comparison of all experimental MS runs with respect to a chosen reference MS run. Subsequent to chromatographic alignment, potentially interesting features were exposed based upon high-intensity peaks found in the aligned collective data set. Individually, these peaks defined frames i.e. well defined rectangular regions in the full scan (*m/z* versus retention time) plane. Reconstructed ion chromatograms were calculated for each frame to assess relative expression ratio and supporting statistics. After framing, MS2 fragment scans associated with each frame were processed with SEQUEST. A consensus protein report was constructed by statistically aggregating frame information to construct peptides and then peptide information to build proteins.

Top panels indicate integrated intensities for each MS raw file for a selection of proteins. Lower panels indicate individual peptide ratios from each protein. Average protein ratios were calculated using variance weighted averaging of each individual peptide measurement. A selection of 19 high abundance differentially expressed proteins with robust pValues were selected for further development into multiplexed SRM assays.

FIGURE 4. Development of SRM assay for target proteins.



SRM Workflow Software was used to mine the discovery data and develop a multiplexed method for quantitative monitoring of 19 target peptides and 172 transitions.

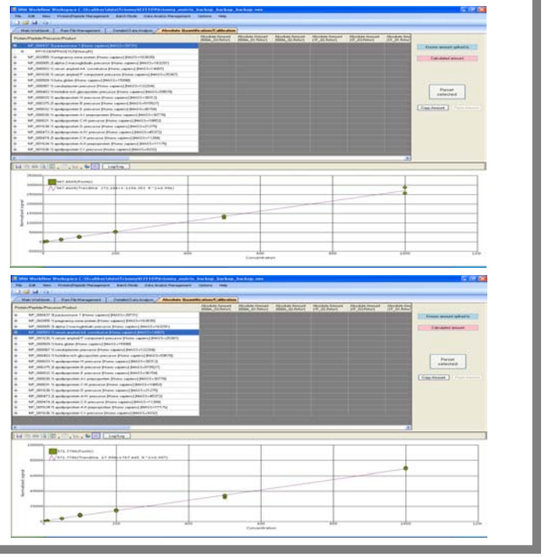
A. Proteins, peptides and transitions assayed.

B. SRM assay data from selected peptides. Pooled serum samples were analyzed on the Vantage as described in the methods.

Panels (arrows) indicate total signal, peak profiles and overlayed chromatographic transition peak areas.

FIGURE 5. SRM assay calibration curves.

Calibration curves were generated for all target peptides, several are shown below. The level of detection (LOD) for all target peptides in serum matrix was 250 attomoles on column. The level of accurate quantitation (LOQ) was 500 attomoles on column.



Conclusions

- Protein/peptide differential expression discovery experiments generate lists of candidate biomarkers.
- SRM assays can provide a high-throughput and efficient vehicle for verification and translation of putative markers into clinical research assays.
- High-resolution MS2 fragmentation data obtained in discovery experiments can facilitate the design of targeted SRM assays.
- The integration of these workflows maximizes the synergy between unbiased and hypothesis-driven biomarker research.
- This application of this workflow was demonstrated with a cohort of maternal serum samples from normal and Trisomy 21 pregnancies.

References

1. Zournatzi V, Daniilidis A, Karidas C, Tantanasis T, Loufopoulos A, Tzafettas J. 2008 Hippocratia 1: 28-32
2. Kagan KO, Wright D, Spencer K, Molina FS, Nicolaides KH. 2008 Ultrasound Obstet Gynecol. 5:493-502
3. Nagalla SR, Canick JA,Jacob T, Schneider KA, Reddy AP, Thomas A, Dasari S, Lu X, Lapidus JA, Lambert-Messerlian GM, Gravett MG, Roberts Jr. CT, Luthy D, Malone FD, D'Alton ME. 2007 Journal of Proteome Research 6:1245-1257

BD Vacutainer is a registered trademark of Becton, Dickinson and Company Corp. SEQUEST is a registered trademark of the University of Washington. All other trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries.