Quantitative, Top-down Identification of Differentially Expressed, Endogenous Peptides in Human Serum: A Robust, Sensitive and Novel Discovery Workflow

David Sarracino[1], Bryan Krastins[1], Michael Athanas[2], Amol Prakash[1], Kypros Nicolaides[3] and Mary F Lopez[1].

[1]Thermo Fisher Scientific BRIMS Center, Cambridge, MA; [2]Vast Scientific, Cambridge, MA; [3]Fetal Medicine Foundation, London, UK

Introduction

Unbiased biomarker discovery workflows for plasma and serum have mostly focused on the identification of differentially expressed candidate proteins and their tryptic peptides. This "shotgun approach" is severely limited by the dynamic range and diversity of proteins in blood. To ameliorate this problem, many protocols require the "depletion" of abundant proteins such as albumin and others by various affinity methods. A major caveat is that endogenous (non-tryptic) peptides are typically tightly bound to albumin and other high abundance "carrier proteins" in vivo, and are therefore lost in the depleted sample.

Endogenous peptides are likely candidate biomarkers for many diseases and pathologies as they are secreted from tissues and enter the bloodstream (1,2). This phenomenon may explain why there has been little success in biomarker discovery using most shotgun methods .

Endogenous peptide recovery from blood poses numerous hurdles. First, the proper collection and storage of blood samples minimizes the generation of artifactual peptides that may be generated *ex vivo*. Second, the large dynamic range in molecular sizes and abundance requires the separation of proteins from peptides and metabolites and the separation must be done under denaturing conditions to ensure the recovery of endogenous peptides bound to carrier proteins. Last, because of the difficulties posed by the identification and sequencing of non-tryptic peptides, very high resolution MS2 and MS3 data must be acquired in a reproducible and robust manner.

In this study, we describe the development of a workflow specifically geared toward the efficient recovery and identification of endogenous peptides from blood utilizing a combination of up-front batch sample prep and on-line liquid chromatography coupled with high resolution tandem MS using HCD and CID fragmentation on an Orbitrap XL. Quantitative differential analysis of the endogenous peptides was carried out using label-free analysis with SIEVE software algorithm. The novel workflow was applied to the investigation of the quantitative differences in endogenous peptides recovered from maternal serum in normal and trisomy 21 pregnancies.

Methods

Maternal serum samples from trisomy 21 and normal first trimester pregnancies were collected from study participants with full consent and approval. The samples were provided by the Fetal Medicine Foundation. Details of experimental design and sample preparation are given in Fig 1, 2. Prepared samples were run on an Orbitrap instrument with CID and HCD as the fragmentation modes. Label-free differential analysis was performed using SIEVE algorithm

FIGURE 1. Schematic of the overall workflow for recovery and analysis of endogenous peptides from blood.

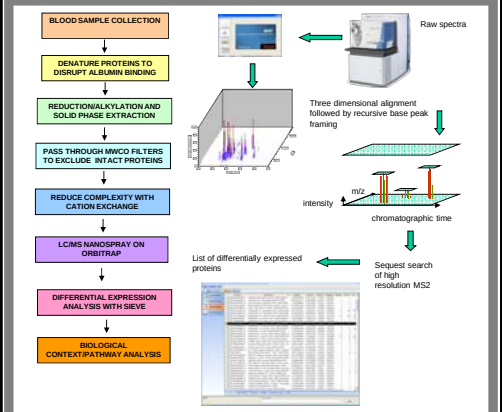


FIGURE 2. Detailed experimental protocol.

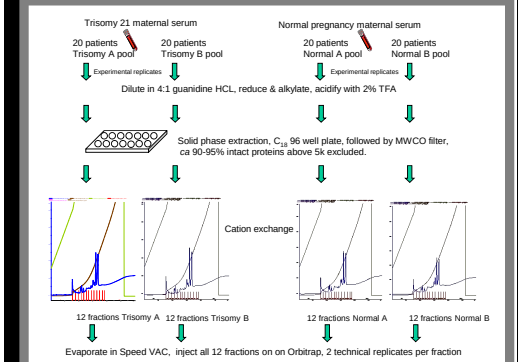
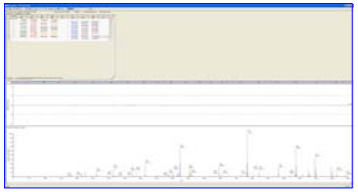


FIGURE 3. Summary of SIEVE data analysis.

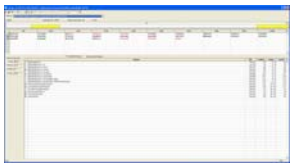
High resolution MS2 spectra were obtained and analyzed with SIEVE



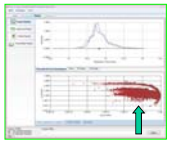
A summary of the analysis results are displayed in the table below. The majority of frames, *i.e.,* peptides were up-regulated in the disease state. Few peptides were down-regulated or equivalent (ratio close to 1). This result is consistent with the hypothesis that endogenous peptide fragments represent a "metabolic snapshot" of proteolytic activities induced by disease or inflammation. These proteolytic fragments are secreted into the bloodstream and may be a rich source of disease biomarkers. **Therefore, only an increase and not a decrease in endogenous peptides would be expected in disease samples.**

Identified Peptides	Ratio up Trisomy 21	Ratio = to 1 Trisomy 21	Ratio Down Trisomy 21
274	257	12	5

FIGURE 4. MS data and SIEVE analysis.



Endogenous peptides typically appear as nested fragments derived from the N and C terminal ends of the protein. This is indicative of endogenous proteolytic activity (as opposed to exogenous tryptic digestion). The volcano plots generated from "top down" (A) endogenous peptide disease *vs* normal data sets display different dynamics than those from "bottom up" (B) tryptic peptide data sets. **One likely explanation is that this is a reflection of increased proteolytic activity in disease samples with the subsequent preponderance of upregulated peptides and few or no downregulated peptides.**



A. endogenous peptides

B. tryptic peptides

FIGURE 5. SIEVE results for a protein ID increased 50 fold (A) versus one with a ratio close to 1 (B). The reconstructed ion chromatograms displayed are typical for most frames with p values of less than 0.1.

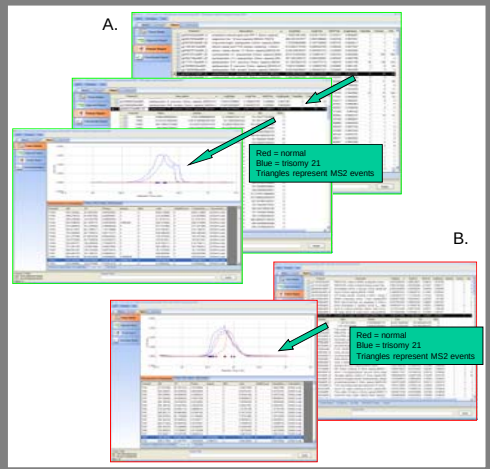
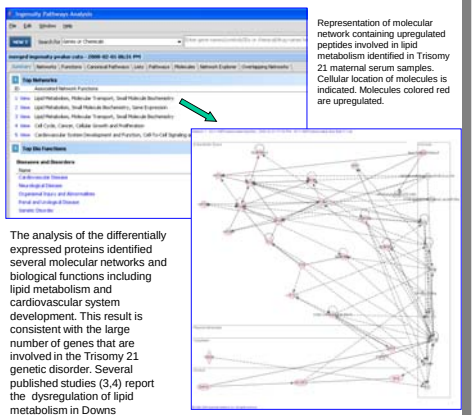


FIGURE 6. Biological context. The data derived from the SIEVE differential analysis with p values less than 0.1 were imported and analyzed in Ingenuity pathway analysis software.



Representation of molecular network containing upregulated peptides involved in lipid metabolism identified in Trisomy 21 maternal serum samples. Cellular location of molecules is indicated. Molecules colored red are upregulated.

The analysis of the differentially expressed proteins identified several molecular networks and biological functions including lipid metabolism and cardiovascular system development. This result is consistent with the large number of genes that are involved in the Trisomy 21 genetic disorder. Several published studies (3,4) report the dysregulation of lipid metabolism in Downs syndrome.

Conclusions

- The described workflow facilitates the recovery of endogenous peptides from blood.
- Endogenous peptides may constitute a viable source of disease biomarkers.
- The results of this study demonstrate that the majority (*ca* 90%) of endogenous peptides were up-regulated in the disease state. This result is consistent with the hypothesis that endogenous peptide fragments represent a "metabolic snapshot" of proteolytic activities induced by disease or inflammation. Therefore, only an increase and not a decrease in endogenous peptides would be expected in disease samples.
- In this study, we present preliminary data that demonstrate an increase in peptides associated with lipid metabolism detected in maternal serum from trisomy 21 pregnancies.
- This result is consistent with previous studies demonstrating dysregulation of lipid metablism in fetal trisomy 21.

References

1. Lopez MF, Kuzdzal S, Mikulskis A, Golenko E, Petricoin III EF, Liotta LA, Patton WF, Lynch C, Gordon R, Whiteley GR, Rosenblatt K, Gurnani P, Nanda A, Neill S, Cullen S, O'Gorman M, Sarracino D, Fishman D. A novel, high-throughput workflow for discovery and identification of serum carrier protein-bound peptide biomarker candidates in ovarian cancer samples. Clinical Chemistry 2007 53; 1067-1074.
2. Lowenthal MS, Mehta AI, Frogale K, Bandle RW, Araujo RP, Hood BL, et al. Analysis of albumin-associated peptides and proteins from ovarian cancer patients. Clin Chem 2005;51:1933–45.
3. Nava S, Fogliani R, Nicolini U. Trisomy 21 is associated with hypercholesterolemia during intrauterine life.Bocconi L, Am J Obstet Gynecol. 1997 Mar;176(3):540-3
4. Zamorano A, Guzmán M, Aspillaga M, Avendaño A, Gatica M. Concentrations of serum lipids in children with Down's syndrome. Arch Biol Med Exp (Santiago). 1991;24(1):49-55.