

Facilitate Biomarker Discovery Using Integrated “Omics” Differential Analysis with High Resolution Accurate LC/MS Approach

Reiko Kiyonami¹, Sergei Snovida², Devin Drew¹, Rosa Viner¹, Julian Saba¹, David Peake¹, Andreas Huhmer¹, Ken Miller¹, ¹Thermo Fisher Scientific, San Jose, CA, USA, ²Thermo Fisher Scientific, Rockford, IL, USA

ABSTRACT

Application of “omics” technologies to unbiased clinical biomarker discovery research of various diseases is emerging because comprehensive molecular profiling using these “omics” technologies could better explain complex disease mechanisms and reach a new level of molecular understanding concerning carcinogenesis, disease progression, and drug-response prediction. Omics approaches are resource-intensive, analytically demanding and require the use of a sophisticated instrument platform and software to identify and quantify hundreds to thousands of molecules with high confidence and accuracy. A complete workflow based on “omics” profiling also requires well developed sample preparation kits. We have worked on developing multiple “omics” profiling workflows using high resolution accurate Orbitrap mass spectrometers in combination with liquid chromatography separations. In this study, we applied lipidomics, proteomics and glycoproteomics profiling workflows to study of Zucker diabetic fatty (ZDF) and lean rats, a model of type 2 diabetes. Plasma samples collected from three lean and three fatty Zucker rats were analyzed. With the multiple “omics” workflows developed here we were able to simultaneously identify and quantify hundreds of individual molecular lipid species, proteins and glycopeptides with high confidence and good analytical precision. Significant changes of individual lipid species, proteins and glycopeptides were observed between the two group of rat plasma samples.

INTRODUCTION

HPLC-MS platforms are increasingly used for large scale “Omics” (proteomics, lipidomics, metabolomics, glycomics) experiments to discover potential biomarkers for early disease diagnosis. Each “omics” data set provides valuable insights into multiple biomarker candidates. However, because properties of a biological system interact with each other, integrating multiple “omics” data will help to understand the system behavior as a whole for unravelling biological regulatory mechanisms to define the emergent properties and help to verify biomarker candidates from each “omics” workflow. Here we conducted lipidomics, proteomics and glycoproteomics studies phenotypical ZDF vs. normal lean rat plasma using a Thermo Scientific™ Orbitrap™ based LC/MS platform. The results from each of the “omics” workflows are reported.

MATERIALS AND METHODS

Sample Preparation

The rat plasma were purchased from Bioreclamation LLC (Westbury, NY). Rat plasma were recovered from whole blood of Zucker lean (3 lots) and Zucker diabetic fatty (ZDF, 3 lots) using EDTA as anti-coagulant by Bioreclamation. Chloroform, methanol and water were used for the lipid extraction. For proteomics/glycoproteomics experiments, each plasma sample was reduced, alkylated and digested with trypsin, respectively. Tryptic digests (1 mg) from each sample (Lean 1,2,3 and Fat 1,2,3) were labeled with TMT sixplex, quenched, combined and glycopeptide enrichment was performed using Thermo Scientific HyperSep™ Retain AX Cartridges. The flow-through from the cartridge (non-glycopeptides) was collected separately.

Methods

Lipid extracts were analyzed by LC-MS/MS in positive and negative mode on Thermo Scientific™ Dionex™ UltiMate™ 3000 Rapid Separation LC (RSLC) system coupled with a Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer over a 30 min gradient. Data were collected using a data dependent MS/MS experiment in which a full MS scan (120,000 resolving power FWHM at m/z 200) was followed by top 15 MS/MS scans (30,000 FWHM resolving power at m/z 200). Data analysis was performed using Thermo Scientific™ LipidSearch™ 4.1 software.

Protein digests labeled with TMT sixplex in which the glycopeptides had been depleted were analyzed by nanoflow LC-MS/MS on EASYnLC™ 1000 coupled with a Thermo Scientific™ Orbitrap Fusion™ Lumos™ mass spectrometer over a 2-hour gradient. Data were collected using top speed data dependent HCD MS/MS experiment in which a full MS scan using 120,000 (FWHM at m/z 200) resolving power was followed by HCD MS/MS scans using 30,000 (FWHM at m/z 200) resolving power with 3 second cycle time. Data analysis were performed using Thermo Scientific™ Proteome Discoverer™ software. Protein annotation was performed using Thermo Scientific™ Protein Center™ software.

Enriched glycopeptides digests labeled with TMT sixplex were analyzed by nano flow LC-MS/MS on EASYnLC 1000 coupled with a Orbitrap Fusion Lumos mass spectrometer over a 2-hour gradient. Data were collected using top speed data dependent EThcD MS/MS, followed by SPS MS³ experiments with 3 second cycle time. Mass resolution (FWHM at m/z 200) used was 120,000 for full MS, 30,000 for EThcD MS/MS and 60,000 resolving power for SPS MS³. MS data analysis were performed using Byonic node (Protein Metrics Inc.) in Proteome Discoverer 2.1.

RESULTS

Lipidomics

Figure 1. Basepeak Chromatograms of a ZDF Plasma

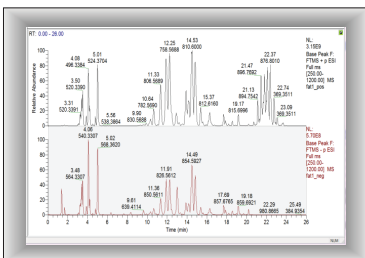
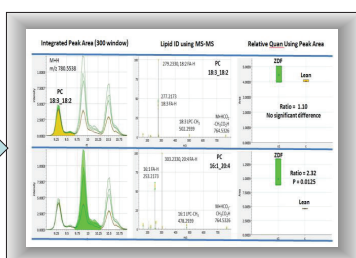


Figure 2. Simultaneous Lipid ID and Relative Quantification Using LipidSearch



Each plasma sample was analyzed in positive and negative mode, respectively. The C30 reversed-phase HPLC column uniquely offers high shape selectivity and provides improved lipid isomer separation efficiency, yielding higher coverage of the rat plasma lipidome.

All raw files were processed using Lipid Search 4.1 SP1 software. The ultra high resolution offered by the QE HF MS enables very narrow mass search windows (3 ppm for MS and 5 ppm for MS/MS) being used for database searching, providing high confidence lipid identification and precise quantification results.

Table 1. Summary of Lipid IDs and Quantitation from the Rat Plasma

Class	Filtered	p < 0.02
PA	1	0
PC	135	93
LPC	110	77
PE	35	0
LPE	4	3
PG	4	2
PI	42	30
PS	3	3
Card	21	-3
Car-G1	13	-10
SM	28	-5
AccA	19	-14
CHE	11	5
SH	6	3
SHL	3	0
DG	30	29
TG	283	62
Total	750	349

The lipid annotations for each individual sample are aligned within a time window and the results are merged into a single report. Positive and negative ion adduct annotations are grouped for lipids eluting at the same retention time. Accurate-mass extracted chromatograms are integrated for each identified lipid precursor ion and peak areas are obtained for quantitation. t-Tests and calculated p-values determine which molecular species are significantly different between sample and control groups. 750 Molecular lipid species across 17 lipid classes were identified after filtering. Fold-changes of the identified lipid species between the lean plasma and the diabetic plasma groups were determined. Significantly different concentration changes (p < 0.05) were observed for 349 lipid species.

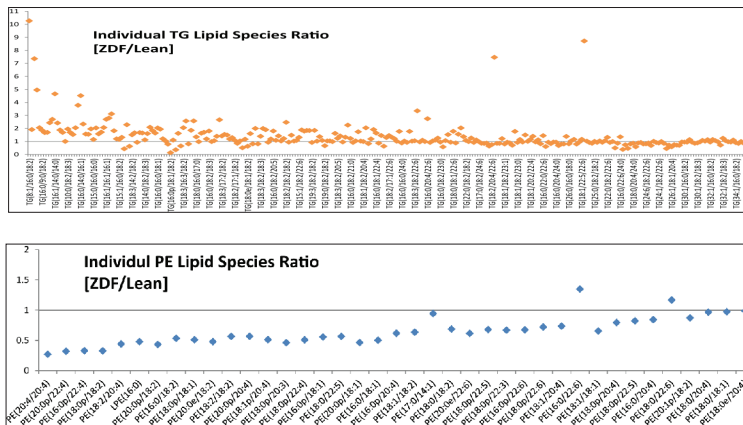
Figure 3. Lipid Class Profiles for TG, LPC, PC, SM, LPE and PE



Significant increases were observed in ZDF plasma for the levels of abundant lipid classes including triglycerides (TG) and phosphatidylcholine (PC). Other lipid classes such as lyso-phosphatidylcholine (LPC) and sphingomyelin (SM) also showed an increasing trend in the ZDF group. In contrast, lyso-phosphatidylethanolamine (LPE) and phosphatidylethanolamine (PE) levels showed significant decreases in the ZDF plasma.

Lipidomics

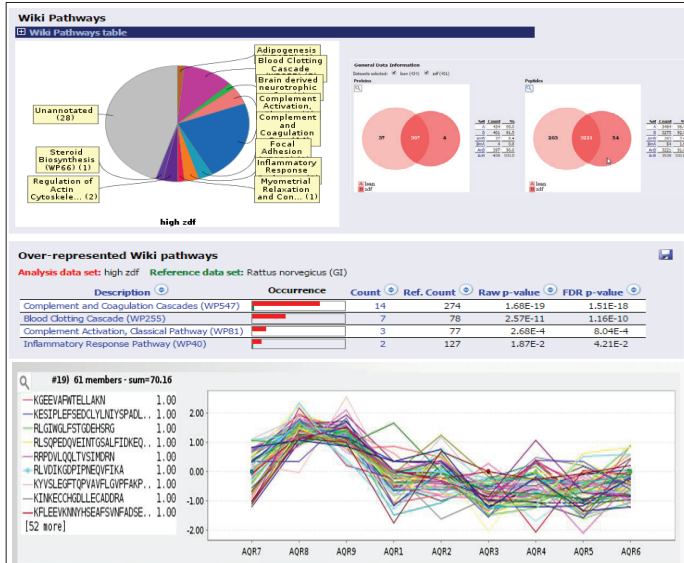
Figure 4. Individual Lipid sub-Class Profiles for TG and PE



It is important to identify which individual lipid species is actually responsible for the observed total lipid composition changes in order to understand biological function each specific lipid. For the TG class, the individual lipid species of TG(8:1/16:0/18:1), TG(18:0/22:6/22:6), TG(16:0/22:6/22:6), TG(8:1/16:0/18:1) and TG(16:0/8:0/18:1) showed more than 5 fold increases in the ZDF plasma and contributed most to the total TG composition increase. About 45% of TG molecular species actually did not show any significant changes. For the PE class, most identified PE molecular species contributed to the total composition decreases in the ZDF plasma in a moderate fashion. These detail information help revealing lipid pathway and biochemical mechanisms.

Proteomics/Glycoproteomics

Figure 5. Identified Protein and Peptides from Unfractionated Rat Plasma Digests



Proteome Discover 2.1 was able to identify more than 400 proteins from unfractionated rat plasma digests. A selected profile group showing elevated levels in ZDF samples (AQR 7-9) relative to lean samples (AQR 1-6). Peptides quantitative values were clustered using Protein Center's Profiling, with group count = 9 and alpha core = 0.5. There over-represented proteins in ZDF rat are involved with several biological pathways, including the inflammatory response pathway and potentially linked with the increased lipid molecular species.

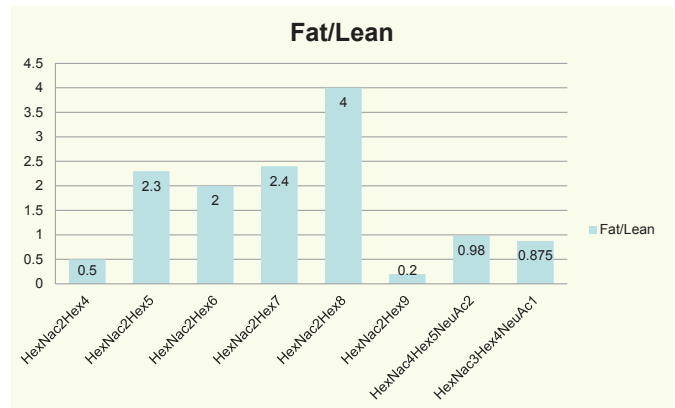
Proteomics/Glycoproteomics

Figure 6. Protein Identification and Quantitation Summary

ETDhcd-HCD data from glycopeptide enriched digest with TMT labeling was processed using By node of PD 2.1. Almost 2000 glycopeptides were identified and quantified from the rat plasma digests. Majority of proteins showed no significant changes between ZDF and Lean plasma group

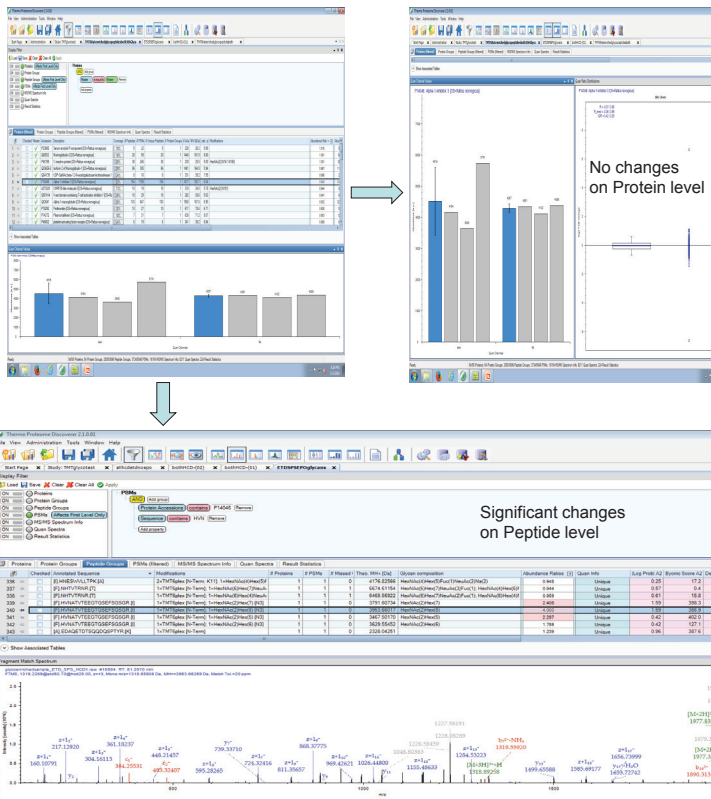
Proteomics/Glycoproteomics

Figure 6. Expression Fold Changes of Peptide HVNATVTEEGTGEFSGSGR- Site N321 in Alpha-1 inhibitor, P14046



The up and down changes in the peptide level with different glycan composition. High mannose species are up and complex/hybrid glycans are down.

Figure 7. Identification and Quantitation Results for Alpha-1 inhibitor, P14046



CONCLUSIONS

- Complete multiple “omics” workflows were developed on high resolution accurate mass Orbitrap – based mass spectrometers coupled with liquid chromatography separation and sophisticated data processing software. The ultra high resolution offered by Orbitrap detector enables another dimension of mass separation, yielding confident unknown molecular identification and precise quantification.
- 750 Molecular lipid species were identified and quantified from the rat plasma samples. The relative concentration changes between the ZDF and lean plasma groups were measured at both total lipid class composition and individual lipid molecular species composition, providing more potential biological insights into understanding the complex processes that lead to the development of diabetes.
- Unique data dependent EThcd MS/MS and SPS MS³ workflows enabled very precise protein/peptide quantitation in a high throughput fashion. Unambiguous glycopeptide characterization was performed using HRAM EThcd spectra. Precise relative quantitation was carried out using HRAM SPS HCD MS³ data.
- Significant increases of glycopeptides containing high mannose species were observed between the ZDF and lean plasma without any changes at the protein level. Further studies will be done to see if these glycan composition changes are linked with lipid molecular species changes.

The protein of the Alpha-1 inhibitor did not show changes in protein levels. However, significant changes in glycopeptide profiles were observed as shown in Figure 7. For example, the amount high mannose glycoforms for site N321 increased 2-4 times in ZDF rat plasma (Figure 8), but there are no changes for hybrid/complex glycosylated species.

For Research Use Only. Not to be used in diagnostic procedures.

Find out more at thermofisher.com

