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"

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 Thermo Scientific™ TMT™ 6-plex (TMTisxplex™) tags, quenched, combined and glycopeptide enrichment was performed using Thermo Scientific™ HyperSep™ Retain AX Cartridges. The flow-through from the cartridge (non-glycopeptide) was collected separately.

Methods

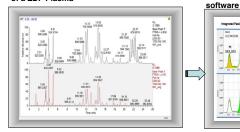
Lipid extracts were analyzed by LC-MS/MS in positive and negative mode on Thermo Scientific™ Dionex™ UltiMate™ 3000 Rapid Separation LC (RSLC) system coupled with a Thermo Scientific™Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer over a 30 min gradient. Data were collected using a data dependent 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 EASY-nLC™ 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 [™] Protein Center[™] software.

Enriched glycopeptides digests labeled with TMT sixplex tags were analyzed by nano flow LC-MS/MS on the Thermo Scientific™ EASY-nLC 1000 coupled with a Thermo Scientific™ Constraints and the transpace of transpace of the transpace of transpace of the transpace of the transpace of transpace of the transpace of transpace o

RESULTS

Lipidomics

Figure 1. Basepeak Chromatograms of a ZDF Plasma



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.

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

Class	Filtered*	p < 0.05
PA	1	0
PC	135	93
LPC	110	77
PE	38	0
LPE	4	3
PG	4	2
PI	42	30
PS	3	3
Cer	21	13
CerG ₁	13	10
SM	28	15
AcCa	19	14
ChE	11	5
SIE	5	3
StE	3	0
DG	30	29
TG	283	52
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 extracted for each identified same are aligned by a 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 determine. Significantly different concentration chances (0 < 0.05) were observed for 349 lipid species.

quantification results.

Figure 2. Simultaneous Lipid ID and

Relative Quantification Using LipidSearch

All raw files were processed using Lipid

Search[™] 4.1 SP1 software. The ultra high

resolution offered by the Q Exactive HF MS

enables very narrow mass search windows

(3 ppm for MS and 5 ppm for MS/MS) being

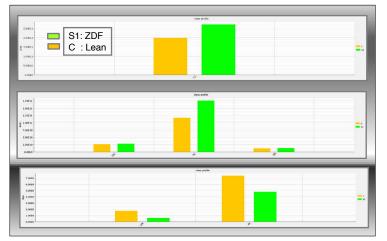
used for database searching, providing high confidence lipid identification and precise

PC 183 182

1830/CO1 04/CO3 302209 7843026 e Lean

Ratio + 2.32

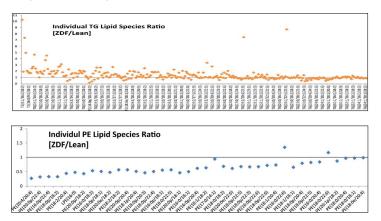




Significant increases were observed in ZDF plasma for the levels of abundant lipid classes including triglycerides (TG) and phosphatidylcholine (PC). Other lipid classes such as lyso-phosphatidylcholine (LPC) and sphingomyelin (SM) also showed an increasing trend in the ZDF group. In contrast, lysophosphatidylethanolamine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

Lipidomics

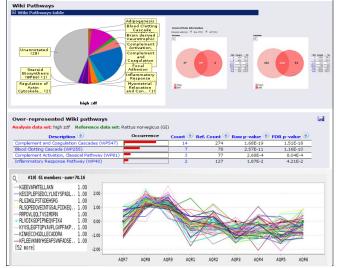
Figure 4. Individual Lipid sub-Class Profiles for TG and PE



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(16:0/18:1) 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



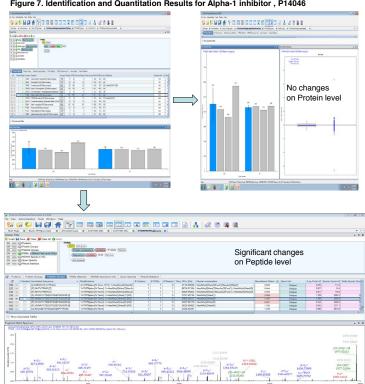
Proteome Discover 2.1 software was able to identify more than 400 proteins from unfractionated rat plasma digests. A selected profile group showing elevated levels in ZDF samples (AQR 7-9) relative to lean samples (AQR 1-6). Peptides quantitative values were clustered using Protein Center's Profiling, 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.

Figure 6. Protein Identification and Quantitation Summary

					× anchelanoops × methodatesethodapycap																•
				oterin Group					to Result Statistic						Advandence Radice 121		n (Cena V)			lhedator	
е.		sectors		Offender	Antoneous A transmission 12125-fitting provided	65	repeate	a roma a	Ungue registers # rist	eor ureas	556	62.6	218	managements	Advictance Papers (+)	10.1	218		75.02	20.0	
			1.2	(200HB)	Manue EDin Fallue non-essanal	115	- 21	108	21		608	- 42.0	6.21		6477	671	3014	25.61		40.4	209
			1.×.	P02764	Mater 1 wid chourster (CS-Patter recessor)	67.		40	14		225	22.6	644		6.617	754.6	2015	625		107.1	201
			1.2	P121764	Alphan 1 waid glycoprotein (CC+Patius representational) Jahran Lampersonaa CC-Patius representational	81		205	45		411	451	6.07		0.011	334.0	2710	45.29		2373	225.8
			1.	P14045	Age 1 and 1 200-Pates revealed	375	754	1102	154		1477	107	6.04		0.042	41111	42017	24.21		41144	2040.2
			1.8	053041	Alpha Lenacegichula (DSeRatua novagicua)	205	105	647	101		2500	147.6	6.90		6.452	VINTA.	2007	24.10		31447	2047
			1.5	924090	Natio 2 His strongenture (20-Mattus non-septual	075	27	252			252	38.0	1.50		0.072	102.5	1011			1120.4	202.5
			1.5	OTTLA	Amphotesic-induced protein 2 125-Rature consequent	in.	27	262	27		120	62.7	1.63	Hard March 10000 M0070	6.717	1011	194.7	10.40		416.7	500.5
			15	GRAIN-1	Arghoare-induced protein 2 (35+hattas ronregious) Intractin (01+Rattus ronreatious)	22%	20	104	20		1432	158.6	7.21	resonance production	6,042	239.3	214.5			406.7	2043
			1.5	P20544	Beta 2 ghrogenteen 1305-Plattus norvegatus)	22%	10	21	10		297	32.2	1.22		1.65	95.0	62.0	18.48		124	79.5
			13	0609115	California and an and a state to a state of the state of	375	20		10		258	28.6	4.12		147	204.0	226.1	10.42		341.0	183.5
			1.5	01000	Cathor/Retenses 10(05-Fothis rotwpous)	125		34			640	92.1	6.72		0.004	194.3	1118	19.52		122.5	124.1
				P13625	Candishame 105-Patha scomanal	105	11	22	11		1010	122.6	5.54		1.04	1072	1014	22.08		115.0	105.7
			1.2	P05371	CLeave 205-Ratus ronapous)			- 11	80		447	41.1	141		0.040	262.1	220.2	10.11		201.2	1924
			1.5	4211228	CMPP 35 Ma malanula 2125 (Rathis remained	1 175	10	19	10		310	34.5	6.72	Herblach Britti	6.044	214	54.4	42.33		72.0	12.5
			-5-	Printers .	Complement Cd 125-Ratio recommond	15	14	44	54		110	110.6		Hen/Net/27 (20154)	6.675	292.5	2184	15.22		162.6	179.4
			1.5	Q62930	someliament component CE105+Eath-a sometricual	115	10	-	10		854	62.2	5.63	- and a state of the state of t	6.621	100.0	105.1	0.0		180.7	71.6
				9914415	Complement factor 1925-Plattue non-monuel	245	- 14	- 37	21		604	67.5	7.42		8.670	128.1	122.4	12.07		192.2	102.2
			1.5	011211	control empirication rights in 101a [199 a reportional	163					334	44.6	4.50		6474	2.8	5.5	22.28		6.1	55
				P42199	Constitute poten (DI-Ratia notanna)	245	- 18	28	28		230	25.5	5.00	Herstyleth IN147, NORT	1.001	2012	2014	23.85		247.3	191.2
			1.5	909079	Fetuar-0 105-Patha romepical	195		66	16		378	41.5	7.00	- and the second second	6.766	421.9	324.5	15.38		455.5	245.6
				P0(399	Factore size that D'affat a received	4)					110	55.6	6.74		6.760	24	14	10.28		6.0	2.4
			1.5	PEARL	Edmogen gamma chain (25)-Faitus non-earoust	75	- 21	20			445	52.5	1.96		6.6%	14.1	85.5	2.4		-	82.2
			15	P04007.0	Receive Confidence and a second		13	- 63	15		2477	272.3	578		1.040	1972	107.5	15.41		163.5	1524
				920059	Remoterie 105-Ratus non-explanat	22%	22	- 67	22		+40	511	7.65		100	283.0	241.9	21.27		2014	213.7
			1.1	Q00FSA	Netidine-rich glucoprotein EDS-Fathus nen-explored	255	12	67	10		125	53.0	7.54		6.673	215.4	198.7	41.24		152.4	175.5
				PO1855	is minim than C many 101+Fature recommend	30	24	-	24		429	42.6	6.24		6.047	1021	64.5	14.14		.891	-014
1	i en	-		PACED		265					222		14.8		6,992		111		10.75	41.0	
			10	P20211	Is service 3 under Conserv Dis Relies permissed	10.5		14			333	26.5	764		0.763	71.0	17.0	12.63		61.0	82.1
			1.5-1	035550	insule-like proveh faceur-binding protein complex acid labile	110.	22	62	23		600	46.8	640		1.014	1714	107.9	19.72		1014	1004
				QUMIS	Inter sists investo white heavy share HI KIS-Patha nor-	0		37			887	99.0	6.25		649	1012	132.0	15.04		105.1	100.4
			1.5	P14485.5	Inclum 2 of Ebringen bets chain (15%-Ratus non-epical)	100		104			600	54.6	214		1.040	800.1	501.0	13.18		001.0	104.0
				903626-3	autom 2 of Musicability 1105-Earlier services and	275		205	15		1441	765.8	5.14		6.007	110.1	1100.0	21.63		1264.8	1011.3
			1.5	P12246-2	Indom 2 of Sectionalism 125-Ratius non-resound	15		147			436	54.5	144		6.000	14171	1756.5	18.78		1356.7	1258.2
			15	ODEN)	Musical da 2125 allana recensional	163	30	89	20		1443	161.6	6.60		1007	100.2	100.5	18.85		260.1	241.1
			1.5	P00011	munipering and debute receptor MI (CE-Flatus norveges)	125		- 12			625	62.1	1.22		6.612	11	11	80.14		4.5	44
			15	967313	Nemano Pick C1-like protein 1 (05+Platue romesicual)	155	24	114	26		1331	145.3	6.62		6.347	100.7	147.6	45.23		111.4	125.7
			1.5	P14272	Plasma kalilenia (CluRatus romagina)	CIG		11	10		638	71.2	1.00		6.403	103.3	1111	24.51	T 28	104.6	106.7
			15	PADE	planet advates provided rendering the non-spous	1205		10			141	29.2	2.54		1.000	86.5	714	16.66		76.6	414
			1.	P11232	Polyonia Olefana nominal	128%	10	27	10		617	72.4	671		6.676	102.4	93.3		14.55	62.5	45.2

ETDhcd-HCD data from glycopeptide enriched digest with TMT labeling was processed using Byonic 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 groups.

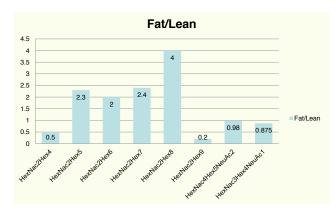
Figure 7. Identification and Quantitation Results for Alpha-1 inhibitor , P14046



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 of 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.

Proteomics/Glycoproteomics

Figure 8. Expression Fold Changes of Peptide HVNATVTEEGTGSEFSGSGR- 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.

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 guantified 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.

www.thermofisher.com

©2016 Thermo Fisher Scientific Inc. All rights reserved. Lipid Search is a registered trademark of MKL TMT and TMTsixplex are trademarks of Proteome Sciences plc.ISO is a trademark of the International Standards Organization. All other trademarks are the property of Thermo Fisi Scientific and its subsidiaries. This information is presented as an example of the capabilities of Thermo Fisher Scientific products. It is not intended to encourage use of these products in any manners that might infringe the intellectual property rights of others. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details.

Africa +43 1 333 50 34 0 Australia +61 3 9757 4300 Austria +43 810 282 206 Belaium +32 53 73 42 41 Brazil +55 11 2730 3006 Canada +1 800 530 8447 China 800 810 5118 (free call domestic) 400 650 5118

Denmark +45 70 23 62 60 Europe-Other +43 1 333 50 34 0 Finland +358 10 3292 200 France +33 1 60 92 48 00 Germany +49 6103 408 1014 India +91 22 6742 9494 Italy +39 02 950 591

Japan +81 6 6885 1213 Korea +82 2 3420 8600 Latin America +1 561 688 8700 Middle East +43 1 333 50 34 0 Netherlands +31 76 579 55 55 New Zealand +64 9 980 6700 Norway +46 8 556 468 00

Russia/CIS +43 1 333 50 34 0 Singapore +65 6289 1190 Sweden +46 8 556 468 00 Switzerland +41 61 716 77 00 Taiwan +886 2 8751 6655 UK/Ireland +44 1442 233555 USA +1 800 532 4752 PN64735-EN 0616S



A Thermo Fisher Scientific Brand