Processing of a Complex Lipid Dataset for the NIST Inter-laboratory Comparison Exercise for Lipidomics Measurements in Human Serum and Plasma

David A Peake,[1] Reiko Kiyonami,[1] Yasuto Yokoi,[2] Yukihiro Fukamachi[2] and Yingying Huang[1]
[1]Thermo Fisher Scientific, San Jose, CA, USA; [2]Mitsui Knowledge Industry, Tokyo, Japan

Poster Note 64471

Overview

Purpose: Demonstrate the use of LipidSearch software to process an entire untargeted data set from NIST human plasma and serum reference materials.

Methods: High resolution Orbitrap LC-MS and dd-MS^2 data were processed with LipidSearch software providing simultaneous identification and quantitation of lipids.

Results: Identification and quantification of almost 1000 lipid species was obtained.

Introduction

Lipids play a key role in cell, tissue and organ physiology with diseases such as cancer and diabetes which involve disruption of metabolic enzyme pathways. Identification of unique lipid biomarkers to distinguish healthy humans compared to those with a disease can have an impact on the early detection of diseases and personalized medicine.

Identification of lipids by untargeted lipidomics requires sophisticated software and the mass spectrometer employed must be capable of separating many overlapping isobaric and isomeric lipid species. We present here the details and challenges of processing data from NIST plasma and serum extracts. New software algorithms were introduced specifically to reduce false positives and automate data review.

Thermo Scientific™ LipidSearch™ software was used for lipid identification through a database search of precursor accurate masses and their predicted fragment ions. The number of lipid species identified in each LC-dd-MS^2 experiment were assessed at sum composition (MS) and isomer (MS^2) levels. Potential lipid species were identified by matching the predicted MS-MS fragments for molecular species observed in positive or negative ion mode.

The data for each run were aligned within a chromatographic time window and positive and negative ion annotations were merged into the results table. This approach provides lipid annotation that reflects the appropriate level of MS^2 fragment ions from the complete dataset giving higher confidence in lipid identifications. The merged results were filtered by main adduct ion, ID quality, signal-to-noise, peak area and relative standard deviation; manual integration was performed if necessary prior to estimating concentration relative to an internal standard for each lipid class. These results demonstrate that in a 60 min LC-MS run that it is possible to identify and quantify approximately 1000 isomeric lipid species from human plasma using an experimental C30 UHPLC column.

Methods

Sample Preparation

Aliquots of human plasma or serum obtained from NIST (80 µL, Table 1) were extracted using the method of Bligh and Dyer.

LC-MS and dd-MS^2

A Thermo Scientific™ Dionex™ UltiMate™ 3000 Rapid Separation LC (RSLC) system and Thermo Scientific™ Q Exactive HF™ MS instrument were employed as described (1). HPLC separation was achieved with a 2.1 x 250 mm, 1.9µm, C30 prototype column and MS analysis was performed at 120K mass resolution and MS^2 at 30K resolution (FWHM at m/z 200).

Data Analysis

Raw data (24 files, 5.9 GB total) were processed using LipidSearch 4.1 software on a Windows 7 64-bit HP Z840 workstation equipped with dual Hex-core Xeon processors (E5-2643v3, 3.4 GHz), 64 GB RAM (DDR4-2133), and 500 GB solid state hard drive.

LipidSearch software was used for lipid identification and relative quantitation using the workflow shown in Figure 1 (2, 3). LC-MS data containing MS and data dependent MS^2 were searched using the parameters listed in Table 2. For each MS^2 spectrum, search results are summarized for lipid species matching the predicted fragmentation pattern from the database with match score and occupancy indicating the fit (Figure 2). The average number of sum composition lipids identified from LC-MS runs collected in positive ion or negative ion mode was 211 and 867, respectively.

Lipid species with the same annotation within a ± 0.1 min time window were merged into the aligned results (Figure 3). The total number of lipids identified in the entire data set are summarized in the Table 3. The unfiltered results includes all lipid species potentially identified within the search parameters. Lower quality annotations and false positives were removed using a combined set of data filters prior to estimation of concentration relative to an internal standard for each class.

Data Processing Workflow using LipidSearch Software (Figure 1).

1) Peak Detection. Read raw files, MS^n and precursor ion accurate masses.

2) Identification. Candidate molecular species are identified by searching a large databas > 10E+7 entries of accurate masses (lipid precursor and fragment ions) predicted from each potential lipid structure and positive/negative ion adducts.

3) Alignment. The search results for each individual sample are aligned within a time window and the results are merged into a single report.

4) Quantification. Accurate-mass extracted ion chromatograms are integrated for each identified lipid precursor and peak areas are obtained. Analyte concentration for each lipid class was estimated relative to the concentration of internal standard.

5) Statistical Analysis. t-Tests determine which species are significantly different betwee sample vs. control groups, and results are displayed in whisker plots.

FIGURE 1. LipidSearch Workflow.

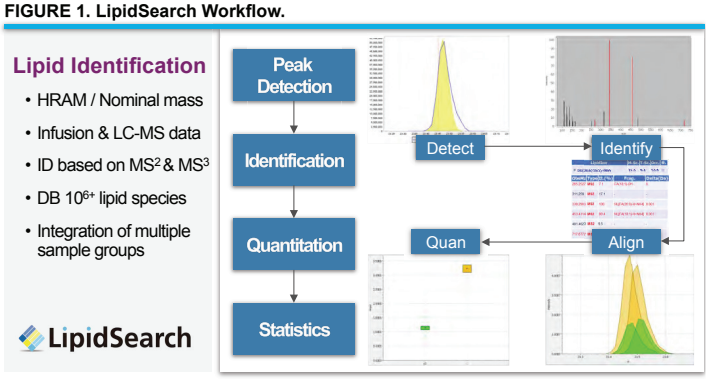

TABLE 1. NIST Human Plasma/Serum Samples

Group	NIST ID	Sample Type	Sample Description
C	SRM 1950	Human plasma	Equal number of men and women
S-1	2378-1	Human serum	Donors did not take fish or flaxseed oil
S-2	2378-2	Human serum	Donors took flaxseed oil supplements
S-3	2378-3	Human serum	Donors took fish oil supplements

Thermo SCIENTIFIC

TABLE 2. LipidSearch Conditions.

Search Parameter	Settings	Units
Precursor mass tolerance	5.0	ppm
Product mass tolerance	5.0	ppm
Prod. Intensity threshold	1.0	%
m-Score threshold/display	2.0 / 5.0	
Quan m/z tolerance	+/- 5.0	ppm
Quan range	+/- 1.0	min
Main isomer peak filter	ON	
ID Quality filter	A,B,C,D	
Adducts (pos ion)	+H, $+NH_4$, +Na	
Adducts (neg ion)	-H, -2H, $+CH_3CO_2$	
Lipid Sub-Classes	LPC, PC, LPE, PE, LPS, PS, LPG, PG, LPI, PI, LPA, PA, CL, SM, Cer, CerG1,G2,G3, DG, TG, ChE, Co	

Merge Parameter	Settings	Units
Retention time tolerance	0.20	min
All isomer peak filter	ON	
m-Score threshold	5.0	
ID Quality filter	A,B,C, D	

Results

Figure 4 shows the relative difference of 3 different lipid species: PC 16:0-20:5, DG 18:1-18:2 and TG 16:0-14:0-14:0. The relative amount of PC 36:5 and TG 44:0 increased in Serum-1, whereas DG 36:3 increased in Serum-2.

Figure 5 shows the relative levels of four different TG 54:6 isomers in the human serum and plasma samples: 1) TG(18:1-18:2-18:3) at 40.04 min, 2) $TG(18:2)_3$ at 40.25 min, 3) TG(16:0-18:1-20:5) at 41.22 min and 4) $TG(16:0)_2(22:6)$ at 41.90 min, retention time. The estimated concentration of TG 54:6 isomers is shown in Table 4. Serum 2 (flaxseed oil) is significantly higher in triglyceride containing 18:2 and 18:3 (ALA) fatty acids whereas Serum 1 (no supplements) contains more TG with polyunsaturated 20:5 (DPA) and 22:6 (DHA).

FIGURE 2. LipidSearch Results for Neg. Ion m/z 762.5289, Rt = 18.05 min.

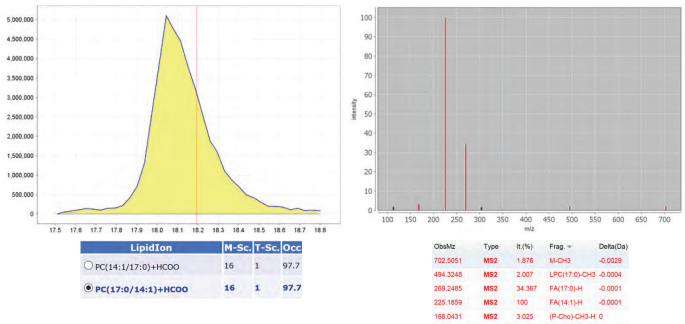

FIGURE 3. Aligned Peak Areas for PC Internal Standard 17:0/14:1, $M+H^+$

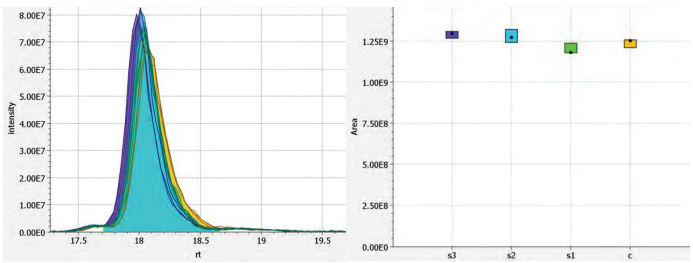

TABLE 3. Number of Lipid Species Identified after Merging Data, Filtered ID's, and Number of Lipid Species Quantified for Each Lipid Sub-class

Lipid Class	No. Species Unfiltered	No. Species Filtered	No. Species Reported
ChE	22	19	18
DG	81	46	45
TG	665	468	452
PC	508	221	220
LPC	106	58	57
PE/LPE	53	33	32
PI	34	25	24
Cer	33	22	21
CerG	20	13	13
SM	138	92	91
Total	**1660**	**997**	**973**

FIGURE 4. Relative Differences of Lipid Species in Human Serum/Plasma

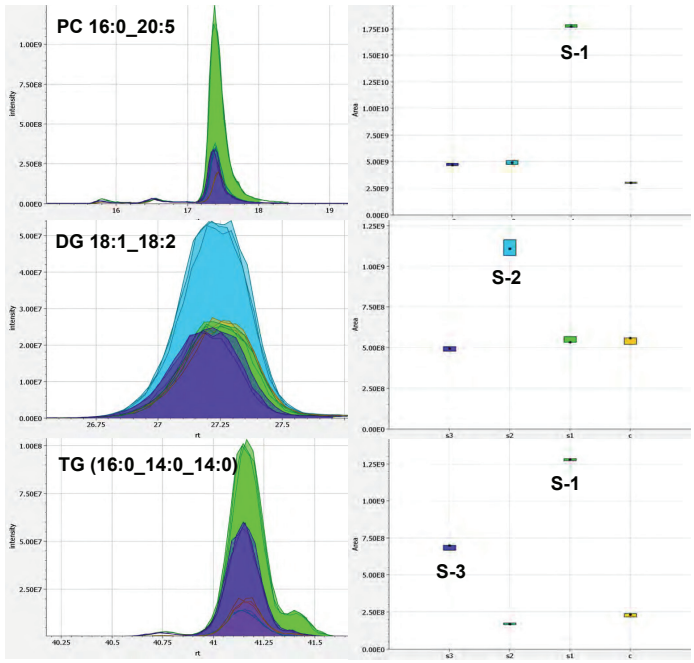

FIGURE 5. Comparison of Isomeric TG 54:6 Species in Human Serum/Plasma

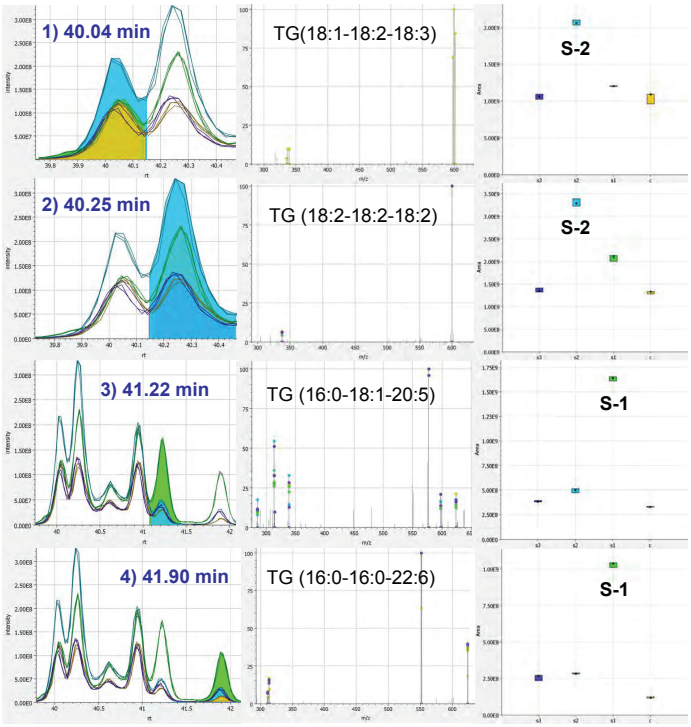

TABLE 4. Concentration of TG 54:6 in NIST Human Serum and Plasma Samples

		Conc., nmol/mL				CV, %			
TG Species	Rt, min	S-1	S-2	S-3	C	S-1	S-2	S-3	C
1) TG (18:1/18:2/18:3)	40.04	3.22	**5.31**	2.26	3.01	3.1	2.6	2.4	5.1
2) TG (18:2/18:2/18:2)	40.25	5.57	**8.49**	2.91	3.79	4.1	3.5	1.7	3.0
TG (54:6)	40.62	3.06	3.31	1.50	2.10	3.2	5.5	7.3	6.0
TG (54:6)	40.94	4.79	4.84	2.43	3.37	4.4	4.1	6.6	4.1
3) TG (16:0/18:1/20:5)	41.21	**4.38**	1.28	0.82	0.95	1.6	5.1	2.7	4.7
4) TG (16:0/16:0/22:6)	41.90	**2.76**	0.73	0.55	0.34	2.2	2.1	6.1	1.1

Criteria for Data Filtering

The overall results for identified and quantified lipids are summarized in Table 3. The number of unfiltered potential lipid identifications (1660) needs to be filtered to remove false positives using a combination of filters provided in LipidSearch. The criteria used to filter TG lipids (665 potential ID's) are summarized below. The number of TG species after filtering was 468 with 452 giving acceptable CV for quantitation.

- Precursor ion is the M+NH_4 adduct
- Lipid ID = Grade A or B; at least 2 out of 3 fatty acyl chains identified by neutral loss of FA+NH_3, monoacylglycerol (MG) and fatty acyl (RCO) fragment ions
- Match-score (Number of predicted peaks matched X weighting factor) ≥ 12
- Occupancy ≥ 25 (% of product ion intensity matched)

Conclusion

- LipidSearch software provides an automated workflow for processing high quality Orbitrap LC-MS/MS untargeted lipidomics data and enables reliable and comprehensive lipid identification and quantification.
- The high mass accuracy in both MS (120K) and MS^2 (30K) obtained with the Q Exactive HF instrument allows confident lipid species identification from the highly complex human serum and plasma extracts. Almost a thousand of species are identified automatically and quantified from a single LC-MS run.
- Orbitrap data combined with LipidSearch software allows simultaneous lipid ID with high coverage and quantitation. Each lipid ID was obtained with a single high quality Orbitrap MS^2 scan over 4 orders of concentration dynamic range.
- The CV of the technical replicates for a majority of the quantified lipids were typically below 15% showing the excellent reproducibility of the Q Exactive HF.

References

1. R Kiyonami, D A Peake, X Liu and Y Huang "Large Scale Lipid Profiling of a Human Serum Lipidome Using a High Resolution Accurate Mass LC/MS/MS Approach" presented at the LIPID MAPS Annual Meeting 2015, May 12-13, 2015, La Jolla, CA.
2. R Taguchi, *et al.,* J. Chrom. A, **2010**, *1217*, 4229–4239. doi: 10.1016/j.chroma. 2010.04.034.
3. T Yamada, *et al.,* J. Chrom. A, **2013**, *1292*, 211-218. doi: 10.1016/j.chroma. 2013.01.078.

www.thermoscientific.com

Africa +43 1 333 50 34 0
Australia +61 3 9757 4300
Austria +43 810 282 206
Belgium +32 53 73 42 41
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 45 453 9100
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
Spain +34 914 845 965
Sweden +46 8 556 468 00
Switzerland +41 61 716 77 00
UK +44 1442 233555
USA +1 800 532 4752

Thermo SCIENTIFIC
A Thermo Fisher Scientific Brand

PN64471-EN 0615S