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], Ken Miller[1]
[1]Thermo Fisher Scientific, San Jose, CA, USA; [2]Mitsui Knowledge Industry, Tokyo, Japan

Poster Note 90115

Introduction

Lipids play a key role in cell, tissue and organ physiology with diseases such as cancer and diabetes which involve disruption of their metabolic enzymes and 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 with an extensive lipid database. In addition, the mass spectrometer employed must be capable of separating many overlapping isobaric and isomeric lipid ions. We present here the details and challenges of the data processing of NIST plasma and serum extracts using the latest version of LipidSearch software. New algorithms were introduced specifically to reduce false positives and to automate the data review.

Thermo Scientific™ LipidSearch™ software was used for lipid identification through a database search of precursor accurate masses and their predicted fragment ions. Each lipid identification is ranked by mass tolerance, match to the theoretical fragment ions and fraction of total MS^2 intensity. 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 using 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

LC-MS Sample Preparation. Aliquots of human plasma or serum (80µL, Table 1) were extracted using the method of Bligh and Dyer.

LC-MS and dd-MS2 Method. A Thermo Scientific™ Dionex™ UltiMate™ 3000 Rapid Separation LC (RSLC) system and Thermo Scientific™ Q Exactive HF™ instrument were employed as described (1). The HPLC separation was achieved with a 2.1 x 250mm, 1.9µm, C30 prototype column and the MS analysis was performed at 120K resolution and MS2 at 30K resolution (FWHM at m/z 200).

Data Analysis. The data were processed using a HP Z840 workstation equipped with dual Hex-core Xeon processors (E5-2643v3, 2133 6C, 3.4 GHz), 64 GB of RAM (DDR4-2133), and a 500GB solid state hard drive.

Data Processing

LipidSearch software was used for lipid identification and relative quantitation using the workflow shown in Figure 1 (2, 3). LC-MS data (24 raw files, 5.9 GB total) containing MS and data dependent-MS2 were searched using the parameters listed in Table 2. For each MS2 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. The same lipid annotations within ± 0.1 min 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. 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.

Thermo SCIENTIFIC

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 took fish oil supplements
S-2	2378-2	Human serum	Donors took flaxseed oil supplements
S-3	2378-3	Human serum	Donors did not take fish or flaxseed oil

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	**Phospholipids**: LPC, PC, LPE, PE, LPS, PS, LPG, PG, LPI, PI, LPA, PA, CL **Sphingolipids**: SM, Cer, CerG1, CerG2, CerG3 **Glycerolipids**: DG, TG **Neutral lipids**: 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

LC-MS/MS 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 database > 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.** The accurate-mass extracted ion chromatograms are integrated for each identified lipid precursor and the peak areas are obtained. Analyte concentration for each lipid class was calculated relative to the concentration of internal standard.
5) **Statistical Analysis.** t-Tests determine which lipid species are significantly different between sample vs. control groups, and results are displayed in a whisker plot.

FIGURE 1. LipidSearch Workflow.

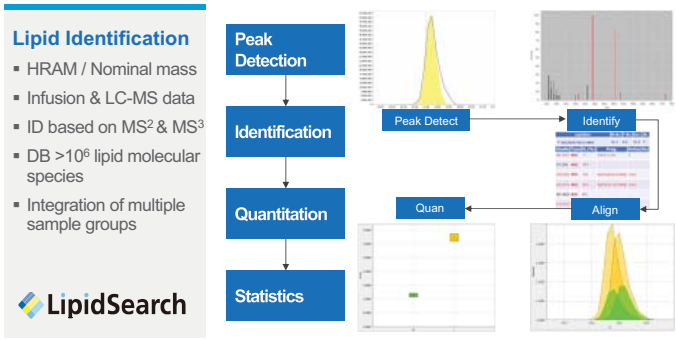

FIGURE 2. LipidSearch Results for Neg. Ion m/z 762.5289, Rt = 18.05 min.

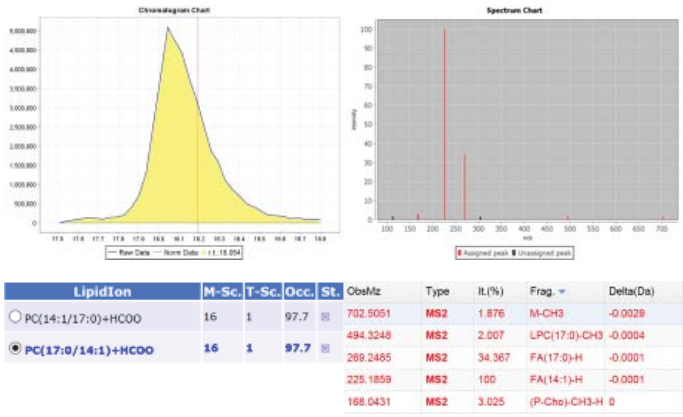

FIGURE 3. Aligned Peak Areas for PC Internal Standard 17:0/14:1, $M+H^+$.

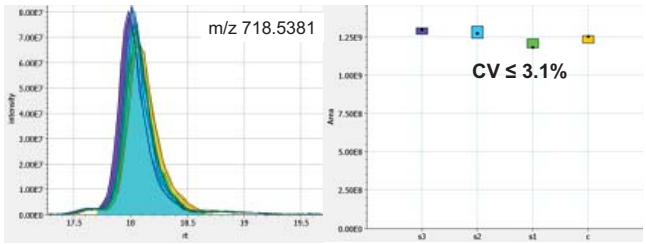

TABLE 3. Number of Lipid Species Identified after Merging Data, Filtered IDs, and Number of Lipid Species Reported and Quantified for Each Lipid Sub-class.

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 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. Serum 2 (donors taking flaxseed oil supplement) is significantly higher in triglyceride containing 18:2 and 18:3 (ALA) fatty acids whereas Serum 1 (donors taking fish oil supplement) contains more TG with polyunsaturated 20:5 (DPA) and 22:6 (DHA).

FIGURE 4. Comparison of Isomeric TG 54:6 Species in Human Serum/Plasma.

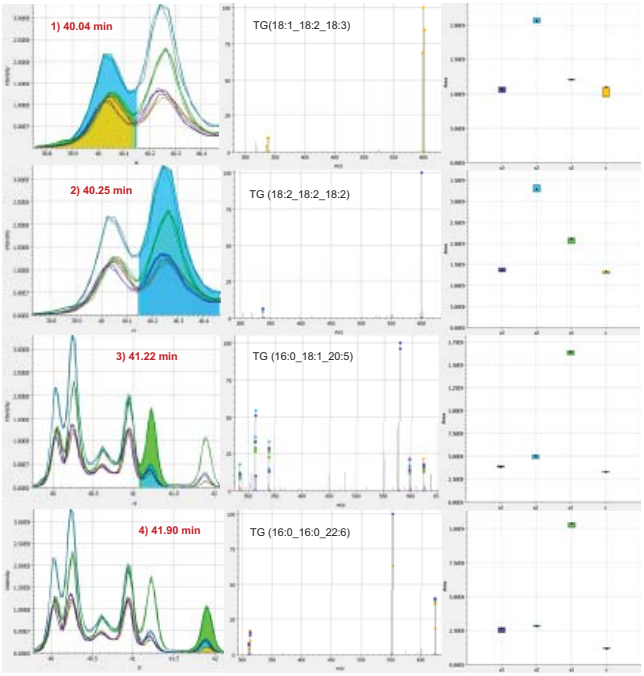

TABLE 4. Concentration of TG 54:6 in NIST Human Serum and Plasma Samples.

TG Species	Rt, min	Conc., nmol/mL				CV, %			
		S-1	S-2	S-3	C	S-1	S-2	S-3	C
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
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
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
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

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.