Application Note

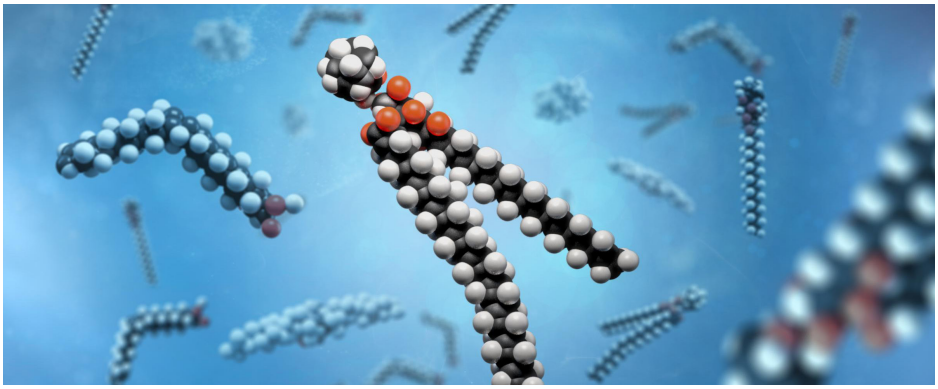
Lipidomics

Agilent

Trusted Answers

Targeted Data Mining and Annotation of Untargeted High-Resolution Lipidomics Data

A comprehensive, high-confidence workflow



Authors

Sheher Banu Mohsin,
Layla Cosovic,
Mark Sartain, Tracy Blethen,
Pietro Morlacchi, and
Daniel Cuthbertson
Agilent Technologies, Inc.

Abstract

Complete and unambiguous characterization of lipids remains a challenging task in untargeted lipidomics approaches. To address this challenge, a highly curated lipidomics database amenable for LC/Q-TOF and Ion Mobility LC/Q-TOF was created with a standardized methodology that is easy to implement in many laboratory environments. To achieve this, a rugged, highly curated, comprehensive LC/TQ method covering a total of 763 lipids across all major lipid classes was used.[1] This set of lipids was the starting point for the creation of a personal compound database and library (PCDL) that includes retention time (RT), MS/MS spectra, and collisional cross section. These serve as orthogonal measurement parameters that ensure higher confidence lipid identification. In this application note, the use of this highly curated PCDL[2] is demonstrated within an untargeted lipidomics workflow designed to help users gain critical biological insights from their lipidomics data.

Introduction

Mass spectrometry-based lipidomics profiling relies on the separation and subsequent identification of many structurally diverse lipids in a single analytical run. Efficient separation of lipids by chromatographic techniques is paramount for their accurate and unambiguous identification, particularly in complex biological extracts. Given the abundance of isomeric and isobaric species within each lipid class, coupled with their similar chromatographic behavior, lipid separation and annotation present notable challenges.

Liquid chromatography techniques for the resolution of lipid isomers and isobars have made significant progress in the last few years. Both reversed- and normal phase methods as well as SFC approaches are available. However, despite these significant improvements, lipid identification by mass spectrometry remains challenging. Complete and unambiguous characterization of lipids (by LC/MS) generally requires the identification of class, elemental composition, R-group position, number, location, and geometry of double bonds along with other parameters. Much of this structural information can be obtained through high-resolution mass spectrometry (HRMS) tandem MS/MS experiments, where the spectral data are analyzed against fragment libraries for compound identification. While small MS/MS spectral libraries from authentic standards are available, many lipid MS/MS spectra are annotated using in silico libraries. These might help provide some knowledge about lipid class and R-group chain length but do not often provide other essential molecular information, making complete lipid annotation an unrealistic goal.

To increase confidence in lipid identification and overcome the ambiguities in their annotation, it is common practice in mass spectrometry to use authentic standards. These standards serve to match mass features based on their RT, accurate mass, and fragmentation patterns when available. Unfortunately, this approach might be limited, as neat standards for thousands of lipids are not available. To address these challenges, alternative strategies for high-confidence lipid annotation are needed.

The task of profiling lipids in plasma was achieved to a high degree of confidence. The approach was to first use a rugged, highly curated, comprehensive LC/TQ method with 763 lipids.[1] This method was designed to cover all the major lipid classes in plasma from small lysolipids to larger nonpolar triacylglycerols. The method underwent a significant level of curation aiming to help facilitate the annotation of numerous lipid isomers, which also required synthesis of lipid standards with specific isomeric characteristics.[3]

This method was transferred to a LC/Q-TOF and ion mobility LC/Q-TOF, which required converting the unit mass resolution masses of each precursor/fragment pair to the correspondent accurate masses. These accurate masses and their corresponding RTs were used as precursor ions in both targeted and untargeted experiments conducted using the LC/Q-TOF system. The HPLC separation method used in these experiments was the same as the original LC/TQ method. The MS/MS spectra generated by the LC/Q-TOF were then exported to create a PCDL, which could be used for the identification of putative lipid entries based on accurate mass, RT, and MS/MS spectra.

RT is one of the most important parameters in annotating lipids with high confidence. Relative standard deviations (RSDs) of RTs from interlaboratory studies at four different sites were used to determine day-to-day and site-to-site variability of RT.[4] RSDs for RT for > 600 lipids were found to be < 0.2%. To further increase confidence in lipid annotation, a set of collision cross section (CCS) values on the ion mobility LC/Q-TOF was generated. By including CCS values, ID Searches can now be done with accurate mass, RT, MS/MS spectral match, and CCS filtering—increasing confidence in annotations.

The process of generating the high-confidence PCDL and the steps taken to ensure a high level of curation are described in detail in Agilent publication number 5994-7627EN.[2]

This paper discusses the workflow implementation for using the high-confidence PCDL for annotating lipids. It covers all the steps from lipid extraction and data acquisition to the final step of lipid annotation using the high-confidence database.

Workflow implementation

This workflow describes a method to profile 677 different lipid species from 44 lipid classes within a 10 µL sample of plasma, or an equivalent amount of cell or tissue homogenate. Lipids are separated by reversed-phase liquid chromatography followed by analysis on a high-resolution mass spectrometer. The total analysis time was 16 minutes per sample for a single column method, or a 13-minute run time for a dual column system. Agilent MassHunter Workstation software was used for data processing.

Experimental

Sample preparation

The procedure for sample preparation, including separation, is described in a previous Agilent publication.[1] Briefly, lipid extractions were performed in 1.5 mL microcentrifuge tubes using polypropylene positive displacement pipettes. Ten microliters of plasma were mixed with 100 µL of extraction solvent, consisting of butanol:methanol (1:1) containing 10 mM ammonium formate and a mixture of internal standards. Each sample was then vortexed for 5 seconds and later bath-sonicated for 1 hour, with the temperature maintained between 21 and 25 °C. Samples were then centrifuged (13,000 rpm, 10 minutes, 20 °C), and the supernatant was transferred into 1.5 mL glass sample vials (part number 5190-9062) with 200 µL glass inserts (part number 5183-2085). Samples were capped using PTFE/S caps (part number 5185-5820) and stored at –80 °C until analysis.

LC instrumentation

Agilent has recently introduced a standard multi-omics LC configuration that enables easy implementation of methods for the analysis of metabolites, lipids, and proteins on any mass spectrometer model. This LC configuration was implemented in the analysis and further showed the RT alignment that is possible with LC standardization.

The LC consisted of an Agilent 1290 Infinity II LC system with the following modules:

- Agilent 1290 Infinity II High-Speed Pump (G7120A)
- Agilent 1290 Infinity II Vialsampler with Thermostat (G7129B)
- Agilent 1290 Infinity II Multicolumn Thermostat (G7116B)

The LC configuration for targeted lipid analysis used an LC/TQ. For untargeted profiling experiments, an Agilent Revident LC/Q-TOF and an Agilent 6560 ion mobility LC/Q-TOF for mass spectral and ion mobility measurements were used. The common chromatographic conditions are shown in Table 1.

Table 1. Chromatographic conditions for lipidomics analysis using an Agilent Revident LC/Q-TOF and an Agilent 6560 ion mobility LC/Q-TOF.

Parameter	Value
Analytical Column	Agilent ZORBAX RRHD Eclipse Plus C18, 2.1 × 100 mm, 1.8 µm (p/n 959758-902)
Inline Filter	Agilent 1290 Infinity II inline filter, 0.3 µm, 2 mm id, 1,300 bar (p/n 5067-6189)
Column Temperature	45 °C
Injection Volume	1 µL (equivalent to 0.1 µL plasma)
Autosampler Temperature	20 °C
Needle Wash	Wash vial, 3 s, 1:1 butanol:methanol Alternatively, flush the port for 10 s with 1:1 mobile phase A:B
Mobile Phase	A) 10 mM ammonium formate, 5 µM Agilent InfinityLab deactivator additive (p/n 5191-3940) in 5:3:2 water:acetonitrile:2-propanol B) 10 mM ammonium formate in 1:9:90 water:acetonitrile:2-propanol (Note: do not add InfinityLab deactivator additive to mobile phase B)
Flow Rate	0.4 mL/min
Gradient Program	Time %B 0 15 2.5 50 2.6 57 9 70 9.1 93 11 96 11.1 100 12 100 12.2 15 16 15
Stop Time	16 min
Post Time	None

Agilent Revident LC/Q-TOF acquisition method

The Revident LC/Q-TOF platform with the Agilent Jet Stream ESI source was used to acquire MS/MS data. Data were collected in both positive and negative mode. Detailed experimental methods for chromatography and Auto MS/MS mass spectrometry parameters are provided in Table 2.

Table 2. Agilent Revident LC/Q-TOF Auto MS/MS parameters.

Parameter	Value
Gas Temperature	300 °C
Gas Flow	10 L/min
Nebulizer Pressure	25 psig
Sheath Gas Temperature	300 °C
Sheath Gas Flow	12 L/min
Capillary Voltage	3,500 V (+), 3,000 V (–)
Nozzle Voltage	0 V
Fragmentor	150 V
Skimmer	45 V
Octopole RF Vpp	750 V
Reference Mass	*m/z* 121.050873, *m/z* 1,221.990637 (+) *m/z* 119.03632, *m/z* 980.016375 (–)
MS and MS/MS Range	*m/z* 50 to 3,000 (+)
Minimum MS and MS/MS Acquisition Rate	3 spectra/s
Isolation Width	Narrow (~1.3 *m/z*)
Collision Energy	25 eV
Maximum Precursors per Cycle	3
Precursor Abundance-Based Scan Speed with Precursor Abundance-Based Acquisition Speed	Yes, target 25,000 counts/spectrum
Use MS/MS Accumulation Time Limit	Yes
Reject Precursors that Cannot Reach Target TIC	No
Threshold for MS/MS	5,000 counts and 0.001%
Active Exclusion Enabled	Yes, one repeat, then exclude for 0.05 minutes
Purity	Stringency 70%, cutoff 0%
Isotope Model	Common organic molecules
Sort Precursors	1, 2, unknown
Static Exclusion Ranges	*m/z* 40 to 151 (+) *m/z* 40 to 210 (–)

Agilent 6560 ion mobility LC/Q-TOF acquisition method

The Agilent 6560 ion mobility LC/Q-TOF platform with the Agilent Jet Stream ESI source was used for determining the CCS values for lipids. Ion mobility (IM) mass spectrometry data were collected in both positive and negative ionization modes from *m/z* 50 to 3,000 with a cycle time of 1 s/spectra, an IM trap fill time of 3,800 µs, and trap release time of 250 µs. Multiplexing was used with a four-bit pulsing sequence length. Lipids were fragmented by collision induced dissociation using a data-independent acquisition All Ions method and alternating frames between 0 and 25 V collision energy. Acquisition parameters for the IM LC/Q-TOF are shown in Table 3.

Table 3. Agilent 6560 ion mobility LC/Q-TOF data-independent acquisition parameters.

Parameter	Value
Gas Temperature	300 °C
Gas Flow	10 L/min
Nebulizer Pressure	25 psig
Sheath Gas Temperature	350 °C
Sheath Gas Flow	11 L/min
Capillary Voltage	3,500 V (+), 3,000 V (–)
Nozzle Voltage	1,000 V
Fragmentor	400 V
Octopole RF Vpp	750 V
Reference Mass	*m/z* 121.050873, *m/z* 1,221.990637 (+) *m/z* 119.03632, *m/z* 980.016375 (–)
MS and MS/MS Range	*m/z* 50 to 3,000 (+)
Frame Rate	0.9 frames/s
IM Transient Rate	17 IM transients/frame
Collision Energy	25 eV
Maximum Drift Time	60 ms
Drift Tube Entrance Voltage	1,750 V

Software for data analysis

Agilent ChemVista software was used to manage and edit the lipid libraries (PCDLs). The data analysis method, Lipids.m, was used for batch-targeted feature extraction using (+H) and ($+NH_4$) checked for positive ions and (−H) and ($+CH_3COO)^-$ for negative ions.

Agilent MassHunter Mass Profiler Professional (MPP) software version 15.1 is used for differential analysis. Two experiments (positive or negative ion) were created with the Lipidomics experiment type. A percentile shift normalization algorithm (75%) was used, and datasets were baselined to the median of all samples. Agilent MassHunter Mass Profiler (MP) software and MPP offer statistical analysis software tools for examining and interpreting complex lipidomics results. MP and MPP enable RT and *m/z* alignment for features across samples, response normalization, and statistics for identification of significant differences between samples. MPP has an added experiment type for annotated lipid analysis, which allows lipid class-based normalization and visualization tools.

Agilent MassHunter ID Browser version 10.0 was used within MPP to make annotations in the untargeted workflow, with masses ± 5 ppm and RTs ± 0.10 minutes as required criteria.

Data analysis for IM LC/Q-TOF and LC/Q-TOF data files

Files were processed using MP software. Note that for multiplexed data files, the files are first demultiplexed before using MP. Briefly, demultiplexed data files from the IM LC/Q-TOF or .d files from the LC/Q-TOF are imported into MP or MPP in a batch mode. After feature extraction, entities are annotated using the high-confidence PCDL. Annotations are based on accurate mass, RT, and MS/MS spectra match. For the IM LC/Q-TOF data files, the CCS value is also matched to provide a high degree of confidence for lipid annotations. A general schematic of the workflow for data acquisition and data analysis is shown in Figure 1. Figure 2 shows the steps and workflow for data analysis.

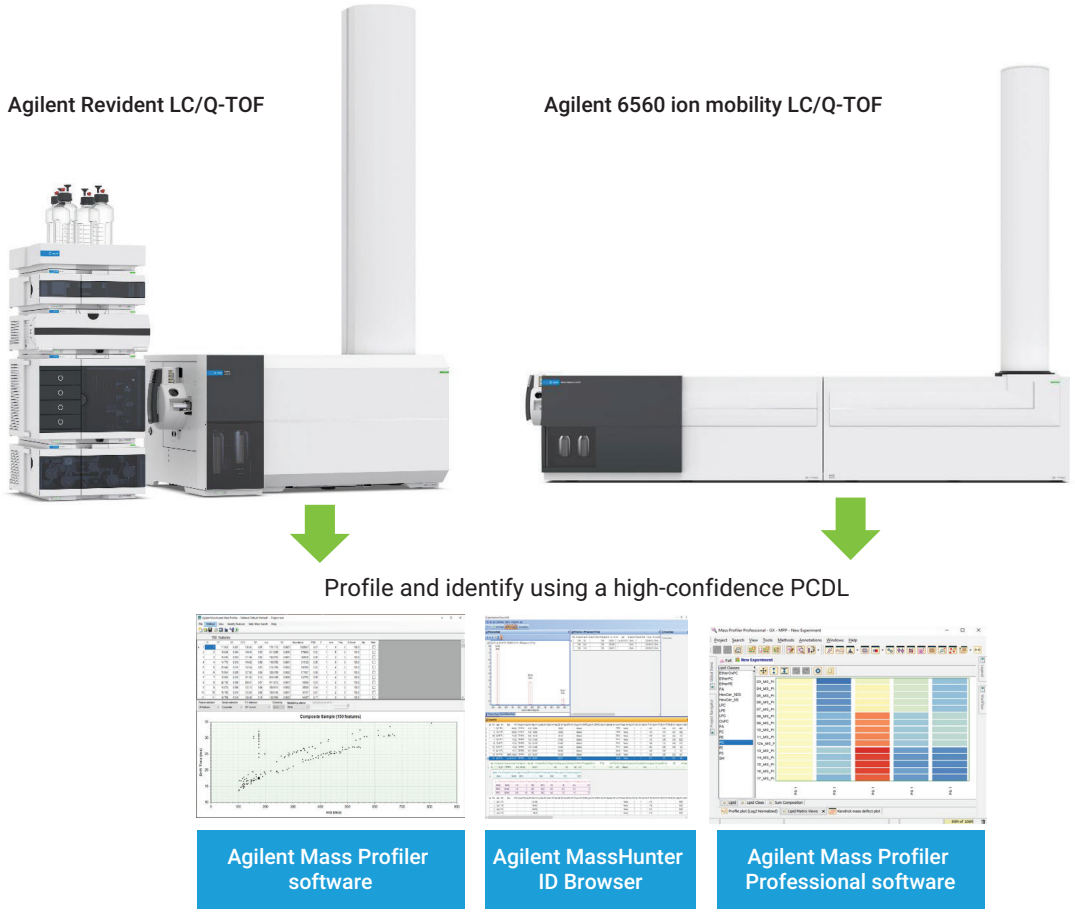


Figure 1. Workflow for data acquisition and data processing.

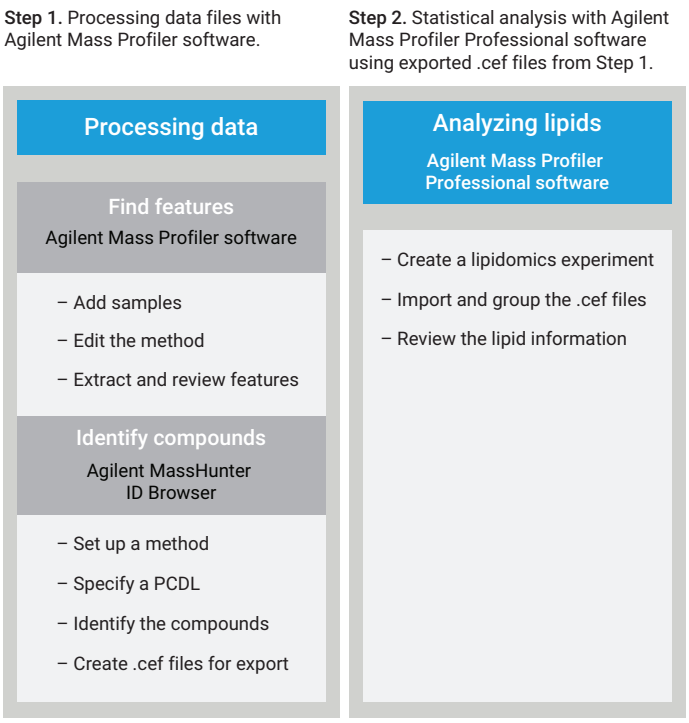


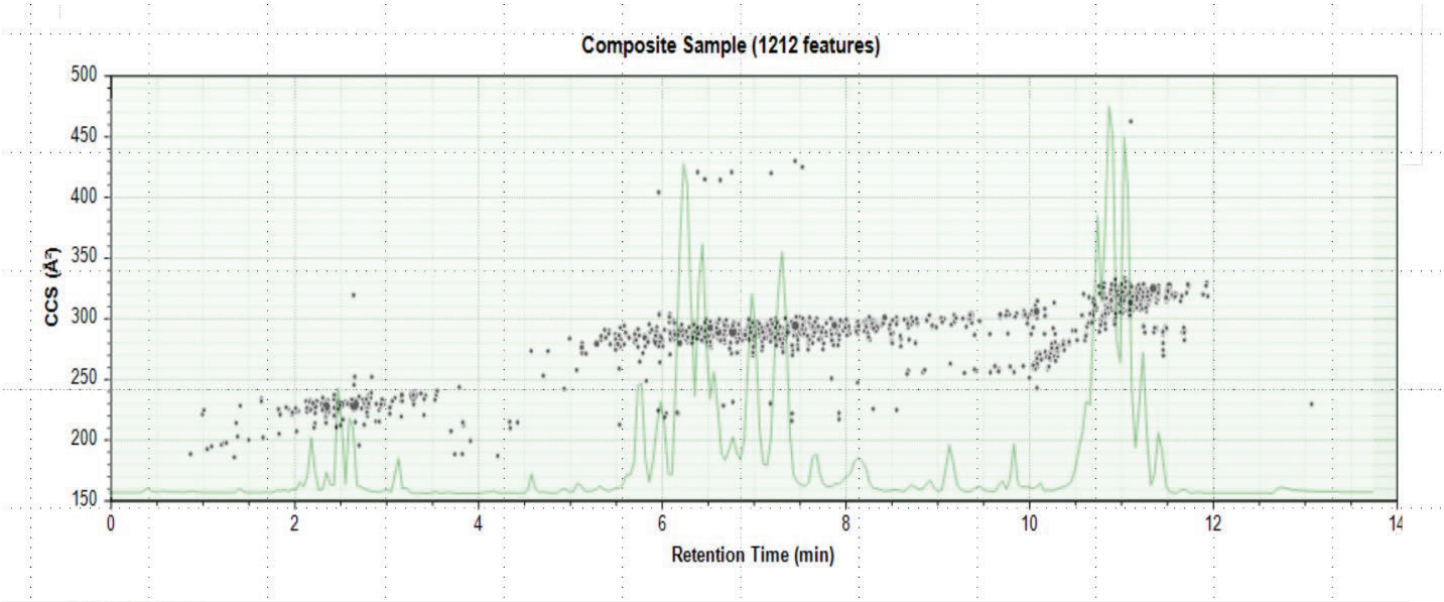
Figure 2. Schematic of the workflow for data analysis.

Results and discussion

Results from Step 1 of the data processing step in Figure 2, which includes feature-finding and annotation, are shown in Figure 3. Each of the rows in the table in Figure 3 can be examined in more detail to show the extracted ion chromatogram and the mass spectrum of the selected feature.

Lipid matrix is one of several visualization tools available for viewing lipid data in MPP. The graphic display shows the different levels of abundance in the samples for each of the lipids in the PC class. Blue means less abundance. Red means strong abundance, and the colors between blue and red mean increasing degrees of abundance, from yellow to orange to red. Figures 4 and 5 show one of several plots for viewing lipid data. Mass versus RT, colored according to lipid class, is shown in Figure 5.

A Kendrick mass defect plot is an excellent tool for visualizing homologous series, as shown in Figure 6.[5] Compounds with the same number of heteroatoms and number of rings plus double bonds, but different numbers of CH_2 groups will have an identical Kendrick mass defect. These compounds line up in a horizontal line on the plot. Any potential database mismatches or missing annotations can easily be identified.



6121 features

	ID	Name	Formula	RT (DB)	Ion Species	Mass (DB)	RT	SD	DT	SD	CCS	SD	m/z	SD	Abundance
1	1	SM (d18:1/16:0)	C39 H79 N2 O...	5.856	(M+H)+	702.5676	5.955	0.008	33.818	0.059	282.42	0.49	703.5746	0.0009	3348672
2	2	LPC (16:0) [sn1]	C24 H50 N O7 P	2.591	(M+H)+	495.3325	2.645	0.007	27.273	0.042	229.55	0.35	496.3407	0.0015	2749995
3	3	PC (18:1_18:1)	C44 H84 N O8 P	7.239	(M+H)+	785.5935	7.523	0.011	34.848	0.039	290.43	0.32	786.5994	0.0011	2521345
4	4	PC (16:0_18:1)	C42 H82 N O8 P	7.055	(M+H)+	759.5778	7.184	0.011	34.331	0.035	286.29	0.29	760.5849	0.0011	2278330
5	5	LPC (18:0) [sn2]	C26 H54 N O7 P	3.111	(M+H)+	523.3638	3.290	0.009	28.224	0.029	237.22	0.25	524.3704	0.0008	2185121
6	6	TG (52:2) [SIM]	C55 H102 O6	11.314	(M+NH4)+	858.7676	11.313	0.003	38.602	0.058	321.16	0.48	876.7991	0.0029	2108318
7	7	PC (18:1_18:2)	C44 H82 N O8 P	6.494	(M+H)+	783.5778	6.766	0.010	34.705	0.045	289.25	0.37	784.5845	0.0008	2080693
8	8	PC (16:0_18:2)	C42 H80 N O8 P	6.355	(M+H)+	757.5622	6.460	0.009	34.044	0.175	283.91	1.46	758.5683	0.0016	2009853
9	9	LPC (18:2) [sn2]	C26 H50 N O7 P	2.174	(M+H)+	519.3325	2.340	0.008	27.095	0.076	227.77	0.64	520.3397	0.0009	1995130
10	10	TG (52:3) [SIM]	C55 H100 O6	11.16	(M+NH4)+	856.7520	11.156	0.006	38.486	0.075	320.21	0.62	874.7841	0.0007	1977839
11	11	PC (15:0_18:1) ...	C41 H73 D7 N ...	6.517	(M+H)+	752.6061	6.634	0.010	34.102	0.191	284.43	1.59	753.6113	0.0012	1967886
12	12		C32 H52 N2 O4		(M+H)+	528.3927	2.774	0.007	27.776	0.091	233.40	0.77	529.3991	0.0005	1943741
13	13	PC (38:6) (a)	C46 H80 N O8 P	5.767	(M+H)+	805.5622	6.206	0.008	34.749	0.043	289.49	0.36	806.5679	0.0019	1915518
14	14	PC (16:0_18:2)	C42 H80 N O8 P	6.355	(M+H)+	757.5622	6.460	0.009	34.319	0.130	286.20	1.09	758.5696	0.0013	1796827

Figure 3. Results from the annotation of a NIST plasma extract using the high-confidence PCDL.

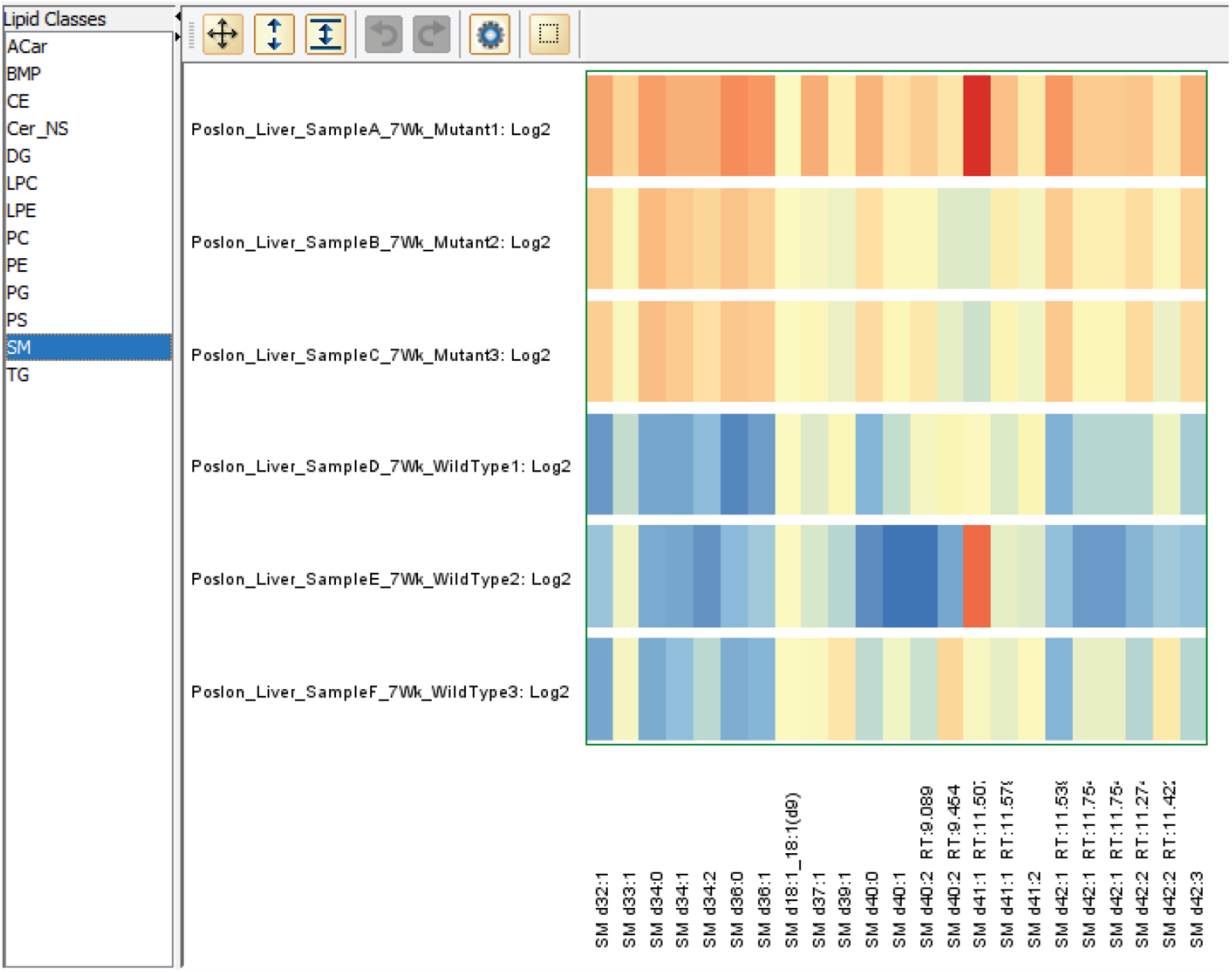


Figure 4. The lipid matrix generated using Agilent Mass Profiler Professional software from Step 2 in Figure 2. Results are shown for PC.

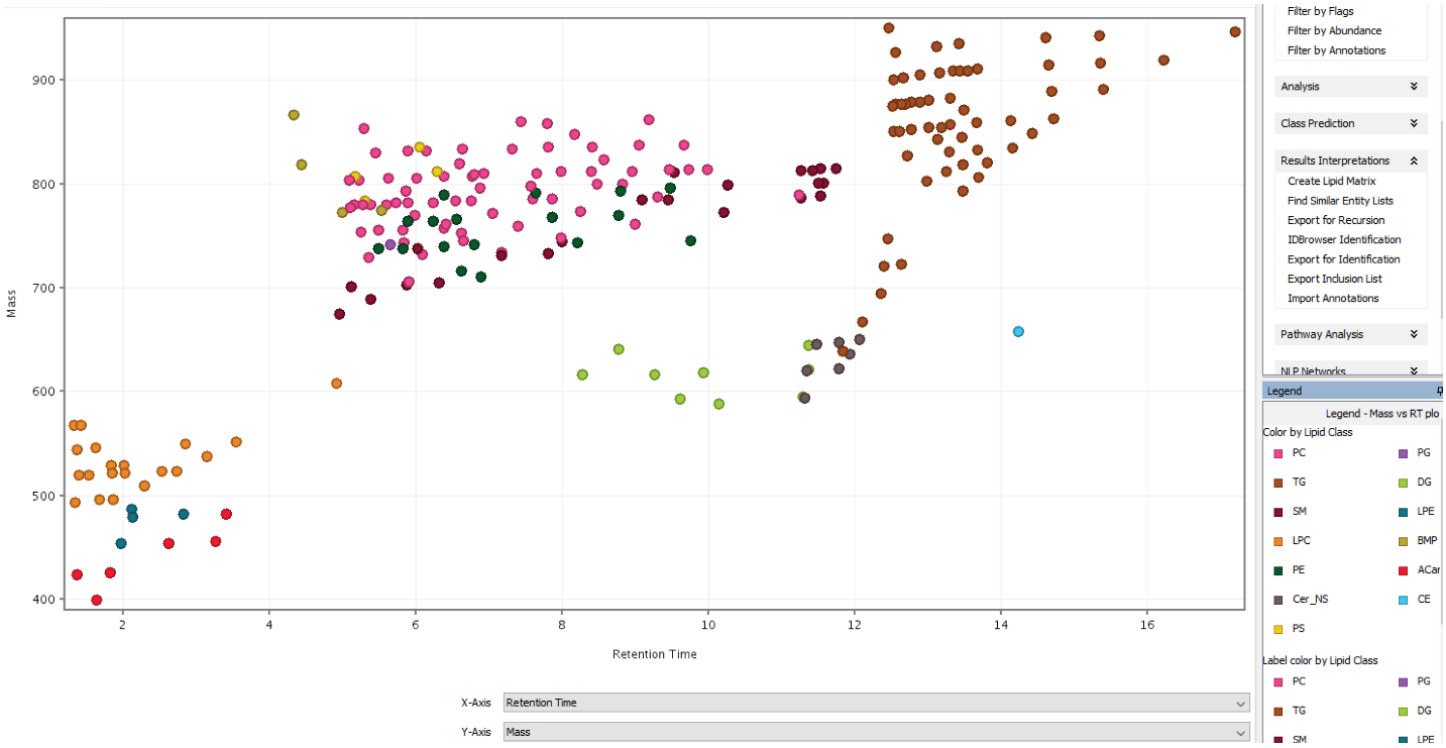


Figure 5. Plot of mass versus RT for different classes of lipids.

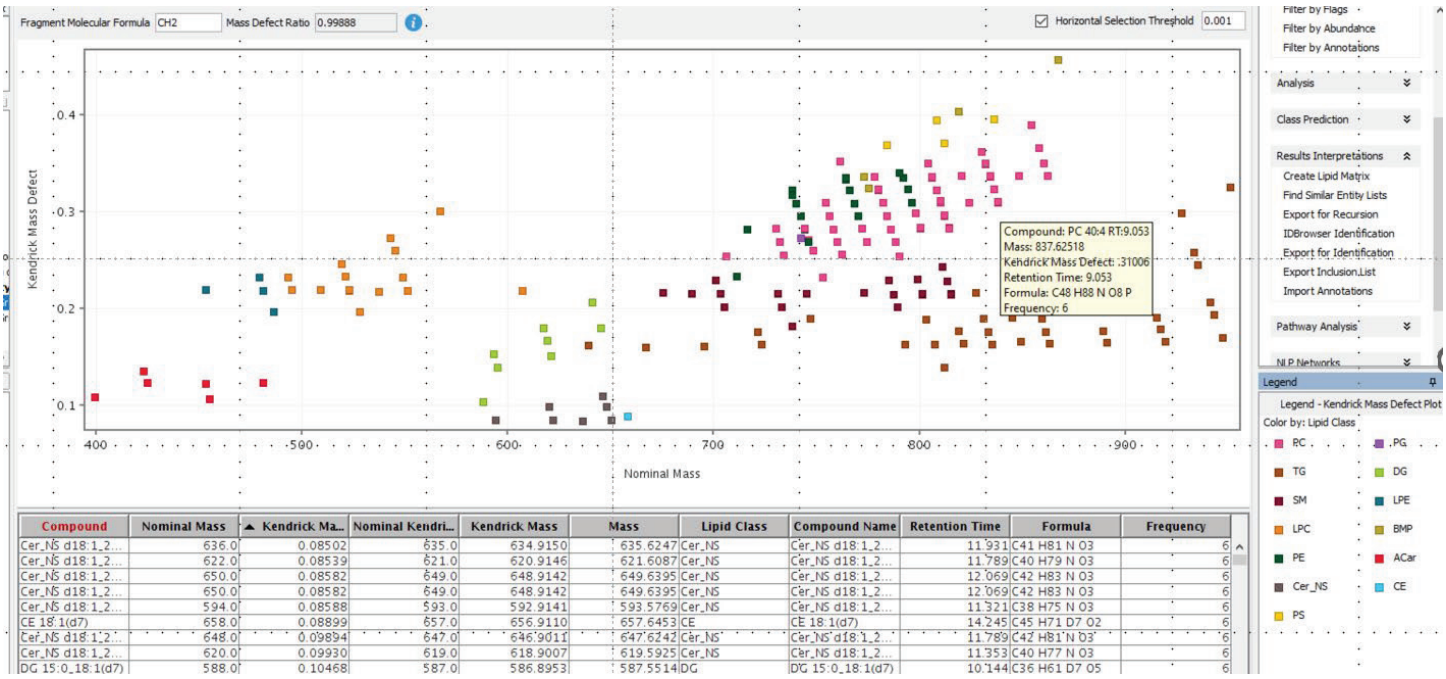


Figure 6. Kendrick mass defect plot showing lipids differing by a CH_2 group.

Conclusion

- A workflow for generating high-quality and high-confidence lipid annotations from untargeted LC/Q-TOF and IM LC/Q-TOF data was described.
- A PCDL was created that covers all the major lipid classes.
- The LC/Q-TOF and IM LC/Q-TOF workflow uses a high-confidence lipid PCDL with accurate mass, RT, MS/MS, and collision cross section entries.
- The workflow covers sample preparation, data acquisition, and data analysis resulting in high-confidence lipid annotations for untargeted LC/Q-TOF and IM LC/Q-TOF data.

References

1. Huynh, K.; Sartain, M.; *et al.* A Comprehensive, Curated, High-Throughput Method for the Detailed Analysis of the Plasma Lipidome. *Agilent Technologies application note*, publication number 5994-3747EN, **2021**.
2. Mohsin, S.; *et al.* Creation of a High-Confidence Lipidomics Personal Compound Database Library – for Targeted Data Mining and Annotation of Untargeted High-Resolution Lipidomics Data. *Agilent Technologies technical overview*, publication number 5994-7627EN, **2024**.
3. Huynh, K.; Barlow, C. K.; Jayawardana, K.S; Weir, J. M.; Mellett, N. A.; Cinel, M.; Magliano, D. J.; Shaw, J. E.; Drew, B. G.; Meikle, P. J. High-Throughput Plasma Lipidomics: Detailed Mapping of the Associations with Cardiometabolic Risk Factors. *Cell Chem. Biol.* **2019**, *26(1)*, 71–84. DOI: 10.1016/j.chembiol.2018.10.008
4. Sartain, M.; *et al.* An Interlaboratory Evaluation of a Targeted Lipidomics Method in Plasma. *Agilent Technologies application note*, publication number 5994-6830EN, **2024**.
5. Korf, A.; Vosse, C.; Schmid, R.; Helmer, P. O.; Jeck, V.; Hayen, H. Three-Dimensional Kendrick Mass Plots as a Tool for Graphical Lipid Identification. *Rapid Commun. Mass Spectrom.* **2018**, *32(12)*, 981–991. DOI: 10.1002/rcm.8117

www.agilent.com

RA45484.6137962963

Printed in the USA, August 12, 2024
5994-7588EN

Agilent
Trusted Answers