analytical chemistry

Article

Cite This: Anal. Chem. 2017, 89, 12360-12368

pubs.acs.org/ac

Validating Quantitative Untargeted Lipidomics Across Nine Liquid Chromatography−High-Resolution Mass Spectrometry Platforms

Tomas Cajka,[†] Jennifer T. Smilowitz,[‡,§] and Oliver Fiehn*[,†,⊥]

[†]West Coast Metabolomics Center, UC Davis Genome Center, University of California, Davis, Davis, California 95616, United States

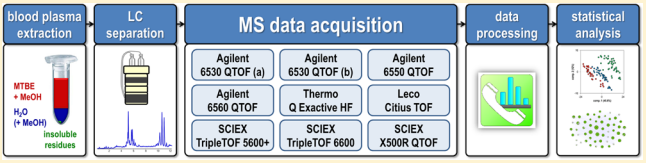
[‡]Department of Food Science and Technology, University of California, Davis, Davis, California 95616, United States

[§]Foods for Health Institute, University of California Davis, Davis, California 95616, United States

[⊥]Biochemistry Department, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Supporting Information

ABSTRACT: Liquid chromatography−mass spectrometry (LC−MS) methods are most often used for untargeted metabolomics and lipidomics. However, methods have not been standardized as accepted "best practice" documents, and reports lack harmonization with respect to quantitative data that enable interstudy comparisons. Researchers use a wide variety of high-resolution mass spectrometers under different operating conditions, and it is unclear if results would yield different biological conclusions depending on the instrument performance. To this end, we used 126 identical human plasma samples and 29 quality control samples from a nutritional intervention study. We investigated lipidomic data acquisitions across nine different MS instruments (1 single TOF, 1 Q/orbital ion trap, and 7 QTOF instruments). Sample preparations, chromatography conditions, and data processing methods were kept identical. Single-point internal standard calibrations were used to estimate absolute concentrations for 307 unique lipids identified by accurate mass, MS/MS spectral match, and retention times. Quantitative results were highly comparable between the LC−MS platforms tested. Using partial least-squares discriminant analysis (PLS-DA) to compare results between platforms, a 92% overlap for the most discriminating lipids based on variable importance in projection (VIP) scores was achieved for all lipids that were detected by at least two instrument platforms. Importantly, even the relative positions of individual samples on the PLS-DA projections were identical. The key for success in harmonizing results was to avoid ion saturation by carefully evaluating linear dynamic ranges using serial dilutions and adjusting the resuspension volume and/or injection volume before running actual study samples.



Liquid chromatography−mass spectrometry (LC−MS) is the preferred technique in metabolomics and lipidomics.[1−4] Advantages are reliable identification of metabolites, even at trace levels, separation of isomers and isobars, and reduced ion-suppression effects compared to direct infusion−MS-based methods.[5] Current LC instruments permit effective compound separations with a throughput of more than 300 samples per week and instrument. Despite this rapid progress in LC−MS platform performances, a key limitation is the current lack of methodological standardization in metabolomics and lipidomics, and harmonization of reporting results for interstudy comparisons in databases such as the Metabolomics Workbench.[6] Few studies have focused on the comparison of multi-instrument or interlaboratory reproducibility of LC−MS-based metabolomics or lipidomics data, usually with a limited number of samples.[7−15] These workflows targeted specific analytes and relied on internal or external standards for quantification, while untargeted workflows compared results from different platforms after statistical analysis using either raw or normalized peak intensities. However, such untargeted peak reports are not easy to compare between laboratories and studies, making validation of biological findings more difficult. Results from untargeted metabolomics and lipidomics analysis are known to be dependent on sampling and extraction methods in addition to details of chromatographic parameters. Estimating absolute chemical concentrations of particular lipid species is an important step to enable direct comparisons of results between studies. Reporting absolute lipid quantities immediately distinguishes major from minor lipid species, allowing biological interpretations of the results in the context of other analytes.[16] However, use of reference standards for each target lipid species is not feasible due to their unavailability or prohibitive price, so compromise is needed for quantification of lipids in biological samples.[17]

Best practice protocols for LC−MS-based untargeted lipidomics include using a series of multiple internal standards (one or more per lipid class), spiked into all the samples at different stages of the sample processing (e.g., extraction, dry extract resuspension).[4] These internal standards can be used to estimate lipid quantifications for each lipid class as one-point peak intensity ratios, multiplied by the concentration of the internal standard. If needed for validation studies, these one-

Received: August 21, 2017
Accepted: October 24, 2017
Published: October 24, 2017

ACS Publications © 2017 American Chemical Society

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368

point calibration measures can be extended to isotope-dilution calibration curves.[17]

As a caveat, one-point calibrations are estimates, not fully validated concentrations. For instance, instrument responses for saturated and unsaturated phospholipid species decrease with increasing acyl-chain length, affecting accurate concentration estimations.[18] Second, peak intensities vary markedly, even within each lipid class, depending on the solvent mixtures at the point of electrospray ionization. For this reason, one-point lipidomic quantifications will yield different results between reversed-phase liquid chromatography (RPLC) and hydrophilic interaction chromatography or normal-phase liquid chromatography (NPLC) methods that separate lipids mainly according to classes of polar head groups.[5,17]

While harmonizing lipidomic reports is under way with respect to naming and annotating lipids,[19,20] it is unclear if results yield different biological conclusions depending on the high-resolution mass spectrometer used in LC−MS-based untargeted lipidomics. Despite strides made to standardize untargeted lipidomics, the type of mass spectrometer used across laboratories can never be fully standardized. To this end, we tested nine different instruments from four mass spectrometry manufacturers, using identical samples, sample preparations, chromatography, and data processing methods. We employed a quadrupole/orbital ion trap, a time-of-flight (TOF) and 7 different quadrupole/TOF mass spectrometers, operating at mass resolving power from 10000 to 62000 full width at half-maximum (fwhm), mass accuracy better than 1−5 ppm, and a linear dynamic range of 3−4 orders of magnitude. Our results are the first to provide a thorough investigation of this issue on a large-scale multi-instrument study.

EXPERIMENTAL SECTION

Chemicals. LC−MS-grade solvents and mobile phase modifiers were obtained from Fisher Scientific, Waltham, MA (water, acetonitrile, and methanol) and Sigma−Aldrich/Fluka, St. Louis, MO (isopropanol, formic acid, ammonium formate, methyl *tert*-butyl ether, and toluene).

Lipid standards [lysophosphatidylethanolamine (LPE) 17:1, lysophosphatidylcholine (LPC) 17:0, phosphatidylcholine (PC) 12:0/13:0, phosphatidylethanolamine (PE) 17:0/17:0, phosphatidylglycerol (PG) 17:0/17:0, d_7-cholesterol, sphingomyelin (SM) d18:1/17:0, ceramide (Cer) d18:1/17:0, sphingosine (d17:1), monoacylglycerol (MG) 17:0/0:0/0:0, diacylglycerols (DG) 12:0/12:0/0:0 and 18:1/2:0/0:0, and d_5-triacylglycerol (TG) 17:0/17:1/17:0] were obtained from Avanti Polar Lipids (Alabaster, AL) with the exceptions of 12-[[(cyclohexylamino)-carbonyl]amino]-dodecanoic acid (CUDA) (Cayman Chemical, Ann Arbor, MI) and cholesteryl ester (CE) 22:1 (Nu-Chek Prep, Elysian, MN).

These internal standards were selected based on previous analysis of nonspiked plasma lipid extracts and monitoring their signal (MS^1 and MS/MS) at the time of their elution. We used this approach to avoid their coelution with lipid species with the same total number of carbons and saturation but differing in fatty acyl constituents.

Human Plasma Samples. For this study, we used a subset of samples from our recent study focused on nutritional phenotyping in response to a test meal containing gamma-linolenic acid.[21] Briefly, in a single blind, placebo-controlled, crossover design, seven healthy subjects consumed a test meal that consisted of GLA fat (borage oil) or a control fat (a mixture of corn, safflower, sunflower, and extra-virgin light olive oils). Compared to the original study, where all subjects were fed on three separate test days for each test meal, a small modification was needed due to sample limitation. Thus, for this study, six subjects were fed on three separate test days for each test meal, while one subject was fed on two separate test days for a control fat meal and four test days for GLA fat (the fourth set was not used in the original study). Plasma samples collected at 0, 2, and 4 h in response to the test meals were used for analysis. In total, 126 samples were analyzed out of which 42 were baseline samples (time 0 h), 40 were control fat samples (time 2 and 4 h), and 44 were GLA fat samples (time 2 and 4 h).

For quality control, a pool sample consisted of a mixture of nonfasting blood plasma (both control and GLA fat) was used. Also, standard reference material SRM 1950 Metabolites in Frozen Human Plasma (NIST, Gaithersburg, MD) was used.

Sample Preparation. Extraction of plasma lipids was carried out using a biphasic solvent system of cold methanol, methyl *tert*-butyl ether (MTBE), and water[22] with some modifications. In more detail, 300 μL of cold methanol containing a mixture of odd chain and deuterated lipid internal standards [LPE(17:1), LPC(17:0), PC(12:0/13:0), PE(17:0/17:0), PG(17:0/17:0), d_7-cholesterol, SM(d18:1/17:0), Cer(d18:1/17:0), sphingosine (d17:1), DG(12:0/12:0/0:0), DG(18:1/2:0/0:0), and d_5-TG-(17:0/17:1/17:0)] was added to a 40 μL blood plasma aliquot in a 2 mL Eppendorf tube and then vortexed (10 s). Then, 1000 μL of cold MTBE containing CE 22:1 (internal standard) was added, followed by vortexing (10 s) and shaking (6 min) at 4 °C. Phase separation was induced by adding 250 μL of LC−MS-grade water followed by centrifugation at 14000 rpm for 2 min. The concentration of each internal standard can be found in Table S1. Ten aliquots (each 100 μL) of the upper organic phase were collected and evaporated. The volumes of plasma samples and extraction solvents were used to ensure an aliquot for each platform and one backup. For a single platform, the method can be scaled as shown in refs 21 and 23. Dried lipid extracts were resuspended using a methanol/toluene (9:1, *v*/*v*) mixture containing an internal standard CUDA (150 ng/mL), vortexed for (10 s), and centrifuged at 14000 rpm for 2 min prior to LC−MS analysis. The resuspension volume was instrument-dependent (Table S2).

LC−MS Analysis. The LC−MS systems used are listed in Table S2. Each LC system consisted of a pump, a column oven, and an autosampler. Lipids were separated on an Acquity UPLC CSH C18 column (100 × 2.1 mm; 1.7 μm) coupled to an Acquity UPLC CSH C18 VanGuard precolumn (5 × 2.1 mm; 1.7 μm) (Waters, Milford, MA). The column was maintained at 65 °C at a flow-rate of 0.6 mL/min. The mobile phases consisted of (A) 60:40 (v/v) acetonitrile:water with ammonium formate (10 mM) and formic acid (0.1%) and (B) 90:10 (v/v) isopropanol:acetonitrile with ammonium formate (10 mM) and formic acid (0.1%). The separation was conducted under the following gradient: 0 min 15% (B); 0−2 min 30% (B); 2−2.5 min 48% (B); 2.5−11 min 82% (B); 11−11.5 min 99% (B); 11.5−12 min 99% (B); 12−12.1 min 15% (B); and 12.1−15 min 15% (B). The injected volume was instrument-dependent (Table S2). Sample temperature was maintained at 4 °C. Detailed instrumental parameters for each MS system are described in the Supporting Information.

Quality Control. Quality control was assured by (i) randomization of the sequence, (ii) injection of 10 pool samples to equilibrate the LC−MS system before actual sequence of samples; (iii) injection of pool samples at the beginning and the end of the sequence and between each 10 actual samples, (iv)

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368

Table 1. Overview of Performance of LC−MS Platforms Evaluated

platform number	MS platform	unique lipids annotated	lipids with RSD ≤ 30% (QC samples, n = 15)[a]	median RSD for QC samples (n = 15)[a]	median RSD for technical replicates (n = 10)[a]	lipids with VIP ≥ 1 for 2-class PLS-DA model[b]
1	Agilent 6530 (a)	299	279	4.5%	7.3%	42
2	Agilent 6530 (b)	288	271	6.4%	7.1%	39
3	Agilent 6550 iFunnel	287	248	5.5%	5.5%	41
4	Agilent 6560 Ion Mobility	280	264	7.0%	9.9%	42
5	Thermo Q Exactive HF	307	293	6.3%	6.2%	43
6	LECO Citius	209	153	19%	10%	33
7	SCIEX TripleTOF 5600+	297	265	9.2%	9.3%	43
8	SCIEX TripleTOF 6600	306	293	4.1%	5.8%	42
9	SCIEX X500R	302	273	7.1%	7.2%	41

[a]Based on raw data with merged adduct species. [b]Based on concentration data using class-specific internal standard normalization.

injection of NIST SRM 1950 at the beginning of the sequence and after injection of 100 actual samples; (v) procedure blank analysis, (vi) replicate analysis of 10 blood plasma samples (covering both the extraction and LC−MS analysis), (vii) checking the peak shape and the intensity of spiked internal standards and the internal standard added prior to injection, and (viii) monitoring mass accuracy of internal standards during the run.

Data Processing. Raw data files were converted to ABF format using Reifycs Abf (Analysis Base File) Converter (accessible at: http://www.reifycs.com/AbfConverter/). In case of the X500R instrument, the wiff2 files were centroided and exported to mzml format using ProteoWizard (v. 3.0.10827) (accessible at http://proteowizard.sourceforge.net/) followed by converting files to ABF format. Raw data files from LECO Citius were exported to mzml format using ChromaTOF-HRT (v 1.74) followed by converting files to ABF format.

For data processing, MS-DIAL (v. 2.52) software program was used.[24] The following parameters were used: retention time begin, 0.3 min; retention time end, 12.6 min; mass range begin, 280 Da; mass range end, 1500 Da; MS1 (centroiding) tolerance, 0.01 Da; smoothing level, 3 scans; minimum peak height, 500 amplitude (QTOFs), 300 amplitude (TOF), 100000 amplitude (Q Exactive HF), 1 amplitude (X500R); mass slice width, 0.05 Da; retention time tolerance for retention time−m/z (t_R−m/z) library, 0.15 min; accurate mass tolerance, 0.03 Da; retention time tolerance for alignment, 0.1 min; MS1 tolerance for alignment, 0.025 Da.

For lipid identification, accurate mass and MS/MS matching was used with the public LipidBlast library of over 200000 MS/MS spectra.[24,25] In total, 676 t_R−m/z pairs were annotated covering 11 lipid classes and various molecular species: AC, CE, cholesterol, Cer (Cer, HexCer, Hex2Cer), DG, LPC, LPE, PC (PC, pPC/oPC), PE (PE, pPE/oPE), SM, and TG. Quantification was performed by combining data for different detected molecular species for each particular lipid (e.g., sum of $[M + NH_4]^+$, $[M + Na]^+$, $[M + K]^+$ adducts for each TG species) followed by normalization using (i) class-specific internal standards and reported "estimated" concentrations (μM) or (ii) sum of all annotated lipids (total ion chromatogram, TIC). Internal standards used for concentration calculation are listed in Tables S1 and S3. For DG species, DG 12:0/12:0/0:0 was used for quantification because of its elution proximity (t_R ∼ 4.3 min) with all DG species (t_R = 5.2−8.2 min) compared to DG 18:1/2:0/0:0 (t_R ∼ 3.2 min). Sphingosine, MG, and PG species were not detected using current protocol. All internal standards, including DG(18:1/2:0/0:0), MG(17:0/0:0/0:0), PG(17:0/17:0), sphingosine d17:1, and CUDA were used for retention time correction for the t_R−m/z lipid library.

Statistical Analysis and Data Visualization. Multivariate analysis was performed using partial least-squares discriminant analysis (PLS-DA) using MetaboAnalyst.[26] Statistical models were created for both concentration and TIC normalized data after logarithmic transformation (base 10) and Pareto scaling. Exported variable importance in projection (VIP) scores were used for evaluation. For locally weighted scatterplot smoothing (LOESS), MetaBox software was used.[27] For metabolic network mapping annotated lipids were imported into the web-based PubChem chemical structural clustering tool (accessible at: https://pubchem.ncbi.nlm.nih.gov) to generate a pairwise chemical similarity matrix. The matrix and PubChem Compound Identifier (CID) pairs for lipids were used as input in MetaBox software for generation of the Cytoscape network file. A threshold of 0.7 Tanimoto score was used to define the similarity cutoff for lipid structures. The final network graph was imported into Cytoscape 2.8.3 (accessible at: http://www.cytoscape.org). Results from statistics generated in MetaboAnalyst were converted into the Cytoscape node attribute file and imported into Cytoscape. The graph was visualized using a yED organic layout algorithm in Cytoscape.

RESULTS AND DISCUSSION

Study Design. A wide range of factors impact the number and nature of detected lipids in untargeted LC−MS-based lipidomics studies. Sample extraction protocols, chromatographic separations, mobile phase compositions and modifiers, ionization modes, mass spectrometry sensitivity, and linear dynamic range, data processing software, and data normalization prior to statistical analysis are considered as key factors.[3,4] In this study, we investigated the impact different mass spectrometers might have on the outcome and biological interpretation of a typical plasma lipidomic study. In order to isolate this one potential key factor, we ensured that all other factors remained identical, namely samples, sample preparations, chromatography, and data processing methods.

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368

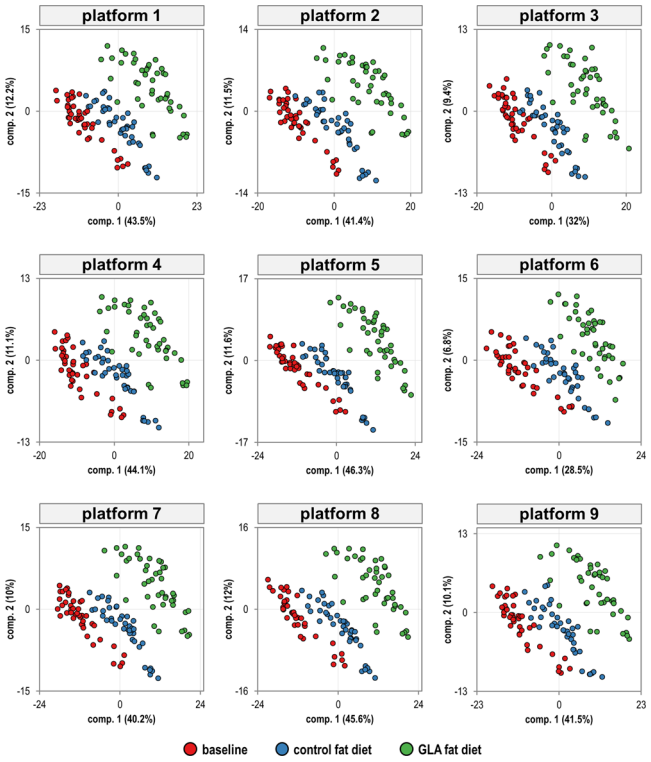


Figure 1. Three-class PLS-DA score plots based on annotated lipids in response to dietary challenges. Prior to statistical analysis, class-specific internal standards were used for normalization followed by calculating estimated concentrations (μM).

We deliberately used samples from a recently published study in order to investigate if statistical and biological conclusions would differ from the published results,[21] depending on the instrumentation used. In a blinded, placebo-controlled, crossover designed study, seven healthy subjects consumed a test meal containing high amounts of gamma-linolenic acid (GLA, 18:3n6) compared to a control meal. Each subject underwent the nutritional test on three separate test days for each test meal, and samples were taken each time over an 8 h period. As a result, this study demonstrated that five of seven subjects had enzymatic capacities to elongate GLA (18:3n6) into dihomo-gamma-linolenic acid (DGLA, 20:3n6), while two subjects did not show such conversions.[21]

All samples were extracted as in the published study before, using methyl *tert*-butyl ether (MTBE)/methanol[21] that has clear advantages over classic chloroform/methanol-based extraction protocols. Because of the low density of the lipid-containing organic phase that forms the upper layer during phase separation, its collection is greatly simplified, and all main lipid classes in plasma are extracted with high recovery.[22] Furthermore, MTBE is nontoxic and noncarcinogenic compared to chloroform. All samples were analyzed using identical chromatographic separation. Specifically, we used reversed-phase ultrahigh-performance liquid chromatography–mass spectrometry (UHPLC–MS) which is the most widely used method, accounting for about 70% of all reported LC–MS lipidomic studies.[4] We employed a short microbore column (100 × 2.1 mm id) with 1.7 μm particle size with C18 as lipophilic sorbent, representing the currently preferred method in LC–MS-based lipidomics.[4] Specifically, we used an Acquity UPLC charged surface hybrid (CSH) C18 column, incorporating a low level surface charge leading to improved peak symmetry. We also used frequently employed LC elution mobile phases, acetonitrile:water (60:40 v/v) and isopropanol:acetonitrile (90:10 v/v). We used a buffer of 10 mM ammonium formate with 0.1% formic acid in ESI(+) that provided the best overall scores over four other commonly used mobile-phase modifier systems.[28] In order to ensure that potential differences in data processing software would not obfuscate statistical or biological results, we used the open source software MS-DIAL[24] for all acquired mass spectrometry data. In addition, we made sure to have a sufficiently high number of samples to enable robust conclusions, using a sequence consisting of total 155 injections per mass spectrometer, including 126 actual study samples, 15 quality control "pool" samples, 2 blanks, 10 technical plasma replicates, and 2 NIST SRM 1950 plasma samples, thus, simulating a large-scale lipidomics study.

In order to focus the comparison of instruments to the most relevant data acquisition, we limited data acquisitions to positive mode electrospray (ESI) because ESI(+) provided more annotated lipids than ESI(−) and also showed the most number of altered lipids in the published gamma-linolenic dietary study.[21] We evaluated nine modern, high-resolution mass spectrometers, comprising 1 single TOF, 1 Q/orbital ion trap, and 7 QTOF instruments (Table 1) operating at a mass resolving power between 10000 and 62000 fwhm.

Initial Evaluation of LC–MS Systems. By nominal performance parameters, each mass spectrometer was expected to be different in sensitivity and linear dynamic range. We performed an initial evaluation of each instrument to avoid saturating the ion source and the detector even for highly abundant, coeluting lipids. To this end, we prepared a pool sample from nonfasting blood plasma extracts and determined optimal dilution and injection volumes for each instrument using the most abundant lipid species, PC(34:2) ($[M + H]^+$, m/z 758.5670) and TG(52:4) ($[M + NH_4]^+$, m/z 872.7650). At the same time, we checked how dilution of the pool sample influenced the detection of lower abundant lipids such as LPC(18:1) ($[M + H]^+$, m/z 522.3554) and TG(54:8) ($[M + NH_4]^+$, m/z 892.7389). Resuspension and injection volumes were considered optimal for linear correlations $R^2 > 0.98$ for both high abundant and low abundant lipids. Simulating samples with lipid concentrations beyond the optimal dilution, we found that 2-fold increase in lipid concentration slightly worsened R^2 (0.95–0.97) for high abundant lipid species on some instruments. For low abundant lipid species, $R^2 > 0.97$ was observed within the optimal working range for all instruments. Overall, instruments had a linear dynamic range from 3 to 4 orders of magnitude. Investigating the linearity of response for multiple analytes represents a crucial factor for the validity of LC–MS-based untargeted workflows, especially, when peak intensities are used to estimate lipid concentrations.

Despite using identical chromatographic methods on all nine LC–MS platforms, slight differences in separation were observed, likely due to different lengths and internal diameters of the tubing used in each LC unit. Figures S1 and S2 show examples of separation of two PC(36:3) isomers and three TG(54:6) isomers, respectively, in blood plasma extracts. Insufficient separation of isomers also poses challenges to data processing, specifically during peak picking, recognizing peak apexes and valleys, and aligning peak across all chromatograms. Overall, however, chromatography results were sufficiently similar between the nine LC–MS platforms that bias in reporting peak heights was largely avoided.

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360–12368

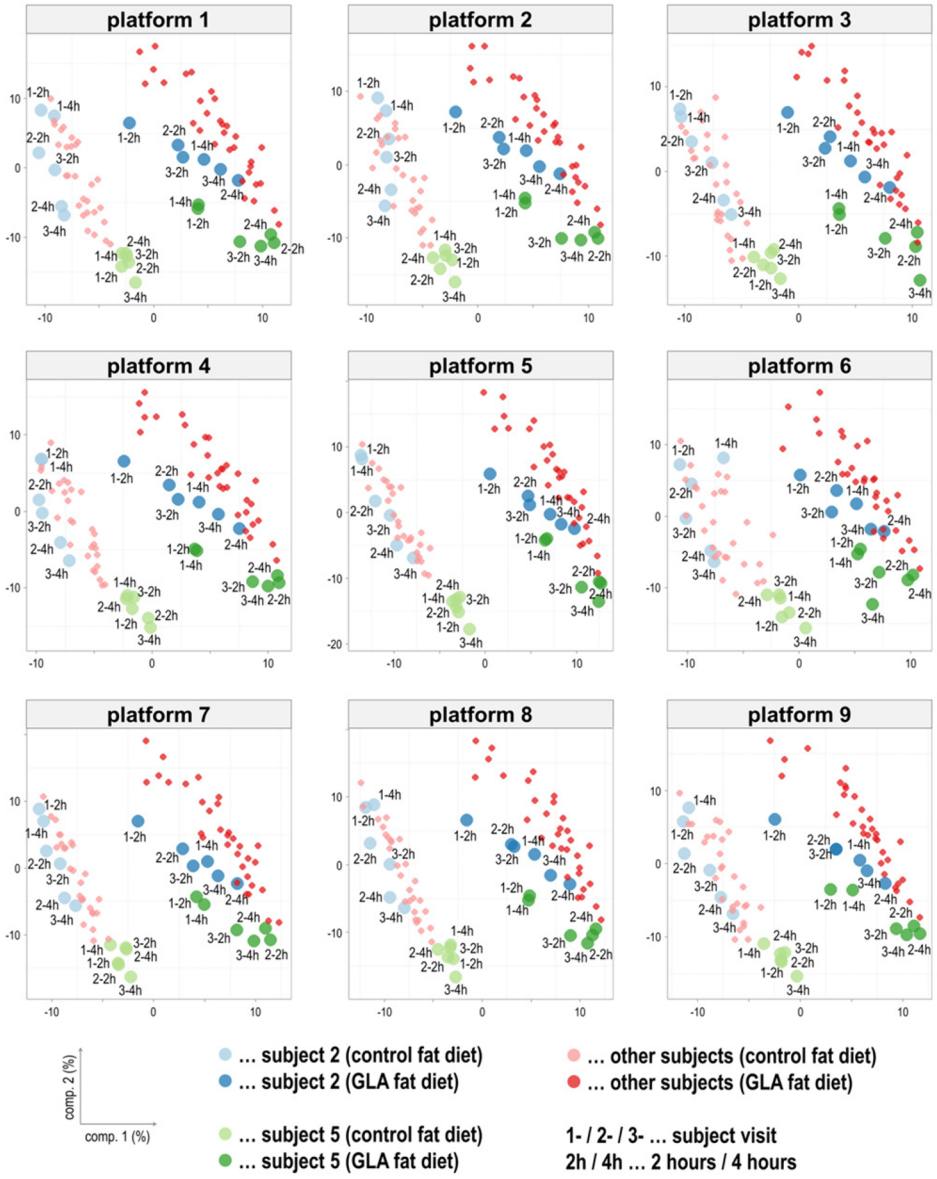


Figure 2. Two-class PLS-DA score plots based on annotated lipids in response to dietary challenges. Prior to statistical analysis, class-specific internal standards were used for normalization followed by calculating "estimated" concentrations (μM). Detailed individual scores are shown for subjects 2 and 5 as an example indicating that individual lipid profiles were highly consistent between the LC−MS platforms.

Statistical Comparison of Nine LC−MS Systems in a Large-Scale Lipidomics Study. For each of the nine LC−MS platforms, a series of 155 injections was conducted including actual samples, QC samples, technical replicates, blank samples, and NIST SRM 1950 plasma samples. After data processing in MS-DIAL, we noted changes in adduct ratios, mainly for di- and triacylglycerols (DG, TG), between the instruments and depending on lipid concentrations. Figure S3 shows box plots for the adduct ratios $[M + Na]^+/[M+NH_4]^+$ and $[M+K]^+/[M+NH_4]^+$ for highly abundant TG(54:4), medium abundant TG(54:2), and low abundant TG(54:1) triacylglycerols in blood plasma samples. Alkali adduct ion ratios increased with decreasing lipid concentrations to the point that for some triacylglycerols, $[M + Na]^+$ and $[M+K]^+$ were observed as the dominating species compared to $[M + NH_4]^+$ adducts. On this account, we merged all three types of adducts for each lipid species to reduce data complexity prior to statistical analysis and also to improve the signal intensity for low abundant species. Subsequently, raw data were converted to estimated absolute lipid concentrations by a single-point calibration using one internal standard for each lipid class (Tables S1 and S3).

Reporting molar concentrations is the preferred way to harmonize lipidomic reports across studies.[16] It would be difficult to validate using internal standards to estimate concentrations for structurally unidentified lipid signals. Furthermore, biological conclusions are best based on identified compounds, not on unknown t_R−m/z signals. Therefore, we evaluated the statistical performance of the nine LC−MS platforms based on annotated lipids and their reported concentrations. We also used an alternative data normalization approach, using the sum of the peak intensities (TIC) of all annotated lipids for each sample, representing data as fraction intensity of the total observed lipidome. In overall, between 209 and 307 unique lipid species (not including internal standards) were annotated on all LC−MS platforms (Table 1). Using the quality control samples, between 153 and 293 lipids showed acceptable reproducibility defined as

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368

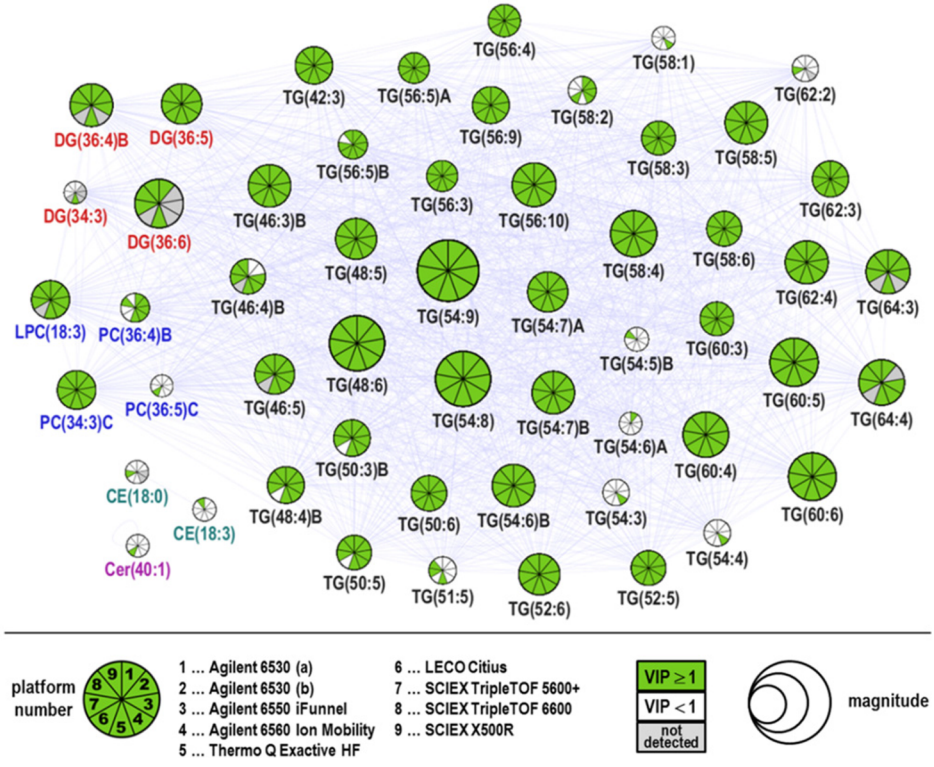


Figure 3. Metabolic network analysis of the postprandial lipid response to dietary challenges based on two-class PLS-DA model using estimated concentrations. The network connects lipid nodes if lipids share high chemical similarity using the PubChem substructure fingerprint Tanimoto scores. Nodes represent lipids which occurred at least once with VIP ≥ 1 over all instruments evaluated. Node sizes scale with average magnitude of VIP values over all instruments. Green nodes represent lipids with VIP ≥ 1, white nodes represent lipids with VIP < 1, gray nodes are used for lipids that were not annotated/detected in particular LC−MS system. Position of each instrument within the node is based on a pie chart.

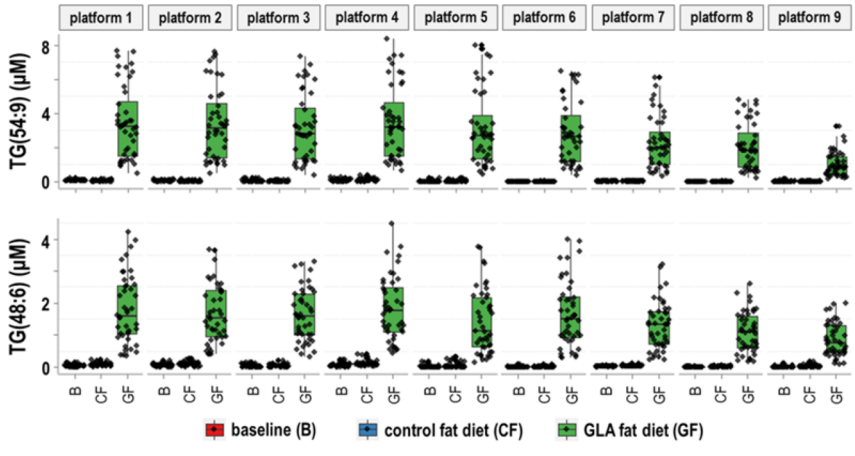


Figure 4. Estimated concentrations by single-point calibrations across nine instruments. For TG(54:9) and TG(48:6), very similar concentration values are obtained for all mass spectrometers, showing a large increase in concentration in response to high gamma-linoleic acid.

relative standard deviations RSD ≤ 30%, with median RSD ranging from 4.1−19%. Technical replicates showed median RSD ranging from 5.5−10% for all platforms.

Partial least-squares discriminant analysis (PLS-DA) was conducted to find differences between plasma lipidomes at baseline and after nutritional intervention with a test meal high in gamma-linolenic acid fat in comparison to the control fat test meal. Figure 1 shows PLS-DA score plots for the three-class model based on annotated lipids in response to dietary challenges using concentration-based data for all LC−MS platforms evaluated. Strikingly, highly similar patterns were observed for all LC−MS platforms. The score plot indicates that "baseline" (0 h) and control fat groups (2 and 4 h) are more closely related than "baseline" and GLA fat groups, which can be explained by their similar fatty acid composition originating from dietary sources rich in oleic acid (18:1). On the other hand, the separation of GLA fat group reflects the differences in plasma lipid profiles in response to the gamma-linolenic acid rich test meal (18:3). Almost identical patterns were also observed for TIC normalized data (Figure S4), showing that indeed TIC normalizations are a valid representation of data.

Next, we tested if all nine LC−MS instruments correctly described the extent and the biochemical profile differences between the seven subjects induced by dietary challenges. To this

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368

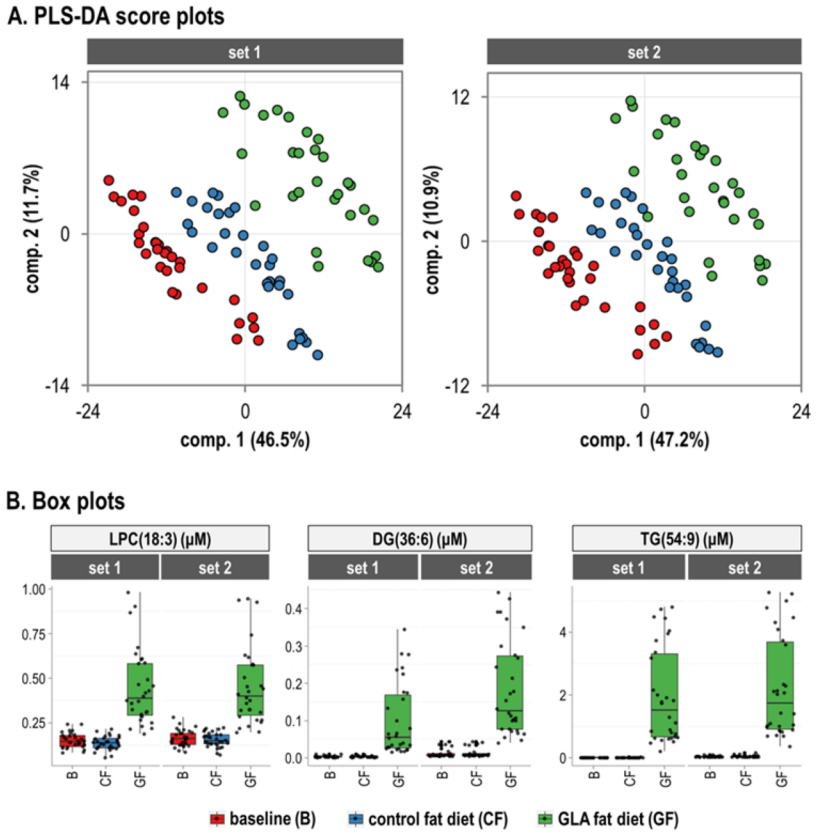


Figure 5. (a) Three-class PLS-DA score plots based on concentrations of annotated lipids in response to dietary challenges acquired on the same LC−MS platform over a 1 year time period (set 1 vs set 2). (b) Estimated concentrations by single-point calibrations between data sets acquired on the same MS system over a 1 year period (set 1 vs set 2). Examples are given for the most discriminating lipid species among different lipid classes LPC(18:3), DG(36:6, 18:3_18:3), and TG(54:9, 18:3_18:3_18:3) showing reproducible quantitative data over a long time period.

end, we focused on the separation of postprandial samples collected from subjects in response to the control and GF test meals. Using a two-class PLS-DA model, lipid profiles showed clear separation of control fat and GLA fat groups (Figures S5 and S6). Importantly, each subject and their postprandial lipidomic response curve could be recognized based on the position on PLS-DA score plots which indicates that individual lipid profiles were highly consistent between the LC−MS platforms (Figure 2).

The use of variable importance of the projection (VIP) of PLS-DA allowed us to identify lipids that most contributed to the separation of plasma samples based on test meals. In order to visualize which lipids were identified as important (VIP ≥ 1), we created a metabolic network that provided a clear picture of each LC−MS platform (Figure 3). There has been a 75% overlap for the 54 most discriminating lipids based on VIP scores higher than 1, independent of the mass spectrometer used. This result showed that the correct biochemical signature was recognized by any mass spectrometer used. All instruments discovered that TG and DG containing 18:3 fatty acid chains were more concentrated in plasma in response to GF test meal. A few lipids were reported with VIP ≥ 1 only on a single platform. When these lipids were not considered, an overall 92% overlap for the 43 discriminating lipids was achieved in the multivariate statistical analysis. For TIC normalized data, a 76% overlap for the 56 most discriminating lipids based on VIP ≥ 1 was observed, with an increase to an overlap of 87% of 48 discriminating lipids that were reported at VIP ≥ 1 for at least two instruments. When comparing TIC normalized to internal-standards normalized data, 96% identical lipids were found for the most discriminating lipids at VIP ≥ 1. Thus, when correct adjustments are made for instrumentation, biological studies can be effectively replicated across different high-resolution MS platforms based on PLS-DA statistical analysis.

Last, we focused on how reported lipid concentrations were consistent among all platforms. Figure 4 shows examples for two most discriminating lipid species, TG(54:9) and TG(48:6). These two triglycerides became highly enriched after gamma-linolenic rich test meals. Estimated concentrations calculated by single-point calibrations provided comparable results between all instruments, with averages ranging from 0.8−1.6 μM for TG(48:6) and 0.9−3.1 μM for TG(54:9). Further investigation showed that reported concentrations can be influenced by the signal intensity of particular lipids as shown in Figure S7. While for low- and medium-abundant TG(54:1) and TG(54:2), respectively, comparable results were obtained between all nine LC−MS instruments, the highly abundant TG(54:4) showed less consistent results between the mass spectrometry platforms. This effect is likely due to nonlinearity of calibrations at high peak intensities that cannot be correctly calculated by single-point calibrations.

In principle, harmonizing metabolomic or lipidomic reports can also be achieved by reporting all peaks in relation to a community quality control, for example, the NIST SRM 1950 human plasma pool. Since identical quality control and NIST SRM 1950 plasma samples were injected on each instrument, we attempted to use these control samples for further improving the normalization of absolute concentration estimates. In Figure S8,

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368

we show the highly abundant TG(54:4) as an example for the effects of different normalization schemes. First, we used locally weighted scatterplot smoothing (LOESS) normalization[27] using the quality control samples that were injected after each set of 10 actual cohort samples. Subsequently, we tested using the NIST SRM 1950 plasma samples injected at the beginning and after 100 actual cohort samples. For NIST SRM 1950 plasma samples, we defined a single normalization factor for each lipid per LC−MS platform. We used an average ($n = 2$) of this reference sample run on each platform and calculated the fold change for each lipid relative to the reference sample value for a reference laboratory. Thus, the normalized concentration c_{ij} for a lipid i, for a platform j was given by

$$\text{normalized } c_{ij} = c_{ij}\frac{\text{mean value (lipid } i\text{) in ref material from ref lab}}{\text{mean value (lipid } i\text{) in ref material from instrument } j}$$

where i is measured lipid and j is the LC−MS platform. For the reference material, we used reference values from ref 29.

Although using QC and NIST SRM 1950 samples for normalization further corrected for instrumental platform effects, insufficient lipid coverage of reference values for NIST SRM 1950 did not permit full utilization of this approach for all lipid species found in biological samples. For instance, reference values were reported for only 18 TG species as isobaric molecular subsets,[29] while our LC−MS lipidomics method permitted annotation of 101 isobaric molecular species (113 lipid species including isomer differentiation). Also, NIST SRM 1950 represents a pooled human plasma obtained from healthy individuals after overnight fasting. Thus, some lipid species occurring due to the nutritional intervention were missing in this reference material.

Reproducibility of Quantitative Data in LC−MS Large-Scale Lipidomics Study. For the TripleTOF 6600 LC−MS platform, we performed a between-series reproducibility test. Ninety samples (for five out of seven subjects) were analyzed again, one year after the original data acquisition. We observed a 97% overlap of annotations of detected lipids between these two data sets. Further, supervised multivariate PLS-DA score plots showed nearly identical quantitative patterns (Figure 5a), and in addition, single-point calibrations also yielded highly similar quantitative data for detected lipids in univariate analyses (Figure 5b). Overall, these data demonstrate very good reproducibility of plasma lipidomics using this protocol.

CONCLUSIONS

Untargeted LC−MS-based lipidomics analysis can yield nearly identical results even when different mass spectrometers are used. As shown by detailed PLS-DA investigations, individual biochemical profiles of subjects after postprandial meal responses are accurately determined independently which type of high-resolution mass spectrometer was used. More importantly, the same top-hit discriminating lipid species were found by all instruments. Estimated concentrations calculated using single-point calibrations provided highly comparable results between LC−MS platforms tested, paving the way toward robust and repeatable reporting of untargeted lipidomic results within and across human plasma studies. Key to harmonization of results is to avoid saturation of both the instrument's ion source and the detector by carefully evaluating linear dynamic ranges using serial dilutions and adjusting the resuspension volume and/or injection volume before running actual study samples.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.7b03404.

Table S1: Concentrations of the internal standards; Table S2: LC−MS systems used for the large-scale multi-instrument study; Table S3: List of annotated lipid species; Section S1: MS detection parameters; Figures S1 and S2: Examples of lipid isomer separation; Figure S3: Box plots for adduct ratios $[M + Na]^+/[M + NH_4]^+$ and $[M + K]^+/[M + NH_4]^+$; Figure S4: Three-class PLS-DA score plots based on annotated lipids in response to dietary challenges (TIC normalization); Figures S5 and S6: Two-class PLS-DA score plots based on annotated lipids in response to dietary challenges (concentration and TIC normalization); Figure S7: Box plots based on quantification results for low, medium, and high abundant TG; and Figure S8: Box plots showing concentrations of TG(54:4) using various normalization strategies (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: ofiehn@ucdavis.edu. Tel: +1-530-754-8258.

ORCID

Tomas Cajka: 0000-0002-9728-3355

Oliver Fiehn: 0000-0002-6261-8928

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was supported by the U.S. National Institutes of Health (NIH) Grants P20 HL113452 and U24 DK097154. Authors thank MS vendors for providing the instruments for evaluation purposes (Agilent 6560 IM QTOF, SCIEX X500R QTOF, Leco Citius TOF).

REFERENCES

(1) Rochat, B. *TrAC, Trends Anal. Chem.* **2016**, *84*, 151−164.

(2) Gika, H. G.; Theodoridis, G. A.; Plumb, R. S.; Wilson, I. D. *J. Pharm. Biomed. Anal.* **2014**, *87*, 12−25.

(3) Hyotylainen, T.; Oresic, M. *Anal. Bioanal. Chem.* **2015**, *407*, 4973−4993.

(4) Cajka, T.; Fiehn, O. *TrAC, Trends Anal. Chem.* **2014**, *61*, 192−206.

(5) Cajka, T.; Fiehn, O. *Anal. Chem.* **2016**, *88*, 524−545.

(6) Sud, M.; Fahy, E.; Cotter, D.; Azam, K.; Vadivelu, I.; Burant, C.; Edison, A.; Fiehn, O.; Higashi, R.; Nair, K. S.; Sumner, S.; Subramaniam, S. *Nucleic Acids Res.* **2016**, *44*, D463−D470.

(7) Martin, J. C.; Maillot, M.; Mazerolles, G.; Verdu, A.; Lyan, B.; Migne, C.; Defoort, C.; Canlet, C.; Junot, C.; Guillou, C.; Manach, C.; Jabob, D.; Bouveresse, D. J. R.; Paris, E.; Pujos-Guillot, E.; Jourdan, F.; Giacomoni, F.; Courant, F.; Fave, G.; Le Gall, G.; Chassaigne, H.; Tabet, J. C.; Martin, J. F.; Antignac, J. P.; Shintu, L.; Defernez, M.; Philo, M.; Alexandre-Gouaubau, M. C.; Amiot-Carlin, M. J.; Bossis, M.; Triba, M. N.; Stojilkovic, N.; Banzet, N.; Molinie, R.; Bott, R.; Goulitquer, S.; Caldarelli, S.; Rutledge, D. N. *Metabolomics* **2015**, *11*, 807−821.

(8) Benton, H. P.; Want, E.; Keun, H. C.; Amberg, A.; Plumb, R. S.; Goldfain-Blanc, F.; Walther, B.; Reily, M. D.; Lindon, J. C.; Holmes, E.; Nicholson, J. K.; Ebbels, T. M. D. *Anal. Chem.* **2012**, *84*, 2424−2432.

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368

(9) Djekic, D.; Pinto, R.; Vorkas, P. A.; Henein, M. Y. *Int. J. Cardiol.* **2016**, *222*, 1042−1048.
(10) Glauser, G.; Veyrat, N.; Rochat, B.; Wolfender, J. L.; Turlings, T. C. J. *J. Chromatogr A* **2013**, *1292*, 151−159.
(11) Cheema, A. K.; Asara, J. M.; Wang, Y.; Neubert, T. A.; Tolstikov, V.; Turck, C. W. *J. Biomol. Technol.* **2015**, *26*, 83−89.
(12) Gika, H. G.; Theodoridis, G. A.; Earll, M.; Snyder, R. W.; Sumner, S. J.; Wilson, I. D. *Anal. Chem.* **2010**, *82*, 8226−8234.
(13) Siskos, A. P.; Jain, P.; Romisch-Margl, W.; Bennett, M.; Achaintre, D.; Asad, Y.; Marney, L.; Richardson, L.; Koulman, A.; Griffin, J. L.; Raynaud, F.; Scalbert, A.; Adamski, J.; Prehn, C.; Keun, H. C. *Anal. Chem.* **2017**, *89*, 656−665.
(14) Pham, H. T.; Arnhard, K.; Asad, Y. J.; Deng, L.; Felder, T. K.; St. John-Williams, L.; Kaever, V.; Leadley, M.; Mitro, N.; Muccio, S.; Prehn, C.; Rauh, M.; Rolle-Kampczyk, U.; Thompson, J. W.; Uhl, O.; Ulaszewska, M.; Vogeser, M.; Wishart, D. S.; Koal, T. *J. Appl. Lab. Med.* **2016**, *1*, 129−142.
(15) Klavins, K.; Neubauer, S.; Al Chalabi, A.; Sonntag, D.; Haberhauer-Troyer, C.; Russmayer, H.; Sauer, M.; Mattanovich, D.; Hann, S.; Koellensperger, G. *Anal. Bioanal. Chem.* **2013**, *405*, 5159−5169.
(16) Liebisch, G.; Ekroos, K.; Hermansson, M.; Ejsing, C. S. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **2017**, *1862*, 747−751.
(17) Wang, M.; Wang, C.; Han, X. *Mass Spectrom. Rev.* **2017**, *36*, 693−714.
(18) Koivusalo, M.; Haimi, P.; Heikinheimo, L.; Kostiainen, R.; Somerharju, P. *J. Lipid Res.* **2001**, *42*, 663−672.
(19) Koelmel, J. P.; Ulmer, C. Z.; Jones, C. M.; Yost, R. A.; Bowden, J. A. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **2017**, *1862*, 766−770.
(20) Bowden, J. A.; Heckert, A.; Ulmer, C. Z.; Jones, C. M.; Koelmel, J. P.; Abdullah, L.; Ahonen, L.; Alnouti, Y.; Armando, A.; Asara, J. M.; Bamba, T.; Barr, J. R.; Bergquist, J.; Borchers, C. H.; Brandsma, J.; Breitkopf, S. B.; Cajka, T.; Cazenave-Gassiot, A.; Checa, A.; Cinel, M. A.; Colas, R. A.; Cremers, S.; Dennis, E. A.; Evans, J. E.; Fauland, A.; Fiehn, O.; Gardner, M. S.; Garrett, T. J.; Gotlinger, K. H.; Han, J.; Huang, Y.; Neo, A. H.; Hyotylainen, T.; Izumi, Y.; Jiang, H.; Jiang, H.; Jiang, J.; Kachman, M.; Kiyonami, R.; Klavins, K.; Klose, C.; Kofeler, H. C.; Kolmert, J.; Koal, T.; Koster, G.; Kuklenyik, Z.; Kurland, I. J.; Leadley, M.; Lin, K.; Maddipati, K. R.; McDougall, D.; Meikle, P. J.; Mellett, N. A.; Monnin, C.; Moseley, M. A.; Nandakumar, R.; Oresic, M.; Patterson, R. E.; Peake, D.; Pierce, J. S.; Post, M.; Postle, A. D.; Pugh, R.; Qui, Y.; Quehenberger, O.; Ramrup, P.; Rees, J.; Rembiesa, B.; Reynaud, D.; Roth, M. R.; Sales, S.; Schuhmann, K.; Schwartzman, M. L.; Serhan, C. N.; Shevchenko, A.; Somerville, S. E.; John-Williams, L. S.; Surma, M. A.; Takeda, H.; Thakare, R.; Thompson, J. W.; Torta, F.; Triebl, A.; Trotzmuller, M.; Ubhayasekera, S. J. K.; Vuckovic, D.; Weir, J. M.; Welti, R.; Wenk, M. R.; Wheelock, C. E.; Yao, L.; Yuan, M.; Zhao, X. H.; Zhou, S. *J. Lip. Res.* 2017, in press, 10.1194/jlr.M079012.
(21) Cajka, T.; Davis, R.; Austin, K. J.; Newman, J. W.; German, J. B.; Fiehn, O.; Smilowitz, J. T. *Metabolomics* **2016**, *12*, 1−16.
(22) Matyash, V.; Liebisch, G.; Kurzchalia, T. V.; Shevchenko, A.; Schwudke, D. *J. Lipid Res.* **2008**, *49*, 1137−1146.
(23) Cajka, T.; Fiehn, O. *Methods Mol. Biol.* **2017**, *1609*, 149−170.
(24) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* **2015**, *12*, 523−526.
(25) Kind, T.; Liu, K. H.; Lee, D. Y.; DeFelice, B.; Meissen, J. K.; Fiehn, O. *Nat. Methods* **2013**, *10*, 755−758.
(26) Xia, J. G.; Sinelnikov, I. V.; Han, B.; Wishart, D. S. *Nucleic Acids Res.* **2015**, *43*, W251−W257.
(27) Wanichthanarak, K.; Fan, S.; Grapov, D.; Barupal, D. K.; Fiehn, O. *Plos One* **2017**, *12*.
(28) Cajka, T.; Fiehn, O. *Metabolomics* **2016**, *12*, article 34, DOI: 10.1007/s11306-015-0929-x.
(29) Quehenberger, O.; Armando, A. M.; Brown, A. H.; Milne, S. B.; Myers, D. S.; Merrill, A. H.; Bandyopadhyay, S.; Jones, K. N.; Kelly, S.; Shaner, R. L.; Sullards, C. M.; Wang, E.; Murphy, R. C.; Barkley, R. M.; Leiker, T. J.; Raetz, C. R. H.; Guan, Z. Q.; Laird, G. M.; Six, D. A.; Russell, D. W.; McDonald, J. G.; Subramaniam, S.; Fahy, E.; Dennis, E. A. *J. Lipid Res.* **2010**, *51*, 3299−3305.

DOI: 10.1021/acs.analchem.7b03404
Anal. Chem. 2017, 89, 12360−12368