This is the published version of a paper published in *TrAC. Trends in analytical chemistry*. 

Citation for the original published paper (version of record):

Lipidomes in health and disease: Analytical strategies and considerations
*TrAC. Trends in analytical chemistry*, 120: 115664
https://doi.org/10.1016/j.trac.2019.115664

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-77194
Lipidomes in health and disease: Analytical strategies and considerations

Fang Wei a, Santosh Lamichhane b, Matej Orešić a,b,c,*, 1, Tuulia Hyötyläinen d, **, 1

a Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, Wuhan, Hubei, 430062, PR China
b Turku Bioscience Centre, University of Turku and Åbo Akademi University, 20520 Turku, Finland
c School of Medical Sciences, Örebro University, 702 81 Örebro, Sweden
d School of Science and Technology, Örebro University, 702 81 Örebro, Sweden

1 Equal contribution.

* Corresponding author. Turku Bioscience, University of Turku and Åbo Akademi University, Tykitokatu 6, FI-20520, Turku, Finland.
** Corresponding author. E-mail addresses: matej.oresic@utu.fi (M. Orešić), tuulia.hyotylainen@oru.se (T. Hyötyläinen).

1. Introduction

Lipids have many essential biological functions such as in cellular signaling and energy storage, as well as by being integral parts of cellular membranes and lipid particles such as lipoproteins and exosomes. Given these important roles in core physiological processes, lipid levels and their production are under tight, homeostatic control. It is thus not surprising that dysregulation of lipid metabolism accompanies the development of many common diseases [1].

Since small changes in concentrations of lipids may have important physiological consequences, the field of lipidomics has need of accurate quantification methods. Nevertheless, currently, the most commonly-used approaches for the characterization of molecular lipids in biological matrices, such as those based on mass spectrometry (MS), still have several limitations in this regard [2]. It should also be noted that the term ‘quantitative’ analysis is often used in an ambiguous manner. Here, we consider an analysis as being accurately ‘quantitative’ in cases where the quantitation is done with authentic standards in combination with the use of isotope-labelled internal standards (ISs).

Another challenge in lipidomic analysis is the identification of lipids. In order to precisely link the assayed lipids with specific biochemical pathways, precise molecular identification is needed, including the determination of, e.g., the exact fatty acyl composition of structural lipids (position on glycerol backbone, double bond positions). Currently, the latter is, in practice, not possible in global lipidomics experiments.

Given a lack of available ‘gold standards’ for accurate, quantitative measurements of all lipids typically covered by lipidomic analyses, with authentic standards being available only for a small subset of lipids, comparison of data from different analytical platforms and laboratories remains a challenge. Community-wide efforts are ongoing, aiming to improve the inter-comparability of lipidomics data. A recent ring study, where the National Institutes of Standards (NIST) plasma reference material SRM-1950 was analyzed by 31 different laboratories, revealed much about the extent of variation of lipidomics data, acquired across different lipidomics platforms [3]. The Lipidomics Standards Initiative (LSI) aims to standardize the reporting of lipidomics data and analyses, and the terminology used [4]. Community-initiated guidelines for
best practices in blood-based lipidomic analyses were also pro-
posed [5]. As the next step in this effort, reference plasma values for
selected lipids are being developed, which will, in time, enable
better assessment of the accuracy of lipid quantification.

In addition to analytical variation, lipid levels are also affected
by many biological factors (Fig. 1). While factors such as diet and
age have commonly been considered when studying lipidomic
profiles, recent data also suggest that several other factors, e.g.,
environmental exposure to chemical pollutants, have a marked
impact on plasma lipid levels [6]. In order to truly understand the
variability of lipids in health and disease, and make data compa-
rable across different studies and populations, it is also essential to
systematically assess the impacts of different internal and external
exposures on lipid levels.

In this review, we discuss recent progress in lipidomic analyses,
with primary emphasis on analytical approaches, as well as on the
different sources of variation affecting the lipid levels, both tech-
nical and biological.

2. Analytical approaches to study lipidomes

The analytical approaches in lipidomics cover the workflow
from sample preparation to data preprocessing, lipid identifi-
cation, and data analysis. Different analytical methodologies have
been developed for lipidomics [7–12]. Below we summarize the key
strategies as applied for the analyses of global lipidomics or specific
lipid classes.

2.1. Analysis of lipids

The approaches for analysis of lipids can be divided into (i)
global lipidomic analyses, covering the major classes of complex
lipids, e.g., triacylglycerols (TGs), glycerophospholipids (GPLs), and
sphingolipids (SLs), and (ii) targeted (quantitative) analyses of
specific lipid classes, which are not well-covered by global meth-
odologies due to their low concentrations or specific physico-
chemical properties. These include, e.g., bile acids, oxylipids, and
endocannabinoids. A recent trend in metabolomics, including lip-
idomics, is to use various combinations of targeted analyses with
global, untargeted approaches [13,14].

Typically, sample preparation for the two analytical approaches
is different. For global lipidomic analyses, liquid-liquid extraction
is commonly used without any further clean-up steps (Table 1), while
for targeted methods, many methods rely on solid-phase extraction
(SPE) or phospholipid removal as a clean-up step after protein
precipitation (Table 2). Several studies have been performed,
comparing different extraction methods, with varying results. Most
of the methods suggested give good recoveries for the most pre-
dominant lipid classes (triacylglycerols, cholesterol esters, and
phosphatidylcholines), however, the extraction of less-abundant
lipids (e.g., phosphatidylinositols, various lysophospholipids,
ceramides, and cholesterol sulfates) is more dependent on the
solvent system used for the extraction [15]. Generally, both the
Folch and Matyash methods have been shown to give robust results
[15,16], while the one-phase extraction method also seem to be a
promising alternative [17].

In most of the global methods, quantitation is performed by
using lipid class-specific internal standards, and these are typically
done either by direct infusion mass spectrometry (DIMS, often
referred to as ‘shotgun lipidomics’), by coupling MS with liquid
chromatography (LC), or for selected lipid classes, to gas chroma-
tography (GC). When reporting lipidomics analyses, the type of
quantitation should be defined, as also recommended by LSI [4].
The quantitation should include (i) normalization, either with

Fig. 1. Overview of various selected factors influencing the lipidomes. Blue lines denote negative associations, red positive associations. These associations are indicative only, based on the main trends as reported in the literature (see Table 4 for more detail and selected references). Abbreviations: BA, bile acid; BPA, bisphenol A; Cer, ceramide; Chol, cholesterol ester; CNS, central nervous system; COPD, chronic obstructive pulmonary disease; DG, diacylglycerol; EC, endocannabinoids; LPC, lysophosphatidylcholine; NAFLD, non-alcoholic fatty liver disease; OL, oxylipids; PC, phosphatidylcholine; PCB, polychlorinated biphenyl; PFAS, per- and polyfluoroalkyl substances; PUFAs, polyunsaturated fatty acid; SM, sphingomyelin; TG, triacylglycerol; T1D, type 1 diabetes; T2D, type 2 diabetes.
Matrix matched calibration however is typically not possible in several other factors, including consideration of matrix effects. Confusion with the accuracy of quantitation, which depends on specific calibration curves of authentic standards (or with lipid class-normal phase separation), or with (level 3) lipid class-specific comprehensive profiling approaches. Some other chromatographic methods, such as multidimensional LC [18,19] and supercritical fluid chromatography [20] have also been suggested for lipidomic analyses. They have, however, not yet been widely adopted, mainly because of the more complex systems that are not readily applied in routine analysis of large sample sets. A majority of the targeted methods for specific lipid classes rely on ultra-high-performance LC coupled to triple-quadrupole tandem MS (UHPLC-QqQ-MS/MS), but recently, global lipidomics. The term ‘absolute quantification’ has sometimes been used in the lipidomics community, however, this is not analytically appropriate terminology when referring to level 2/3 quantitation.

### Table 1
Comparison of methods for global lipidomic analysis.

<table>
<thead>
<tr>
<th>Method</th>
<th>Comments</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXTRACTION METHODS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified Folch extraction</td>
<td>chloroform/methanol/water (2/2/1.8, v/v/v) biphasic solvent system</td>
<td>Efficient extraction of the main lipid classes, e.g., most phospholipids, sphingolipids, DGS and TGs, cholesterol and CE s.</td>
<td>Recoveries of polar lipids is less efficient, extract is in lower layer, making automatization less convenient. Carry-over of water causes increased ion suppression and adduct formation due to the co-extraction of salts and other polar metabolites.</td>
</tr>
<tr>
<td>MTBE:MeOH extraction</td>
<td>MTBE/methanol/water</td>
<td>Efficient extraction of the main classes of lipids, extraction phase is in the upper layer, good recovery also for more polar lipids.</td>
<td>Requires overnight incubation for sufficient removal of proteins, shorter incubations may result in higher analytical variation.</td>
</tr>
<tr>
<td>IPA</td>
<td>Protein precipitation, single solvent</td>
<td>Good recoveries for most abundant lipid classes.</td>
<td>Not yet widely used, not much data on the robustness of the methodology.</td>
</tr>
<tr>
<td>MeOH/MTBE/CHCl3</td>
<td>One-phase extraction</td>
<td>Good recoveries for most abundant lipid classes, simple extraction.</td>
<td></td>
</tr>
<tr>
<td><strong>SEPARATION METHODS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPLC-MS</td>
<td>Separation based on hydrophobicity, both group and with-in group separation of lipids</td>
<td>Good separation for most lipids, accurate retention times (typically under 15 RSD), facilitating lipid identification</td>
<td>Internal standards not co-eluting with most lipids, which may compromise normalization. Need for larger number of internal standards.</td>
</tr>
<tr>
<td>HILIC-MS</td>
<td>Lipid group separation, neutral lipids not well retained</td>
<td>Lipid class specific internal standards can be used for normalization.</td>
<td>Individual lipids are not well separated; severe matrix suppression; broad peaks may compromise the normalization as internal standards do not fully coelute.</td>
</tr>
<tr>
<td>NPLC-MS</td>
<td>Lipid group separation</td>
<td>Lipid class specific internal standards can be used for normalization</td>
<td>Individual lipids are not separated: severe matrix suppression, broad peaks may compromise the normalization</td>
</tr>
<tr>
<td><strong>MS METHODS WITH LC</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC-QTOFMS</td>
<td>Fast scanning</td>
<td>Enables fast separation and short LC analysis.</td>
<td>Resolution is typically 30,000–60000.</td>
</tr>
<tr>
<td>LC-IM-QTOFMS</td>
<td>IM allows additional separation of ionized molecules based on their different ion mobilities in low or high electric fields due to differences in size, shape, charge and mass</td>
<td>Enhanced separation and identification of isomeric and isobaric lipid species; much cleaner fragment ion spectra</td>
<td>Loss of sensitivity due to the IM step, depending in the instrument setup, more complex data processing.</td>
</tr>
<tr>
<td>LC-Orbitrap</td>
<td>Ultra-high resolution</td>
<td>Increased separation due to the ultra-high resolution</td>
<td>Resolution up to 120,000, ultra-high resolution requires slower LC analysis due to slower scanning rates than in QTOFMS.</td>
</tr>
<tr>
<td><strong>DIRECT INFUSION (SHOTGUN) MS METHODS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>Uses unique MS/MS fragment for the lipid class of interest</td>
<td>Relatively simple and fast analysis</td>
<td>Matrix suppression reduces the sensitivity, no separation of isomeric species having identical fragmentation patterns, quantification affected by differences in FA chain and saturation.</td>
</tr>
<tr>
<td>HRMS</td>
<td>Multi-product-ion spectra (PIS) or neutral losses (NLS) extracted from the acquired data of the product ion spectra to identify individual species; or data-dependent acquisition to acquire product ion spectra for identification full mass spectra</td>
<td>Broader lipid coverage compared to tandem MS</td>
<td>Matrix suppression reduces the sensitivity, no separation of isomeric species having identical fragmentation patterns, quantification affected by differences in FA chain and saturation particularly among non-polar lipid classes.</td>
</tr>
<tr>
<td>Multidimensional MS</td>
<td>Utilizes the intra-source separation; identification of individual species by array analysis using tandem MS techniques (PIS and NLS) to determine building blocks of molecular ions;</td>
<td>Broader lipid coverage than in other shotgun methods, intra-source separation reduces the matrix suppression effects; more accurate quantitation</td>
<td>Requires knowledge of lipid fragmentation, not well suited for the analysis of unknown or uncharacterized lipid classes, no separation of isomeric species having identical fragmentation patterns, Loss of sensitivity due to the IM step, depending in the instrument setup</td>
</tr>
<tr>
<td>IM-QTOFMS</td>
<td>IM allows additional separation of ionized molecules based on their different ion mobilities in low or high electric fields due to differences in size, shape, charge and mass; can be used in combination with other workflows</td>
<td>Enhanced separation and identification of isomeric and isobaric lipid species, reduced matrix suppression effects</td>
<td></td>
</tr>
</tbody>
</table>
profiling methods for bioactive lipids, such as bile acids, oxylipins, endocannabinoids and fatty acid esters of hydroxy fatty acids have been developed, based on UHPLC coupled to high resolution MS (UHPLC-HRMS) [21]. These lipids are typically not covered by global lipidomics analyses, as they are present in very low concentrations and require more specific sample preparation. In addition, specific lipids, such as free fatty acids, total esterified fatty acids (after saponification) as well as some steroids are typically done with GC combined with MS, although also LC-MS methods have been developed for these compounds.

Advantages and limitations of the most common methods are summarized in Table 1 for global lipidomics, and in Table 2 for targeted analysis of commonly-analyzed lipid classes. While the global profiling methods give good coverage of lipids in general, there is still a major challenge as regards separation and identification of lipid regioisomers such as those that differ in sites of unsaturation or positions of substitution on a glycerol backbone. LC combined with ion mobility mass spectrometry (IM-MS) adds an orthogonal separation dimension between LC and MS without impacting the analysis time as the IM separation occurs on a timescale of milliseconds, making it compatible with even UHPLC-HRMS methodologies. The additional separation by IM can be used to enhance separations based on lipid classes, double bond positional isomers, and glycerolipids differing in acyl chain connectivity to the glycerol backbone [22–26].

2.2. Lipid identification

One of the challenges in global lipidomics is lipid identification, given the huge number of different lipids, which vary in lipid class, geometric isomerism (cis versus trans), and the position, lengths and degrees of unsaturation of fatty acyl constituents. Identification can be conducted for lipid class by DIMS or LC-MS if this lipid class can be sensitively ionized, and/or possesses sensitive and specific fragments within their linear dynamic range. In untargeted LC-MS, lipid peaks are typically resolved by LC-MS/MS, performed separately, or by using automated MS/MS approaches. Typically, lipid species are annotated by the number of total carbons and their bond position and geometric isomer level requires specific approaches, as discussed later.

Another challenge is the identification and quantification of lipid classes which are poorly ionizable, or do not possess sensitive and characteristic fragment ions. For these lipids, many recent efforts including chemical derivatization are frequently used to enhance MS ionization efficiency and selectivity, to facilitate structure elucidation, or to improve chromatographic separation [27,28]. Similar approaches can be used for more specific identification of other lipids as well. Chemical derivatization aims to introduce a polar group to modify the target lipid molecule, convert fatty acids and eicosanoids to a cationic AMPP amide improved sensitivity. Conversion of the carboxylic acid of FFAs and eicosanoids to a cation, at a time of detection up to 60,000-fold compared to negative mode efficiencies of individual molecular species of a nonpolar lipid class to a polar lipid class of which ionization efficiencies of individual species are essentially identical and remarkably improved.

2.2.1. Identification of fatty acids with derivatization

Fatty acids (FAs) represent a broad family of lipids containing at least one carboxylic acid group and a long aliphatic chain, which includes free fatty acids (FFAs) and modified fatty acids. Modification of these FA species through enzymatic or nonenzymatic processes generates large numbers of modified FA species, such as oxidized FFAs, fatty acid esters of hydroxy fatty acids (FAHFs), nitrosylated fatty acids, halogenated fatty acids and so on. These lipid species fulfill multiple critical roles in cellular function in mammalian organisms. Due to their low ionization efficiency in negative ion mode, wide concentration range and shortage of specific fragments, identification and quantification of these lipid species is still a challenging task.

In recent years, LC-MS or DIMS in conjunction with chemical derivatization has been a widespread strategy for the analysis of FAs and modified FAs. The derivatization reagent N-(4-aminomethylphenyl) pyridinium (AMPP) is usually used to derivatize FFAs and eicosanoids via an amide linkage [29,30] (Fig. 2a). This charge reversal derivatization allows detection of cations rather than anions in electrospray ionization mass spectrometry, which enhances sensitivity. Conversion of the carboxylic acid of FFAs and eicosanoids to a cationic AMPP amide improved sensitivity of detection up to 60,000-fold compared to negative mode electrospray ionization detection of underivatized analytes. 2-dimethylaminoethanolamine (DMED) has also been widely-utilized to readily label carboxyl groups of FAs, eicosanoids and FAHFs [31] (Fig. 2b). The detection sensitivities of DMED-labelled eicosanoids improved by 3–104 fold compared with unlabelled analytes, and the detection sensitivities of FAHFs increased by 7–72 fold upon DMED labeling.
Besides AMPP and DMED, several other reagents, including N,N-diethyl-1,2-ethanediame (DEEA) [32] (Fig. 2b), 2,4-bis(diethylamino)-6-hydrazino-1,3,5-triazine (T3) [33] (Fig. 2c), 5-(dimethylamino) naphthalene-1-sulfonyl piperazine (Dns-PP) [34] (Fig. 2d), have also been used for FA derivatization. In particular, increasing interest has been directed towards the interactions between gut microbes and short-chain fatty acids (SCFAs, ≤6 C atoms). Specifically, N-(4-(aminomethyl) benzyl) aniline (4-AMBA) and N-(4-(aminomethyl) benzyl) aniline-d5 (4-AMBA-d5) were used to label SCFAs (Fig. 2e), and this method was able to simultaneously quantify 34 SCFAs carrying alkyl, carbonyl, hydroxy, and alkenyl groups with high sensitivity [35].

2.2.2. Identification of glycerophospholipids with derivatization

GPLs are characterized by two fatty acyl groups at sn-1 and sn-2 positions and a polar headgroup attached to a common glycerol scaffold at the sn-3 position; subclassification of GPLs is based on polar headgroup structure, such as phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidyserine (PS), phosphatidic acid (PA), phosphatidylglycerol (PG), phosphatidylinositol (PI) and their lyso variants (lysoPLs). The phosphate ester group of GPLs is readily fragmented and could yield characteristic fragment ions in ESI-MS/MS. DIMS holds promise for the characterization of GPLs, due to effective intra-source separation of predetermined groups of lipid classes according to their intrinsic electrical propensities; lipid extracts can be directly analyzed without pre-separation. For example, collision-induced decomposition of the positive protonated molecule ion ([M+H]^+) of PC yields the expected ion of m/z 184 (headgroup of PC), and takes advantage of the unique feature of tandem-quadrupole mass spectrometers for detecting specific compounds, in that precursor-ion scanning (PIS) of 184 Da can be used for identification of all PC species. However, with DIMS, some GPLs are difficult to ionize and the analysis suffers from ion suppression in complex biological matrices. It is therefore proposed that enhancing ionization efficiencies of molecular species of different GPL classes by chemical derivatization would be particularly valuable to eliminate this endogenous interference. Besides, enhancement of ionization efficiencies of targeted lipid classes would broaden their linear dynamic range. Furthermore, some derivatization strategies can facilitate structural characterization of GPLs [36].

Methylation of primary amines as well as the hydroxyl group of phosphate(s) of GPLs by using diazomethane or trimethylsilyl diazomethane (TMS-diazomethane) is a widely-applied strategy for enhancing ionization efficiencies [37,38] (Fig. 3). This reaction concomitantly produces a fixed positive charge on the lipid through the formation of a quaternary ammonium ion and by neutralizing the negative charge of the phosphate group. Sensitivity
enhancements were observed for PEs, PCs, sphingomyelins (SMs), and PSs, particularly in MS/MS experiments where ion fragmentation was consolidated to only one or two channels. Thus, these fragment ions can be employed for sensitive identification and quantification of these lipid classes.

In addition to the advantages of enhancement of detection sensitivity, chemical derivatization can also assist in elimination of the ambiguity introduced by the presence of isomeric species, particularly for isomeric species of lipids of different classes but which have the same ionic elemental compositions [39]. For instance, d6-S,S'dimethylthiobutanoylhydroxysuccinimide ester (d6-DMBNHS) reagent was used to selectively derivatize PE and PS, eliminating the possibility of isobaric mass overlap of these species with the precursor ions of all other lipids in the crude extracts, thereby enabling their unambiguous assignment. Moreover, DMBNHS and iodine/methanol were sequentially used to derivatize plasmalogen ether-containing PE and PS lipids when diacyl- and O-alkyl-ether-containing PE and PS lipids occurred simultaneously [39]. This sequential derivatization enables discrimination of plasmalogens and unsaturated O-alkyl ether-containing lipids. Other methods including acetone stable-isotope derivatization for profiling and accurate quantification of aminophospholipids in biological samples have also been reported [40–42].

### 2.2.3. Advances in identification of double bond location(s) of unsaturated lipids

The specific locations of double bonds in mammalian lipids have profound effects on biological membrane structure, dynamics and lipid second messenger production. Identification and quantitation of lipid C=C location isomers with adequate sensitivity is a long-standing challenge in lipidomics. The main obstacle with determining C=C locations is that low-energy CID are not effective at breaking the C=C bond due to the high dissociation energy required for cleavage. A number of novel approaches including ozonolysis, epoxidation and derivatization strategies have been explored to overcome these obstacles in recent years (Table 3).

**Table 3** Summary of recently advanced strategies for identification of double bond location.

<table>
<thead>
<tr>
<th>Targeted lipid class</th>
<th>Methods</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGs, GPLs, LPLs, FFAs, SMs, ChoEs</td>
<td>Ozonolysis</td>
<td>[43,44,136,137]</td>
</tr>
<tr>
<td>FAs</td>
<td>AMPP derivatization</td>
<td>[30,45]</td>
</tr>
<tr>
<td>FA, FAME, GPLs</td>
<td>Epoxidation reaction</td>
<td>[49–51]</td>
</tr>
<tr>
<td>GPLs, SMs</td>
<td>UVPD</td>
<td>[52–54]</td>
</tr>
<tr>
<td>TG, PC, SM</td>
<td>RDD</td>
<td>[55]</td>
</tr>
<tr>
<td>TG, SM</td>
<td>EIEIO</td>
<td>[56,57]</td>
</tr>
</tbody>
</table>

Ozonolysis utilizes the reaction of ozone with carbon-carbon double bonds to produce characteristic fragment ions that are specific to C=C bonds. In this reaction, ozone molecule attacks the C=C bond, inducing the formation of a primary ozonide which is very unstable and dissociates into aldehyde and Criegee ions (likely to be either a vinyl hydroperoxide or carboxylic acid generated from the rearrangement of the Criegee intermediate). Since the aldehyde and Criegee ions have a 16 Da difference, they can be used as characteristic ions for unambiguous identification of C=C locations of lipid [43] (Fig. 4a). However, long acquisition times and spectral complexity arising from the presence of isomeric or isobaric lipids present in direct infusion workflows limit the application of this approach when dealing with complex biological samples. More recently, Mitchell et al. reported a strategy combining LC with ozone-induced dissociation (OzID) which reduces spectral complexity and enables discrimination of lipid isomers and isobars [44]. The positions of C=C bonds in unsaturated lipids across 11 classes, including GPLs, TGs and cholesterol esters, were unequivocally assigned based on predictable OzID transitions.

As previously stated, charge-switch derivatization of the carboxylic acid functional group of FAs with AMPP has also been utilized to generate characteristic fragment ions for specific double bond locations [45]. Specifically, the locations of proximal double bonds in AMPP-derivatized fatty acids are identified by diagnostic fragment ions resulting from the markedly-reduced 1,4-hydrogen...
elimination from the proximal olefinic carbons. Additional fragmentation patterns resulting from allylic cleavages further substantiate the double bond position. This strategy for pinpointing FA double bond locations has the distinct advantages of both high specificity and high sensitivity.

Recently, a C=C specific derivatization approach based on coupling the Paternò–Büchi (PB) reaction with MS/MS has been developed to confidently assign C=C locations in various classes of lipid species [46–48] (Fig. 4b). The reaction mechanism involves activation of the carbonyl group within an aldehyde or ketone to a diradical upon UV excitation, which subsequently reacts with the C=C bond. Two isomeric oxetane ring compounds can then be formed depending on the relative positions of the carbonyl and the C=C bond. After CID, acetone-tagged, unsaturated lipids generate an aldehyde/ketone and olefin diagnostic pair with a mass difference of 26 Da from the cleavage of the C-C bond at the initial C=C bond position and the C-O bond of the initial carbonyl group. Therefore, these paired double bond diagnostic ions can be used for identification of double bond locations and quantification of these unsaturated lipid isomers. This method enabled profiling of unsaturated lipids from various biological samples (blood, plasma, and cell lines) [47].

The use of an epoxidation reaction is another emerging and powerful approach to identify lipid C=C locations [49–51]. Epoxidation of the C=C bond leads to the production of an epoxide, which produces abundant diagnostic ions indicative of the C=C location upon CID. This strategy has been successfully applied to a range of unsaturated FAs, fatty acid methyl/ethyl esters and GPLs to assign their double bond locations.

Another activation method that has demonstrated success for characterization of lipids is ultraviolet photodissociation (UVPD). Brodbelt et al. used UVPD implemented on an Orbitrap mass spectrometer to localize double bond positions within GPLs [52,53]. Cleavage of the C=C adjacent to the double bond provides a diagnostic mass difference of 24 Da and enables differentiation of C=C positional isomers. UVPD was also shown to yield a series of novel, structurally-diagnostic product ions resulting from the cleavage of both sphingosine carbon—carbon and acyl chain carbon—carbon double bonds for direct localization of the site(s) of unsaturation, as well as via diagnostic cleavages of the sphingosine backbone and N-C amide bond linkages [54]. Other tandem MS techniques, such as radical directed dissociation (RDD) and electron excitation of ions from organics (EIEIO), have also been employed for locating lipid C=C bonds. In the RDD approach, bifunctional molecules that contain a photocaged radical initiator and a lipid-adding group are used to form noncovalent complexes with a lipid during electrospray ionization. Irradiation of these noncovalent complexes at 266 nm UV cleave the carbon–iodine bond to release a highly reactive phenyl radical. Subsequent CID of the nascent radical ions would generate diagnostic ions for the C=C bond and chain-branching positions in GPLs, sphingomyelins and TGs [55]. In EIEIO, singly-protonated precursor lipid ions generated by an ESI are introduced into a branched ion trap wherein they react with a 10 eV electron beam. This produces information-rich fragment ions that have been shown to reveal lipid class, respective acyl chain lengths, the number and locations of double bonds in the fatty acid chains, and regiosomer specificity. This technique was applied to provide in-depth structural elucidation for PCs, SMs, and TG molecular species [56,57].

2.3. Quality control

Strict quality assurance (QA) and quality control (QC) are two quality management processes that are integral to lipidomics in order to produce high-quality data in any high-throughput analytical chemistry laboratory. The most commonly-applied guidelines are those published by the US Food and Drug Administration (FDA) [58], but also other guidelines, such as EMA Bioanalytical Method Validation Guidelines for bioanalytical method validation [59] and the Japanese Ministry of Health, Labour and Welfare (MHLW) Guidelines on Bioanalytical Method Validation in Pharmaceutical Development [60] are adaptable for lipidomics. The current status of the development of different guidelines for bioanalysis has recently been reviewed [61].

In global lipidomics approaches, use of appropriate internal standards is recommended, which are added to the sample prior to sample preparation, along with analysis of different types of QC standards together with the samples. Pooled QC samples, with a small aliquot of each sample analyzed, are recommended as a control to assess system stability. In addition, reference standard material, such as NIST SRM 1950, should be analyzed in each sample batch. Recently, the NIST SRM 1950 plasma was analyzed in 31 individual laboratories globally, using both LC-MS/MS and shotgun MS on different mass spectrometers [3]. As a result,
consensus values of molar concentrations of specific species from different lipid classes that rely on concordance between independent measurements were reported. Nevertheless, the current limitation of such ring studies in lipidomics is that ‘gold standards’, i.e., accurate reference concentrations, are not available for all lipids measured. Such ring studies can, however, provide invaluable information concerning the extent of variability in measured concentrations between different laboratories, and may thus help identify areas for further improvement of specific methods.

In order to improve the quality of lipidomic analyses, acceptance criteria for precision and accuracy of the lipid data should be defined. Currently, there are no predetermined QC acceptance criteria for lipidomics. For untargeted assays, it is only possible to provide a relative measure of precision—the random error in quantification over repeated measurement of a biologically-identical sample. Based on the aforementioned “Bioanalytical Method Validation; Guidance for Industry” [58], similar thresholds have been suggested with some modification for lipidomics as well. For non-targeted metabolomics analysis by LC—MS, values close to 30% RSD could be allowed.

The choice of correct internal standards is essential. Preferably, isotopically-labelled lipid standards should be used, with concentrations close to biologically-relevant ranges, or, when not available, lipids with, e.g., odd-chain fatty acid composition can be used. It is not possible to have standards for every single lipid, however, and sufficient selection of standards should be made so as to not only cover the different subclasses of lipids, but also fatty acid composition (chain length, saturation level) should cover a sufficient range of lipids. Combining internal standards with external standard calibration will allow better (semi)quantitative results than single point calibration.

Pooled QC sample data can also be used for modelling and correcting for systematic measurement bias, commonly called batch effects. Correction of drift effects is challenging, as the direction of change and degree of nonlinearity is unique to each lipid. While some compounds show only minimal drift, others can have a significant linear or even nonlinear change in response over time. Several methods have been suggested for the correction of such batch effects, including support vector regression [62], the wavelet transform method [63], locally-estimated scatterplot smoothing (LOESS) [64], using local minimum values coupled with robust statistical analysis [65]. Most of the methods utilize the data from (pooled) QC samples for batch correction while others use the data from all samples, or are based on advanced quantile normalization [66,67]. For example, the wavelet transform method [63] uses the time trend of samples over the injection order, decomposes the original data into multi-scale data with different features, extracts and removes the batch effect information in the multi-scale data. Several software packages, mainly implemented in the R statistical programming language, have been developed which allow batch correction either as a part of a complete data analysis workflow or for stand-alone use [63,66]. Also, web-based tools, which include batch correction, have recently been introduced [68,69]. The SERFF method (systematic error removal using random forest), developed by Fiehn and colleagues, has been specifically developed for large-scale untargeted lipidomics studies [68]. However, there are currently no guidelines as to when it is advantageous/required for these batch correction methods to be applied. The challenge of batch correction is that the applied correction may, in fact, overcorrect the data (i.e., remove true biological signal along with the analytical batch effect). Based on our experience, batch effect correction should be used with caution, and should not be used if the QC data (namely, both the raw variation of the ISTDs and the data from the QC, i.e., pooled samples, in-house QC) shows low variation (RSD < 15–20%). Typically, this is the case in batches with less than 300–400 samples.

2.4. Data processing

Data processing transforms raw lipidomics data from the analytical instrument into a dataset in tabular format, which can then be further analyzed by statistical and pathway analysis approaches. Several data processing tools have been developed by the lipidomics community, both for LC-MS and shotgun-MS approaches [70,71]. Also, instrument vendors offer solutions for processing and analysis of metabolomics (including lipidomics) data. Popular metabolomics open-source packages such as XCMS [72,73], MZmine [74,75] and MS-DIAL [76], are also commonly used in lipidomic studies. In a recent systematic evaluation of five, commonly-used metabolomics data processing software packages, MZmine outperformed other tested software in terms of quantification accuracy, and it reported the greatest number of true discriminating markers together with the fewest false markers [77]. The LipidFinder tool was developed by LIPID MAPS, which facilitates peak filtering and lipid identification after processing of lipidomics data with XCMS, thus also reducing the number of false positive peaks [78].

Several databases and in-silico tools have also been developed to facilitate lipid identification in lipidomics experiments, including, e.g., several tools and resources from LIPID MAPS [79], LipidMatch [80], LipidBlast [81] and LIQUID [82]. LIPID MAPS also developed tools which enable bulk structure searches, thus enabling reporting of measured lipids at an appropriate level of structural detail [79]. Use of IM-MS further improves the ability to identify lipids due to its high separation efficiency. Zhu et al. have recently developed a workflow for the identification of lipids by IM-MS, LipidIMMS Analyzer, which includes a database of 260,000 lipids, with four-dimensional information for each lipid (m/z, retention time, collision cross-section and MS/MS spectra) [83]. McLean and colleagues have developed an online resource, which includes over 3800 experimentally-acquired collision cross-section (CCS) values, including for over 800 lipids [26,84].

Ultimately, the choice of data processing method may depend on the analytical method used, local expertise in metabolomics data processing as well as on local bioinformatics workflows. Before being put to use, any data processing solution should be evaluated and tested as part of the lipidomics method development and validation. The above-mentioned open source tools all require proper optimization of parameters for appropriate use.

3. Factors influencing lipidomes

The concentrations of circulating lipids are highly dynamic, and several factors will impact the measured lipid profiles. While there are several biological factors having an impact, there are also many other factors, such as sampling method and even sample handling. Below we highlight and discuss various biological and analytical factors influencing lipid levels.

3.1. Biological factors

Lipid profiles are affected by both genetic and environmental factors, such as diet, the host’s microbiome, gender, age, ethnicity and lifestyle [85–87]. In addition, circadian variation and epigenetics have an impact on lipid profiles [88–90]. The major factors affecting lipidomes are summarized in Table 4 and illustrated in Fig. 1.

Diet markedly affects the lipid levels. In a twin study, Frahnnow et al. evaluated systemic lipidomic adaptation after switching from
a low-fat to a high-fat diet (HFD) [91]. The results showed that lipid profile of twins changed already within the first week of HFD. Short-term [92,93] and long-term [86,94] clinical trials also provide clear evidence that diet has a marked impact on lipid profiles in healthy individuals. Therefore, in order to understand the physiological roles of lipids and to draw meaningful conclusions, diet should be considered as a key confounder in clinical lipidomics study settings.

The gut microbiome is another important regulator of host lipid metabolism [85,95,96]. Various lipids, including those of microbial origin, directly influence the lipid profile of the host [97,98]. Recent study by Brown et al. showed that gut microbe-derived sphingolipids are critical for maintaining a balance in host-produced sphingolipid levels [99]. Microbial catabolism and host lipid metabolism are highly inter-dependent [95,97]. One example is microbial bile acid transformation (taurine and glycine conjugation), which subsequently influences lipid absorption, distribution and the metabolic status of the host [85,100]. Since the gut microbiome has significant effects on systemic lipid profiles, understanding of the gut-microbiome-lipidome axis may contribute to better evaluation of lipidomic biomarkers in health and disease.

Body mass index (BMI), sex and age are also considered as major factors affecting lipidomes. We have recently shown that levels of lipids in the blood profoundly changed during early infancy as well as during the first few years of life, however it is currently unclear whether this is a direct effect of age or due to diet-age interaction [101]. Similarly, blood lipid profiles have been reported to be selectively altered by aging [87]. A recent study of elderly subjects (100 healthy subjects, 56–100 years of age) showed a universal reduction in lipid concentrations in older subjects, independent of sex and BMI, which was especially apparent in the oldest subjects over 95 years of age. Furthermore, the lipid profile itself is substantially different in the elderly population compared to previous data from young and middle-aged adults. Interestingly, the reduction of most lipids in older age was most striking in males [102]. However, irrespective of age, sex-associated differences in lipid profiles were also identified. Ishikawa et al. have highlighted that sphingomyelin species remained significantly higher in males than in females, regardless of age [103]. There have been many reports confirming gender-specific lipidomic differences in humans, in particular with circulating GPL and SL species appearing to vary between males and females [104–107].

Recent studies have shown that several classes of lipid species, including TGs, GPLs, SLs and sterols are influenced by circadian rhythm [108–110]. A human circadian metabolome study revealed that among all the metabolites that were under circadian control, more than 70% were lipids, independent of the patterns of eating and sleeping [110]. In addition, a report from Chua et al. showed that healthy individuals exhibit distinct circadian lipid phenotypes [109]. Thus, in lipidomic study settings, diurnal sample matching seems an appropriate approach for improving the predictive validity of lipid biomarkers.

Exogenous chemicals, including both drugs and environmental chemicals, can also modify the lipidome. Several drugs target lipid metabolism (e.g., statins) or are known to have a major impact on lipids (NSAIDs, anticoagulants) [111–113], and there are several drugs with insufficiently-characterized metabolic impact (i.e., hormones, including contraceptives, steroids, diuretics). Recently, exposure to environmental chemicals, such as endocrine disrupting chemicals (e.g., per- and polyfluorinated compounds (PFAS), pesticides, PCBs, phthalates) has shown to have a marked impact on the lipidome [6,114].

### Table 4

Factors impacting the lipidomes, a summary.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Lipid affected</th>
<th>Level of impact</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diet</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fat</td>
<td>Upregulation of short, saturated and monounsaturated TGs and ChoEs</td>
<td>NA</td>
<td>[138]</td>
</tr>
<tr>
<td>High Fat diet (Western-style diet)</td>
<td>↑ Cer, ↑ SM, ↑ PC(0)</td>
<td>102 out of 150 lipid species (58%)</td>
<td>[91]</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>↑ TGs, HDL cholesterol ↓</td>
<td>Liver fat increased by 27%</td>
<td>[92]</td>
</tr>
<tr>
<td>Nordic diet (whole-grain products, berries, three fish meals per week)</td>
<td>↓ γ-linolenic acid</td>
<td>NA</td>
<td>[88]</td>
</tr>
<tr>
<td>Mediterranean diet</td>
<td>↓ LDL-C to HDL-C, Apo B to Apo A1 ratios, PC(38:4), PE(38:5), ChoEs (16:1), TG(48:0), TG(50:0), DAg</td>
<td>NA</td>
<td>[94]</td>
</tr>
<tr>
<td><strong>Genetic and non-genetic factors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>SM(d18:1/24:2) F↓, SM(d18:1/23:1) F↑, M↑, ChoE (M1), TG (M1), Cer (M1), GM (M1), LPC (M1), SM (F1), THG (F1), and PS (F1)</td>
<td>5–17%</td>
<td>[104,139]</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>↓ Total cholesterol (TC) and low-density lipoprotein cholesterol</td>
<td>NA</td>
<td>[140,141]</td>
</tr>
<tr>
<td>Iranians vs. Europeans and Indians</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic vs. non-Hispanic</td>
<td>HDL-cholesterol, and higher TG/HDL-cholesterol, Cer (↑), MHC (↑), THG (↑), GM (↑), SM (↑), PC (↑), PC(O)↑, LPC(O)↑, PE↑, LPE (↑), PI (↑), PG (↑), ChoE (↑), free cholesterol (↑), DG (↑), and TG (↑)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Lower lipids at very old age (&gt;95 y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Circadian variation</td>
<td></td>
<td>13% across individuals</td>
<td>[109]</td>
</tr>
<tr>
<td>Epigenetics</td>
<td>TGs and PCs</td>
<td></td>
<td>[88–90]</td>
</tr>
<tr>
<td>Methylations</td>
<td>Cholesterol, sphingolipids, and glycerophospholipids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gut microbes</td>
<td>Sphingolipids, Bile acids</td>
<td>NA</td>
<td>[85,99,100]</td>
</tr>
<tr>
<td><strong>Exogenous chemicals including pharmacological agents</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statins</td>
<td>↓ TG, cholesterol, LDL-C</td>
<td>41%, 55%, 39% respectively</td>
<td>[111]</td>
</tr>
<tr>
<td>Antipsychotics (olanzapine, risperidone and aripiprazole)</td>
<td>↑ PC, PE and TG</td>
<td>NA</td>
<td>[142]</td>
</tr>
<tr>
<td>Environmental chemicals: POP, PFAS</td>
<td>SM, PC, SM, LPC</td>
<td>NA</td>
<td>[6,114]</td>
</tr>
</tbody>
</table>
There are reports suggesting that the human lipid profile is heritable [91,115]. Specifically, the San Antonio Family Heart Study showed that the genetic profile of individuals influenced their lipidome, which also strongly associated with various cardiovascular risk factors [115]. Genome-wide association studies (GWAS) have also revealed 157 loci significantly associated with blood lipid metabolism [116]. Data from both observational settings and clinical trials suggest that genetic factors may interact with diet and lifestyle factors in determining circulating lipid profiles [117]. However, it is also likely that interactions with epigenetic factors, such as DNA methylation, contribute to the modulation of host lipid metabolism [89,118].

Taken together, understanding how genetic, non-genetic and environmental factors together influence lipid metabolism will be essential in the study of lipids in health and disease. While current community-wide efforts are focused on establishing 'normal' healthy levels of lipids, such knowledge may be of little use, without sufficient data on the impacts of different internal and external exposures on lipid levels.

3.2. Effects of sample handling

Blood-based samples such as plasma and serum are the most frequently used sample matrices in lipidomics research. There are differences in lipidomes from these two blood-based sample types [8,103,119–122], with plasma being generally preferred for lipidomics analyses [5]. Particularly in clinical lipidomics, the quality of the samples is of utmost importance for valid diagnostics and appropriate therapeutic decisions, as well as for reliable, conclusive research findings. However, even highly-standardized processes are not foolproof, and both human and systematic errors (e.g., delays during blood preprocessing, interruptions in the cold chain) may occur. Less-than-ideal pre-analytical sample handling in relation to clotting time, storage time, freeze-thaw cycles, and collection techniques can each alter lipidomics results significantly [123]. Of the lipid classes commonly covered in global lipidomics studies, lysPCs in particular change significantly due to handling at room temperature — with the single exception of lysPC(18:2) [124]. Whilst other lipids may also be affected by sample handling conditions, however, most of the major lipids have been shown to be sufficiently stable if, crucially, samples are cooled quickly and processed within two hours of collection [119,120,125–127]. Also, several studies have shown that lipids are stable for between two to four freeze-thaw cycles [124]. Hemolytic samples usually contain increased concentrations of lipids mainly originating from the cellular membranes of erythrocytes [128]. Here, it should be noted that several bioactive lipid classes, such as oxylipins, are highly unstable, and for these lipid classes, specific procedures during both sampling and preanalytical procedures are required (Table 2).

Ideally, preanalytical quality should be evaluated by using specific markers that measured during the analytical workflow. However, currently there are no fully-validated markers for assessing preanalytical quality of samples. Due to the large individual variation in lipid levels, it is not really feasible to use cut-off levels for individual lipids, e.g., lysPCs, for determining whether an individual sample has acceptable pre-analytical quality. In addition, critical evaluation of candidate markers in the context of other parameters that may result in abnormal levels of the potential markers should be considered. A recent study suggested the use of the ratio of total lysPCs and total PCs as a reasonable marker of preanalytical quality [124]. For single markers, increased levels of sphingosine-1-phosphate d18:2 (S1P-d18:2) has also been suggested as a marker of preanalytical quality [129], as the levels of this lipid are not affected by, e.g., hepatocellular carcinoma, liver cirrhosis, hepatitis, severe sepsis, prior cardiopulmonary resuscitation, or strenuous exercise. S1P-related analytes, like sphingomyelins, did not show pronounced changes owing to improper sample handling.

4. Conclusions

The field of lipidomics has witnessed rapid progress over the past years, becoming one of the fastest growing ‘omics’ fields. With rapid growth comes growing pains, here in the form of the inherent limitations of lipidomic analyses, making it challenging to compare data acquired from different platforms and laboratories. Loose application of terminology, such as when referring to quantification of lipids, further complicates the matter. In analytical terms, global lipidomic methods are not fully (that is: ‘absolutely’) quantitative, because it is not possible to obtain authentic standards or labelled internal standards for each lipid measured. Furthermore, reporting of lipid identities is insufficiently harmonized and often the lipids are not reported at a sufficient level of molecular detail, as afforded by the experimental data. Nevertheless, progress is being made on several fronts. Community-wide efforts are ongoing, aiming to harmonize the reporting of lipidomics experiments and the data generated, as well as to provide guidelines for best practices in lipidomic analyses. Development of plasma reference values for selected lipids will facilitate assessment of accuracy of quantitation in lipidomic studies.

Whilst it is generally recommended to report only identified lipids from lipidomics experiments, new studies are also revealing that there are many lipids yet to be discovered. Recent discoveries of novel eicosanoids and related oxylipids [130] and FAHFA-containing TGs [131] are two such recent examples. Improvements in analytical and data processing solutions for lipid identification and quantitation are likely to substantially expand the coverage of lipidomes, which will transform the field of lipid research. Studies of interactions of lipids with various internal (e.g., gut microbiome, genome) and external (e.g., diet, environmental exposures such as chemical pollutants) exposures are also new frontiers in lipid research. Understanding how such exposures impact lipid levels is essential if one is to study lipids in health and disease.

Acknowledgments

M.O. and T.H. were supported from the European Union Horizon 2020 Programme project EPOs (no. 634413). The authors thank to Dr. Aidan McInlinchey for assistance with language editing.

References


