



Christoph Krettler, BSc

# Identification of Galacto- and Oxidized Lipids with Lipid Data Analyzer

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Biomedical Engineering

submitted to

**Graz University of Technology**

Supervisor

Gerhard G. Thallinger, PhD

Institute of Neural Engineering

Graz, March 2020

## AFFIDAVIT<sup>1</sup>

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Graz, \_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

---

<sup>1</sup>Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008; Genehmigung des Senates am 1.12.2008

## Abstract

Life as we know it is not possible without lipids. Changes in lipid homeostasis can lead to various diseases, raising the importance of reliable identification and measurement of lipids. For example, oxidized lipids have been connected to age-related and chronic diseases, atherosclerosis, inflammation and immune responses and galactolipids play important parts in structural integrity of plant membranes. Due to its sensitivity, liquid chromatography coupled mass spectrometry (LC-MS) has become the *de facto* standard in lipidomics research, leading to a rise in bioinformatics tools to analyze the recorded data. However, due to the enormous diversity of lipids, most tools cover only a marginal range of lipid classes. In an effort to reduce such a shortcoming, this work aims to extend the lipid species covered by the Lipid Data Analyzer (LDA). Tools including LipidMatch, LipidMatch Flow and LPPtiger were used to survey lipids over a total of four different datasets to get an overview of LDA implementation gaps. Appropriate mass lists were generated for MS<sup>1</sup> identifications and the proprietary decision rulesets were extended for MS<sup>2</sup> identifications of important galactolipids and oxidized lipids, according to lipid species-specific fragmentation patterns. Furthermore, LDA source code was extended to enable identification of (oxidatively) modified fatty acyl chains. LDA now reliably identifies monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG), trigalactosyldiacylglycerol (TriGDG), tetragalactosyldiacylglycerol (TetraGDG) and sulfoquinovosyldiacylglycerol (SQDG), as well as oxidatively modified versions of all thirty implemented lipid classes with very low false positive rates. Comparison with LipidMatch, LipidMatch Flow and LPPtiger showed that LDA has a better coverage of the newly implemented lipids, providing researchers with a powerful platform to elucidate diseases caused by perturbations in the oxidized lipidome.

**Keywords:** *lipidomics, bioinformatics, oxidized lipids, galactolipids, mass spectrometry, cheminformatics*

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Lipids . . . . .	1
1.2	Lipid Notation . . . . .	2
1.3	Lipid Analysis . . . . .	2
1.4	Galactolipids . . . . .	2
1.5	Oxidative Lipidomics . . . . .	3
1.5.1	Lipid Peroxidation . . . . .	3
1.5.2	Oxidized Lipids . . . . .	4
1.5.3	Occurrence . . . . .	5
1.6	Bioinformatics Tools . . . . .	5
1.7	Motivation and Aims . . . . .	6
<b>2</b>	<b>Methods</b>	<b>7</b>
2.1	Tools . . . . .	7
2.2	Datasets . . . . .	7
2.3	Analysis Settings . . . . .	8
2.4	Hardware Environment . . . . .	8
<b>3</b>	<b>Results</b>	<b>9</b>
3.1	LipidMatch Results . . . . .	9
3.1.1	Dataset 1 - Mouse Liver . . . . .	9
3.1.2	Dataset 2 - Ryegrass Leaves . . . . .	10
3.1.3	Dataset 3 - Oxidized PC Standard . . . . .	12
3.1.4	Dataset 4 - Oxidized Wheat Seeds . . . . .	12
3.2	LPptiger Results . . . . .	12
3.2.1	Dataset 3 - Oxidized PC Standard . . . . .	12
3.2.2	Dataset 4 - Oxidized Wheat Seeds . . . . .	12
3.3	Generation of LDA Mass Lists . . . . .	12
3.3.1	Galactolipids . . . . .	13
3.3.2	Oxidized Lipids . . . . .	15
3.4	Extending LDA Rulesets . . . . .	15
3.4.1	MGDG . . . . .	15
3.4.2	DGDG . . . . .	17
3.4.3	TriGDG . . . . .	19
3.4.4	TetraGDG . . . . .	20
3.4.5	SQDG . . . . .	20
3.4.6	DGDG-Estolides . . . . .	21
3.4.7	Single Chain (MG) Galactolipid Species . . . . .	21
3.4.8	Oxidized Lipids . . . . .	21
3.5	Notation of Oxidized Lipids in LDA . . . . .	22
3.6	Novel LDA Features . . . . .	22
3.7	Lipid Data Analyzer Results . . . . .	24
3.7.1	Dataset 1 - Mouse Liver . . . . .	24
3.7.2	Dataset 2 - Ryegrass Leaves . . . . .	25
3.7.3	Dataset 3 - Oxidized PC Standard . . . . .	27
3.7.4	Dataset 4 - Oxidized Wheat Seeds . . . . .	27
3.8	Galactolipids - Comparison . . . . .	28
3.8.1	Validation of LipidMatch-only Identifications . . . . .	29
3.9	Oxidized Lipids - Comparison . . . . .	29
<b>4</b>	<b>Discussion</b>	<b>31</b>
4.1	Tools . . . . .	31
4.2	Datasets . . . . .	31
4.3	Notation . . . . .	32
4.4	Galactolipids . . . . .	33
4.5	Oxidized Lipids . . . . .	34
4.6	Noisy Spectra . . . . .	35

4.7	Outlook . . . . .	35
<b>5</b>	<b>Appendix</b>	<b>42</b>
5.1	LipidMatch Results . . . . .	42
5.1.1	LM - Dataset 1 . . . . .	42
5.1.2	LMF - Dataset 1 . . . . .	44
5.1.3	LM - Dataset 2 . . . . .	45
5.1.4	LMF - Dataset 2 . . . . .	47
5.2	LDA Mass Lists . . . . .	48
5.3	Novel LDA Features . . . . .	64
5.4	Lipid Data Analyzer Results . . . . .	69
5.4.1	Dataset 1 . . . . .	69
5.4.2	Dataset 2 . . . . .	71
5.4.3	Dataset 4 . . . . .	72
5.4.4	Dataset 2: Selected Spectra . . . . .	74
5.4.5	Dataset 3: Selected Spectra . . . . .	82
5.4.6	Dataset 4: Selected Spectra . . . . .	85
5.5	Galactolipids - Comparison . . . . .	89
5.5.1	LipidMatch-only Identifications . . . . .	91
5.6	Noisy Spectra . . . . .	93

## Abbreviations

---

AcylGlcADG	Acylglucuronosyldiacylglycerol
Cer-AP	Ceramide alpha-hydroxy fatty acid-phytospingosine
Cer-NS	Ceramide non-hydroxyfatty acid-sphingosine
Cer	Ceramide
CL	Cardiolipin
AcCar	Acylcarnitine
DG	Diacylglycerol
DGDG	Digalactosyldiacylglycerol
DGMG	Digalactosylmonoacylglycerol
ESI	Electrospray ionization
EpO	Epoxy-
EpOO	Epidioxy-
FAHFA	Fatty acid ester of hydroxyl fatty acid
HexCer-AP	Hexosylceramide alpha-hydroxy fatty acid-phytospingosine
Ke	Keto-
LC-MS	Liquid chromatography coupled mass spectrometry
LDA	Lipid Data Analyzer
LM	LipidMatch
LMF	LipidMatch Flow
LPA	Lysophosphatidic acid
LPC	Lysophosphatidylcholine
LPE	Lysophosphatidylethanolamine
LPG	Lysophosphatidylglycerol
MG	Monoacylglycerol
MGDG	Monogalactosyldiacylglycerol
MGMG	Monogalactosylmonoacylglycerol
oxCL	Oxidized cardiolipin
oxDG	Oxidized diacylglycerol
oxDGDG	Oxidized digalactosyldiacylglycerol
oxLPC	Oxidized lysophosphatidylcholine
oxMGDG	Oxidized monogalactosyldiacylglycerol
oxPG	Oxidized phosphatidylglycerol
oxPL	Oxidized phospholipids
oxPS	Oxidized phosphatidylserine
oxTG	Oxidized triacylglycerol
P-	Ether bond, not further defined
PA	Phosphatidic acid
PC	Phosphatidylcholine

PE	Phosphatidylethanolamine
PEtOH	Phosphatidylethanol
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
Plasmany-	Ether bond in position <i>sn</i> -1 to an alkyl group
Plasmenyl-	Ether bond in position <i>sn</i> -1 to an alkenyl group
PS	Phosphatidylserine
SM	Sphingomyelin
SQDG	Sulfoquinovosyldiacylglycerol
SQMG	Sulfoquinovosylmonoacylglycerol
TetraGDG	Tetragalactosyldiacylglycerol
TG	Triacylglycerol
TriGDG	Trigalactosyldiacylglycerol
UHPLC	Ultra-high performance liquid chromatography

---

# 1 Introduction

## 1.1 Lipids

Being loosely defined as biomolecules that are soluble in nonpolar solvents [1], lipids form a diverse group of organic compounds upon which life as we know it depends. Major functions of lipids include membrane structural components, energy and heat sources as well as signaling processes [2]. The totality of lipids in cells is called the lipidome. Diseases including diabetes, obesity, Alzheimer's disease, liver disease, hypertension and schizophrenia have been linked to perturbations in the lipidome [3]. Structural diversity stems from various combinations of the two basic building blocks – ketoacyl groups and isoprene groups [4] – as well as additional modifications, based upon which scientists categorize and classify lipids. The LIPID MAPS<sup>®</sup> consortium [5–7] categorizes lipids in fatty acyls (figure 1a), glycerophospholipids (figure 1b), sterol lipids (figure 1c), saccharolipids (figure 1d), glycerolipids (figure 1e), sphingolipids (figure 1f), prenol lipids (figure 1g) and polyketides (figure 1h), whereby each category is further divided into classes and subclasses [7,8]. The subclass of galactolipids, for example, comprises glycerolipids that are attached to galactose moieties. The number of lipids in a cellular lipidome is estimated to be in the tens of thousands to millions [9] depending on the level of structural resolution; the LipidHome database [3] contains currently about 20 thousand lipid species (i.e. lipids where no structural resolution on the composite fatty acids is available) and about 36 million different lipid molecular species (i.e. lipids where structural resolution on the composite fatty acids is available). This already high number gets diminished when considering isomers (i.e. lipids where structural resolution on the double bond position of composite fatty acids is available) and oxidatively modified chains (i.e. lipids where the composite fatty acids are oxidatively modified).

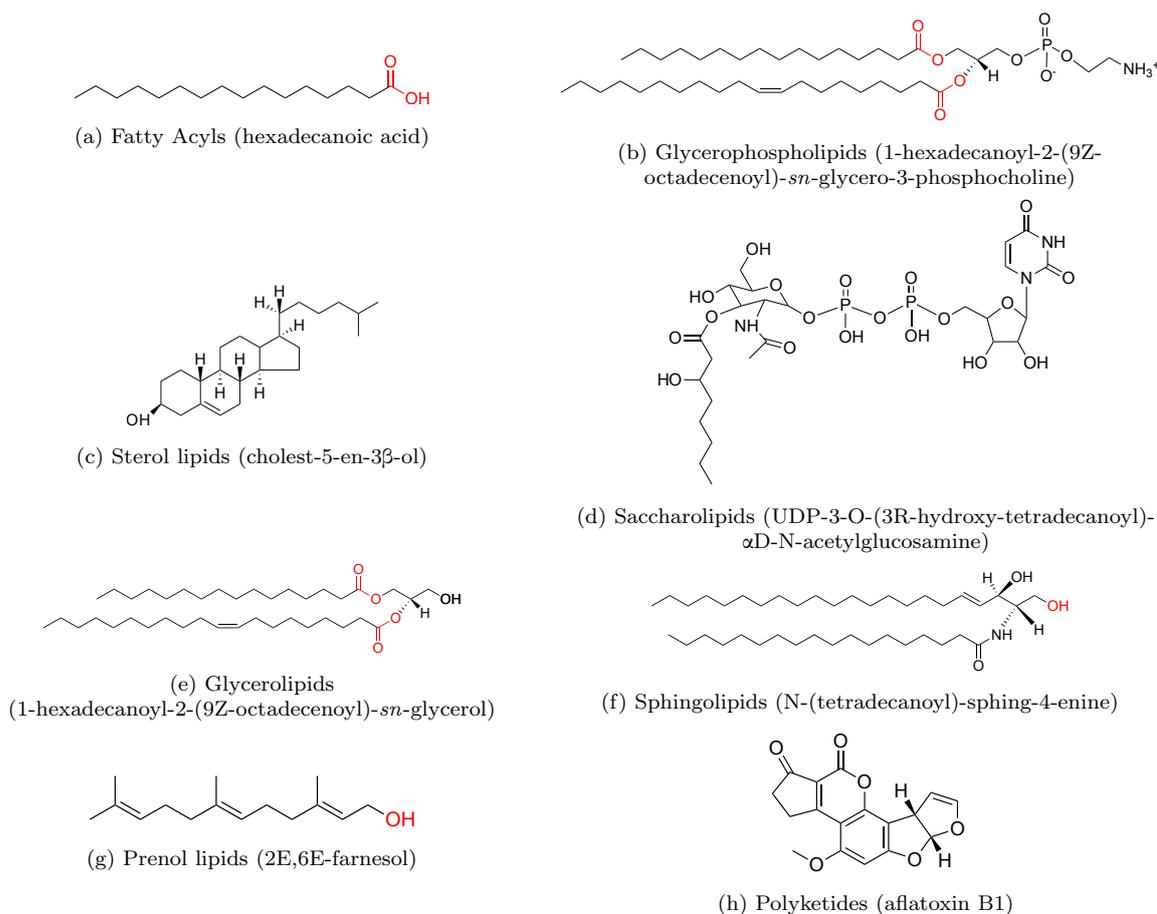


Figure 1: Representative structures for each lipid class; adapted from [8]

## 1.2 Lipid Notation

The shorthand notation described by Liebisch *et al.* [10] is building upon the LIPID MAPS<sup>®</sup> terminology. For lipids where the specific fatty acids are unknown, the notation *lipid class abbreviation followed by number of C-atoms:number of double bonds* is used (for example PC(36:6)). If the fatty acids linked to the glycerol are known but the *sn*-position is unknown, the separator `_` is used between the fatty acids (for example PC(18:1\_18:2)) and if the *sn*-position is known, the separator `/` is used (for example PC(18:1/18:2)) [10]. For the annotation of mass spectra, a nomenclature for lipids and lipid fragments was proposed by Pauling *et al.* [11]. Intact precursor ions are annotated based on the general agreement that uncharged molecules are represented by the symbol M and that charged derivatives corresponding to loss or gain of an adduct are denoted as [M-adduct]<sup>charge</sup> or [M+adduct]<sup>charge</sup>, respectively. In case of lipids, this convention is often adapted by substituting M with the shorthand notation for the lipid molecules (for example [PC(34:1)+H]<sup>+</sup>). Further, the authors propose a three-step procedure for shorthand notation of lipid fragment ions [11]. The steps comprise:

1. Detected fragment ion *m/z* values are first recapitulated using mass-balanced chemical reactions showing putative structures of both charged and neutral fragments
2. These fragments are then annotated using fragment type-specific annotation rules
3. Prioritizing the nomenclature to use for shorthand notation of detected fragment ion *m/z* values is based on fragment type, charge and mass difference between charged fragments and composites of neutral fragments

## 1.3 Lipid Analysis

The analysis of lipids has fascinated researchers for many decades. Back when analysis was mostly restricted to gas chromatography and thin-layer chromatography, large scale analysis had been very tedious [12]. Since then, breakthroughs in mass spectrometry (MS) technologies have led to an increase in mass accuracy and resolution as well as advances in ionization modalities, allowing contemporary lipidomics approaches to qualitatively and quantitatively analyze the entire complement of lipids in biological samples [13], consequently empowering researchers to determine structures, functions and interactions of lipids as well as dynamics and changes of the entire lipidome.

Lipidomics approaches should adhere to rigorous protocols: Once obtained, the samples should be snap-frozen and stored at ultra-low temperatures until use [14]. Sample preparation has a major impact both on the quality as well as on the sample throughput in lipidomics, whereby different types of sample matrices may require different types of sample preparation protocols [15]. Traditionally, isolation of lipids has relied on extraction procedures established by Folch *et al.* [16] or Bligh and Dyer [17]; however, the development of protocols using the less toxic methyl-tert-butyl ether (MTBE) are now providing an alternative [18].

As stated before, mass spectrometry is the workhorse of lipidomics, whereby two different approaches are widely used: Shotgun MS and MS coupled with chromatography. Hu and Zang [19] give an excellent overview of mass spectrometry-based lipidomics. In case of shotgun MS, no previous separation is performed and the lipid extract is directly infused into the mass spectrometer for analysis. It is less time consuming, more convenient and reproducible than other methods. In contrast, chromatographic coupled methods are necessary for extensive lipidomics analysis of complex biological samples – reducing matrix effects, separating lipid isomers and enriching low-abundance lipid molecules. Commonly used chromatographic methods are gas chromatography and ultra-high performance liquid chromatography. State of the art technologies include high-resolution MS, which provides accurate mass values of lipid molecules and their product ions, tandem mass spectrometry (MS<sup>2</sup>), which makes use of precursor ions, product ions and neutral loss scans, and the soft ionization technique ESI [19].

## 1.4 Galactolipids

Galactolipids are a part of the category of glycerolipids (figure 1e), in which the lipid carries one or more galactoses respectively, attached to diacylglycerol (DG) [20] at the *sn*-3 position via acetal linkage and two fatty acyl groups attached to the *sn*-1 and *sn*-2 positions [21]. They are the most abundant lipids in nature with monogalactosyldiacylglycerol (figure 2a) and digalactosyldiacylglycerol (DGDG) (figure 2b) constituting about 75% of total membrane lipids in plants [22]. In case of sulfoquinovosyldiacylglycerol (SQDG) (figure 2d), the chemical structure is characterized by two nonpolar fatty acyl chains, with

various degrees of unsaturation, bonded to the glycerol backbone's *sn*-1 and *sn*-2 positions, and a polar head group represented by a sulfoquinovose molecule [23]. SQDGs are relatively abundant sulfolipids specifically associated with photosynthetic membranes of higher plants, mosses, ferns, algae and most photosynthetic bacteria [23, 24]. The importance of galactolipids in green tissue is highlighted by the fact that those plants are less dependent on the precious nutrient phosphate than other eukaryotes with predominantly phospholipid-containing membranes [22]. In addition to galactoses, galactolipids may comprise estolides, extending the class considerably [25]. As stated by Isbell [26], estolides are natural and synthetic compounds derived from fats and oils. The estolide structure is identified by the secondary ester linkage of one fatty acyl molecule to the alkyl backbone of another fatty acid fragment. The estolide number (EN) is indicating the extent of oligomerization of the molecule. Estolides can be free acids, esters or found within a triglyceride structure [26]. Examples of estolides of DGDG can be seen in figure 2e, where one additional fatty acyl molecule is attached to DGDG (DGDG-EN1) and in figure 2f where two additional fatty acyl molecules are attached to DGDG (DGDG-EN2).

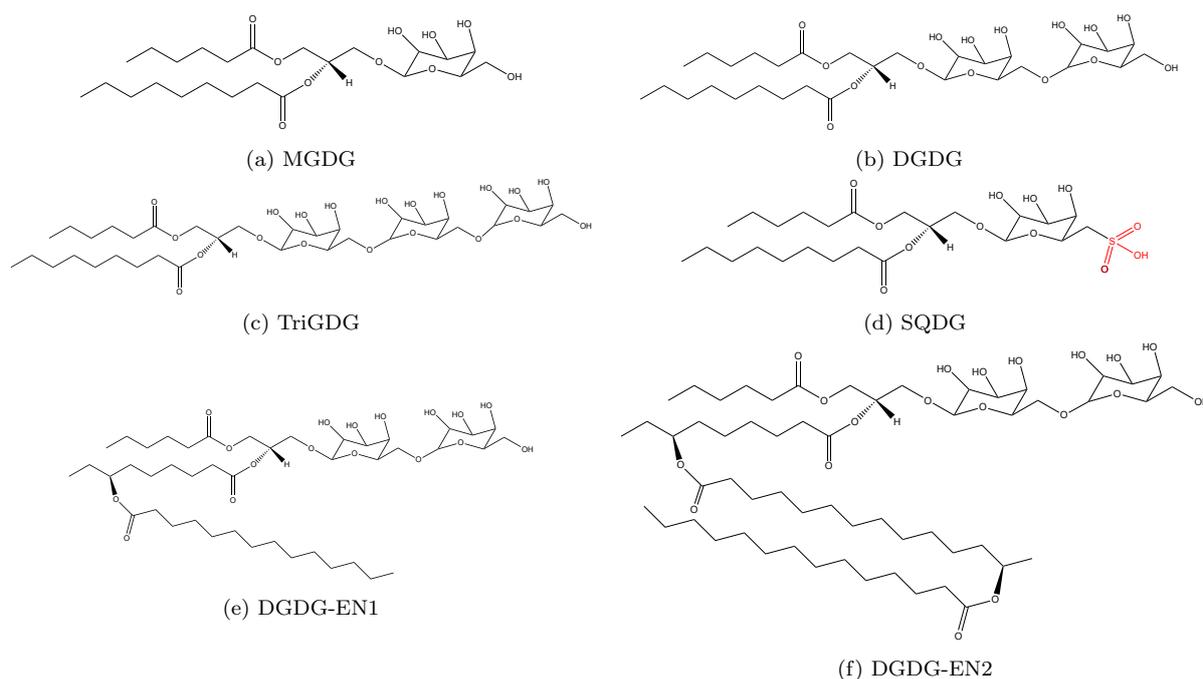


Figure 2: Structures of important galactolipids; adapted from [25]

## 1.5 Oxidative Lipidomics

Present-day challenges in lipid research include the newly emerging field of oxidative lipidomics, where lipids that were modified under oxidative stress - by a process called lipid peroxidation - are investigated. It has become evident that these modifications of lipids are critical to a number of cellular functions and disease states [27]. Up to now, analysis remains elusive due the vast scope of potential oxidized lipid species, low abundances and scarcity of modification-specific studies of MS<sup>2</sup> fragmentation spectra.

### 1.5.1 Lipid Peroxidation

Lipid peroxidation is a complex mechanism which can be separated into three distinct phases: initiation, propagation and termination [28]. Initiation is mediated either by free radicals that cause hydrogen abstraction, or by enzymes including lipoxygenases, cyclooxygenases, and various cytochromes [29], whereby the non-enzymatic (autocatalytic) pathway is potentially more devastating [30]. During the next step, propagation, the lipids themselves become radicals and mediate new peroxidation reactions and finally, during termination, stable molecules are formed [28]. Due to the low C-H bonding energy adjacent to double bonds, unsaturated fatty acids are major targets for modification under oxidative stress [31], and the rate of oxidation depends on the degree of unsaturation (i.e rate increases with number of double bonds) [32]. The reaction results in a wide range of oxidized products that depend on the nature of the oxidant species; nitric oxide radicals for example, form nitrated lipids [31, 33]. Products can be further

divided into long-chain products, which are products that preserve the original chain length, truncated products, formed by cleavage of fatty acyl chains [34] as well as full chain length products with rearrangements like cyclizations [35]. The oxidation of unsaturated fatty acids is accelerated by exposure to light, as ultraviolet light irradiation can produce free radicals [28]. In biological systems, photooxidation takes an important role in tissues like the retina, where light damage induces the peroxidation reaction [36]. Lipid peroxidation following the free radical mechanism has been observed for free fatty acids as well as esterified fatty acids [37, 38]. Prevalent oxidation moieties include oxo-, hydroxy- and hydroperoxy-modifications, but many others are possible [31] (table 1).

Table 1: Common oxidation moieties

Prefix	Structure
oxo-	
keto-	
epoxy-	
furan-	
hydroxy-	
epidioxy-	
carboxy-	
hydroperoxy-	
cyclopentane-	
bromo-	
chloro-	
fluoro-	
iodo-	
nitro-	

R - rest of the acyl chain; in case of esterified FA, esterified part

R' - rest of the acyl chain; non-esterified carboxy part

### 1.5.2 Oxidized Lipids

In plants, elevated levels of oxidized lipids are known to be connected to drought-induced leaf damage [39], photooxidative stress [40], metal-induced root damage [41], pathogen stress [42], and seed aging [30]. The biological role of oxidized (phospho)lipids in mammals depends on the location and nature of changes [34], but roles in age-related and chronic diseases [43], atherosclerosis [44], and inflammation and immune response [45] are topics of research and mutagenic, carcinogenic and cytotoxic properties are thought to be connected to elevated levels of oxidized lipids [46]. In addition, the presence of oxidized (phospho)lipids in biological membranes induces changes in physical properties such as viscosity, which can have an impact on the integrity of the membrane [34, 47], causing apoptotic events [34, 48]. In foods, oxTGs are a major cause of deterioration in quality and nutritive value [49].

At present, the number of oxidized lipids represented by entries in databases like LIPID MAPS<sup>®</sup> is still marginal: Structures of only 47 oxidized phosphatidylcholines (oxPCs), 52 oxidized phosphatidylethanolamines (oxPEs), 36 oxidized phosphatidylserines (oxPSSs), 36 oxidized phosphatidylinositols (ox-

PIs), 36 oxidized phosphatidylglycerols (oxPGs), 36 oxidized phosphatidic acids (oxPAs) and three oxidized cardiolipins (oxCLs) are currently available. Oxidized galactolipids (oxGLs), triacylglycerols (oxTGs), diacylglycerols (oxDGs) and other oxidized lipids, are currently not present.

**Oxidized Phospholipids** Analysis of the oxidation products of phospholipids is a more challenging task than the analysis of corresponding parent phospholipids due to the complexity of the oxidation products, the sensitivity of the methods, and instability of some of the oxidation products [50]. In a study on phosphatidylcholines, lipid species were artificially oxidized *in vitro* and analyzed by LC-MS<sup>2</sup> analysis, as well as techniques for higher structural resolution [51]. In the study, the formation of oxidized products resulted from the oxidation reaction of the *sn*-2 substituent as the *sn*-1 substituent was a saturated fatty acid (saturated fatty acids remain unaffected during radical peroxidation reaction due to the lack of allylic hydrogen atoms) [51]. Controlled oxidation studies with other PC species have been conducted and came to similar conclusions [52–57], but have the same inherent limitation, as they use PCs with saturated fatty acids at *sn*-1 position. Studies on *in vitro* phospholipid oxidation have been conducted as well, with biological samples including humans [58] and mice [59], whereby the oxidized phospholipids (oxPLs) couldn't be structurally determined in the former study, because of the low abundance of oxPLs generated by human platelets [58,59], and about 20 molecular species of oxPLs have been detected in the study on mice. Further, it has been shown that the high chemical stability of oxidized short-chain aldehydes of PCs allow additional reactions with amino groups present in peptides, proteins and PEs leading to even more complex products [34,60].

**Oxidized Triacylglycerols** Oxidized triacylglycerols have been found to include hydroperoxides, hydroxides, epoxides, epidioxides, hydroperoxy-epidioxides hydroxyl-epidioxides, as well as mono-, bis- and tris-hydroperoxides (hydroperoxide modifications on one, two and all three chains respectively) and keto-derivatives [38].

### 1.5.3 Occurrence

Given that oxidized lipids often activate metabolic pathways, they are usually present in very low quantities [29]. Analysis of low abundance oxidized lipids in the presence of high abundance non-oxidized structural lipids is a daunting task [29], also, unlike oxidized fatty acids, oxidized lipids are hardly available commercially [30]. Still, a study on naturally aged wheat seeds found clear trends regarding the number and abundance of oxidized lipids: Over all lipid species, they found a quite constant relation of native to oxidized lipids: If there was a large number of TGs, there were also many oxTGs, and, vice versa, a small number of lysophosphatidylglycerols (LPCs) coincided with a small number of oxLPCs [30]. More complex modifications are less frequent (i.e. 4 additional oxygen molecules) than less complex ones (i.e. 1 additional oxygen molecule). Note that, while the results stem from naturally occurring *in vivo* oxidation, the levels might not represent physiological conditions, as the seeds have been stored (partly in bad condition) for 15 years. Another limiting factor is that some fatty acid hydroperoxides such as 18:2[OOH] are not stable *in vivo*, but are the substrates of various enzymes like glutathione peroxidase and phospholipase, and readily metabolized and excreted, meaning, that the amount of fatty acid hydroperoxides does not necessarily reflect the extent of lipid peroxidation [61–63]. Nonetheless, oxidative lipidomics provides a platform for the identification of new (oxidized) lipid mediators and allows the study of the underlying mechanism in lipid-mediated cellular signaling [29].

## 1.6 Bioinformatics Tools

One of many remaining challenges concerns the high-throughput annotation of the recorded MS-data. While there are some high-quality bioinformatics tools available, many of them lack a convincing approach to annotate a vast variety of lipids, while keeping false positives low and being adaptable to different MS-platforms.

In light of other tools shortcomings, Lipid Data Analyzer (LDA) was developed, utilizing a novel 3D algorithm. LDA employs decision rulesets to enable automated and reliable annotation of lipid species and their molecular structures in high-throughput data from chromatography coupled tandem mass spectrometry [64]. LDA uses a targeted approach: It first scans the MS<sup>1</sup> spectra searching only precursor ions, previously defined in a mass list containing exact masses of the desired lipids. In a second step, the MS<sup>2</sup> fragmentation patterns of the precursor ions are analyzed, whereby identification is done by utilizing so-called decision rulesets. In LDA, identification is performed at different levels, the first being

by precursor mass only. Though, as lipids of completely different structures might have the same mass, this method may lead to a high number of false positives, because no structural information can be derived from the MS<sup>1</sup> identification. Here, the fragmentation patterns of the MS<sup>2</sup> spectra come into play, whereby identification of headgroup, headgroup and fatty acyl constituents, and headgroup, fatty acyl constituents and fatty acyl position may be possible.

In contrast to LDA, LipidMatch (LM), an automated workflow for rule-based lipid identification, is using an untargeted approach, where the MS raw data is tested against the whole lipid fragmentation library of the tool, containing the most comprehensive lipid fragmentation libraries of freely available software, when ranked by the number of lipid types [65]. LipidMatch Flow (LMF) [66] builds upon LipidMatch and adds automatized file conversation and peak-picking as well as blank filtering to it's list of features. All three tools allow user-generated rules, for coverage of additional lipids. While LDA provides a built-in interface to edit the rules, saved as *txt* files, LipidMatch doesn't provide an interface, and uses *csv* files.

Another tool, LPPtiger, specializes in the analysis of oxidized lipids. It's an open-source software tool for identification of oxidized phospholipids (oxPL) from data-dependent LC-MS datasets and combines three unique algorithms to predict oxidized lipidome, generate oxPL spectra libraries, and identify oxPLs from tandem MS data using parallel processing and a multi-scoring identification workflow [67]. LPPtiger relies on sample-specific phospholipid (PL) lipidome with a defined, discrete fatty acid composition [67], to predict possible oxidized lipids and generate *in silico* spectra of oxidized lipids.

## 1.7 Motivation and Aims

The importance of qualitative and quantitative analysis of lipids is highlighted by considering the vast implications a change in the lipidome can have. However, due to the enormous diversity of lipids, most data analysis tools cover only a marginal range of lipid classes. In an effort to reduce this shortcoming, this work extends the lipid species covered by Lipid Data Analyzer to galactolipids and oxidized lipids by following the specific aims:

1. Analyzing datasets with LipidMatch, LipidMatch Flow and LPPtiger to get an overview of lipid species in the datasets that are not implemented by Lipid Data Analyzer
2. Generating LDA appropriate mass lists for galacto- and oxidized lipids
3. Extending LDA decision ruleset to cover galacto- and oxidized lipids
4. Extending the LDA source code to allow identification of oxidatively modified fatty acyl chains
5. Analyzing datasets with the extended LDA version
6. Comparison of LipidMatch-, LipidMatch Flow-, LPPtiger- and LDA results for galacto- and oxidized lipids

## 2 Methods

### 2.1 Tools

MSConvertGui 3.0.1899.0 (part of the ProteoWizard package [68]) was used for conversion of raw data to *ms2*, *mzXML* and *mzML* files. MZmine 2.23 [69] was used for generating a peak-area-list according to the LipidMatch 2.0.2 [65] manual, with the help of the MZmine batch file included in the LipidMatch directory. The files were then analyzed using the LipidMatch R script [70]. LipidMatch Flow [66, 71] 0.0.2 was used to analyze the raw files utilizing its automated workflow, with an additional built-in blank filtering step. The mass list for Lipid Data Analyzer 2.6.3\_3 [64] was generated using R 3.3.3 [72] in RStudio 1.1.383 [73] with the packages XLConnect [74], janitor [75], plyr [76], pracma [77] and data.table [78]. LipidPioneer 1.0 [79] was used to partly verify the correct masses. SeeMS 3.0.18264.0 (part of the ProteoWizard package [68]) was used to verify LipidMatch identifications. DRAWBOARDpdf 5.8.210.0 [80] was used to manually annotate spectra when needed. Eclipse IDE 4.11.0 was used to edit the LDA’s Java source code. LPPtiger (hotfix2019-version) [67] was used to benchmark LDA’s results of oxidized datasets.

### 2.2 Datasets

**Dataset 1: Mouse Liver** As the first dataset, a total of 28 files from three biological samples including files *020.liver2-1.Orbitrap\_HCD\_pos* to *036.liver2-3.Orbitrap\_HCD\_pos* (positive ion mode) and *020.liver2-1.Orbitrap\_HCD\_neg* to *036.liver2-3.Orbitrap\_HCD\_neg* (negative ion mode) were taken from a biological study on mouse liver [64] with the MetaboLights identifier MTBLS396 [81]. Here, chromatographic separation of lipids was performed on a Waters BEH C8 column (Waters Corporation, Milford, MA, USA), thermostated to 50 °C in a Dionex Ultimate XRS UHPLC system, and the used mass spectrometer was an Orbitrap Velos operated in HCD mode. The positive data has about 5500 and the negative data about 5900 scans per file.

**Dataset 2: Ryegrass Leaves** Files *20130726\_NP+ve\_003.raw* to *20130726\_NP+ve\_016.raw* were taken from a large-scale metabolomics study on ryegrass [82], with the MetaboLights identifier MTBLS66 for positive mode [83]. The MetaboLights identifier for the negative mode data is MTBLS68, here files *20130729\_NP-ve\_003.raw* to *20130729\_NP-ve\_003.raw* were taken for the analysis [84]. In total 28 files from 13 biological samples were analyzed. The positive ion mode data has about 7200 MS scans per file and the negative ion data about 7700 scans per file. The data was recorded with a Thermo LC-MS system (Thermo Fisher Scientific, Waltham, MA, USA) consisting of an Accela 1250 quaternary UHPLC pump, a PAL auto-sampler fitted with a 15,000 psi injection valve (CTC Analytics AG., Zwingen, Switzerland) a 20 µl injection loop, and a Q-Exactive Orbitrap mass spectrometer with electrospray ionization.

**Dataset 3: Oxidized PC Standard** A single file originating from one sample (waters raw; negative ion mode) was taken from a study on oxidized PC standards [67], which was kindly provided by Maria Fedorova and Zhixu Ni. The investigated sample contained oxidized products of PC(16:0\_18:1) and PC(16:0\_18:2) that were manually verified. Here, an acquity UPLC M-class (Waters Corporation, Milford, MA, USA) was coupled online to a Synapt G2-Si mass spectrometer equipped with an ESI source (Waters Corporation, Milford, MA, USA). The data includes 3234 scans.

**Dataset 4: Oxidized Wheat Seeds** The raw data (81 files in *mzXML* format originating from 10 biological samples; positive ion mode) was taken from a study on oxidized wheat seeds [30] already mentioned in section 1.5.2 and kindly provided by David Riewe. Analytes were separated by a 1290 UHPLC device (Agilent Technologies, Santa Clara, CA, USA) using a C8 reverse-phase column and mass spectral analysis was conducted using a Bruker Maxis HD device upgraded with a Maxis II detector (Bruker Corporation, Billerica, MA, USA). The scans per file range from about 1500 to 2000.

## 2.3 Analysis Settings

**Dataset 1: Mouse Liver** To analyze dataset 1, LDA was run with the settings *Orbitrab\_velos\_pro\_HCD*, *Fragmentation Selection 1: +25*, and *Fragmentation Selection 2: -50*, while LipidMatch was operated with its standard settings. The LDA mass list in positive ion mode included the classes TG, DG, PC, PE, PS, PI, LPC, LPE, LPS, SM, Cer, P-PC, P-PE, PG, MGMT, MGDG, DGMG, DGDG, DGDG-EN1, DGDG-EN2, TriGDG, TetraGDG, SQMG and SQDG. The LDA mass list in negative ion mode included the classes PC, LPC, PE, LPE, PS, LPS, PI, PG, SM, Cer, P-PC, P-PE, MGMT, MGDG, DGMG, DGDG, DGDG-EN1, DGDG-EN2, TriGDG, TetraGDG, SQMG and SQDG. All mentioned classes were covered from (10:0) to (24:6) per fatty acyl chain. An exemplary mass list for MGDG can be seen in supplementary tables S9-S14 for positive ion mode and supplementary tables S15-S20 for negative ion mode.

**Dataset 2: Ryegrass Leaves** For dataset 2, the settings *Orbitrab\_exactive*, *Fragmentation Selection 1: 25* and *Fragmentation Selection 2: -30* were used for LDA analysis. LipidMatch was operated with its standard settings. The LDA mass list for dataset 2 was the same as the one used for dataset 1.

**Dataset 3: Oxidized PC Standard** For LDA analysis the setting *Orbitrab\_velos\_pro\_HCD*, *noIntensity* with *ms2PrecursorTolerance* set to 0.1 proved to be the most reliable option, but is by no means optimized for this dataset. A LDA mass list was generated for oxidation products between PC(34:0) and PC(34:2) with up to three additional OH-modifications and up to two additional O-modifications (supplementary table S21). For generating the in-silico oxidation in LPPtiger, all settings were used at their maximum option (i.e. *Oxidation-level: 3*, *Max modification sites: 8*, with *OAP*, *OCP*, *Lyso OAP* and *Lyso OCP* as well as *Prostanes* boxes checked). For the identification itself, all standard settings were kept as they were, except for Overall score filter which was set to >70%. In Lipid Match, *Retention Time Window* was set to 0.6, *ppm window* to 200, *Mass accuracy window* to 0.2 and *MS/MS Isolation Window* to 2.

**Dataset 4: Oxidized Wheat Seeds** The setting *Orbitrab\_velos\_pro\_HCD*, *noIntensity* proved to be the most reliable option for LDA analysis, with *ms2PrecursorTolerance* set to the standard value of 0.01. The LDA mass list was created from the identifications reported in the paper by Riewe *et al.* [30], thus included 173 TG species, 87 DG species, seven MG species, seven PC species, ten PE species, seven PG species, five PI species, 17 LPC species, seven LPE species, two LPG species, 29 DGDG species and 13 MGDG species. The full mass list can be seen in supplementary tables S22-S34.

## 2.4 Hardware Environment

All tasks including programming and analysis were conducted using the latest version of *Microsoft Windows 10* running on a *Microsoft Surface Pro 3* device with a two core (1.7 GHz) *Intel Core i7-4658U* CPU, 8GB of LPDDR3 RAM (1600 MHz) and an integrated *Intel HD 5000* graphics card.

## 3 Results

As a starting point, dataset 1 was analyzed using an untargeted approach with the help of the LipidMatch tool [65]. The workflow consists of converting the raw data files to *ms2* and *mzXML* files with the help of MSConvertGui, extracting a peak-area-list out of the *mzXML* files utilizing MZmine and a batch file contained in the LipidMatch directory, and eventually processing of the peak-area-list and the *ms2* files using the LipidMatch R script. In the next step, interesting lipid subclasses included in the LipidMatch results – but currently not covered by Lipid Data Analyzer – were chosen for further investigation. Subclasses oxTG and DGDG were chosen for further work. For these classes, an R script was written, creating a theoretical mass list of the precursor ions in LDA appropriate form. Sum formulas of headgroups were deduced from structure drawings (see fig 2) and partly verified with the LipidPioneer tool [79], whereby the accurate masses for the atoms constituting the lipids were taken from an appropriate paper on atomic weights [85]. In the course of the work, the subclasses MGDG, TriGDG, TetraGDG and SQDG as well as MGMG, DGMG, SQMG and DGDG-EN1, DGDG-EN2, were also implemented. Extensive literature research was conducted, to understand the MS<sup>2</sup> fragmentation patterns, which were consequently used to formulate the rules. In the next step the dataset was analyzed using LDA with the new mass list and extended ruleset. As no MS<sup>2</sup> patterns matched those for galactolipids, a second dataset was found by searching the MetaboLights [86] repository, and was subsequently analyzed.

For the next phase, oxidative lipidomics was chosen as the topic of interest. Literature research led to an understanding of the oxidation mechanism and the nature of the most widespread oxidation modifications, and a concept for LDA analysis was developed. Ultimately, new mass lists were implemented both for target lipids and their constituent fatty acyls and LDA source code was modified. To prove the validity of the implementation, two new datasets were organized from research groups working on oxidized lipids and a benchmark test with LPPtiger was conducted.

To standardize results for this report, a script was written to remove all non-MS<sup>2</sup> matches from LDA results, another script was written that cleans the LipidMatch results to only show unique lipid species confirmed by MS<sup>2</sup> (on fatty acyl level) and yet another script was written to summarize LDA results, of all analyzed files, and clean them to only show unique lipid species confirmed by MS<sup>2</sup> (on fatty acyl level). All of the scripts were implemented using R.

### 3.1 LipidMatch Results

#### 3.1.1 Dataset 1 - Mouse Liver

LipidMatch identified a total of 186 different lipid species in dataset 1. The most frequently reported class was TG with a total of 45 distinct annotations (see table 2), where the [M+NH<sub>4</sub>]<sup>+</sup> adduct was more often reported (45 times) than the [M+Na]<sup>+</sup> adduct (17 times). All mentioned species were identified by precursor ion mass and confirmed by MS<sup>2</sup> fragmentation patterns including fatty acyl chain information (fatty acyl level). The 115 identifications in positive ion mode, include the subclasses DG, LPC, LPE, PC, PE, PEtOH, P-PC, SM and TG as well as Coenzyme Q9 (supplementary table S1). In negative ion mode, LipidMatch identified 71 lipid species including Cer-NS, CL, LPC, LPE, oxPG, PC, PE, PI, Plasmenyl-PE and PS (supplementary table S2). LipidMatch Flow analysis of the same dataset reported a significantly lower number of lipids than LipidMatch: 85 in positive ion mode and 37 in negative ion mode.

Table 2: Identified lipid classes by LipidMatch in dataset 1; LipidMatch Flow identifications inside round brackets

Lipid Class	Adduct	Unique Species per Adduct	Unique Species
Cer-NS	$[M+HCO_2]^-$	1 (0)	1 (0)
CL	$[M-2H]^-$	2 (0)	2 (0)
Co(Q9)	$[M+NH_4]^+$	1 (0)	1 (0)
DG	$[M+NH_4]^+$	4 (3)	4 (3)
LPC	$[M+H]^+$ $[M+HCO_2]^-$	5 (4) 5 (4)	7 (5)
LPE	$[M+H]^+$ $[M-H]^-$	1 (1) 4 (2)	4 (3)
oxPG	$[M-H]^-$	1 (0)	1 (0)
PC	$[M+H]^+$ $[M+Na]^+$ $[M+HCO_2]^-$	33 (21) 11 (8) 26 (16)	35 (31)
PE	$[M+H]^+$ $[M+Na]^+$ $[M-H]^-$	16 (12) 1 (1) 19 (9)	19 (17)
PEtOH	$[M+NH_4]^+$	1 (0)	1 (0)
PI	$[M-H]^-$	7 (3)	7 (3)
Plasmanyln-PC	$[M+H]^+$	2 (2)	2 (2)
Plasmenyl-PE	$[M-H]^-$	4 (2)	4 (2)
PS	$[M+H]^+$ $[M-H]^-$	0 (1) 2 (1)	2 (2)
SM	$[M+H]^+$	5 (1)	5 (1)
TG	$[M+Na]^+$ $[M+NH_4]^+$	17 (32) 45 (12)	45 (34)

### 3.1.2 Dataset 2 - Ryegrass Leaves

A total of 175 lipids were reported in dataset 2 by LipidMatch. LipidMatch Flow reported significantly less lipids in both positive and negative ion mode (31 and 66 vs. 76 and 100, respectively), whereby some species like AcCar were reported by LipidMatch but not LipidMatch Flow (table 3). The LipidMatch identifications in positive ion mode included the subclasses AcCar, DG, DGDG, LPC, LPE, MG, MGDG, oxLPC, oxTG, PA, PC, PE, PEtOH, PG, PS, So and TG (supplementary table S5). The analysis in negative ion mode led to 101 identifications including the subclasses AcylGlcADG, Cer-AP, CL, DGDG, FAHFA, HexCer-AP, LPA, LPC, LPE, MGDG, oxCL, oxPC, oxPE, oxPG, PA, PC, PE, PG, PI, PS and SQDG (supplementary table S6).

Table 3: Identified lipid classes by LipidMatch in dataset 2; LipidMatch Flow identifications inside round brackets

Lipid Class	Adduct	Unique Species per Adduct	Unique Species
AcCar	$[M+H]^+$	2 (0)	2 (0)
AcylGlcADG	$[M-H]^-$	1 (2)	1 (2)
Cer-AP	$[M+HCO_2]^-$	5 (3)	5 (3)
CL	$[M-2H]^-$	2 (1)	2 (1)
Co(Q9)	$[M+NH_4]^+$	1 (0)	1 (0)
DG	$[M+NH_4]^+$	7 (6)	7 (6)
DGDG	$[M+NH_4]^+$	2 (1)	8 (5)
	$[M+HCO_2]^-$	8 (5)	
FAHFA	$[M-H]^-$	5 (1)	5 (1)
HexCer-AP	$[M+HCO_2]^-$	4 (3)	4 (3)
LPA	$[M-H]^-$	3 (2)	3 (2)
LPC	$[M+H]^+$	3 (3)	3 (3)
	$[M+HCO_2]^-$	2 (2)	
LPE	$[M+H]^+$	4 (3)	4 (3)
	$[M-H]^-$	3 (3)	
MG	$[M+NH_4]^+$	1 (1)	1 (1)
MGDG	$[M+Na]^+$	4 (4)	7 (6)
	$[M+NH_4]^+$	2 (1)	
	$[M+HCO_2]^-$	4 (3)	
oxCL	$[M-2H]^-$	1 (0)	1 (0)
oxLPC	$[M+H]^+$	2 (0)	2 (1)
	$[M+Na]^+$	1 (1)	
oxPC	$[M+HCO_2]^-$	4 (1)	4 (1)
oxPE	$[M-H]^-$	3 (3)	3 (3)
oxPG	$[M-H]^-$	1 (1)	1 (1)
oxTG	$[M+NH_4]^+$	4 (2)	4 (2)
PA	$[M+H]^+$	7 (0)	11 (9)
	$[M-H]^-$	11 (9)	
PC	$[M+H]^+$	2 (0)	9 (5)
	$[M+HCO_2]^-$	9 (5)	
	$[M+Na]^+$	1 (1)	
PE	$[M+H]^+$	8 (1)	12 (8)
	$[M+Na]^+$	2 (1)	
	$[M-H]^-$	10 (8)	
PEtOH	$[M-H]^+$	1 (0)	1 (0)
PG	$[M+H]^+$	4 (2)	14 (9)
	$[M+NH_4]^+$	3 (1)	
	$[M-H]^-$	13 (8)	
PI	$[M-H]^-$	4 (2)	4 (2)
PS	$[M+H]^+$	2 (0)	4 (1)
	$[M-H]^-$	3 (1)	
So	$[M+H]^+$	2 (1)	2 (1)
SQDG	$[M-H]^-$	4 (4)	4 (4)
TG	$[M+Na]^+$	15 (3)	17 (3)
	$[M+NH_4]^+$	15 (1)	

### 3.1.3 Dataset 3 - Oxidized PC Standard

LipidMatch was able to identify three oxidized lipid species in dataset 3 (table 4), though it doesn't officially support Waters MS-systems. Here, one of the three identifications seems to be a false positive, as the m/z value doesn't match the molecule's weight properly. As no blank files were present, the dataset couldn't be analyzed using LipidMatch Flow.

Table 4: Lipid molecular species identified in dataset 3 by LipidMatch; putative false positives are marked in gray; species names adapted to LDA shorthand notation

Lipid Species	Adduct	Lipid Molecular Species
oxPC(34:2[OH])		oxPC(16:0-18:2[OH])
oxPC(34:1[OH])	[M+HCO <sub>2</sub> ] <sup>-</sup>	oxPC(16:0-18:1[OH])
oxPC(34:2[3O])		oxPC(16:0-18:2[3O])

### 3.1.4 Dataset 4 - Oxidized Wheat Seeds

As LipidMatch only lists support for Thermo and Agilent MS-systems, it exited preemptively without returning results, trying to analyze dataset 4 which was recorded using a Bruker instrument. LipidMatch Flow couldn't handle the dataset either, as no blank files were present.

## 3.2 LPPtiger Results

LPPtiger was used to benchmark LDA's results of oxidized lipids and was consequently not used on datasets 1 and 2, and instead only on datasets 3 and 4.

### 3.2.1 Dataset 3 - Oxidized PC Standard

With the preexisting knowledge of the sample containing PC(16:0/18:1) and PC(16:0/18:2), the sample was analyzed using LPPtiger. The tool reported eight lipids in dataset 3 (table 5), with seven of the identified lipids carrying oxidative modifications.

Table 5: Lipid molecular species identified in dataset 3 by LPPtiger; species names adapted to LDA shorthand notation

Lipid Species	Adduct	Lipid Molecular Species
oxPC(34:0[O])		oxPC(16:0/18:0[O])
oxPC(34:0[O,OH])		oxPC(16:0/18:0[O,OH])
oxPC(34:0[2OH])		oxPC(16:0/18:0[2OH])
oxPC(34:1)		oxPC(16:0/18:1)
oxPC(34:1[O])	[M+HCO <sub>2</sub> ] <sup>-</sup>	oxPC(16:0/18:1[O])
oxPC(34:1[2O])		oxPC(16:0/18:1[2O])
oxPC(34:1[OH,OOH])		oxPC(16:0/18:1[OH,OOH])
oxPC(34:2[O,OOH])		oxPC(16:0/18:2[O,OOH])

### 3.2.2 Dataset 4 - Oxidized Wheat Seeds

Dataset 4 couldn't be analyzed by LPPtiger as it only works for data recorded in negative ion mode.

## 3.3 Generation of LDA Mass Lists

LDA uses Excel spreadsheets to define masses of molecules to be searched for in the MS<sup>1</sup> spectrum. A valid mass list contains columns for name, double bonds and chemical formula as well as columns for each possible adduct of the lipid subclass; each subclass is defined in a separate worksheet. In an optional column, a retention time can be defined for the molecule. LDA uses separate mass lists for positive- and negative ion mode respectively. Examples of mass lists can be found in section 5.2. In this project, an R script was written, that creates an appropriate mass list for the new lipid subclasses. The masses for atoms constituting the lipids were taken from an appropriate paper on atomic weights [85]. For lipid species with a single chain, a minimum chain length of 6 carbon atoms was assumed. It was further

assumed, that chains comprised of up to nine carbon atoms can have no double bonds (9:0), chains comprised of ten or eleven carbon atoms can have one double bond (11:1), chains comprised of 12 to 15 carbon atoms can have up to 4 double bonds (15:4) and chains between 16 and 28 carbon atoms can have up to 6 double bonds (28:6). For lipid species with two chains it was assumed that the two chains together must be comprised of at least twenty carbon atoms (20:0). It was further assumed that following species exist:

I (20:x) to (29:x) with  $x = 1 - 4$

II (30:x) to (37:x) with  $x = 1 - 8$

III (38:x) to (39:x) with  $x = 1 - 10$

IV (40:x) to (48:x) with  $x = 1 - 12$

For lipid species with three chains it was assumed that the three chains together must be comprised of at least twenty-eight carbon atoms (28:0). It was further assumed that following species exist:

I (28:x) to (29:x) with  $x = 1 - 2$

II (30:x) to (39:x) with  $x = 1 - 4$

III (40:x) to (47:x) with  $x = 1 - 6$

IV (48:x) to (49:x) with  $x = 1 - 8$

V (50:x) to (53:x) with  $x = 1 - 10$

VI (54:x) to (57:x) with  $x = 1 - 12$

VII (58:x) to (59:x) with  $x = 1 - 14$

VIII (60:x) to (63:x) with  $x = 1 - 16$

IX (63:x) to (70:x) with  $x = 1 - 18$

Aforementioned assumptions were made as the same patterns are given in the mass list provided by the developers of LDA, which was attached to dataset 1. The masses were then calculated by adding the molecular weight of the respective headgroup atoms to the ones of the acyl chains, looping over each possible fatty acyl chain and double bond combination.

### 3.3.1 Galactolipids

Table 6 shows the structural composition of the implemented galactolipids and their possible adducts. For TriGDG and TetraGDG only ammonium adducts were reported in literature, but out of curiosity, also masses for sodium, formate and acetate adducts were calculated. It was further assumed, that the estolide- and MG-species can form the same adducts as their DG counterparts.

Table 6: Galactolipids - structure and adducts

Lipid Class	Fatty Aclys	Galactose Rings	C-atoms	O-atoms	S-atoms	H-atoms	Adducts
MGDG	2	1	chain-atoms + 9	10	0	C-atoms*2 - double-bonds*2 - 4	[M+NH <sub>4</sub> ] <sup>+</sup> , [M+Na] <sup>+</sup> , [M+Li] <sup>+</sup> [9, 25, 87, 88] [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup> [89]
DGDG	2	2	chain-atoms + 15	15	0	C-atoms*2 - double-bonds*2 - 6	[M+NH <sub>4</sub> ] <sup>+</sup> [25], [M+Na] <sup>+</sup> , [M+Li] <sup>+</sup> , [M+K] <sup>+</sup> [87] [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup> [89]
TriGDG	2	3	chain-atoms + 20	21	0	C-atoms*2 - double-bonds*2 - 8	[M+NH <sub>4</sub> ] <sup>+</sup> [25], [M+Na] <sup>+</sup> [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup>
TetraGDG	2	4	chain-atoms + 27	25	0	C-atoms*2 - double-bonds*2 - 10	[M+NH <sub>4</sub> ] <sup>+</sup> [25], [M+Na] <sup>+</sup> [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup>
SQDG	2	1	chain-atoms + 9	12	1	C-atoms*2 - double-bonds*2 - 4	[M-H+Na <sub>2</sub> ] <sup>+</sup> [90] [M-H] <sup>-</sup> [23]
DGDG-EN1	3	2	chain-atoms + 15	15	0	C-atoms*2 - double-bonds*2 - 8	[M+NH <sub>4</sub> ] <sup>+</sup> , [M+Na] <sup>+</sup> , [M+Li] <sup>+</sup> [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup>
DGDG-EN2	4	2	chain-atoms + 15	15	0	C-atoms*2 - double-bonds*2 - 10	[M+NH <sub>4</sub> ] <sup>+</sup> , [M+Na] <sup>+</sup> , [M+Li] <sup>+</sup> [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup>
MGMG	1	1	chain-atoms + 9	9	0	C-atoms*2 - double-bonds*2 - 2	[M+NH <sub>4</sub> ] <sup>+</sup> , [M+Na] <sup>+</sup> , [M+Li] <sup>+</sup> [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup>
DGMG	1	2	chain-atoms + 15	14	0	C-atoms*2 - double-bonds*2 - 4	[M+NH <sub>4</sub> ] <sup>+</sup> , [M+Na] <sup>+</sup> , [M+Li] <sup>+</sup> [M-H] <sup>-</sup> , [M+CH <sub>3</sub> CO <sub>2</sub> ] <sup>-</sup> , [M+HCO <sub>2</sub> ] <sup>-</sup>
SQMG	1	1	chain-atoms + 9	11	1	C-atoms*2 - double-bonds*2 - 2	[M-H+Na <sub>2</sub> ] <sup>+</sup> [M-H] <sup>-</sup>

### 3.3.2 Oxidized Lipids

New masses and chemical formulas for lipids with additional oxidative modifications are calculated on the fly during analysis according to table 7.

Table 7: Oxidation moieties: Mass formulas

Modification	Fatty Acyl (FA) Mass Formula
oxo-, keto-, epoxy-,furan-	FA + O - H <sub>2</sub>
hydroxy-	FA + O
hydroperoxy-	FA + O <sub>2</sub>
bromo-	FA + Br - H
chloro-	FA + Cl - H
fluoro-	FA + F - H
iodo-	FA + I - H
nitro-	FA + NO <sub>2</sub> - H

### 3.4 Extending LDA Rulesets

LDA uses simple text files for the formulation of the fragmentation rules. Each possible adduct defined in the mass list needs its own rule file, in order to make a MS<sup>2</sup> identification possible. The file is structured into four parts - *GENERAL*, *HEAD*, *CHAINS* and *POSITION*. The section *GENERAL* lists the amount of chains and other properties used for the whole of the molecule. The sections *HEAD* and *CHAINS* are further structured into a *!FRAGMENTS* and a *!INTENSITIES* part. The *!FRAGMENTS* part lists the fragment ions used for the identification of the headgroup and of the fatty acyl chains respectively. The *!INTENSITIES* part can be used to describe intensity relationships between different fragments. Each fragment and intensity equation can be set as mandatory (or not), in order for an identification to be made. Additionally, the option *other* is available for fragments, where fragments that are not supposed to occur in the desired class can be defined, thus this option can be used to keep false positive identifications lower. The last section *POSITION* can be used to define intensity relationships between chain fragments in order to identify which chain belongs to which *sn*-position.

#### 3.4.1 MGDG

**Ammoniated Adduct** Han [9] states, that in case of the ammoniated form of MGDG a complex fragmentation pattern evolves, with ions of  $[M+NH_4-NH_3]^+$  (denoted as *NL\_NH<sub>3</sub>*) corresponding to the loss of ammonia,  $[M+NH_4-NH_3-H_2O]^+$  (denoted as *NL\_NH<sub>3</sub>-H<sub>2</sub>O*) denoting the sequential loss of ammonia and water, as well as  $[M+NH_4-NH_3-Hex]^+$  (denoted as *NL\_NH<sub>3</sub>-Hex*) through the loss of ammonia and the hexose ring, and  $[M+NH_4-NH_3-H_2O-Hex]^+$  (denoted as *NL\_NH<sub>3</sub>-H<sub>2</sub>O-Hex*) through the loss of ammonia, water and the hexose ring. Further, two ions of  $[M+NH_4-NH_3-Hex-R_xCOOH]^+$  (*x*=1,2) (denoted as *NL\_Carboxy\_NH<sub>3</sub>-Hex*) corresponding to the loss of ammonia, the hexose ring and the respective fatty acid, are to be expected [9]. In the rule listed below, for the identification of the head group only *NL\_NH<sub>3</sub>-H<sub>2</sub>O-Hex* is mandatory, as the other fragments are low abundant. For identification of the chains, *NL\_Carboxy\_NH<sub>3</sub>-Hex* was used. The intensity equation is used to dismiss false positives, when a very high abundant PC-head fragment is present.

```
MGDG_NH4.frag
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=D*d+:d+
DoubleBondsFromName=D*d+:d+

[HEAD]
!FRAGMENTS
Name=Precursor           Formula=$PRECURSOR           Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3              Formula=$PRECURSOR-NH3       Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H2O          Formula=$PRECURSOR-NOH5      Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex          Formula=$PRECURSOR-C6H13O5N  Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H2O_Hex      Formula=$PRECURSOR-C6H15O6N  Charge=1 MSLevel=2 mandatory=true
Name=PChead_184_wrong    Formula=C5H15N04P            Charge=1 MSLevel=2 mandatory=other

!INTENSITIES
Equation=PChead_184_wrong<0.5*$BASEPEAK mandatory=true

[CHAINS]
```

```

!FRAGMENTS
Name=NL_Carboxy_NH3_Hex      Formula=$PRECURSOR-$CHAIN-C6H13O5N  Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH         Formula=$CHAIN-H3O2                  Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH              Formula=$CHAIN-OH                     Charge=1 MSLevel=2 mandatory=false
Name=Carboxy                  Formula=$CHAIN                          Charge=1 MSLevel=2 mandatory=false

```

**Lithiated Adduct** The lithiated form of MGDG yields comparatively simple fragmentation patterns, consisting of low abundant ions corresponding to lithiated hexose (denoted as *HexLi*) and two ions of  $[M+Li-R_xCOOH]^+$  ( $x=1,2$ ) (denoted as *NL\_Carboxy*) through the loss of the respective fatty acids [9]. Consequently, the lithiated hexose ion was used for head group identification and the fatty acid loss for the identification of the chains. The following code represents the rules for the identification of  $[MGDG+Li]^+$ , where the most important ones are colored in green.

```

MGDG_Li.frag

[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor      Formula=$PRECURSOR      Charge=1 MSLevel=2 mandatory=false
Name=HexLi          Formula=C6H10O5Li      Charge=1 MSLevel=2 mandatory=true

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy     Formula=$PRECURSOR-$CHAIN  Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH Formula=$CHAIN-H3O2        Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH     Formula=$CHAIN-OH         Charge=1 MSLevel=2 mandatory=false

```

**Sodiated Adduct** The fragmentation patterns are analogous to the lithiated adduct where Li gets substituted by Na [9]. It follows that the sodiated hexose ion (denoted as *HexNa*) was used for head group identification. Additionally, experiments regarding regioisomerism of the sodiated form came to the conclusion that the preferred fragmentation process is the loss of the acyl side chain from the *sn*-1 position compared to the one of the *sn*-2 position [91], with an intensity ratio of  $[M+Na-R_1COOH]^+$  to  $[M+Na-R_2COOH]^+$  ranging between 1.9 and 3.2 [92]. The following code shows the additionally formulated intensity rule for positional identification in green text.

```

MGDG_Na.frag

[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+
ChainCutoff=1%
RetentionTimeParallelSeries=true

[HEAD]
!FRAGMENTS
Name=Precursor      Formula=$PRECURSOR      Charge=1 MSLevel=2 mandatory=false
Name=HexNa          Formula=C6H10O5Na      Charge=1 MSLevel=2 mandatory=true

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy     Formula=$PRECURSOR-$CHAIN  Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH Formula=$CHAIN-H3O2        Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH     Formula=$CHAIN-OH         Charge=1 MSLevel=2 mandatory=false

[POSITION]
!INTENSITIES
Equation=NL_Carboxy[1]>NL_Carboxy[2]*1.9  mandatory=false

```

**Deprotonated Adduct** The deprotonated adduct  $[M-H]^-$  forms prominent peaks at the masses 235.0818 ( $C_9H_{15}O_7$ ) and 253.0923 ( $C_9H_{17}O_8$ ) [93], corresponding to  $[Glycerol+Hex+H_2O]^+$  (denoted as *GlycerolHexH<sub>2</sub>O*) and  $[Glycerol+Hex+H_4O_2]^+$  (denoted as *GlycerolHexH<sub>4</sub>O<sub>2</sub>*) respectively. The fragments  $[Glycerol+Hex_2+H_2O]^+$  (denoted as *GlycerolHex<sub>2</sub>H<sub>2</sub>O*) and  $[Glycerol+Hex_2+H_4O_2]^+$  (denoted as *GlycerolHex<sub>2</sub>H<sub>4</sub>O<sub>2</sub>*) were introduced as fragments of DGDG and used for filtering false positives via the intensity rules.

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+
ChainCutoff=1%

[HEAD]
!FRAGMENTS
Name=Precursor           Formula=$PRECURSOR           Charge=1 MSLevel=2 mandatory=false
Name=GlycerolHexH20      Formula=C9H1507              Charge=1 MSLevel=2 mandatory=true
Name=GlycerolHexH402     Formula=C9H1708              Charge=1 MSLevel=2 mandatory=true
Name=GlycerolHex2H20     Formula=C15H25012           Charge=1 MSLevel=2 mandatory=other
Name=GlycerolHex2H402    Formula=C15H27013           Charge=1 MSLevel=2 mandatory=other

!INTENSITIES
Equation=GlycerolHex2H20*1.5<GlycerolHexH20      mandatory=true
Equation=GlycerolHex2H402*1.5<GlycerolHexH402    mandatory=true

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy           Formula=$PRECURSOR-$CHAIN    Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_H20_OH       Formula=$CHAIN-H302          Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH           Formula=$CHAIN-OH            Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_H            Formula=$CHAIN-H              Charge=1 MSLevel=2 mandatory=true
```

**Formate and Acetate Adducts** For formate  $[M+HCO_2]^-$  and acetate  $[M+CH_3CO_2]^-$  adducts the same fragmentation rules as for the deprotonated adduct were assumed.

### 3.4.2 DGDG

**Ammoniated Adduct** The fragmentation patterns of the ammoniated adduct of DGDGs were reported as having two low abundant ions of  $[M+NH_4-NH_3-Hex]^+$  (denoted as *NL-NH<sub>3</sub>-Hex*) and  $[M+NH_4-NH_3-Hex-H_2O]^+$  (denoted as *NL-NH<sub>3</sub>-H<sub>2</sub>O-Hex*), as well as two fragment ions of  $[M+NH_4-NH_3-Hex_2]^+$  (denoted as *NL-NH<sub>3</sub>-Hex<sub>2</sub>*) and  $[M+NH_4-NH_3-Hex_2-H_2O]^+$  (denoted as *NL-NH<sub>3</sub>-H<sub>2</sub>O-Hex<sub>2</sub>*), and ions corresponding to  $[Glycerol+R_xCOOH-H_2O]^+$  ( $x=1,2$ ) (denoted as *CarboxyGlycerol-H<sub>2</sub>O*) [9,25]. In light of this, *NL-NH<sub>3</sub>-H<sub>2</sub>O-Hex<sub>2</sub>* was set as being mandatory for head group identification and *CarboxyGlycerol-H<sub>2</sub>O* was used for identification of the fatty acyl chains, as can be seen in the following code.

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor           Formula=$PRECURSOR           Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex         Formula=$PRECURSOR-C6H1305N  Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H20_Hex     Formula=$PRECURSOR-C6H1506N  Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex2        Formula=$PRECURSOR-C12H23010N Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H20_Hex2    Formula=$PRECURSOR-C12H25011N Charge=1 MSLevel=2 mandatory=true

[CHAINS]
!FRAGMENTS
Name=CarboxyGlycerol_H20 Formula=$CHAIN+C3H50         Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H20_OH       Formula=$CHAIN-H302          Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH           Formula=$CHAIN-OH            Charge=1 MSLevel=2 mandatory=false
```

**Lithiated Adduct** For the lithiated adduct of DGDG, fragment ions of  $[M+Li-Hex_2]^+$  (denoted as *NL-Hex<sub>2</sub>*) corresponding to the neutral loss of the di-hexose moiety, one or two ions of  $[M+Li-R_xCOOH]^+$  ( $x=1,2$ ) (denoted as *NL\_Carboxy*) corresponding to the loss of the respective fatty acid and one or two ions of  $[M+Li-Hex-R_xCOOH]^+$  ( $x=1,2$ ) (denoted as *NL\_Carboxy-Hex*) corresponding to the sequential loss of hexose and respective fatty acid as well as a cluster of lithiated hexose derivatives –  $[Li+Hex]^+$  (denoted as *HexLi*),  $[Li+Hex_2]^+$  (denoted as *Hex<sub>2</sub>Li*) and  $[Li+Hex_2+Glyceride]^+$  – have been observed [9]. Consequently, the fragments *Hex<sub>2</sub>Li* and *NL-Hex* have been chosen for identification of the headgroup, as seen in the following code. The fragments *NL\_Carboxy* and *NL\_Carboxy-Hex* have been chosen for the identification of fatty acyl chains. (For the Na adduct, these rules have been the most robust, while testing – and the DGDG adducts of Na yield the same fragmentation patterns as the Li adducts).

## DGDG.Li.frag

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor          Formula=$PRECURSOR          Charge=1 MSLevel=2 mandatory=false
Name=HexLi              Formula=C6H1005Li          Charge=1 MSLevel=2 mandatory=false
Name=Hex2Li             Formula=C12H20010Li        Charge=1 MSLevel=2 mandatory=true
Name=NL_Hex             Formula=$PRECURSOR-C6H1005 Charge=1 MSLevel=2 mandatory=true
Name=NL_Hex2            Formula=$PRECURSOR-C12H20010 Charge=1 MSLevel=2 mandatory=false

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy         Formula=$PRECURSOR-$CHAIN   Charge=1 MSLevel=2 mandatory=true
Name=NL_Carboxy_Hex    Formula=$PRECURSOR-$CHAIN-C6H1005 Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH    Formula=$CHAIN-H3O2        Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH        Formula=$CHAIN-OH           Charge=1 MSLevel=2 mandatory=false
```

**Sodiated Adduct** The sodiated adduct leads to the same fragmentation patterns as the lithiated adduct, where Li is substituted by Na. Additionally, studies concerning the *sn*-position of the acyl chains came to the conclusion that the loss of the chain at the *sn*-1 position is the preferred process, with intensity ratios of  $[M+Na-R_1COOH]^+$  to  $[M+Na-R_2COOH]^+$  ranging between 1.6 and 3.7 [92]. The intensity rule is shown in green text in the following code.

## DGDG.Na.frag

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor          Formula=$PRECURSOR          Charge=1 MSLevel=2 mandatory=false
Name=HexNa              Formula=C6H1005Na          Charge=1 MSLevel=2 mandatory=false
Name=Hex2Na             Formula=C12H20010Na        Charge=1 MSLevel=2 mandatory=true
Name=NL_Hex             Formula=$PRECURSOR-C6H1005 Charge=1 MSLevel=2 mandatory=true
Name=NL_Hex2            Formula=$PRECURSOR-C12H20010 Charge=1 MSLevel=2 mandatory=false

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy         Formula=$PRECURSOR-$CHAIN   Charge=1 MSLevel=2 mandatory=true
Name=NL_Carboxy_Hex    Formula=$PRECURSOR-$CHAIN-C6H1005 Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH    Formula=$CHAIN-H3O2        Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH        Formula=$CHAIN-OH           Charge=1 MSLevel=2 mandatory=false
Name=Carboxy            Formula=$CHAIN               Charge=1 MSLevel=2 mandatory=false
Name=CarboxyGlycerol_H2O Formula=$CHAIN+C3H5O        Charge=1 MSLevel=2 mandatory=false

[POSITION]
!INTENSITIES
Equation=NL_Carboxy[1]>NL_Carboxy[2]*1.6 mandatory=true
```

**Potassium Adduct** To implement the potassium adduct of DGDG, first the chemical element K had to be added to the *elementconfig.xml* file of LDA which can be found in the tools' root directory. The following code shows how the element was added. The used information was taken from a paper on atomic weights [85].

## elementconfig.xml

```
<?xml version="1.0" encoding="UTF-8"?>

<configurations>
  <config id="default">
    <elements>
      <element symbol="K" valency="1">
        <isotope mass="38.9637069" abund="0.9329" />
        <isotope mass="39.96399867" abund="0.0001" />
        <isotope mass="40.96182597" abund="0.0673" />
      </element>
    </elements>
  </config>
</configurations>
```

The fragmentation process of the potassium adduct isn't well understood, but yields fragment ions of  $[M+K-R_xCOOH]^+$  ( $x=1,2$ ) (denoted as *NL\_Carboxy*) and  $[M+K-R_1COOH-CH_2]^+$  (denoted as *NL\_Carboxy.CH<sub>2</sub>*) [9,87]. As Lipid Data Analyzer didn't provide an option for chain fragments that are

only mandatory for one of the chains, both fragments were declared to be mandatory for an identification; declaring the *NL-Carboxy-CH<sub>2</sub>* fragment as not mandatory, would have led to too many false positive identifications.

---

DGDG.K.frag

---

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor          Formula=$PRECURSOR          Charge=1 MSLevel=2 mandatory=false
Name=HexK               Formula=C6H1005K           Charge=1 MSLevel=2 mandatory=false
Name=Hex2K              Formula=C12H20010K        Charge=1 MSLevel=2 mandatory=false
Name=NL_Hex            Formula=$PRECURSOR-C6H1005 Charge=1 MSLevel=2 mandatory=false

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy        Formula=$PRECURSOR-$CHAIN  Charge=1 MSLevel=2 mandatory=true
Name=NL_Carboxy_CH2    Formula=$PRECURSOR-$CHAIN-CH2 Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH    Formula=$CHAIN-H3O2        Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH        Formula=$CHAIN-OH          Charge=1 MSLevel=2 mandatory=false
```

---

**Deprotonated Adduct** Analogously to the deprotonated adduct of MGDG, the deprotonated adduct of DGDG yields fragment ions of  $[\text{Glycerol}+\text{Hex}+\text{H}_2\text{O}]^+$  (denoted as *GlycerolHexH<sub>2</sub>O*) and  $[\text{Glycerol}+\text{Hex}+\text{H}_4\text{O}_2]^+$  (denoted as *GlycerolHexH<sub>4</sub>O<sub>2</sub>*) [93]. The fragments  $[\text{Glycerol}+\text{Hex}_2+\text{H}_2\text{O}]^+$  (denoted as *GlycerolHex<sub>2</sub>H<sub>2</sub>O*) and  $[\text{Glycerol}+\text{Hex}_2+\text{H}_4\text{O}_2]^+$  (denoted as *GlycerolHex<sub>2</sub>H<sub>4</sub>O<sub>2</sub>*) were introduced, as DGDG features a *Hex<sub>2</sub>* moiety, and declared to be mandatory for identification. Additionally, possible  $[\text{Glycerol}+\text{Hex}_3+\text{H}_2\text{O}]^+$  (denoted as *GlycerolHex<sub>3</sub>H<sub>2</sub>O*) and  $[\text{Glycerol}+\text{Hex}_3+\text{H}_4\text{O}_2]^+$  (denoted as *GlycerolHex<sub>3</sub>H<sub>4</sub>O<sub>2</sub>*) ions were assumed for TriGDG and used for filtering false positives via the intensity rules. The ruleset for the deprotonated adduct is shown below.

---

DGDG.-H.frag

---

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor          Formula=$PRECURSOR          Charge=1 MSLevel=2 mandatory=false
Name=GlycerolHex2H2O    Formula=C15H25012          Charge=1 MSLevel=2 mandatory=true
Name=GlycerolHex2H4O2   Formula=C15H27013          Charge=1 MSLevel=2 mandatory=true
Name=GlycerolHexH2O     Formula=C9H1507             Charge=1 MSLevel=2 mandatory=false
Name=GlycerolHexH4O2    Formula=C9H1708             Charge=1 MSLevel=2 mandatory=false
Name=GlycerolHex3H2O    Formula=C14H25012          Charge=1 MSLevel=2 mandatory=other
Name=GlycerolHex3H4O2   Formula=C14H27013          Charge=1 MSLevel=2 mandatory=other

!INTENSITIES
Equation=GlycerolHex3H2O*1.5<GlycerolHex2H2O    mandatory=true
Equation=GlycerolHex3H4O2*1.5<GlycerolHex2H4O2  mandatory=true

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy        Formula=$PRECURSOR-$CHAIN  Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_H2O_OH    Formula=$CHAIN-H3O2        Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH        Formula=$CHAIN-OH          Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_H         Formula=$CHAIN-H           Charge=1 MSLevel=2 mandatory=true
```

---

**Formate and Acetate Adducts** For formate and acetate adducts of DGDG the same fragmentation patterns as for the deprotonated adduct were assumed.

### 3.4.3 TriGDG

**Ammoniated Adduct** The ammoniated adduct of TriGDG yields fragment ions of  $[\text{M}+\text{NH}_4-\text{NH}_3-\text{Hex}_3]^+$  (denoted as *NL-NH<sub>3</sub>-Hex<sub>3</sub>*),  $[\text{M}+\text{NH}_4-\text{NH}_3-\text{Hex}_3-\text{H}_2\text{O}]^+$  (denoted as *NL-NH<sub>3</sub>-H<sub>2</sub>O-Hex<sub>3</sub>*) and  $[\text{Glycerol}+\text{R}_x\text{COOH}-\text{H}_2\text{O}]^+$  ( $x=1,2$ ) (denoted as *CarboxyGlycerol-H<sub>2</sub>O*) [9, 25]. The resulting ruleset can be seen below.

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=D*\d+:\d+
DoubleBondsFromName=D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor           Formula=$PRECURSOR           Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex         Formula=$PRECURSOR-C6H1305N  Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H2O_Hex     Formula=$PRECURSOR-C6H1506N  Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex2        Formula=$PRECURSOR-C12H23010N Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H2O_Hex2    Formula=$PRECURSOR-C12H25011N Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex3        Formula=$PRECURSOR-C18H33015N Charge=1 MSLevel=2 mandatory=true
Name=NL_NH3_H2O_Hex3    Formula=$PRECURSOR-C18H35016N Charge=1 MSLevel=2 mandatory=true

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy_NH3_Hex Formula=$PRECURSOR-$CHAIN-C6H1305N Charge=1 MSLevel=2 mandatory=false
Name=CarboxyGlycerol_H2O Formula=$CHAIN+C3H5O              Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH      Formula=$CHAIN-H3O2               Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH          Formula=$CHAIN-OH                 Charge=1 MSLevel=2 mandatory=false
Name=Carboxy              Formula=$CHAIN                     Charge=1 MSLevel=2 mandatory=false
```

**Other Adducts** The rules for all other adducts of TriGDG were assumed to be the same as for DGDG, with the difference, that *Hex*<sub>2</sub> was substituted by *Hex*<sub>3</sub> in all mandatory fragments.

### 3.4.4 TetraGDG

**Ammoniated Adduct** Analogously to DGDG and TriGDG, the ammoniated adduct of TetraGDG was observed to yield fragment ions of  $[M+NH_4-NH_3-Hex_4]^+$  (denoted as *NL\_NH<sub>3</sub>-Hex<sub>4</sub>*),  $[M+NH_4-NH_3-Hex_4-H_2O]^+$  (denoted as *NL\_NH<sub>3</sub>-H<sub>2</sub>O-Hex<sub>4</sub>*) and  $[Glycerol+R_xCOOH-H_2O]^+$  ( $x=1,2$ ) (denoted as *CarboxyGlycerol-H<sub>2</sub>O*) [9, 25]. The following code shows the resulting rule file.

```
[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=D*\d+:\d+
DoubleBondsFromName=D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor           Formula=$PRECURSOR           Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex         Formula=$PRECURSOR-C6H1305N  Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H2O_Hex     Formula=$PRECURSOR-C6H1506N  Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex2        Formula=$PRECURSOR-C12H23010N Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H2O_Hex2    Formula=$PRECURSOR-C12H25011N Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex3        Formula=$PRECURSOR-C18H33015N Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_H2O_Hex3    Formula=$PRECURSOR-C18H35016N Charge=1 MSLevel=2 mandatory=false
Name=NL_NH3_Hex4        Formula=$PRECURSOR-C24H43020N Charge=1 MSLevel=2 mandatory=true
Name=NL_NH3_H2O_Hex4    Formula=$PRECURSOR-C24H45021N Charge=1 MSLevel=2 mandatory=true

[CHAINS]
!FRAGMENTS
Name=NL_Carboxy_NH3_Hex Formula=$PRECURSOR-$CHAIN-C6H1305N Charge=1 MSLevel=2 mandatory=false
Name=CarboxyGlycerol_H2O Formula=$CHAIN+C3H5O              Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH      Formula=$CHAIN-H3O2               Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH          Formula=$CHAIN-OH                 Charge=1 MSLevel=2 mandatory=false
Name=Carboxy              Formula=$CHAIN                     Charge=1 MSLevel=2 mandatory=false
```

**Other Adducts** The rules for all other adducts of TetraGDG were assumed to be the same as for DGDG, with the difference, that *Hex*<sub>2</sub> was substituted by *Hex*<sub>4</sub> in all mandatory fragments.

### 3.4.5 SQDG

**Sodiated Adduct** For SQDG precursor ions of  $[M-H+Na_2]^+$  were observed, with fragment ions of  $[M-H+Na_2-R_xCOOH]^+$  ( $x=1,2$ ) (denoted as *NL\_Carboxy*) and  $[M-H+Na_2-R_1COOH-R_2COOH]^+$  (denoted as *SQ\_head*) [90]. Further, these adducts show an ion at  $[SO_3Na_2]^+$  [94] as seen in the following code.

```

[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor           Formula=$PRECURSOR           Charge=1 MSLevel=2 mandatory=false
Name=SQ_head             Formula=C9H15O8SNa2         Charge=1 MSLevel=2 mandatory=true
Name=SO3Na2              Formula=SO3Na2              Charge=1 MSLevel=2 mandatory=true

[CHAINS]
!FRAGMENTS
Name=Carboxy             Formula=$CHAIN               Charge=1 MSLevel=2 mandatory=false
Name=NL_Carboxy         Formula=$PRECURSOR-$CHAIN   Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH     Formula=$CHAIN-H3O2         Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH         Formula=$CHAIN-OH           Charge=1 MSLevel=2 mandatory=false

```

**Deprotonated Adduct** For the  $[M-H]^-$  adduct an abundant dehydrosulfoglycosyl anion ( $[C_6H_9O_7S]^-$ ) is observed in the fragmentation spectrum [95]. A study on spinach observed a systematic prevalence of fatty acid loss from the *sn*-1 position of glycerol and exploited this information for the regiochemical *sn*-1/*sn*-2 assignment; the intensity ratios of  $[M-H-R_1COOH]^-$  to  $[M-H-R_2COOH]^-$  was between 1.22 and 3.14 [23] and was subsequently used as an intensity rule for positional identification, as seen in the code below. Additionally, fragments of  $[SO_3H]^-$ ,  $[CH_3O_3S]^-$ ,  $[C_6H_5O_3]^-$ ,  $[C_4H_5O_4S]^-$ ,  $[C_3H_5O_5S]^-$ ,  $[C_4H_5O_5S]^-$  and  $[C_6H_7O_6S]^-$  are frequently observed [96].

```

[GENERAL]
AmountOfChains=2
ChainLibrary=fattyAcidChains.xlsx
CAtomsFromName=\D*\d+:\d+
DoubleBondsFromName=\D*\d+:\d+

[HEAD]
!FRAGMENTS
Name=Precursor           Formula=$PRECURSOR           Charge=1 MSLevel=2 mandatory=false
Name=Dehydrosulfoglycosyl Formula=C6H9O7S             Charge=1 MSLevel=2 mandatory=true
Name=SO3H                Formula=SO3H                 Charge=1 MSLevel=2 mandatory=false
Name=C6H7O6S             Formula=C6H7O6S             Charge=1 MSLevel=2 mandatory=false
Name=C4H5O5S             Formula=C4H5O5S             Charge=1 MSLevel=2 mandatory=false
Name=C3H5O5S             Formula=C3H5O5S             Charge=1 MSLevel=2 mandatory=false
Name=C4H5O4S             Formula=C4H5O4S             Charge=1 MSLevel=2 mandatory=false
Name=C6H5O3              Formula=C6H5O3              Charge=1 MSLevel=2 mandatory=false
Name=CH3O3S              Formula=CH3O3S              Charge=1 MSLevel=2 mandatory=false

[CHAINS]
!FRAGMENTS
Name=Carboxy             Formula=$CHAIN               Charge=1 MSLevel=2 mandatory=false
Name=NL_Carboxy         Formula=$PRECURSOR-$CHAIN   Charge=1 MSLevel=2 mandatory=true
Name=Carboxy_H2O_OH     Formula=$CHAIN-H3O2         Charge=1 MSLevel=2 mandatory=false
Name=Carboxy_OH         Formula=$CHAIN-OH           Charge=1 MSLevel=2 mandatory=false

[POSITION]
! INTENSITIES
Equation=NL_Carboxy[1]>NL_Carboxy[2]*1.22 mandatory=true

```

### 3.4.6 DGDG-Estolides

Rules for DGDG-estolides were formulated analogously to the rules for DGDG, with the only difference, that one (DGDG-EN1) or two (DGDG-EN2) additional fatty acyl chains are attached to the DGDG molecule.

### 3.4.7 Single Chain (MG) Galactolipid Species

The rules for MGMG, DGMG and SQMG were formulated analogously to the rules of MGDG, DGDG and SQDG respectively. The only difference being that, instead of two fatty acyl chains these species contain only one fatty acyl chain.

### 3.4.8 Oxidized Lipids

It was assumed that the  $MS^2$  fragments of oxidized lipids are very similar to those of unmodified lipids. However, it became evident that the intensity relationships for *sn*-positional identification of the  $[M-CH_3]^-$

adduct of oxPC proved unreliable and were consequently removed from the rule set. Other than that, no new fragmentation rules were implemented for oxidized lipids.

### 3.5 Notation of Oxidized Lipids in LDA

As of now, notation guidelines for oxidized lipids are scarce and incomplete. Hence, many different approaches are being used [10,30,67,97]. We settled on adding an *ox*-prefix to the lipid abbreviation to mark oxidized classes. The shorthand notation of the modification (table 8) is enclosed in square brackets and written directly after the chain it concerns. Multiple modifications of different types are jointly enclosed in square brackets and separated by a comma. For multiple modifications of the same type, their number is simply added in front of the modification. Accordingly, the lipid oxMGDG(16:2[O,OH]\_18:4[4OH]) has an additional O-atom and an additional OH-group on its 16:2 chain, and four additional OH-groups on its 18:4 chain.

Table 8: Oxidation moieties: Shorthand notation used in LDA; gray symbols are not being used as of now, due to notation ambiguities (see section 4.3 for more details)

Modification	Shorthand Notation	Modification	Shorthand Notation
oxo-	O	hydroperoxy-	OOH
keto-	Ke	cyclopentane-	Cy
epoxy-	EpO	bromo-	Br
furan-	Fu	chloro-	Cl
hydroxy-	OH	fluoro-	Fl
epidioxo-	EpOO	iodo-	I
carboxy-	COOH	nitro-	NO <sub>2</sub>

### 3.6 Novel LDA Features

To allow LDA to identify fragments of oxidized lipids in MS<sup>2</sup> spectra, the list of white-listed fatty acyls and the target mass list had to be extended and LDA source code had to be modified accordingly. First, a new column *oxidation-state* was introduced in the fatty acids mass list in which the respective chain-modification is listed as seen as in table 9. In the target mass list, the same column *oxidation-state* was introduced (see table 10), to allow LDA to filter out combinations of fatty acids that don't have the same sum of modifications as the parent-molecule. To illustrate this point further, the lipid species oxMGDG(36[O,OH]) could possibly be the molecular species oxMGDG(18:3[O]\_18:3[OH]) or the molecular species oxMGDG(18:3[O,OH]\_18:3) but not oxMGDG(18:3[O,OH]\_18:3[O,OH]). The *ox*-prefix is programmatically added to the lipid whenever the respective *oxidation-state* cell isn't empty. The modifications and the corresponding compositional changes were defined in a *modconfig.xml* file as seen below. LDA will then automatically calculate new masses and chemical formulas for all modifications listed in the *oxidation-state* column. In the column itself, modifications concerning the same molecule are divided by a comma and modifications concerning new molecules are divided by a semicolon. For example, the entry ;OH;O,OH;2OH;3OH;4OH for oxMGDG(36:6) will look for oxMGDG(36:6), oxMGDG(36:6[OH]), oxMGDG(36:6[O,OH]), oxMGDG(36:6[2OH]), oxMGDG(36:6[3OH]) and oxMGDG(36:6[4OH]). For the decision rules, the algorithm is first looking for the respective *ox*-rule (for example, *oxMGDG-NH4.frag.txt*), but will fall back to the standard-rule (for example, *MGDG-NH4.frag.txt*), if no *ox*-rule is available. The approach is backward compatible (i.e. can be used with fatty acid mass lists and target mass lists without the *oxidation-state* column).

```
<?xml version="1.0" encoding="UTF-8"?>
<modifications>
  <modification symbol="OH">
    <element name="O">+1</element>
  </modification>
  <modification symbol="O">
    <element name="O">+1</element>
    <element name="H">-2</element>
  </modification>
  <modification symbol="OOH">
    <element name="O">+2</element>
  </modification>
  <modification symbol="NO2">
    <element name="O">+2</element>
    <element name="N">+1</element>
  </modification>
</modifications>
```

modconfig.xml

```

<element name="H">-1</element>
</modification>
<modification symbol="I">
<element name="I">+1</element>
<element name="H">-1</element>
</modification>
<modification symbol="Cl">
<element name="Cl">+1</element>
<element name="H">-1</element>
</modification>
<modification symbol="Br">
<element name="Br">+1</element>
<element name="H">-1</element>
</modification>
<modification symbol="Fl">
<element name="Fl">+1</element>
<element name="H">-1</element>
</modification>
<modification symbol="">
</modification>
</modifications>

```

In the *Statistical Analysis* module of LDA, new check boxes *Show MS<sup>n</sup> Only*, *Chain Evidence Only* and *Combine Classes with ox-Classes* have been introduced (supplementary figure S3). *Show MS<sup>n</sup> Only* will lead to LDA displaying only identifications, with at least MS<sup>n</sup> headgroup evidence. When *Chain Evidence Only* is checked, MS<sup>n</sup> evidence for chains is needed too, to show the identification. *Combine Classes with ox-Classes* will merge all classes with their respective *ox-classes*, allowing comparison of abundances between the classes (supplementary figure S4 vs. S5). The check boxes *Show MS<sup>n</sup> Only* and *Chain Evidence Only* have been implemented in the *Display Results* section as well (supplementary figure S1), where, when checked, MS<sup>n</sup> evidence is needed in order to list the identification (supplementary figure S2).

Table 9: Excerpt of the modified FA mass list

Name	dfs	C	H	O	mass	oxidation-state
2	:	0	2	4	2	60.021120547999999 ;OH;O,OH;2OH;3OH;4OH
3	:	0	3	6	2	74.036770582000003 ;OH;O,OH;2OH;3OH;4OH
4	:	0	4	8	2	88.052420616000006 ;OH;O,OH;2OH;3OH;4OH
5	:	0	5	10	2	102.06807070000001 ;OH;O,OH;2OH;3OH;4OH
6	:	0	6	12	2	116.083720683999999 ;OH;O,OH;2OH;3OH;4OH
7	:	0	7	14	2	130.099370717999999 ;OH;O,OH;2OH;3OH;4OH
8	:	0	8	16	2	144.115020751999999 ;OH;O,OH;2OH;3OH;4OH
9	:	0	9	18	2	158.130670786 ;OH;O,OH;2OH;3OH;4OH
10	:	0	10	20	2	172.14632082 ;OH;O,OH;2OH;3OH;4OH
11	:	0	11	22	2	186.161970854 ;OH;O,OH;2OH;3OH;4OH
12	:	0	12	24	2	200.177620887999998 ;OH;O,OH;2OH;3OH;4OH
13	:	0	13	26	2	214.193270921999998 ;OH;O,OH;2OH;3OH;4OH
14	:	0	14	28	2	228.208920955999999 ;OH;O,OH;2OH;3OH;4OH
15	:	0	15	30	2	242.224570989999999 ;OH;O,OH;2OH;3OH;4OH
16	:	0	16	32	2	256.240221023999999 ;OH;O,OH;2OH;3OH;4OH
17	:	0	17	34	2	270.255871058000003 ;OH;O,OH;2OH;3OH;4OH
18	:	0	18	36	2	284.271521092 ;OH;O,OH;2OH;3OH;4OH
19	:	0	19	38	2	298.287171125999998 ;OH;O,OH;2OH;3OH;4OH
20	:	0	20	40	2	312.302821160000001 ;OH;O,OH;2OH;3OH;4OH
21	:	0	21	42	2	326.318471193999998 ;OH;O,OH;2OH;3OH;4OH
23	:	0	23	46	2	354.349771261999999 ;OH;O,OH;2OH;3OH;4OH
24	:	0	24	48	2	368.365421295999997 ;OH;O,OH;2OH;3OH;4OH
25	:	0	25	50	2	382.38107133 ;OH;O,OH;2OH;3OH;4OH
26	:	0	26	52	2	396.396721363999997 ;OH;O,OH;2OH;3OH;4OH
27	:	0	27	54	2	410.412371398 ;OH;O,OH;2OH;3OH;4OH
28	:	0	28	56	2	424.428021431999998 ;OH;O,OH;2OH;3OH;4OH
29	:	0	29	58	2	438.443671466000001 ;OH;O,OH;2OH;3OH;4OH
30	:	0	30	60	2	452.459321499999999 ;OH;O,OH;2OH;3OH;4OH

Table 10: Excerpt of the modified target mass list for TGs; all three fatty acyl chains combined can carry a total of 0-4 additional OH-groups

Name	dbs	C	H	O	P	N	D	mass(form[+NH4]	name[NH4])	oxidation-state
28	:	0	31	58	6	0	0	544.45716515501204		;OH;O,OH;2OH;3OH;4OH
28	:	1	31	56	6	0	0	542.44151509061203		;OH;O,OH;2OH;3OH;4OH
28	:	2	31	54	6	0	0	540.42586502621202		;OH;O,OH;2OH;3OH;4OH
29	:	0	32	60	6	0	0	558.47281521941204		;OH;O,OH;2OH;3OH;4OH
29	:	1	32	58	6	0	0	556.45716515501204		;OH;O,OH;2OH;3OH;4OH
29	:	2	32	56	6	0	0	554.44151509061203		;OH;O,OH;2OH;3OH;4OH
30	:	0	33	62	6	0	0	572.48846528381205		;OH;O,OH;2OH;3OH;4OH
30	:	1	33	60	6	0	0	570.47281521941204		;OH;O,OH;2OH;3OH;4OH
30	:	2	33	58	6	0	0	568.45716515501204		;OH;O,OH;2OH;3OH;4OH
30	:	3	33	56	6	0	0	566.44151509061203		;OH;O,OH;2OH;3OH;4OH
30	:	4	33	54	6	0	0	564.42586502621202		;OH;O,OH;2OH;3OH;4OH
31	:	0	34	64	6	0	0	586.50411534821194		;OH;O,OH;2OH;3OH;4OH
31	:	1	34	62	6	0	0	584.48846528381205		;OH;O,OH;2OH;3OH;4OH
31	:	2	34	60	6	0	0	582.47281521941204		;OH;O,OH;2OH;3OH;4OH
31	:	3	34	58	6	0	0	580.45716515501204		;OH;O,OH;2OH;3OH;4OH
31	:	4	34	56	6	0	0	578.44151509061203		;OH;O,OH;2OH;3OH;4OH
32	:	0	35	66	6	0	0	600.51976541261195		;OH;O,OH;2OH;3OH;4OH
32	:	1	35	64	6	0	0	598.50411534821194		;OH;O,OH;2OH;3OH;4OH
32	:	2	35	62	6	0	0	596.48846528381205		;OH;O,OH;2OH;3OH;4OH
32	:	3	35	60	6	0	0	594.47281521941204		;OH;O,OH;2OH;3OH;4OH
32	:	4	35	58	6	0	0	592.45716515501204		;OH;O,OH;2OH;3OH;4OH
33	:	0	36	68	6	0	0	614.53541547701195		;OH;O,OH;2OH;3OH;4OH
33	:	1	36	66	6	0	0	612.51976541261195		;OH;O,OH;2OH;3OH;4OH
33	:	2	36	64	6	0	0	610.50411534821194		;OH;O,OH;2OH;3OH;4OH
33	:	3	36	62	6	0	0	608.48846528381205		;OH;O,OH;2OH;3OH;4OH
33	:	4	36	60	6	0	0	606.47281521941204		;OH;O,OH;2OH;3OH;4OH
34	:	0	37	70	6	0	0	628.55106554141196		;OH;O,OH;2OH;3OH;4OH
34	:	1	37	68	6	0	0	626.53541547701195		;OH;O,OH;2OH;3OH;4OH
34	:	2	37	66	6	0	0	624.51976541261195		;OH;O,OH;2OH;3OH;4OH

### 3.7 Lipid Data Analyzer Results

#### 3.7.1 Dataset 1 - Mouse Liver

A total of 221 lipid species were identified by LDA in dataset 1. The most frequently reported class was TG with a total of 43 identifications (table 11). With 43 reported  $[M+NH_4]^+$  adducts of TG, it was more prevalent than the  $[M+Na]^+$  adduct (12 identifications). In positive ion mode, LDA identified 109 different lipid species (supplementary table S35), including the subclasses DG, LPC, LPE, P-PE, PC, PE and TG. In negative ion mode, 112 different lipid species were identified, whereby the subclasses Cer, LPC, LPE, P-PE, PC, PE, PG, PI and PS were included (supplementary table S36). All listed lipids were identified on fatty acyl level.

Table 11: Identified lipid classes by LDA in dataset 1

Lipid Class	Adduct	Unique Species per Adduct	Unique Species
Cer	$[M+HCO_2]^-$	5	5
DG	$[M+Na]^+$	7	7
	$[M+NH_4]^+$	6	
LPC	$[M+H]^+$	6	10
	$[M-CH_3]^-$	4	
	$[M+HCO_2]^-$	10	
LPE	$[M+H]^+$	1	6
	$[M-H]^-$	6	
P-PE	$[M+H]^+$	2	8
	$[M-H]^-$	8	
PC	$[M+H]^+$	34	40
	$[M+Na]^+$	9	
	$[M+HCO_2]^-$	38	
PE	$[M+H]^+$	16	24
	$[M+Na]^+$	1	
	$[M-H]^-$	24	
PG	$[M-H]^-$	9	9
PI	$[M-H]^-$	9	9
PS	$[M-H]^-$	3	3
TG	$[M+Na]^+$	12	43
	$[M+NH_4]^+$	43	

### 3.7.2 Dataset 2 - Ryegrass Leaves

In dataset 2, 113 lipid species have been identified by LDA. In positive ion mode, 54 different lipid species were reported, spanning subclasses DG, DGDG, DGMG, LPC, MGDG, MGMG, PC, PE, PG, PI, TG and TriGDG (supplementary table S37). In negative ion mode, 59 species were identified, including the subclasses DGDG, LPC, LPE, MGDG, P-PE, PC, PE, PG, PI, PS, SQDG and SQMG (supplementary table S38). All listed species, were confirmed by MS<sup>2</sup> fragmentation patterns on fatty acyl level.

**MGDG** Five different species of MGDG have been identified by Lipid Data Analyzer (table 12), with adducts of  $[M+Na]^+$ ,  $[M+NH_4]^+$  or  $[M+HCO_2]^-$ . Supplementary Figure S6 shows the spectrum of the ammonium adduct of MGDG(36:6), where the fragment *NL\_NH<sub>3</sub>\_H<sub>2</sub>O\_Hex* was used for the headgroup identification and two *NL\_Carboxy\_NH<sub>3</sub>\_Hex* ions for chain identification. Additionally, *Carboxy\_OH* and *Carboxy\_H<sub>2</sub>O\_OH* ions for chains (18:2), (18:3) and (18:4), further suggest a mix of MGDG(18:3/18:3) and MGDG(18:2\_18:4). In supplementary figure S7 the spectrum for the sodium adduct of MGDG(36:6) can be seen, with a fragment ion corresponding to *HexNa*, which was used for headgroup identification and an ion for (*NL\_Carboxy*) used for identification of the two (18:3) chains. The two ions *Carboxy\_OH* and *Carboxy\_H<sub>2</sub>O\_OH* further support the identification of the two chains being (18:3). The spectrum for the  $[M+HCO_2]^-$  adduct of MGDG(36:6) (supplementary figure S8), shows the two ions *GlycerolHexH<sub>2</sub>O* and *GlycerolHexH<sub>4</sub>O<sub>2</sub>* needed for headgroup identification, and two *Carboxy\_H* ions, corresponding to chains (18:3).

**DGDG** For DGDG, LDA identified nine different species (table 12), with adducts of  $[M+Na]^+$ ,  $[M+NH_4]^+$ ,  $[M+HCO_2]^-$ , or  $[M-H]^-$ . The spectrum of the ammonium adduct of DGDG(36:6), as seen as in supplementary figure S9, shows characteristic headgroup fragment ions of *NL\_NH<sub>3</sub>\_Hex<sub>2</sub>* and *NL\_NH<sub>3</sub>\_H<sub>2</sub>O\_Hex<sub>2</sub>* as well as fragment ions for chains of (18:2),(18:3) and (18:4), suggesting a mix of DGDG(18:3/18:3) and DGDG(18:2\_18:4). Supplementary Figure S10 shows the spectrum of the sodium adduct of DGDG(36:6), where the *Hex<sub>2</sub>Na* ion was used for headgroup identification which is further supported by a *NL\_Hex* ion. Three other ions correspond to ions of the fatty acids, which were identified as (18:3). The spectrum of the deprotonated adduct of DGDG(36:6) can be seen in supplementary figure S11, where headgroup ions *GlycerolHexH<sub>2</sub>O*, *GlycerolHex<sub>2</sub>H<sub>2</sub>O*, *GlycerolHex<sub>2</sub>H<sub>4</sub>O<sub>2</sub>* and an *Carboxy\_H* ion belonging to the fatty acid (18:3) can be seen. The spectrum for the  $[M+HCO_2]^-$  adduct of DGDG(35:6), shows the two characteristic headgroup ions *GlycerolHex<sub>2</sub>H<sub>2</sub>O* and *GlycerolHex<sub>2</sub>H<sub>4</sub>O<sub>2</sub>* as well as *Carboxy\_H* ions suggesting two (18:3) chains. Two ions corresponding to *GlycerolHexH<sub>2</sub>O* and *GlycerolHexH<sub>4</sub>O<sub>2</sub>* are present as well (supplementary figure S12).

Table 12: Identified lipid classes by LDA in dataset 2

Lipid Class	Adduct	Unique Species per Adduct	Unique Species
DG	$[M+Na]^+$	7	10
	$[M+NH_4]^+$	7	
DGDG	$[M-H]^-$	1	9
	$[M+Na]^+$	3	
	$[M+HCO_2]^-$	9	
	$[M+NH_4]^+$	3	
DGMG	$[M+Na]^+$	1	2
	$[M+NH_4]^+$	1	
	$[M+HCO_2]^-$	2	
LPC	$[M+H]^+$	3	3
	$[M+HCO_2]^-$	3	
LPE	$[M-H]^-$	3	3
MGDG	$[M+Na]^+$	1	5
	$[M+NH_4]^+$	3	
	$[M+HCO_2]^-$	4	
MGMG	$[M+Na]^+$	1	1
	$[M+NH_4]^+$	1	
	$[M+HCO_2]^-$	1	
P-PE	$[M-H]^-$	3	3
PC	$[M+H]^+$	4	8
	$[M+HCO_2]^-$	8	
	$[M+Na]^+$	3	
	$[M-CH_3]^-$	1	
PE	$[M-H]^-$	8	8
	$[M+H]^+$	4	
	$[M+Na]^+$	1	
PG	$[M-H]^-$	3	6
	$[M+H]^+$	3	
	$[M+Na]^+$	1	
PI	$[M-H]^-$	5	5
	$[M+H]^+$	2	
PS	$[M-H]^-$	5	5
SQDG	$[M-H]^-$	4	4
SQMG	$[M-H]^-$	1	1
TG	$[M+Na]^+$	18	18
	$[M+NH_4]^+$	15	
TriGDG	$[M+NH_4]^+$	1	1

**TriGDG** One  $[M+NH_4]^+$  adduct of TriGDG was identified by LDA in positive ion mode (table 12). The spectrum of this ammoniated adduct of TriGDG(36:6) is shown in supplementary figure S13. Ions *NL-NH<sub>3</sub>-Hex<sub>3</sub>* and *NL-NH<sub>3</sub>-H<sub>2</sub>O-Hex<sub>3</sub>* identify the headgroup, whereby the other ions point to the fatty acids (18:2),(18:3) and (18:4), suggesting a mix of TriGDG(18:3/18:3) and TriGDG(18:2.18:4).

**SQDG** For SQDG, LDA identified four different species in negative ion mode for the deprotonated adduct (table 12). In supplementary figure S14 the spectrum of the deprotonated adduct of SQDG(36:6) is shown. The characteristic dehydrosulfoglycosyl ion was used for head group identification and the *NL-Carboxy* ion to identify the chains (18:3).

**MGMG** LDA identified one MGMG species with  $[M+Na]^+$ ,  $[M+NH_4]^+$  and  $[M+HCO_2]^-$  adducts (table 12). Supplementary Figure S15 shows the spectrum of the ammonium adduct of MGMG(18:3), where the headgroup specific ions *NL-NH<sub>3</sub>-Hex* and *NL-NH<sub>3</sub>-H<sub>2</sub>O-Hex* as well as the fatty acid (18:3) specific ions *Carboxy-H<sub>2</sub>O-OH* and *Carboxy-OH* can be seen. The spectrum of the sodiated adduct of MGMG(18:3) shows an *NL-Carboxy* ion corresponding to the fatty acid chain (18:3) as well as the characteristic headgroup fragment *HexNa* (supplementary figure S16). In supplementary figure S17 the spectrum of the  $[M+HCO_2]^-$  adduct of MGMG(18:3) is shown, with the characteristic head group ions *GlycerolHexH<sub>4</sub>O<sub>2</sub>* and *GlycerolHexH<sub>2</sub>O* as well as the chain fragment ion *Carboxy-H(18:3)*.

**DGMG** Two DGMG species were identified by LDA including  $[M+Na]^+$ ,  $[M+NH_4]^+$  and  $[M+HCO_2]^-$  adducts (table 12). Supplementary Figure S18 shows the ammonium adduct of DGMG(18:3) with ions  $NL\_NH_3\_H_2O\_Hex_2$  and  $NL\_NH_3\_Hex_2$  identifying the headgroup and  $Carboxy\_OH$  and  $Carboxy\_H_2O\_OH$  ions used for identification of the chains (18:3). In supplementary figure S19 the sodiated adduct of DGMG(18:3) is shown. Here the ion  $Hex_2Na$  and  $NL\_Hex$  identify the headgroup and  $NL\_Carboxy$  identifies the chain (18:3). The spectrum of the  $[M+HCO_2]^-$  adduct of DGMG(18:3) is shown in supplementary figure S20, with the characteristic head group ions  $GlycerolHexH_4O_2$  and  $GlycerolHexH_2O$  as well as the chain fragment ion  $Carboxy\_H(18:3)$ .

**SQMG** LDA identified one SQMG species in negative ion mode (table 12). In the spectrum of the deprotonated adduct of SQMG(18:3) the characteristic dehydrosulfoglycosyl ion can be seen, as well as an ion corresponding to  $[SO_3]^-$  and other characteristic ions. The  $NL\_Carboxy$  ion corresponds to the fatty acid (18:3) (supplementary figure S21).

### 3.7.3 Dataset 3 - Oxidized PC Standard

For LDA analysis, a mass list was generated for oxidation products between PC(34:0) and PC(34:2) with up to three additional OH-modifications and up to two additional O-modifications. LDA was able to identify 19 lipid species in dataset 3 (table 13). After manual inspection of the MS<sup>2</sup> spectra, four identifications were considered to be false positives, as only fragments of low abundance were annotated and major fragments were left unannotated.

Table 13: Lipid molecular species identified in dataset 3 by LDA; putative false positives are marked in gray

Lipid Species	Adduct	Lipid Molecular Species
oxPC(34:0[OH])	$[M+HCO_2]^-$	oxPC(16:0_18:0[OH])
oxPC(34:0[2OH])	$[M-CH_3]^-$ $[M+HCO_2]^-$	oxPC(16:0_18:0[2OH]) oxPC(16:0/18:0[2OH])
PC(34:1)	$[M-CH_3]^-$ $[M+HCO_2]^-$	PC(16:0/18:1) PC(16:0/18:1)
oxPC(34:1[OH])	$[M-CH_3]^-$ $[M+HCO_2]^-$	oxPC(16:0_18:1[OH]) oxPC(16:0/18:1[OH])
oxPC(34:1[2O])	$[M-CH_3]^-$ $[M+HCO_2]^-$	oxPC(16:0_18:1[2O]) oxPC(16:0/18:1[2O])
oxPC(34:1[2OH])	$[M-CH_3]^-$	oxPC(16:0_18:1[2OH])
oxPC(34:1[3OH])	$[M-CH_3]^-$ $[M+HCO_2]^-$	oxPC(16:0_18:1[3OH]) oxPC(16:0/18:1[3OH])
PC(34:2)	$[M+HCO_2]^-$	PC(16:1_18:1)
oxPC(34:2[OH])	$[M-CH_3]^-$ $[M+HCO_2]^-$	oxPC(16:0_18:2[OH]) oxPC(16:0/18:2[OH])
oxPC(34:2[2OH])	$[M-CH_3]^-$ $[M+HCO_2]^-$	oxPC(16:0_18:2[2OH]) oxPC(16:0/18:2[2OH])
oxPC(34:2[3OH])	$[M-CH_3]^-$ $[M+HCO_2]^-$	oxPC(16:0_18:2[3OH]) oxPC(16:0/18:2[3OH])

### 3.7.4 Dataset 4 - Oxidized Wheat Seeds

With the help of an R script, an LDA target mass list was created directly out of the identifications listed in the supplementary table of the paper by Riewe *et al.* [30] and subsequently, the data was analyzed. The most frequently reported class was TG with a total of 133 identifications, 68 of which were not oxidatively modified (table 14). The relative number of lipids identified shows a distinct pattern: Non-oxidized species are most frequent and species get less frequent the more OH-modifications they carry. The same pattern holds true for the abundances of reported lipids, as can be seen for DGDG (as DGDG(36:4) and oxDGDG(36:4[OH]) in supplementary figure S34), and TG (as TG(54:7), oxTG(54:7[OH]), oxTG(54:7[2OH]) and oxTG(54:7[3OH]) in supplementary figure S35). The respective spectra for those lipids can be seen in supplementary figures S28 and S29 for DGDG and supplementary figures S30, S31, S32 and S33 for TG.

Table 14: Unique lipid species identified by LDA in dataset 4

Lipid Class	Total	0 OH	1 OH	2 OH	3 OH	4 OH
TG	133	68	41	14	8	2
DG	48	27	9	10	1	1
PC	2	2	0	0	0	0
PE	7	7	0	0	0	0
PI	5	5	0	0	0	0
MGDG	11	6	3	2	0	0
DGDG	27	20	3	4	0	0

### 3.8 Galactolipids - Comparison

**Datasets 1 - Mouse Liver, 3 - Oxidized PC Standard and 4 - Oxidized Wheat Seeds** As galactolipids are only present in plants and some bacteria [22], no galactolipids were expected in datasets 1 and 3. LDA identified a total of 11 MGDG species and 27 DGDG species in dataset 4, whereby neither LipidMatch nor LipidMatch Flow could handle this dataset.

**Dataset 2 - Ryegrass Leaves** Lipid Data Analyzer identified six species of MGDG, nine species of DGDG, one species of TriGDG, four species of SQDG, one species of MGMG, two species of DGMG and one species of SQMG. LipidMatch on the other hand identified seven species of MGDG, eight species of DGDG, and four species of SQDG. Identification of TriGDG, MGMG, DGMG and SQMG was not supported by LipidMatch at that time. Throughout both, positive and negative ion mode, LipidMatch Flow reported less galactolipids than LDA and LipidMatch. A Venn diagram of the identified galactolipids (molecular species plus adducts) by Lipid Data Analyzer, LipidMatch and LipidMatch Flow is shown in figure 3. Table 15 shows an overview of the identified lipids by LDA, LipidMatch and LipidMatch Flow, whereby supplementary tables S41, S42 and S43 show a more in-depth comparison of the galactolipid species identified by the three tools.

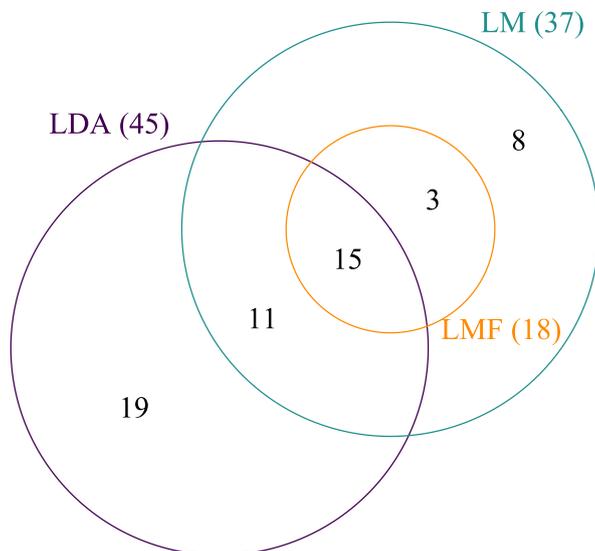


Figure 3: Identified galactolipids - LDA vs. LipidMatch vs. LipidMatch Flow; out of a total of 56 identified lipid molecular species with distinct adducts

Table 15: Identified galactolipids in dataset 2; LDA vs. LipidMatch vs. LipidMatch Flow

Lipid Class	Adduct	Unique Identifications by LDA	Unique Identifications by LM (LMF)
DGDG	[M-H] <sup>-</sup>	1	not implemented
	[M+Na] <sup>+</sup>	3	
	[M+HCO <sub>2</sub> ] <sup>-</sup>	9	
	[M+NH <sub>4</sub> ] <sup>+</sup>	3	
DGMG	[M+Na] <sup>+</sup>	1	not implemented
	[M+NH <sub>4</sub> ] <sup>+</sup>	1	
	[M+HCO <sub>2</sub> ] <sup>-</sup>	2	
MGDG	[M+Na] <sup>+</sup>	1	4 (4)
	[M+NH <sub>4</sub> ] <sup>+</sup>	3	2 (1)
	[M+HCO <sub>2</sub> ] <sup>-</sup>	4	4 (3)
MGMG	[M+Na] <sup>+</sup>	1	not implemented
	[M+NH <sub>4</sub> ] <sup>+</sup>	1	
	[M+HCO <sub>2</sub> ] <sup>-</sup>	1	
SQDG	[M-H] <sup>-</sup>	4	4 (4)
SQMG	[M-H] <sup>-</sup>	1	not implemented
TriGDG	[M+NH <sub>4</sub> ] <sup>+</sup>	1	not implemented

### 3.8.1 Validation of LipidMatch-only Identifications

Eleven lipid molecular species were identified by LipidMatch but not by Lipid Data Analyzer caused by various reasons (supplementary table S44). Two selected spectra of LipidMatch-only identifications can be seen in the appendix. Supplementary Figure S36 shows the spectrum in which LipidMatch identified [MGDG(28:3)+Na]<sup>+</sup>. Here, the MS<sup>2</sup> identification was achieved by annotating the m/z values 519.2929 (denoted as *NL-T1&Na*) corresponding to the loss of one fatty acid and Na, 413.2153 (denoted as *NL-T2&Na*) corresponding to the loss of the other fatty acid and Na, and 243.084 (denoted as *NL-T1&T2&Na*) corresponding to the loss of both fatty acids and Na. In supplementary figure S37, the spectrum is shown, where LipidMatch identified [MGDG(36:7)+Na]<sup>+</sup>. The identification was obtained by the m/z values 243.21 corresponding to the loss of both fatty acyl chains and Na (denoted as *NL-T1&T2&Na*), 517.28 corresponding to the loss of one fatty acyl chain and Na (denoted as *NL-T1&Na*), and 519.29 corresponding to the loss of the other fatty acyl chain and Na (denoted as *NL-T2&Na*). Most of the false positive identifications made by LipidMatch were filtered out by LipidMatch Flow.

## 3.9 Oxidized Lipids - Comparison

**Dataset 1 - Mouse Liver and Dataset 2 Wheat Seeds** As neither dataset 1 nor 2 explicitly concern oxidized lipids, oxidative analysis with LDA or LPPTiger had been skipped for these datasets.

**Dataset 3 - Oxidized PC Standard** The [M-CH<sub>3</sub>]<sup>-</sup> adduct was not implemented in the LPPTiger version used in this work, but other than that, the identifications made by LPPTiger and LDA in dataset 3 are very similar (table 16). LDA identified 19 lipid species in dataset 3, whereby four of the identified lipids are thought to be false positives. LPPTiger and LipidMatch identified eight and three lipid species respectively (table 16). One of the LipidMatch identifications is thought to be a false positive. LipidMatch Flow couldn't handle the dataset, as no blank files were present. Differences in reported lipids at identical m/z values between LDA and LPPTiger are due to ambiguities in notation only, where LPPTiger is sometimes able to rank modifications by their probability (due to so called fingerprint spectra). Supplementary figure S22 shows the MS<sup>2</sup> spectrum of oxPC(16:0/18:2[OH])+HCO<sub>2</sub><sup>+</sup> identified by LDA and the identification at the same m/z value made by LPPTiger (as oxPC(16:0/18:1[O])) can be seen in supplementary figure S23. Again, the difference is only due to notation ambiguities. Supplementary figure S24 shows a MS<sup>2</sup> spectra of an adduct that has been missed by LPPTiger but was identified by LDA.

Table 16: Lipid molecular species identified in dataset 3 by LDA, LPPtiger and LipidMatch; putative false positives are marked in gray; species names adapted to LDA shorthand notation

m/z	LDA	LPPtiger	LipidMatch	Adduct
744.5	PC(16:0/18:1)			
758.5	oxPC(16:0_18:2[OH])			
760.5	oxPC(16:0_18:1[OH])			
772.5	oxPC(16:0_18:1[2O])			
774.5	oxPC(16:0_18:2[2OH])	adduct not implemented	adduct not implemented	[M-CH <sub>3</sub> ] <sup>-</sup>
776.5	oxPC(16:0_18:1[2OH])			
778.5	oxPC(16:0_18:0[2OH])			
790.5	oxPC(16:0_18:2[3OH])			
792.5	oxPC(16:0_18:1[3OH])			
802.5	PC(16:1_18:1)			
804.5	PC(16:0/18:1)	PC(16:0/18:1)		
818.5	oxPC(16:0/18:2[OH])	oxPC(16:0/18:1[O])	oxPC(16:0_18:2[OH])	
820.5	oxPC(16:0/18:1[OH])	oxPC(16:0/18:0[O])	oxPC(16:0_18:1[OH])	
822.6	oxPC(16:0_18:0[OH])			
832.5	oxPC(16:0/18:1[2O])	oxPC(16:0/18:1[2O])		[M+HCO <sub>2</sub> ] <sup>-</sup>
834.5	oxPC(16:0/18:2[2OH])			
836.5		oxPC(16:0/18:0[O,OH])		
838.5	oxPC(16:0/18:0[2OH])	oxPC(16:0/18:0[2OH])		
850.5	oxPC(16:0/18:2[3OH])	oxPC(16:0/18:2[O,OOH])	oxPC(16:0_18:2[3O])	
852.5	oxPC(16:0/18:1[3OH])	oxPC(16:0/18:1[OH,OOH])		

**Dataset 4 - Oxidized Wheat Seeds** As neither LPPtiger nor LipidMatch or LipidMatch Flow could handle dataset 4, a comparison of lipids identified by LDA and the reported lipids in the original paper by Riewe *et al.* [30] was conducted (table 17). Note, that lipids with different retention times are not counted separately and lipids without MS<sup>2</sup> evidence to support chain identification are not listed. The number of identifications per class by LDA resemble the ones made by Riewe *et al.* for DGDG, MGDG and PI, but differ for other lipid classes. Patterns for the frequency of lipids vs. oxidized lipids (their relative number) and their abundances are in agreement between LDA and Riewe *et al.*

Table 17: Unique lipid species identified by LDA vs. reported by Riewe *et al.*

Lipid Class	LDA					Riewe <i>et al.</i>				
	0 OH	1 OH	2 OH	3 OH	4 OH	0 OH	1 OH	2 OH	3 OH	4 OH
TG	68	41	14	8	2	71	58	23	12	9
DG	27	9	10	1	1	35	17	18	11	6
MG	0	0	0	0	0	6	1	0	0	0
PC	2	0	0	0	0	3	1	0	2	1
PE	7	0	0	0	0	8	2	0	0	0
PG	0	0	0	0	0	7	0	0	0	0
PI	5	0	0	0	0	5	0	0	0	0
LPC	0	0	0	0	0	12	4	1	0	0
LPE	0	0	0	0	0	7	0	0	0	0
LPG	0	0	0	0	0	2	0	0	0	0
MGDG	6	3	2	0	0	6	4	3	0	0
DGDG	20	3	4	0	0	20	4	5	0	0

## 4 Discussion

In this work, Lipid Data Analyzer [64] was extended to cover identification of galactolipids and oxidized lipids. Novel rulesets were implemented to accomplish support for galactolipids. The reliability of these rules was tested and a comparison with LipidMatch [65] and LipidMatch Flow [66] was conducted. The new decision rules proved to be a reliable addition to LDA for the identification of MGDG, DGDG, TriGDG, SQDG, MGMG and DGMG, as demonstrated with dataset 2. Further changes to mass lists, FA lists and LDA source code allow analysis of lipids with oxidized fatty acyls. Two datasets with heavily oxidized lipids were investigated and benchmarking the results with LPPtiger [67] proved the validity of the implementation; LDA could even identify more oxidized lipids and handle more datasets.

### 4.1 Tools

New and changing technologies like data-independent acquisition modes of MS data (DIA/MS<sup>E</sup>) and different, non-standardized data formats make it hard to provide a one-fits-all solution for analyzing MS lipidomics data. LipidMatch, LipidMatch Flow, LPPtiger and LDA are widely used lipidomics tools with their individual strengths and weaknesses.

**LipidMatch** In general LipidMatch covers an extensive library of lipids, though TriGDG, TetraGDG, MGMG, DGMG and SQMG are not supported. LipidMatch is implemented in R and in order to use the script, first manual file conversion and peak-picking need to be performed. In contrast to an automated workflow, this process introduces more possibilities to make a mistake while also rendering it tedious. The script runs only on R version 3.3.3 and the provided batch file for MZmine works only under version 2.23. The verification of annotations is only possible via the resulting excel files and cannot be viewed directly in a spectrum which makes it cumbersome. The libraries don't support intensity relationships, giving rise to possible false negatives. An additional downside of LipidMatch is the lack of official support of data originating from MS-systems that are neither from Thermo nor Agilent. This might explain the poor performance analyzing dataset 3 and that LipidMatch prematurely exited analyzing dataset 4.

**LipidMatch Flow** LipidMatch Flow makes the processing of raw files easier, as no manual file conversions and peak-picking are necessary. However, LipidMatch Flow works only if there are blank files present in the study. Consequently, dataset 3 and 4 couldn't be analyzed using LipidMatch Flow. In the other datasets, LipidMatch Flow identified significantly less lipids than LipidMatch, although this could be due to better false positive filtering.

**LPPtiger** LPPtiger uses a unique approach to analyze oxidized lipids. In order to use the tool, pre-existing knowledge about the non-oxidized lipids in the sample is required. From this, the tool will derive possible oxidized products and generate *in silico* spectra to check against the data. The user interface was easy to understand and the results include the annotated spectra, making it easy to verify annotation. A big drawback is that the tool currently only works for PLs and oxPLs in negative ion mode.

**Lipid Data Analyzer** The interface of LDA was easy to understand and intuitive to use. A built-in user-interface makes creating new rules easy and verification of identifications is uncomplicated, as the annotated spectrum can be viewed directly in LDA. The workflow itself is straight-forward as well, no prior file conversions or peak-picking needs to be performed. In contrast to LipidMatch Flow, no blank filtering can be performed, which would be an useful addition as it removes signals from non-biological origins and reduces false positives [98].

### 4.2 Datasets

As galactolipids are prominent in plants and some bacteria [22], dataset 1 from murine liver was analyzed but didn't result in any MS<sup>2</sup> matches. To test the validity of the decision rules, a second dataset was taken from a large-scale metabolomics study on ryegrass [82], found in the MetaboLights [86] repository. In an attempt to get additional datasets, authors of some of the referenced papers were contacted: Cornelia Herrfurth and Ivo Feussner [91] of the Department of Plant Biochemistry of the Georg August University in Göttingen were very cooperative, but could unfortunately only provide MS<sup>1</sup> data. Even in the MetaboLights repository it was hard to find studies with appropriate design (i.e. aimed at lipidomics), method (i.e. LC-MS) and object (i.e. plants or bacteria). For oxidized lipids, two datasets with heavily

oxidized lipids were found: Dataset 3 was taken from a study on oxidized PC standards [67], kindly provided by Maria Fedorova and Zhixu Ni. Dataset 4 originated from a study on oxidized wheat seeds [30], and was kindly provided by David Riewe.

### 4.3 Notation

Even after immense efforts to create a more unified nomenclature (for example by the research team behind LipidHome [3]), and owing to the vast number of lipid species, it’s often confusing to sift through literature from different authors. This holds particularly true for more ”exotic” lipids. Examples being that the term glycerogalactolipids is mostly used synonymously with galactolipids or HexDAGs or that different authors use DAG and TAG while others use DG and TG to denote diacylglycerides and triacylglycerides respectively. To respect the citations, here, fatty acids and fatty acyls are mostly used interchangeably, although they differ in one oxygen atom and one hydrogen atom. Furthermore, over-reporting structural resolution of lipids seems to be a common problem (i.e. reporting the lipid with structural details that are not conferred by fragmentation data) [65]. In this work, the denotation TriGDG and TetraGDG were used, as proposed by Benning and Dörmann [22], whereby species with additional estolides were denoted as, for example DGDG-EN1, leaning on the nomenclature according to Isbell [26].

**Oxidized Lipids** Biochemists often use terms notating high structural resolution of the molecule, for example 1-Palmitoyl-2-(9-keto-12-oxo-10-dodecenoic acid)-PC. It seems that using *ox* as a prefix has prevailed for shorthand notation of lipid classes, e.g. oxPC as used in the works of Reis *et al.* [31]. The shorthand notation for fatty acyls with additional functional groups, whose positions are not known, according to Liebisch *et al.* [10] is that functional groups are shown after the number of double bonds separated by an underscore and followed by the number of groups if there are more than one. This could make for some odd to read notation for molecular species of lipids in cases like oxTG(18:2\_12:1\_O\_18:3\_O\_OOH) or oxPC(16:0\_20:4\_O\_OH\_OOH). In LDA we settled for a more easy to read notation scheme, where the modifications are inside square brackets, like oxPC(28:1[2O]) for MS<sup>1</sup> identifications and oxPC(16:0.12:1[2O]) for MS<sup>2</sup> identifications. The two aforementioned lipid species in Liebisch *et al.* notation would thus become oxTG(18:2.12:1[O]\_18:3[O,OOH]) and oxPC(16:0.20:4[O,OH,OOH]) in LDA notation. Others have been using this type of notation as well [97].

**Ambiguity** Oxo-, keto-, epoxy- and furan- modifications will not be discernible by typical mass-spectrometry analysis, as they have the same masses, and oxo- (keto-, epoxy-, furan-) and hydroxy-modifications differ only in two H-atoms, which could be interpreted as the mass difference of a single double bond (e.g. oxPC(34:1[OH]) has the same mass as oxPC(34:0[O])). A similar issue emerges when the fatty acyl includes a hydroperoxy modification which has the same mass as an hydroxyl-hydroxy modification, and when the chain includes epidioxide which has the same mass as an oxo-hydroxy modification (table 18). The ambiguity in notation is vast and gets more problematic the more oxidation modifications a lipid carries (table 19). Taking an extreme case, oxTG(54:6[4OH]) has the exact same mass as oxTG(54:2[4O]). An optimal notation scheme would take into account these ambiguities, as the molecule’s structural resolution could otherwise be over-reported. Indeed, distinctions between several modifications without proven, distinct fragmentation patterns would be misleading. Research concerning these modification-specific patterns is scarce and needs ultra-high resolution data and high collision energies. For differentiation between modifications, loss of water (-H<sub>2</sub>O), loss of hydrogen peroxide (-H<sub>2</sub>O<sub>2</sub>) or combined loss of two or three water molecules, have been used to propose the presence of hydroxy, peroxy and di-hydroxy or polyhydroxy groups in oxidized PCs [34,99] and deoxygenation (-O) has been observed for PCs with additional O-groups [100]. Separation by ion mobility mass spectrometry could provide excellent resolution for differentiation of modifications. Additionally, investigation of elution profiles of oxidized lipids could lead to distinction between short- and long-chain peroxidation products and modifications [51].

Table 18: LC-MS notation ambiguities

Modification	Incorrect Annotation	Possible Workaround
oxo- [O]	[Ke], [EpO], [Fu] or [+1db OH]	-
keto- [Ke]	[O], [EpO], [Fu] or [+1db OH]	-
epoxy- [EpO]	[O], [Ke], [Fu] or [+1db OH]	-
furan- [Fu]	[O], [Ke], [EpO] or [+1db OH]	-
hydroxy- [OH]	[-1db O], [-1db Ke], [-1db EpO] or [-1db Fu]	Look for additional hydroxy fragments
epidioxy- [EpOO]	[O,OH] and permutations due to previous ambiguities	-
carboxy- [COOH]	[O,OH] and permutations due to previous ambiguities	Look for additional carboxy fragments
hydroperoxy- [OOH]	[2OH] and permutations due to previous ambiguities	Look for additional hydroperoxy fragments
cyclopentane- [Cy]	+ 1db	-

Table 19: LC-MS notation ambiguities - examples

Chemical Formula	Notation	Chemical Formula	Notation
C <sub>42</sub> H <sub>82</sub> O <sub>9</sub> NP	oxPC(34:0[O])	C <sub>42</sub> H <sub>68</sub> O <sub>10</sub> NP	oxPC(34:6[Ke,Fu])
	oxPC(34:0[Ke])		oxPC(34:6[EpO,Fu])
	oxPC(34:0[EpO])		oxPC(37:7[O,OH])
	oxPC(34:1[OH])		oxPC(37:7[Ke,OH])
C <sub>42</sub> H <sub>68</sub> O <sub>10</sub> NP	oxPC(34:6[2O])		oxPC(37:7[EpO,OH])
	oxPC(34:6[2Ke])		oxPC(37:7[Fu,OH])
	oxPC(34:6[2EpO])		oxPC(37:7[COOH])
	oxPC(34:6[O,Ke])		oxPC(37:7[EpOO])
	oxPC(34:6[O,EpO])		oxPC(37:8[2OH])
	oxPC(34:6[O,Fu])		oxPC(37:8[OOH])
	oxPC(34:6[Ke,EpO])		

#### 4.4 Galactolipids

LDA and LipidMatch results concerning galactolipids are mostly in agreement, proving the validity of the new decision rules. LipidMatch-only identifications mostly turned out as false positives and were not reported by LipidMatch Flow. The reliability of the rules for TetraGDG, DGDG-EN1 and DGDG-EN2 remains to be seen, as no such species were detected by LDA in any of the datasets. For ammoniated adducts of galactolipids (in positive ion mode), a low mass cluster of ions seems to be characteristic (supplementary figure S39), which may stem from further breakdown of fatty acid residues. The same cluster can't be observed for the sodium adducts.

**LipidMatch-only Identifications** Supplementary Figure S36 shows one spectrum in which LipidMatch identified [MGDG(10:0.18:3)+Na]<sup>+</sup>. LDA did not report this lipid, because the fragment *HexNa* corresponding to a m/z value 185.041 is missing from the spectrum. In supplementary figure S37 the spectrum is shown, in which LipidMatch identified [MGDG(18:3.18:4)+Na]<sup>+</sup>. Here, the identification was made by ions of minute intensity, whereas highly abundant ions were left unannotated. Consequently, the two lipid identifications are likely false positive identifications. In another case, where LipidMatch identified [MGDG(16:0.18:3)+Na]<sup>+</sup>, the *HexNa* fragment was missing as well, so that LDA did not identify it, but this lipid molecular species was identified with a different adduct by LDA. In all other cases of LipidMatch-only identifications, LDA did identify the same species, but not the same chain combination. These LipidMatch identifications are highly unlikely as well, as the chain fragments are of very small abundance. Most of the false positive identifications were filtered out during analysis with LipidMatch Flow (S44).

## 4.5 Oxidized Lipids

The scope of possible oxidized lipids is enormous as can be seen from the results of a study [51] on oxPCs (table 20). It can be noticed, that the chains can carry more than one oxidative modification, and that combinations of different oxidative modifications are possible. All the moieties shown in table 1 have been observed in fatty acids esterified to phospholipids, where chlorinated and brominated species are products of reactions with hypohalous acids [31]. Likewise, studies investigating oxTGs have shown a broad variety of different oxidative modifications of fatty acyl chains (table 21).

Table 20: PCs and their oxidation products; adapted in LDA shorthand notation from [51]

Lipid Molecular Species	Oxidation Product Molecular Species	Lipid Molecular Species	Oxidation Product Molecular Species
PC(16:0_18:2)	oxPC(16:0.8:0[O])	PC(16:0_20:4)	oxPC(16:0_5:0[OH])
	oxPC(16:0.9:0[O])		oxPC(16:0.8:1[O])
	oxPC(16:0.12:1[O])		oxPC(16:0.8:1[OH])
	oxPC(16:0.12:1[2O])		oxPC(16:0.8:1[O,OH])
	oxPC(16:0.12:1[OH])		oxPC(16:0.9:2[2O])
	oxPC(16:0.18:2[O])		oxPC(16:0.11:2[2OH])
	oxPC(16:0.18:2[OH])		oxPC(16:0.13:3[OH])
	oxPC(16:0.18:2[OOH])		oxPC(16:0.20:4[OH])
	oxPC(16:0.18:2[OH,OOH])		oxPC(16:0.20:4[OOH])
	oxPC(16:0.18:2[3OH])		oxPC(16:0.20:4[2OOH])
	oxPC(16:0.18:2[4OH])		oxPC(16:0.20:4[3OOH])
			oxPC(16:0.20:4[OH,OOH])
			oxPC(16:0.20:4[OH,2OOH])
	oxPC(16:0.20:4[OH,3OOH])		
	oxPC(16:0.20:4[O,OH,OOH])		
	oxPC(16:0.20:2[Cy,O,2OH])		

Table 21: Excerpt of molecular species of oxTG found in studies, adapted in LDA shorthand notation

Study	Oxidation Product Molecular Species
Commercial sunflower seed oil, stored for three month beyond the "best before" date, then kept at 60°C for 18 days in open bottles [101]	oxTG(18:3_13:2[O,OH]_13:2[O,OH])
	oxTG(22:0_13:2[2O]_13:2[2O])
	oxTG(18:3_9:0[O]_18:3)
	oxTG(18:3_13:2[2O]_18:3)
	oxTG(18:3_13:2[O,OH]_18:3)
	oxTG(18:3[OH]_18:3_9:0[O])
	oxTG(18:3[OH]_9:0[O]_9:0[O])
Commercial corn oil and sunflower seed oil, TGs were purified by thin-layer chromatography and oxidation was accelerated by adding tert-Butyl hydroperoxide to the purified TGs [102]	oxTG(18:3_9:0[O]_18:3)
	oxTG(18:2_9:0[O]_18:2[OOH])
	oxTG(18:2_12:1[O]_18:3[O,OOH])
	oxTG(18:1_18:2_12:1[3O])
	oxTG(18:1_12:1[O]_18:1[2O])

**Dataset 1 - Mouse Liver and Dataset 2 Wheat Seeds** Neither dataset 1 nor 2 explicitly concern oxidized lipids, hence, oxidative analysis with LDA had been skipped for these datasets. More importantly, analysis of these datasets with the vast mass lists needed for oxidized lipids wouldn't be feasible on the used desktop home computer.

**Dataset 3 - Oxidized PC Standard** In dataset 3, the identifications of LPPTiger and LDA are very similar, though LDA covers more adducts. After the initial analysis as described in section 3.7.3, a full mass list of oxPCs between oxPC(8:0) and oxPC(40:0) for up to four additional OH-groups was generated and the dataset analyzed. During that analysis, many lipids were reported that turned out to be false positives, and hence, they are not included here. Nonetheless, some promising looking spectra for oxPC(33:2[2OH]) (supplementary figure S25), oxPC(33:2[3OH]) (supplementary figure S26) and oxPC(33:4[3OH]) (supplementary figure S27) were annotated. The false positives can be clearly attributed

to the inaccuracy of the data (masses differ up to 0.1Da) and non-optimized settings for handling the data.

Even by utilizing modification specific fragmentation spectra, identification of functional groups in this dataset remains difficult: Looking at, for example, the identification of PC(16:0/18:0[2OH]) at  $m/z$  of 838.5 (table 16), fragments for  $[M-(16:0)-O]^-$  and  $[M-PChead60-(16:0)-O]^-$  were found, pointing towards an additional O-atom, whereby the identified fragments  $[18:2[2OH]-H_2O]^-$  and  $[18:2[2OH]-H_2O_2]^-$  give reason to suspect an additional OOH-group. In contrast, the reported 2OH-group is improbable, as no specific fragments were identified (like  $[M-H_4O_2]^-$ ,  $[M-(16:0)-H_4O_2]^-$  or  $[18:2[2OH]-H_4O_2]^-$ ). Also, no  $[M-O]^-$ ,  $[M-H_2O]^-$  or  $[M-H_2O_2]^-$  fragments were found, and all reported fragments are of very low abundance, making classification almost impossible.

**Dataset 4 - Oxidized Wheat Seeds** The difference between identifications made by Riewe *et al.* and by LDA stem either from stricter rules and mass tolerances used by LDA, or non-optimized parameters of LDA's 3D algorithm, or both. Either way, the disparity is not due to a problem with the implementation that allows LDA to identify oxidized lipids. Quite the contrary, the analysis further proves the concept: The  $MS^2$  spectra look promising and the abundance patterns match the findings of Riewe *et al.* [30]. The target mass list for analysis was created directly out of the identifications listed in the original paper. A more extensive mass list might have led to more identifications than in the original study by Riewe *et al.*, but analysis wouldn't have been feasible for all 81 raw files of the dataset.

## 4.6 Noisy Spectra

A couple of very noisy spectra (supplementary figure S38) led to false positive identifications in dataset 3 by LDA. Further, it has been observed that  $NH_4$  adducts of various lipid classes (MGDG, DGDG, TG, possibly more) lead to a characteristic low mass cluster of fragments (supplementary figure S39). During analysis of oxidized lipids by LDA, sometimes, oxidized fragments were mistakenly annotated inside this cluster, leading to false positive identifications. Adequate filtering algorithms might be developed to reduce these types of misannotations.

## 4.7 Outlook

The scope of possibilities of oxidized fatty acyls and resulting lipids is enormous, but with this extension of LDA, an important step towards high-throughput oxidative lipidomics was taken. Comparison with other up to date lipidomics tools showed that LDA has a better coverage of the newly implemented lipids, including galactolipids. This extended version of LDA provides researches with a powerful platform to elucidate diseases caused by perturbations in the oxidized lipidome. Future versions of LDA could implement modification-specific fragmentation spectra and (blank) filtering methods, offering yet another layer of verification and false positive safeguarding respectively.

## References

- [1] Moss GP, Smith PAS, Tavernier D: **Glossary of class names of organic compounds and reactivity intermediates based on structure (IUPAC Recommendations 1995)**. *Pure and Applied Chemistry* 1995, **67**(8-9):1307–1375.
- [2] Harayama T, Riezman H: **Understanding the diversity of membrane lipid composition**. *Nature Reviews Molecular Cell Biology* 2018, **19**(5):281–296.
- [3] Foster JM, Moreno P, Fabregat A, Hermjakob H, Steinbeck C, Apweiler R, Wakelam MJO, Vizcaíno JA: **LipidHome: A Database of Theoretical Lipids Optimized for High Throughput Mass Spectrometry Lipidomics**. *PLoS ONE* 2013, **8**(5):e61951.
- [4] Fahy E, Cotter D, Sud M, Subramaniam S: **Lipid classification, structures and tools**. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2011, **1811**(11):637–647.
- [5] **The LIPID MAPS Lipidomics Gateway**. <https://www.lipidmaps.org/>. [Accessed October 4, 2019].
- [6] Fahy E, Sud M, Cotter D, Subramaniam S: **LIPID MAPS online tools for lipid research**. *Nucleic Acids Research* 2007, **35**(Web Server):W606–W612.
- [7] Fahy E, Subramaniam S, Murphy RC, Nishijima M, Raetz CRH, Shimizu T, Spener F, van Meer G, Wakelam MJO, Dennis EA: **Update of the LIPID MAPS comprehensive classification system for lipids**. *Journal of Lipid Research* 2008, **50**(Supplement):S9–S14.
- [8] Fahy E, Subramaniam S, Brown HA, Glass CK, Merrill AH, Murphy RC, Raetz CRH, Russell DW, Seyama Y, Shaw W, Shimizu T, Spener F, van Meer G, van Nieuwenhze MS, White SH, Witztum JL, Dennis EA: **A comprehensive classification system for lipids**. *Journal of Lipid Research* 2005, **46**(5):839–862.
- [9] Han X: *Lipidomics: Comprehensive Mass Spectrometry of Lipids*. Hoboken, NJ, USA : John Wiley & Sons, Inc., 1 edition 2016.
- [10] Liebisch G, Vizcaíno JA, Köfeler H, Trötz Müller M, Griffiths WJ, Schmitz G, Spener F, Wakelam MJO: **Shorthand notation for lipid structures derived from mass spectrometry**. *Journal of Lipid Research* 2013, **54**(6):1523–1530.
- [11] Pauling JK, Hermansson M, Hartler J, Christiansen K, Gallego SF, Peng B, Ahrends R, Ejsing CS: **Proposal for a common nomenclature for fragment ions in mass spectra of lipids**. *PLoS ONE* 2017, **12**(11):e0188394.
- [12] Kohlwein SD: **Opinion articles on lipidomics — A critical assessment of the state-of-the-art**. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2017, **1862**(8):729–730.
- [13] Gross RW: **The evolution of lipidomics through space and time**. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* 2017, **1862**(8):731–739.
- [14] Kita Y, Tokuoka SM, Shimizu T: **Mediator lipidomics by liquid chromatography-tandem mass spectrometry**. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2017, **1862**(8):777–781.
- [15] Hyötyläinen T, Ahonen L, Pöhö P, Orešič M: **Lipidomics in biomedical research-practical considerations**. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2017, **1862**(8):800–803.
- [16] Folch J, Lees MB, Stanley GHS: **A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues**. *The Journal of Biological Chemistry* 1957, **226**:497–509.
- [17] Bligh EG, Dyer WJ: **A Rapid Method of Total Lipid Extraction and Purification**. *Canadian Journal of Biochemistry and Physiology* 1959, **37**(8):911–917.
- [18] Haslam RP, Feussner I: **Green light for lipid fingerprinting**. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2017, **1862**(8):782–785.

- [19] Hu T, Zhang JL: **Mass-spectrometry-based lipidomics**. *Journal of Separation Science* 2017, **41**:351–372.
- [20] Hölzl G, Dörmann P: **Chloroplast Lipids and Their Biosynthesis**. *Annual Review of Plant Biology* 2019, **70**:51–81.
- [21] Napolitano A, Carbone V, Saggese P, Takagaki K, Pizza C: **Novel Galactolipids from the Leaves of *Ipomoea batatas* L.: Characterization by Liquid Chromatography Coupled with Electrospray Ionization-Quadrupole Time-of-Flight Tandem Mass Spectrometry**. *Journal of Agricultural and Food Chemistry* 2007, **55**(25):10289–10297.
- [22] Dörmann P, Benning C: **Galactolipids rule in seed plants**. *Trends in Plant Science* 2002, **7**:112–118.
- [23] Zianni R, Bianco G, Lelario F, Losito I, Palmisano F, Cataldi TR: **Fatty acid neutral losses observed in tandem mass spectrometry with collision-induced dissociation allows regiochemical assignment of sulfoquinovosyl-diacylglycerols**. *Journal of Mass Spectrometry* 2013, **48**(2):205–215.
- [24] Benning C: **Biosynthesis and Function of the Sulfolipid Sulfoquinovosyl Diacylglycerol**. *Annual Review of Plant Physiology and Plant Molecular Biology* 1998, **49**:53–75.
- [25] Moreau RA, Doehlert DC, Welti R, Isaac G, Roth M, Tamura P, Nuñez A: **The Identification of Mono-, Di-, Tri-, and Tetragalactosyl-diacylglycerols and their Natural Estolides in Oat Kernels**. *Lipids* 2008, **43**(6):533–548.
- [26] Isbell TA: **Chemistry and physical properties of estolides**. *Grasas y Aceites* 2011, **62**:8–20.
- [27] Sparvero LJ, Amoscato AA, Kochanek PM, Pitt BR, Kagan VE, Bayır H: **Mass-spectrometry based oxidative lipidomics and lipid imaging: applications in traumatic brain injury**. *Journal of Neurochemistry* 2010, **115**(6):1322–1336.
- [28] Frankel EN: *Lipid Oxidation (eBook)*. Cambridge, Cambs, UK : Woodhead Publishing, 2 edition 2014.
- [29] Anthonymuthu TS, Kim-Campbell N, Bayır H: **Oxidative lipidomics: applications in critical care**. *Current Opinion in Critical Care* 2017, **23**(4):251–256.
- [30] Riewe D, Wiebach J, Altmann T: **Structure Annotation and Quantification of Wheat Seed Oxidized Lipids by High-Resolution LC-MS/MS**. *Plant Physiology* 2017, **175**(2):600–618.
- [31] Reis A, Spickett CM: **Chemistry of phospholipid oxidation**. *Biochimica et Biophysica Acta - Biomembranes* 2012, **1818**(10):2374–2387.
- [32] Ahmed M, Pickova J, Ahmad T, Liaquat M, Farid A, Jahangir M: **Oxidation of Lipids in Foods**. *Sarhad Journal of Agriculture* 2016, **32**(3):230–238.
- [33] Milic I, Griesser E, Vemula V, Ieda N, Nakagawa H, Miyata N, Galano JM, Oger C, Durand T, Fedorova M: **Profiling and relative quantification of multiply nitrated and oxidized fatty acids**. *Analytical and Bioanalytical Chemistry* 2015, **407**(19):5587–5602.
- [34] Domingues MRM, Reis A, Domingues P: **Mass spectrometry analysis of oxidized phospholipids**. *Chemistry and Physics of Lipids* 2008, **156**(1-2):1–12.
- [35] Spickett CM, Pitt AR: **Oxidative Lipidomics Coming of Age: Advances in Analysis of Oxidized Phospholipids in Physiology and Pathology**. *Antioxidants & Redox Signaling* 2015, **22**(18):1646–1666.
- [36] Gugiu BG, Mesáros CA, Sun M, Gu X, Crabb JW, Salomon RG: **Identification of Oxidatively Truncated Ethanolamine Phospholipids in Retina and Their Generation from Polyunsaturated Phosphatidylethanolamines**. *Chemical Research in Toxicology* 2006, **19**(2):262–271.
- [37] Parsons BJ, Spickett CM: **Special issue on "Analytical methods for the detection of oxidized biomolecules and antioxidants"**. *Free Radical Research* 2015, **49**(5):473–476.

- [38] Zeb A: **Chemistry and liquid chromatography methods for the analyses of primary oxidation products of triacylglycerols.** *Free Radical Research* 2015, **49**(5):549–564.
- [39] Avramova V, Abdelgawad H, Vasileva I, Petrova AS, Holec A, Mariën J, Asard H, Beemster GTS: **High Antioxidant Activity Facilitates Maintenance of Cell Division in Leaves of Drought Tolerant Maize Hybrids.** *Frontiers in Plant Science* 2017, **8**:84–103.
- [40] Broin M, Rey P: **Potato Plants Lacking the CDSP<sub>32</sub> Plastidic Thioredoxin Exhibit Overoxidation of the BAS<sub>1 2</sub>-Cysteine Peroxiredoxin and Increased Lipid Peroxidation in Thylakoids under Photooxidative Stress.** *Plant Physiology* 2003, **132**(3):1335–1343.
- [41] Feigl G, Lehotai N, Molnár Á, Ördög A, Rodríguez-Ruiz M, Palma JM, Corpas FJ, Erdei L, Kolbert Z: **Zinc induces distinct changes in the metabolism of reactive oxygen and nitrogen species (ROS and RNS) in the roots of two *Brassica* species with different sensitivity to zinc stress.** *Annals of Botany* 2014, **116**(4):613–625.
- [42] Pandey S, Fartyal D, Agarwal A, Shukla T, James D, Kaul T, Negi YK, Arora S, Reddy MK: **Abiotic Stress Tolerance in Plants: Myriad Roles of Ascorbate Peroxidase.** *Frontiers in Plant Science* 2017, **8**:581–594.
- [43] Pierre JL, Favier AE, Cadet J, Kalyanaraman B, Fontecave M: *Analysis of Free Radicals in Biological Systems.* Basel, BS, CH: Birkhäuser, 1 edition 1995.
- [44] Shrestha R, Hui SP, Miura Y, Yagi A, Takahashi Y, Takeda S, Fuda H, Chiba H: **Identification of molecular species of oxidized triglyceride in plasma and its distribution in lipoproteins.** *Clinical Chemistry and Laboratory Medicine* 2015, **53**(11):1859–1869.
- [45] Leitinger N: **The Role of Phospholipid Oxidation Products in Inflammatory and Autoimmune Diseases.** In *Lipids in Health and Disease.* Edited by Quinn PJ, Wang X, Dordrecht, ZH, NL: Springer Netherlands 2008:325–350.
- [46] Keller J, Camaré C, Bernis C, Astello-García M, de la Rosa APB, Rossignol M, del Socorro Santos Díaz M, Salvayre R, Negre-Salvayre A, Guéraud F: **Antiatherogenic and antitumoral properties of *Opuntia cladodes*: inhibition of low density lipoprotein oxidation by vascular cells, and protection against the cytotoxicity of lipid oxidation product 4-hydroxynonenal in a colorectal cancer cellular model.** *Journal of Physiology and Biochemistry* 2015, **71**(3):577–587.
- [47] Megli FM, Russo L: **Different oxidized phospholipid molecules unequally affect bilayer packing.** *Biochimica et Biophysica Acta - Biomembranes* 2008, **1778**:143–152.
- [48] Borst JW, Visser NV, Kouptsova O, Visser AJ: **Oxidation of unsaturated phospholipids in membrane bilayer mixtures is accompanied by membrane fluidity changes.** *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* 2000, **1487**:61–73.
- [49] Kato S, Shimizu N, Hanzawa Y, Otoki Y, Ito J, Kimura F, Takekoshi S, Sakaino M, Sano T, Eitsuka T, Miyazawa T, Nakagawa K: **Determination of triacylglycerol oxidation mechanisms in canola oil using liquid chromatography-tandem mass spectrometry.** *npj Science of Food* 2018, **2**:1.
- [50] Yin H, Cox BE, Liu W, Porter NA, Morrow JD, Milne GL: **Identification of intact oxidation products of glycerophospholipids *in vitro* and *in vivo* using negative ion electrospray iontrap mass spectrometry.** *Journal of Mass Spectrometry* 2009, **44**(5):672–680.
- [51] Reis A, Domingues MRM, Amado FML, Ferrer-Correia AJV, Domingues P: **Separation of peroxidation products of diacyl-phosphatidylcholines by reversed-phase liquid chromatography-mass spectrometry.** *Biomedical Chromatography* 2005, **19**(2):129–137.
- [52] Podrez EA, Poliakov E, Shen Z, Zhang R, Deng Y, Sun M, Finton PJ, Shan L, Gugiu B, Fox PL, Hoff HF, Salomon RG, Hazen SL: **Identification of a Novel Family of Oxidized Phospholipids That Serve as Ligands for the Macrophage Scavenger Receptor CD36.** *Journal of Biological Chemistry* 2002, **277**(41):38503–38516.

- [53] Berry KAZ, Murphy RC: **Free Radical Oxidation of Plasmalogen Glycerophosphocholine Containing Esterified Docosahexaenoic Acid: Structure Determination by Mass Spectrometry.** *Antioxidants & Redox Signaling* 2005, **7**(1-2):157–169.
- [54] Göpfert MS, Siedler F, Siess W, Sellmayer A: **Structural Identification of Oxidized Acyl-Phosphatidylcholines That Induce Platelet Activation.** *Journal of Vascular Research* 2005, **42**(2):120–132.
- [55] Uhlson C, Harrison K, Allen CB, Ahmad S, White CW, Murphy RC: **Oxidized Phospholipids Derived from Ozone-Treated Lung Surfactant Extract Reduce Macrophage and Epithelial Cell Viability.** *Chemical Research in Toxicology* 2002, **15**(7):896–906.
- [56] Watson AD, Leitinger N, Navab M, Faull KF, Hörkkö S, Witztum JL, Palinski W, Schwenke D, Salomon RG, Sha W, Subbanagounder G, Fogelman AM, Berliner JA: **Structural Identification by Mass Spectrometry of Oxidized Phospholipids in Minimally Oxidized Low Density Lipoprotein That Induce Monocyte/Endothelial Interactions and Evidence for Their Presence *in Vivo*.** *Journal of Biological Chemistry* 1997, **272**(21):13597–13607.
- [57] Khaselev N, Murphy RC: **Structural characterization of oxidized phospholipid products derived from arachidonate-containing plasmenyl glycerophosphocholine.** *Journal of Lipid Research* 2000, **41**(4):564–572.
- [58] Slatter DA, Aldrovandi M, O’Connor A, Allen SM, Brasher CJ, Murphy RC, Mecklemann S, Ravi S, Darley-Usmar V, O’Donnell VB: **Mapping the Human Platelet Lipidome Reveals Cytosolic Phospholipase A<sub>2</sub> as a Regulator of Mitochondrial Bioenergetics during Activation.** *Cell Metabolism* 2016, **23**(5):930–944.
- [59] Aoyagi R, Ikeda K, Isobe Y, Arita M: **Comprehensive analyses of oxidized phospholipids using a measured MS/MS spectra library.** *Journal of Lipid Research* 2017, **58**(11):2229–2237.
- [60] Reis A, Domingues P, Ferrer-Correia AJV, Domingues MRM: **Peptide-phospholipid cross-linking reactions: Identification of leucine enkephalin-alka(e)nal-glycerophosphatidylcholine adducts by tandem mass spectrometry.** *Journal of the American Society for Mass Spectrometry* 2006, **17**(5):657–660.
- [61] Terao J, Shibata SS, Matsushita S: **Selective quantification of arachidonic acid hydroperoxides and their hydroxy derivatives in reverse-phase high performance liquid chromatography.** *Analytical Biochemistry* 1988, **169**(2):415–423.
- [62] Yoshida Y, Hayakawa M, Niki E: **Total hydroxyoctadecadienoic acid as a marker for lipid peroxidation *in vivo*.** *BioFactors* 2005, **24**(1-4):7–15.
- [63] Yoshida Y, Umeno A, Shichiri M: **Lipid peroxidation biomarkers for evaluating oxidative stress and assessing antioxidant capacity *in vivo*.** *Journal of Clinical Biochemistry and Nutrition* 2013, **52**:9–16.
- [64] Hartler J, Triebel A, Ziegl A, Trötzlmüller M, Rechberger GN, Zeleznik OA, Zierler KA, Torta F, Cazenave-Gassiot A, Wenk MR, Fauland A, Wheelock CE, Armando AM, Quehenberger O, Zhang Q, Wakelam MJO, Haemmerle G, Spener F, Köfeler HC, Thallinger GG: **Deciphering lipid structures based on platform-independent decision rules.** *Nature Methods* 2017, **14**:1171–1180.
- [65] Koelmel JP, Kroeger NM, Ulmer CZ, Bowden JA, Patterson RE, Cochran JA, Beecher CWW, Garrett TJ, Yost RA: **LipidMatch: an automated workflow for rule-based lipid identification using untargeted high-resolution tandem mass spectrometry data.** *BMC Bioinformatics* 2017, **18**:331.
- [66] Koelmel JP, Li Y, Cochran JA, Patt AC, Kroeger NM, Bowden JA, Ulmer CZ, Mathé E, Yost RA, Garrett TJ: **LipidMatch Flow: A UserFriendly Software Covering the Entire Lipidomics Workflow.** In *Proceedings of the 66th ASMS Conference on Mass Spectrometry and Allied Topics*, 3-6-2018, San Diego, CA, USA 2018.

- [67] Ni Z, Angelidou G, Hoffmann R, Fedorova M: **LPPtiger software for lipidome-specific prediction and identification of oxidized phospholipids from LC-MS datasets**. *Scientific Reports* 2017, **7**:15138.
- [68] Kessner D, Chambers M, Burke R, Agus D, Mallick P: **ProteoWizard: open source software for rapid proteomics tools development**. *Bioinformatics* 2008, **24**(21):2534–2536.
- [69] Pluskal T, Castillo S, Villar-Briones A, Orešič M: **MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data**. *BMC Bioinformatics* 2010, **11**:395–406.
- [70] **LipidMatch Download**. <http://secim.ufl.edu/secim-tools/lipidmatch/>.
- [71] **LipidMatch Flow Download**. <http://secim.ufl.edu/secim-tools/lipidmatchflow/>.
- [72] R Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria 2017. [<https://www.R-project.org/>].
- [73] RStudio Team: *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA 2016.
- [74] Mirai Solutions GmbH: *XLConnect: Excel Connector for R* 2018. [R package version 0.2-15].
- [75] Firke S: *janitor: Simple Tools for Examining and Cleaning Dirty Data* 2019. [R package version 1.2.0].
- [76] Wickham H: **The Split-Apply-Combine Strategy for Data Analysis**. *Journal of Statistical Software* 2011, **40**:1–29.
- [77] Borchers HW: *pracma: Practical Numerical Math Functions* 2019. [R package version 2.2.5].
- [78] Dowle M, Srinivasan A: *data.table: Extension of ‘data.frame’* 2017. [R package version 1.10.4-3].
- [79] Ulmer CZ, Koelmel JP, Ragland JM, Garrett TJ, Bowden JA: **LipidPioneer : A Comprehensive User-Generated Exact Mass Template for Lipidomics**. *Journal of The American Society for Mass Spectrometry* 2017, **28**(3):562–565.
- [80] **Drawboard: Easier PDF markup software**. <https://www.drawboard.com/pdf/>.
- [81] **Dataset 1 - Mouse Liver Download**. [http://genome.tugraz.at/lda2/data/Biological/Bio\\_OrbitrapVelosProHCD.zip](http://genome.tugraz.at/lda2/data/Biological/Bio_OrbitrapVelosProHCD.zip).
- [82] Subbaraj AK, Huege J, Fraser K, Cao M, Rasmussen S, Faville M, Harrison SJ, Jones CS: **A large-scale metabolomics study to harness chemical diversity and explore biochemical mechanisms in ryegrass**. *Communications Biology* 2019, **2**:87.
- [83] **Dataset 2 - Ryegrass Leaves Download; Positive Ion Mode**. <https://www.ebi.ac.uk/metabolights/MTBLS66>.
- [84] **Dataset 2 - Ryegrass Leaves Download; Negative Ion Mode**. <https://www.ebi.ac.uk/metabolights/MTBLS68>.
- [85] de Laeter JR, Böhlke JK, Bièvre PD, Hidaka H, Peiser HS, Rosman KJR, Taylor PDP: **Atomic weights of the elements. Review 2000 (IUPAC Technical Report)**. *Pure and Applied Chemistry* 2003, **75**(6):683–800.
- [86] Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, Mahendrakar T, Williams M, Neumann S, Rocca-Serra P, Maguire E, González-Beltrán A, Sansone SA, Griffin JL, Steinbeck C: **MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data**. *Nucleic Acids Research* 2012, **41**(D1):D781–D786.
- [87] Wang W, Liu Z, Ma L, Hao C, Liu S, Voinov VG, Kalinovskaya NI: **Electrospray ionization multiple-stage tandem mass spectrometric analysis of diglycosyldiacylglycerol glycolipids from the bacteria *Bacillus pumilus***. *Rapid Communications in Mass Spectrometry* 1999, **13**(12):1189–1196.

- [88] Welti R, Wang X, Williams TD: **Electrospray ionization tandem mass spectrometry scan modes for plant chloroplast lipids.** *Analytical Biochemistry* 2003, **314**:149 – 152.
- [89] Tatituri RVV, Brenner MB, Turk J, Hsu FF: **Structural elucidation of diglycosyl diacylglycerol and monoglycosyl diacylglycerol from *Streptococcus pneumoniae* by multiple-stage linear ion-trap mass spectrometry with electrospray ionization.** *Journal of Mass Spectrometry* 2012, **47**:115–123.
- [90] Sasaki GL, Gorin PA, Tischer CA, Iacomini M: **Sulfonoglycolipids from the lichenized basidiomycete *Dictyonema glabratum*: isolation, NMR, and ESI-MS approaches.** *Glycobiology* 2001, **11**(4):345–351.
- [91] Ibrahim A, Schütz AL, Galano JM, Herrfurth C, Feussner K, Durand T, Brodhun F, Feussner I: **The Alphabet of Galactolipids in *Arabidopsis thaliana*.** *Frontiers in Plant Science* 2011, **2**:95.
- [92] Guella G, Frassanito R, Mancini I: **A new solution for an old problem: the regiochemical distribution of the acyl chains in galactolipids can be established by electrospray ionization tandem mass spectrometry.** *Rapid Communications in Mass Spectrometry* 2003, **17**(17):1982–1994.
- [93] Buseman CM: **Wounding Stimulates the Accumulation of Glycerolipids Containing Oxophytodienoic Acid and Dinor-Oxophytodienoic Acid in *Arabidopsis* Leaves.** *Plant Physiology* 2006, **142**:28–39.
- [94] Kim YH, Yoo JS, Kim MS: **Structural characterization of sulfoquinovosyl, monogalactosyl and digalactosyl diacylglycerols by FAB-CID-MS/MS.** *Journal of Mass Spectrometry* 1997, **32**(9):968–977.
- [95] Cedergren RA, Hollingsworth RI: **Occurrence of sulfoquinovosyl diacylglycerol in some members of the family Rhizobiaceae.** *The Journal of Lipid Research* 1994, **35**:1452–1461.
- [96] Zhang X, Phaner CJ, Ferguson-Miller SM, Reid GE: **Evaluation of ion activation strategies and mechanisms for the gas-phase fragmentation of sulfoquinovosyldiacylglycerol lipids from *Rhodobacter sphaeroides*.** *International Journal of Mass Spectrometry* 2012, **316-318**:100–107.
- [97] Sala P, Pötz S, Brunner M, Trötz Müller M, Fauland A, Triebel A, Hartler J, Lankmayr E, Köfeler H: **Determination of Oxidized Phosphatidylcholines by Hydrophilic Interaction Liquid Chromatography Coupled to Fourier Transform Mass Spectrometry.** *International Journal of Molecular Sciences* 2015, **16**(12):8351–8363.
- [98] Patterson RE, Kirpich AS, Koelmel JP, Kalavalapalli S, Morse AM, Cusi K, Sunny NE, McIntyre LM, Garrett TJ, Yost RA: **Improved experimental data processing for UH-PLC–HRMS/MS lipidomics applied to nonalcoholic fatty liver disease.** *Metabolomics* 2017, **13**(11):142–153.
- [99] Reis A, Domingues M, Amado F, Ferrer-Correia A, Domingues P: **Radical peroxidation of palmitoyl-linoleoyl-glycerophosphocholine liposomes: Identification of long-chain oxidised products by liquid chromatography–tandem mass spectrometry.** *Journal of Chromatography B* 2007, **855**(2):186–199.
- [100] Adachi J, Yoshioka N, Sato M, Nakagawa K, Yamamoto Y, Ueno Y: **Detection of phosphatidylcholine oxidation products in rat heart using quadrupole time-of-flight mass spectrometry.** *Journal of Chromatography B* 2005, **823**:37–43.
- [101] Sjövall O, Kuksis A, Kallio H: **Tentative identification and quantification of TAG core aldehydes as dinitrophenylhydrazones in autoxidized sunflowerseed oil using reversed-phase HPLC with electrospray ionization MS.** *Lipids* 2003, **38**(11):1179–1190.
- [102] Sjövall O, Kuksis A, Kallio H: **Formation of triacylglycerol core aldehydes during rapid oxidation of corn and sunflower oils with *tert*-butyl hydroperoxide/Fe<sup>2+</sup>.** *Lipids* 2002, **37**:81–94.

## 5 Appendix

### 5.1 LipidMatch Results

#### 5.1.1 LM - Dataset 1

Table S1: MS<sup>2</sup> confirmed lipids in dataset 1 (positive ion mode) by LipidMatch

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	Co(Q9)	30	PC(35:1)	59	PE(40:5)	88	TG(52:1)
2	DG(34:2)	31	PC(35:3)	60	PE(38:5)	89	TG(51:3)
3	DG(36:3)	32	PC(38:4)	61	PE(40:7)	90	TG(54:7)
4	DG(36:2)	33	PC(40:6)	62	PE(36:3)	91	TG(51:2)
5	DG(36:4)	34	PC(38:3)	63	PEtOH(38:6)	92	TG(50:0)
6	LPC(18:1)	35	PC(36:1)	64	Plasmany1-PC(34:1)	93	TG(56:6)
7	LPC(18:2)	36	PC(36:2)	65	Plasmany1-PC(38:5)	94	TG(50:4)
8	LPC(20:3)	37	PC(40:5)	66	SM(42:2)	95	TG(56:9)
9	LPC(20:4)	38	PC(38:2)	67	SM(34:1)	96	TG(52:6)
10	LPC(22:6)	39	PC(40:4)	68	SM(40:1)	97	TG(53:3)
11	LPE(18:0)	40	PC(38:5)	69	SM(42:1)	98	TG(53:2)
12	PC(32:2)	41	PC(36:3)	70	SM(41:1)	99	TG(53:4)
13	PC(34:4)	42	PC(40:7)	71	TG(48:1)	100	TG(54:3)
14	PC(33:2)	43	PC(40:8)	72	TG(48:2)	101	TG(54:2)
15	PC(35:4)	44	PC(37:2)	73	TG(54:0)	102	TG(56:2)
16	PC(36:4)	45	PC(39:4)	74	TG(48:3)	103	TG(54:1)
17	PC(38:6)	46	PC(42:10)	75	TG(50:5)	104	TG(58:5)
18	PC(34:1)	47	PE(24:0)	76	TG(49:3)	105	TG(58:7)
19	PC(32:0)	48	PE(38:6)	77	TG(52:3)	106	TG(54:4)
20	PC(34:2)	49	PE(36:4)	78	TG(52:2)	107	TG(56:4)
21	PC(32:1)	50	PE(34:2)	79	TG(50:2)	108	TG(58:9)
22	PC(34:0)	51	PE(34:1)	80	TG(56:8)	109	TG(58:8)
23	PC(34:3)	52	PE(36:5)	81	TG(56:7)	110	TG(56:5)
24	PC(33:1)	53	PE(38:7)	82	TG(54:6)	111	TG(58:6)
25	PC(36:5)	54	PE(40:6)	83	TG(50:1)	112	TG(56:3)
26	PC(38:7)	55	PE(36:2)	84	TG(54:5)	113	TG(55:3)
27	PC(37:4)	56	PE(38:4)	85	TG(50:3)	114	TG(58:10)
28	PC(35:2)	57	PE(38:3)	86	TG(52:5)	115	TG(62:14)
29	PC(39:6)	58	PE(36:1)	87	TG(52:4)		

Table S2: MS<sup>2</sup> confirmed lipids in dataset 1 (negative ion mode) by LipidMatch

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	Cer-NS(42:2)	19	PC(32:0)	37	PC(40:8)	55	PE(38:5)
2	CL(78:25)	20	PC(36:5)	38	PC(37:2)	56	PE(40:6)
3	CL(74:22)	21	PC(32:1)	39	PC(39:4)	57	PE(40:7)
4	LPC(16:0)	22	PC(34:3)	40	PE(24:0)	58	PE(36:3)
5	LPC(18:0)	23	PC(34:0)	41	PE(38:6)	59	PI(36:4)
6	LPC(18:1)	24	PC(38:7)	42	PE(36:4)	60	PI(34:2)
7	LPC(18:2)	25	PC(37:4)	43	PE(34:2)	61	PI(37:4)
8	LPC(20:4)	26	PC(35:2)	44	PE(34:1)	62	PI(38:4)
9	LPE(16:0)	27	PC(38:4)	45	PE(36:5)	63	PI(38:3)
10	LPE(18:0)	28	PC(36:2)	46	PE(38:7)	64	PI(38:5)
11	LPE(18:1)	29	PC(38:3)	47	PE(34:3)	65	PI(36:3)
12	LPE(20:4)	30	PC(36:1)	48	PE(37:4)	66	Plasmenyl-PE(36:4)
13	oxPG(34:1)	31	PC(40:5)	49	PE(38:4)	67	Plasmenyl-PE(38:4)
14	PC(32:2)	32	PC(38:2)	50	PE(36:2)	68	Plasmenyl-PE(38:5)
15	PC(34:2)	33	PC(40:6)	51	PE(38:3)	69	Plasmenyl-PE(38:6)
16	PC(36:4)	34	PC(38:5)	52	PE(40:5)	70	PS(40:6)
17	PC(38:6)	35	PC(36:3)	53	PE(36:1)	71	PS(38:4)
18	PC(34:1)	36	PC(40:7)	54	PE(40:4)		

### 5.1.2 LMF - Dataset 1

Table S3: MS<sup>2</sup> confirmed lipids in dataset 1 (positive ion mode) by LipidMatch Flow

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	DG(34:2)	23	PC(38:4)	45	PE(38:4)	67	TG(56:5)
2	DG(36:3)	24	PC(40:6)	46	PE(36:1)	68	TG(52:3)
3	DG(36:4)	25	PC(36:1)	47	PE(38:5)	69	TG(52:4)
4	LPC(18:1)	26	PC(40:5)	48	PE(36:3)	70	TG(52:2)
5	LPC(18:2)	27	PC(36:2)	49	PE(40:7)	71	TG(52:1)
6	LPC(20:4)	28	PC(40:4)	50	Plasmenyl-PC(34:1)	72	TG(54:7)
7	LPC(22:6)	29	PC(38:3)	51	Plasmenyl-PC(38:5)	73	TG(51:2)
8	LPE(18:0)	30	PC(38:5)	52	PS(38:4)	74	TG(51:3)
9	PC(34:4)	31	PC(40:7)	53	SM(40:1)	75	TG(50:4)
10	PC(36:5)	32	PC(36:3)	54	TG(48:1)	76	TG(56:9)
11	PC(34:3)	33	PC(40:8)	55	TG(48:2)	77	TG(52:6)
12	PC(36:4)	34	PC(37:2)	56	TG(48:3)	78	TG(53:3)
13	PC(34:2)	35	PC(39:4)	57	TG(49:3)	79	TG(53:2)
14	PC(38:6)	36	PC(42:10)	58	TG(50:2)	80	TG(53:4)
15	PC(34:0)	37	PE(24:0)	59	TG(56:6)	81	TG(54:3)
16	PC(34:1)	38	PE(38:6)	60	TG(50:3)	82	TG(54:2)
17	PC(33:1)	39	PE(36:4)	61	TG(50:1)	83	TG(54:4)
18	PC(37:4)	40	PE(34:2)	62	TG(56:8)	84	TG(58:9)
19	PC(35:2)	41	PE(34:1)	63	TG(54:5)	85	TG(58:8)
20	PC(39:6)	42	PE(36:5)	64	TG(56:7)	86	TG(56:4)
21	PC(35:1)	43	PE(38:7)	65	TG(52:5)	87	TG(58:7)
22	PC(35:3)	44	PE(40:6)	66	TG(54:6)		

Table S4: MS<sup>2</sup> confirmed lipids in dataset 1 (negative ion mode) by LipidMatch Flow

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	LPC(16:0)	11	PC(34:3)	21	PC(37:2)	31	PE(38:4)
2	LPC(18:1)	12	PC(34:0)	22	PC(39:4)	32	PI(37:4)
3	LPC(18:2)	13	PC(38:7)	23	PE(38:6)	33	PI(38:3)
4	LPC(20:4)	14	PC(37:4)	24	PE(36:5)	34	PI(36:3)
5	LPE(18:1)	15	PC(35:2)	25	PE(34:3)	35	Plasmenyl-PE((38:4)
6	LPE(20:4)	16	PC(36:2)	26	PE(38:7)	36	Plasmenyl-PE((38:6)
7	PC(34:2)	17	PC(36:1)	27	PE(37:4)	37	PS(38:4)
8	PC(38:6)	18	PC(40:5)	28	PE(40:5)		
9	PC(32:1)	19	PC(38:2)	29	PE(36:1)		
10	PC(36:5)	20	PC(38:5)	30	PE(40:4)		

### 5.1.3 LM - Dataset 2

Table S5: MS<sup>2</sup> confirmed lipids in dataset 2 (positive ion mode) by LipidMatch

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	AcCar(5:0)	20	MG(18:3)	39	PC(34:1)	58	So(18:0)
2	AcCar(6:0)	21	MGDG(28:3)	40	PC(34:3)	59	So(18:1)
3	Co(Q9)	22	MGDG(34:3)	41	PC(36:6)	60	TG(34:0)
4	DG(34:3)	23	MGDG(36:6)	42	PE(34:3)	61	TG(52:5)
5	DG(34:2)	24	MGDG(36:4)	43	PE(34:2)	62	TG(52:6)
6	DG(32:0)	25	MGDG(36:7)	44	PE(34:4)	63	TG(52:4)
7	DG(34:4)	26	oxLPC(24:1[Ke,OH])	45	PE(36:3)	64	TG(50:2)
8	DG(36:5)	27	oxLPC(24:1[2O])	46	PE(36:5)	65	TG(50:3)
9	DG(36:4)	28	oxTG(50:1[OH])	47	PE(36:4)	66	TG(52:3)
10	DG(36:6)	29	oxTG(52:2[OH])	48	PE(36:6)	67	TG(52:2)
11	DGDG(34:3)	30	oxTG(54:9[OH])	49	PE(39:3)	68	TG(50:1)
12	DGDG(36:6)	31	oxTG(52:2[2OH])	50	PEtOH(40:6)	69	TG(52:1)
13	LPC(16:0)	32	PA(34:2)	51	PG(26:1)	70	TG(54:4)
14	LPC(18:2)	33	PA(34:3)	52	PG(34:2)	71	TG(54:3)
15	LPC(18:3)	34	PA(34:4)	53	PG(34:3)	72	TG(54:2)
16	LPE(16:0)	35	PA(36:3)	54	PG(34:4)	73	TG(54:6)
17	LPE(18:2)	36	PA(36:5)	55	PG(36:7)	74	TG(54:8)
18	LPE(18:3)	37	PA(36:4)	56	PS(42:3)	75	TG(54:7)
19	LPE(8:0)	38	PA(36:6)	57	PS(40:3)	76	TG(54:9)

Table S6: MS<sup>2</sup> confirmed lipids in dataset 2 (negative ion mode) by LipidMatch

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	AcylGlcADG(52:5)	26	LPA(16:0)	51	PA(34:4)	76	PE(36:6)
2	CerAP(42:2)	27	LPA(18:0)	52	PA(36:2)	77	PG(32:3)
3	CerAP(42:1)	28	LPA(18:3)	53	PA(36:3)	78	PG(32:2)
4	CerAP(40:1)	29	LPC(16:0)	54	PA(36:5)	79	PG(32:4)
5	CerAP(42:0)	30	LPC(18:2)	55	PA(36:4)	80	PG(34:2)
6	CerAP(43:1)	31	LPE(16:0)	56	PA(36:6)	81	PG(32:1)
7	CL(70:2)	32	LPE(18:2)	57	PA(36:7)	82	PG(34:3)
8	CL(70:4)	33	LPE(18:3)	58	PC(33:4)	83	PG(32:0)
9	DGDG(34:3)	34	MGDG(34:3)	59	PC(34:3)	84	PG(34:4)
10	DGDG(34:2)	35	MGDG(34:2)	60	PC(34:2)	85	PG(34:5)
11	DGDG(34:4)	36	MGDG(34:4)	61	PC(34:1)	86	PG(36:5)
12	DGDG(34:1)	37	MGDG(36:6)	62	PC(40:1)	87	PG(36:4)
13	DGDG(35:3)	38	oxCL(74:8[OH])	63	PC(36:3)	88	PG(36:7)
14	DGDG(36:6)	39	oxPC(34:3[3O])	64	PC(36:5)	89	PG(36:6)
15	DGDG(36:3)	40	oxPC(34:2[OOH])	65	PC(36:4)	90	PI(34:3)
16	DGDG(36:4)	41	oxPC(29:2[COOH])	66	PC(36:6)	91	PI(34:2)
17	FAHFA(28:0)	42	oxPC(36:6[3O])	67	PE(31:1)	92	PI(36:5)
18	FAHFA(27:3)	43	oxPE(34:3[3O])	68	PE(34:3)	93	PI(36:4)
19	FAHFA(21:0)	44	oxPE(36:6[3O])	69	PE(34:2)	94	PS(36:5)
20	FAHFA(36:5)	45	oxPE(36:5[3O])	70	PE(34:1)	95	PS(36:4)
21	FAHFA(36:3)	46	oxPG(34:1[O])	71	PE(32:2)	96	PS(40:3)
22	HexCerAP(42:2)	47	PA(33:3)	72	PE(33:2)	97	SQDG(34:3)
23	HexCerAP(42:1)	48	PA(33:4)	73	PE(36:3)	98	SQDG(32:0)
24	HexCerAP(41:1)	49	PA(34:2)	74	PE(36:5)	99	SQDG(36:6)
25	HexCerAP(40:1)	50	PA(34:3)	75	PE(36:4)	100	SQDG(36:3)

### 5.1.4 LMF - Dataset 2

Table S7: MS<sup>2</sup> confirmed lipids in dataset 2 (positive ion mode) by LipidMatch Flow

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	DG(34:3)	9	LPC(18:2)	17	MGDG(36:6)	25	PE(36:3)
2	DG(34:2)	10	LPC(18:3)	18	MGDG(36:4)	26	PG(26:1)
3	DG(34:4)	11	LPE(16:0)	19	MGDG(36:7)	27	PG(34:4)
4	DG(36:5)	12	LPE(18:2)	20	oxLPC(24:1[2O])	28	So(18:0)
5	DG(36:4)	13	LPE(18:3)	21	oxTG(50:1[OH])	29	TG(52:6)
6	DG(36:6)	14	MG(18:3)	22	oxTG(52:2[2OH])	30	TG(50:3)
7	DGDG(36:6)	15	MGDG(28:3)	23	PC(34:3)	31	TG(50:2)
8	LPC(16:0)	16	MGDG(34:3)	24	PE(34:3)		

Table S8: MS<sup>2</sup> confirmed lipids in dataset 2 (negative ion mode) by LipidMatch Flow

Lipid Species		Lipid Species		Lipid Species		Lipid Species	
1	AcylGlcADG(52:5)	18	LPC(18:2)	35	PA(36:5)	52	PG(32:3)
2	Cer-AP(42:2)	19	LPE(16:0)	36	PA(36:4)	53	PG(32:4)
3	Cer-AP(42:1)	20	LPE(18:2)	37	PA(36:6)	54	PG(34:2)
4	Cer-AP(43:1)	21	LPE(18:3)	38	PA(36:7)	55	PG(32:1)
5	CL(70:4)	22	MGDG(34:3)	39	PC(33:4)	56	PG(34:4)
6	DGDG(34:3)	23	MGDG(34:2)	40	PC(34:3)	57	PG(34:5)
7	DGDG(34:4)	24	MGDG(36:6)	41	PC(34:2)	58	PG(32:2)
8	DGDG(34:1)	25	oxPC(29:2[COOH])	42	PC(36:5)	59	PG(36:7)
9	DGDG(35:3)	26	oxPE(34:3[3O])	43	PC(36:6)	60	PI(34:3)
10	DGDG(36:6)	27	oxPE(36:6[3O])	44	PE(31:1)	61	PI(34:2)
11	FAHFA(36:5)	28	oxPE(36:5[3O])	45	PE(34:3)	62	PS(40:3)
12	HexCer-AP(42:2)	29	oxPG(34:3[O])	46	PE(34:2)	63	SQDG(34:3)
13	HexCer-AP(42:1)	30	PA(33:3)	47	PE(32:2)	64	SQDG(32:0)
14	HexCer-AP(41:1)	31	PA(33:4)	48	PE(36:3)	65	SQDG(36:6)
15	LPA(18:0)	32	PA(34:3)	49	PE(36:5)	66	SQDG(36:3)
16	LPA(18:3)	33	PA(34:2)	50	PE(36:4)		
17	LPC(16:0)	34	PA(34:4)	51	PE(36:6)		

## 5.2 LDA Mass Lists

### Dataset 1 and 2

Table S9: Mass list dataset 1 and 2 - MGDG positive ion mode (part 1/6)

Name	dfs	C	H	O	mass(form[+NH4] name[NH4])	mass(form[+Na] name[Na])	mass(form[+H] name[H])
20	: 0	29	54	10	580.4055235	585.360919	563.3789744
20	: 1	29	52	10	578.3898734	583.345269	561.3633243
20	: 2	29	50	10	576.3742234	581.3296189	559.3476743
20	: 3	29	48	10	574.3585733	579.3139688	557.3320242
20	: 4	29	46	10	572.3429232	577.2983188	555.3163741
21	: 0	30	56	10	594.4211736	599.3765691	577.3946245
21	: 1	30	54	10	592.4055235	597.360919	575.3789744
21	: 2	30	52	10	590.3898734	595.345269	573.3633243
21	: 3	30	50	10	588.3742234	593.3296189	571.3476743
21	: 4	30	48	10	586.3585733	591.3139688	569.3320242
22	: 0	31	58	10	608.4368236	613.3922192	591.4102745
22	: 1	31	56	10	606.4211736	611.3765691	589.3946245
22	: 2	31	54	10	604.4055235	609.360919	587.3789744
22	: 3	31	52	10	602.3898734	607.345269	585.3633243
22	: 4	31	50	10	600.3742234	605.3296189	583.3476743
23	: 0	32	60	10	622.4524737	627.4078692	605.4259246
23	: 1	32	58	10	620.4368236	625.3922192	603.4102745
23	: 2	32	56	10	618.4211736	623.3765691	601.3946245
23	: 3	32	54	10	616.4055235	621.360919	599.3789744
23	: 4	32	52	10	614.3898734	619.345269	597.3633243
24	: 0	33	62	10	636.4681238	641.4235193	619.4415747
24	: 1	33	60	10	634.4524737	639.4078692	617.4259246
24	: 2	33	58	10	632.4368236	637.3922192	615.4102745
24	: 3	33	56	10	630.4211736	635.3765691	613.3946245
24	: 4	33	54	10	628.4055235	633.360919	611.3789744
25	: 0	34	64	10	650.4837738	655.4391693	633.4572247
25	: 1	34	62	10	648.4681238	653.4235193	631.4415747
25	: 2	34	60	10	646.4524737	651.4078692	629.4259246
25	: 3	34	58	10	644.4368236	649.3922192	627.4102745
25	: 4	34	56	10	642.4211736	647.3765691	625.3946245
26	: 0	35	66	10	664.4994239	669.4548194	647.4728748
26	: 1	35	64	10	662.4837738	667.4391693	645.4572247
26	: 2	35	62	10	660.4681238	665.4235193	643.4415747
26	: 3	35	60	10	658.4524737	663.4078692	641.4259246
26	: 4	35	58	10	656.4368236	661.3922192	639.4102745
27	: 0	36	68	10	678.5150739	683.4704695	661.4885248
27	: 1	36	66	10	676.4994239	681.4548194	659.4728748
27	: 2	36	64	10	674.4837738	679.4391693	657.4572247
27	: 3	36	62	10	672.4681238	677.4235193	655.4415747
27	: 4	36	60	10	670.4524737	675.4078692	653.4259246
28	: 0	37	70	10	692.530724	697.4861195	675.5041749
28	: 1	37	68	10	690.5150739	695.4704695	673.4885248
28	: 2	37	66	10	688.4994239	693.4548194	671.4728748
28	: 3	37	64	10	686.4837738	691.4391693	669.4572247
28	: 4	37	62	10	684.4681238	689.4235193	667.4415747

Table S10: Mass list dataset 1 and 2 - MGDG positive ion mode (part 2/6)

Name	dbs	C	H	O	mass(form[+NH4] name[NH4])	mass(form[+Na] name[Na])	mass(form[+H] name[H])
29	: 0	38	72	10	706.5463741	711.5017696	689.519825
29	: 1	38	70	10	704.530724	709.4861195	687.5041749
29	: 2	38	68	10	702.5150739	707.4704695	685.4885248
29	: 3	38	66	10	700.4994239	705.4548194	683.4728748
29	: 4	38	64	10	698.4837738	703.4391693	681.4572247
30	: 0	39	74	10	720.5620241	725.5174197	703.535475
30	: 1	39	72	10	718.5463741	723.5017696	701.519825
30	: 2	39	70	10	716.530724	721.4861195	699.5041749
30	: 3	39	68	10	714.5150739	719.4704695	697.4885248
30	: 4	39	66	10	712.4994239	717.4548194	695.4728748
30	: 5	39	64	10	710.4837738	715.4391693	693.4572247
30	: 6	39	62	10	708.4681238	713.4235193	691.4415747
30	: 7	39	60	10	706.4524737	711.4078692	689.4259246
30	: 8	39	58	10	704.4368236	709.3922192	687.4102745
31	: 0	40	76	10	734.5776742	739.5330697	717.5511251
31	: 1	40	74	10	732.5620241	737.5174197	715.535475
31	: 2	40	72	10	730.5463741	735.5017696	713.519825
31	: 3	40	70	10	728.530724	733.4861195	711.5041749
31	: 4	40	68	10	726.5150739	731.4704695	709.4885248
31	: 5	40	66	10	724.4994239	729.4548194	707.4728748
31	: 6	40	64	10	722.4837738	727.4391693	705.4572247
31	: 7	40	62	10	720.4681238	725.4235193	703.4415747
31	: 8	40	60	10	718.4524737	723.4078692	701.4259246
32	: 0	41	78	10	748.5933243	753.5487198	731.5667752
32	: 1	41	76	10	746.5776742	751.5330697	729.5511251
32	: 2	41	74	10	744.5620241	749.5174197	727.535475
32	: 3	41	72	10	742.5463741	747.5017696	725.519825
32	: 4	41	70	10	740.530724	745.4861195	723.5041749
32	: 5	41	68	10	738.5150739	743.4704695	721.4885248
32	: 6	41	66	10	736.4994239	741.4548194	719.4728748
32	: 7	41	64	10	734.4837738	739.4391693	717.4572247
32	: 8	41	62	10	732.4681238	737.4235193	715.4415747
33	: 0	42	80	10	762.6089743	767.5643699	745.5824252
33	: 1	42	78	10	760.5933243	765.5487198	743.5667752
33	: 2	42	76	10	758.5776742	763.5330697	741.5511251
33	: 3	42	74	10	756.5620241	761.5174197	739.535475
33	: 4	42	72	10	754.5463741	759.5017696	737.519825
33	: 5	42	70	10	752.530724	757.4861195	735.5041749
33	: 6	42	68	10	750.5150739	755.4704695	733.4885248
33	: 7	42	66	10	748.4994239	753.4548194	731.4728748
33	: 8	42	64	10	746.4837738	751.4391693	729.4572247
34	: 0	43	82	10	776.6246244	781.5800199	759.5980753
34	: 1	43	80	10	774.6089743	779.5643699	757.5824252
34	: 2	43	78	10	772.5933243	777.5487198	755.5667752
34	: 3	43	76	10	770.5776742	775.5330697	753.5511251
34	: 4	43	74	10	768.5620241	773.5174197	751.535475
34	: 5	43	72	10	766.5463741	771.5017696	749.519825
34	: 6	43	70	10	764.530724	769.4861195	747.5041749
34	: 7	43	68	10	762.5150739	767.4704695	745.4885248
34	: 8	43	66	10	760.4994239	765.4548194	743.4728748

Table S11: Mass list dataset 1 and 2 - MGDG positive ion mode (part 3/6)

Name	dfs	C	H	O	mass(form[+NH4] name[NH4])	mass(form[+Na] name[Na])	mass(form[+H] name[H])
35	: 0	44	84	10	790.6402745	795.59567	773.6137254
35	: 1	44	82	10	788.6246244	793.5800199	771.5980753
35	: 2	44	80	10	786.6089743	791.5643699	769.5824252
35	: 3	44	78	10	784.5933243	789.5487198	767.5667752
35	: 4	44	76	10	782.5776742	787.5330697	765.5511251
35	: 5	44	74	10	780.5620241	785.5174197	763.535475
35	: 6	44	72	10	778.5463741	783.5017696	761.519825
35	: 7	44	70	10	776.530724	781.4861195	759.5041749
35	: 8	44	68	10	774.5150739	779.4704695	757.4885248
36	: 0	45	86	10	804.6559245	809.61132	787.6293754
36	: 1	45	84	10	802.6402745	807.59567	785.6137254
36	: 2	45	82	10	800.6246244	805.5800199	783.5980753
36	: 3	45	80	10	798.6089743	803.5643699	781.5824252
36	: 4	45	78	10	796.5933243	801.5487198	779.5667752
36	: 5	45	76	10	794.5776742	799.5330697	777.5511251
36	: 6	45	74	10	792.5620241	797.5174197	775.535475
36	: 7	45	72	10	790.5463741	795.5017696	773.519825
36	: 8	45	70	10	788.530724	793.4861195	771.5041749
37	: 0	46	88	10	818.6715746	823.6269701	801.6450255
37	: 1	46	86	10	816.6559245	821.61132	799.6293754
37	: 2	46	84	10	814.6402745	819.59567	797.6137254
37	: 3	46	82	10	812.6246244	817.5800199	795.5980753
37	: 4	46	80	10	810.6089743	815.5643699	793.5824252
37	: 5	46	78	10	808.5933243	813.5487198	791.5667752
37	: 6	46	76	10	806.5776742	811.5330697	789.5511251
37	: 7	46	74	10	804.5620241	809.5174197	787.535475
37	: 8	46	72	10	802.5463741	807.5017696	785.519825
38	: 0	47	90	10	832.6872246	837.6426202	815.6606755
38	: 1	47	88	10	830.6715746	835.6269701	813.6450255
38	: 2	47	86	10	828.6559245	833.61132	811.6293754
38	: 3	47	84	10	826.6402745	831.59567	809.6137254
38	: 4	47	82	10	824.6246244	829.5800199	807.5980753
38	: 5	47	80	10	822.6089743	827.5643699	805.5824252
38	: 6	47	78	10	820.5933243	825.5487198	803.5667752
38	: 7	47	76	10	818.5776742	823.5330697	801.5511251
38	: 8	47	74	10	816.5620241	821.5174197	799.535475
38	: 9	47	72	10	814.5463741	819.5017696	797.519825
38	: 10	47	70	10	812.530724	817.4861195	795.5041749
39	: 0	48	92	10	846.7028747	851.6582702	829.6763256
39	: 1	48	90	10	844.6872246	849.6426202	827.6606755
39	: 2	48	88	10	842.6715746	847.6269701	825.6450255
39	: 3	48	86	10	840.6559245	845.61132	823.6293754
39	: 4	48	84	10	838.6402745	843.59567	821.6137254
39	: 5	48	82	10	836.6246244	841.5800199	819.5980753
39	: 6	48	80	10	834.6089743	839.5643699	817.5824252
39	: 7	48	78	10	832.5933243	837.5487198	815.5667752
39	: 8	48	76	10	830.5776742	835.5330697	813.5511251
39	: 9	48	74	10	828.5620241	833.5174197	811.535475
39	: 10	48	72	10	826.5463741	831.5017696	809.519825

Table S12: Mass list dataset 1 and 2 - MGDG positive ion mode (part 4/6)

Name	dbs	C	H	O	mass(form[+NH4] name[NH4])	mass(form[+Na] name[Na])	mass(form[+H] name[H])
40	: 0	49	94	10	860.7185248	865.6739203	843.6919757
40	: 1	49	92	10	858.7028747	863.6582702	841.6763256
40	: 2	49	90	10	856.6872246	861.6426202	839.6606755
40	: 3	49	88	10	854.6715746	859.6269701	837.6450255
40	: 4	49	86	10	852.6559245	857.61132	835.6293754
40	: 5	49	84	10	850.6402745	855.59567	833.6137254
40	: 6	49	82	10	848.6246244	853.5800199	831.5980753
40	: 7	49	80	10	846.6089743	851.5643699	829.5824252
40	: 8	49	78	10	844.5933243	849.5487198	827.5667752
40	: 9	49	76	10	842.5776742	847.5330697	825.5511251
40	: 10	49	74	10	840.5620241	845.5174197	823.535475
40	: 11	49	72	10	838.5463741	843.5017696	821.519825
40	: 12	49	70	10	836.530724	841.4861195	819.5041749
41	: 0	50	96	10	874.7341748	879.6895704	857.7076257
41	: 1	50	94	10	872.7185248	877.6739203	855.6919757
41	: 2	50	92	10	870.7028747	875.6582702	853.6763256
41	: 3	50	90	10	868.6872246	873.6426202	851.6606755
41	: 4	50	88	10	866.6715746	871.6269701	849.6450255
41	: 5	50	86	10	864.6559245	869.61132	847.6293754
41	: 6	50	84	10	862.6402745	867.59567	845.6137254
41	: 7	50	82	10	860.6246244	865.5800199	843.5980753
41	: 8	50	80	10	858.6089743	863.5643699	841.5824252
41	: 9	50	78	10	856.5933243	861.5487198	839.5667752
41	: 10	50	76	10	854.5776742	859.5330697	837.5511251
41	: 11	50	74	10	852.5620241	857.5174197	835.535475
41	: 12	50	72	10	850.5463741	855.5017696	833.519825
42	: 0	51	98	10	888.7498249	893.7052204	871.7232758
42	: 1	51	96	10	886.7341748	891.6895704	869.7076257
42	: 2	51	94	10	884.7185248	889.6739203	867.6919757
42	: 3	51	92	10	882.7028747	887.6582702	865.6763256
42	: 4	51	90	10	880.6872246	885.6426202	863.6606755
42	: 5	51	88	10	878.6715746	883.6269701	861.6450255
42	: 6	51	86	10	876.6559245	881.61132	859.6293754
42	: 7	51	84	10	874.6402745	879.59567	857.6137254
42	: 8	51	82	10	872.6246244	877.5800199	855.5980753
42	: 9	51	80	10	870.6089743	875.5643699	853.5824252
42	: 10	51	78	10	868.5933243	873.5487198	851.5667752
42	: 11	51	76	10	866.5776742	871.5330697	849.5511251
42	: 12	51	74	10	864.5620241	869.5174197	847.535475
43	: 0	52	100	10	902.765475	907.7208705	885.7389259
43	: 1	52	98	10	900.7498249	905.7052204	883.7232758
43	: 2	52	96	10	898.7341748	903.6895704	881.7076257
43	: 3	52	94	10	896.7185248	901.6739203	879.6919757
43	: 4	52	92	10	894.7028747	899.6582702	877.6763256
43	: 5	52	90	10	892.6872246	897.6426202	875.6606755
43	: 6	52	88	10	890.6715746	895.6269701	873.6450255
43	: 7	52	86	10	888.6559245	893.61132	871.6293754
43	: 8	52	84	10	886.6402745	891.59567	869.6137254
43	: 9	52	82	10	884.6246244	889.5800199	867.5980753
43	: 10	52	80	10	882.6089743	887.5643699	865.5824252
43	: 11	52	78	10	880.5933243	885.5487198	863.5667752
43	: 12	52	76	10	878.5776742	883.5330697	861.5511251

Table S13: Mass list dataset 1 and 2 - MGDG positive ion mode (part 5/6)

Name	dbs	C	H	O	mass(form[+NH4] name[NH4])	mass(form[+Na] name[Na])	mass(form[+H] name[H])
44	: 0	53	102	10	916.781125	921.7365206	899.7545759
44	: 1	53	100	10	914.765475	919.7208705	897.7389259
44	: 2	53	98	10	912.7498249	917.7052204	895.7232758
44	: 3	53	96	10	910.7341748	915.6895704	893.7076257
44	: 4	53	94	10	908.7185248	913.6739203	891.6919757
44	: 5	53	92	10	906.7028747	911.6582702	889.6763256
44	: 6	53	90	10	904.6872246	909.6426202	887.6606755
44	: 7	53	88	10	902.6715746	907.6269701	885.6450255
44	: 8	53	86	10	900.6559245	905.61132	883.6293754
44	: 9	53	84	10	898.6402745	903.59567	881.6137254
44	: 10	53	82	10	896.6246244	901.5800199	879.5980753
44	: 11	53	80	10	894.6089743	899.5643699	877.5824252
44	: 12	53	78	10	892.5933243	897.5487198	875.5667752
45	: 0	54	104	10	930.7967751	935.7521706	913.770226
45	: 1	54	102	10	928.781125	933.7365206	911.7545759
45	: 2	54	100	10	926.765475	931.7208705	909.7389259
45	: 3	54	98	10	924.7498249	929.7052204	907.7232758
45	: 4	54	96	10	922.7341748	927.6895704	905.7076257
45	: 5	54	94	10	920.7185248	925.6739203	903.6919757
45	: 6	54	92	10	918.7028747	923.6582702	901.6763256
45	: 7	54	90	10	916.6872246	921.6426202	899.6606755
45	: 8	54	88	10	914.6715746	919.6269701	897.6450255
45	: 9	54	86	10	912.6559245	917.61132	895.6293754
45	: 10	54	84	10	910.6402745	915.59567	893.6137254
45	: 11	54	82	10	908.6246244	913.5800199	891.5980753
45	: 12	54	80	10	906.6089743	911.5643699	889.5824252
46	: 0	55	106	10	944.8124252	949.7678207	927.7858761
46	: 1	55	104	10	942.7967751	947.7521706	925.770226
46	: 2	55	102	10	940.781125	945.7365206	923.7545759
46	: 3	55	100	10	938.765475	943.7208705	921.7389259
46	: 4	55	98	10	936.7498249	941.7052204	919.7232758
46	: 5	55	96	10	934.7341748	939.6895704	917.7076257
46	: 6	55	94	10	932.7185248	937.6739203	915.6919757
46	: 7	55	92	10	930.7028747	935.6582702	913.6763256
46	: 8	55	90	10	928.6872246	933.6426202	911.6606755
46	: 9	55	88	10	926.6715746	931.6269701	909.6450255
46	: 10	55	86	10	924.6559245	929.61132	907.6293754
46	: 11	55	84	10	922.6402745	927.59567	905.6137254
46	: 12	55	82	10	920.6246244	925.5800199	903.5980753
47	: 0	56	108	10	958.8280752	963.7834707	941.8015261
47	: 1	56	106	10	956.8124252	961.7678207	939.7858761
47	: 2	56	104	10	954.7967751	959.7521706	937.770226
47	: 3	56	102	10	952.781125	957.7365206	935.7545759
47	: 4	56	100	10	950.765475	955.7208705	933.7389259
47	: 5	56	98	10	948.7498249	953.7052204	931.7232758
47	: 6	56	96	10	946.7341748	951.6895704	929.7076257
47	: 7	56	94	10	944.7185248	949.6739203	927.6919757
47	: 8	56	92	10	942.7028747	947.6582702	925.6763256
47	: 9	56	90	10	940.6872246	945.6426202	923.6606755
47	: 10	56	88	10	938.6715746	943.6269701	921.6450255
47	: 11	56	86	10	936.6559245	941.61132	919.6293754
47	: 12	56	84	10	934.6402745	939.59567	917.6137254

Table S14: Mass list dataset 1 and 2 - MGDG positive ion mode (part 6/6)

Name	dbs	C	H	O	mass(form[+NH4] name[NH4])	mass(form[+Na] name[Na])	mass(form[+H] name[H])	
48	:	0	57	110	10	972.8437253	977.7991208	955.8171762
48	:	1	57	108	10	970.8280752	975.7834707	953.8015261
48	:	2	57	106	10	968.8124252	973.7678207	951.7858761
48	:	3	57	104	10	966.7967751	971.7521706	949.770226
48	:	4	57	102	10	964.781125	969.7365206	947.7545759
48	:	5	57	100	10	962.765475	967.7208705	945.7389259
48	:	6	57	98	10	960.7498249	965.7052204	943.7232758
48	:	7	57	96	10	958.7341748	963.6895704	941.7076257
48	:	8	57	94	10	956.7185248	961.6739203	939.6919757
48	:	9	57	92	10	954.7028747	959.6582702	937.6763256
48	:	10	57	90	10	952.6872246	957.6426202	935.6606755
48	:	11	57	88	10	950.6715746	955.6269701	933.6450255
48	:	12	57	86	10	948.6559245	953.61132	931.6293754

Table S15: Mass list dataset 1 and 2 - MGDG negative ion mode (part 1/6)

Name	dbs	C	H	O	mass(form[+C2H3O2] name[C2H3O2])	mass(form[+HCOO] name[HCOO])	mass(form[-H] name[-H])	
20	:	0	29	54	10	621.3855509	607.3699008	561.3644215
20	:	1	29	52	10	619.3699008	605.3542507	559.3487714
20	:	2	29	50	10	617.3542507	603.3386007	557.3331214
20	:	3	29	48	10	615.3386007	601.3229506	555.3174713
20	:	4	29	46	10	613.3229506	599.3073005	553.3018212
21	:	0	30	56	10	635.4012009	621.3855509	575.3800716
21	:	1	30	54	10	633.3855509	619.3699008	573.3644215
21	:	2	30	52	10	631.3699008	617.3542507	571.3487714
21	:	3	30	50	10	629.3542507	615.3386007	569.3331214
21	:	4	30	48	10	627.3386007	613.3229506	567.3174713
22	:	0	31	58	10	649.416851	635.4012009	589.3957216
22	:	1	31	56	10	647.4012009	633.3855509	587.3800716
22	:	2	31	54	10	645.3855509	631.3699008	585.3644215
22	:	3	31	52	10	643.3699008	629.3542507	583.3487714
22	:	4	31	50	10	641.3542507	627.3386007	581.3331214
23	:	0	32	60	10	663.4325011	649.416851	603.4113717
23	:	1	32	58	10	661.416851	647.4012009	601.3957216
23	:	2	32	56	10	659.4012009	645.3855509	599.3800716
23	:	3	32	54	10	657.3855509	643.3699008	597.3644215
23	:	4	32	52	10	655.3699008	641.3542507	595.3487714
24	:	0	33	62	10	677.4481511	663.4325011	617.4270217
24	:	1	33	60	10	675.4325011	661.416851	615.4113717
24	:	2	33	58	10	673.416851	659.4012009	613.3957216
24	:	3	33	56	10	671.4012009	657.3855509	611.3800716
24	:	4	33	54	10	669.3855509	655.3699008	609.3644215
25	:	0	34	64	10	691.4638012	677.4481511	631.4426718
25	:	1	34	62	10	689.4481511	675.4325011	629.4270217
25	:	2	34	60	10	687.4325011	673.416851	627.4113717
25	:	3	34	58	10	685.416851	671.4012009	625.3957216
25	:	4	34	56	10	683.4012009	669.3855509	623.3800716

Table S16: Mass list dataset 1 and 2 - MGDG negative ion mode (part 2/6)

Name	dbs	C	H	O	mass(form[+C2H3O2] name[C2H3O2])	mass(form[+HCOO] name[HCOO])	mass(form[-H] name[-H])
26	: 0	35	66	10	705.4794512	691.4638012	645.4583219
26	: 1	35	64	10	703.4638012	689.4481511	643.4426718
26	: 2	35	62	10	701.4481511	687.4325011	641.4270217
26	: 3	35	60	10	699.4325011	685.416851	639.4113717
26	: 4	35	58	10	697.416851	683.4012009	637.3957216
27	: 0	36	68	10	719.4951013	705.4794512	659.4739719
27	: 1	36	66	10	717.4794512	703.4638012	657.4583219
27	: 2	36	64	10	715.4638012	701.4481511	655.4426718
27	: 3	36	62	10	713.4481511	699.4325011	653.4270217
27	: 4	36	60	10	711.4325011	697.416851	651.4113717
28	: 0	37	70	10	733.5107514	719.4951013	673.489622
28	: 1	37	68	10	731.4951013	717.4794512	671.4739719
28	: 2	37	66	10	729.4794512	715.4638012	669.4583219
28	: 3	37	64	10	727.4638012	713.4481511	667.4426718
28	: 4	37	62	10	725.4481511	711.4325011	665.4270217
29	: 0	38	72	10	747.5264014	733.5107514	687.5052721
29	: 1	38	70	10	745.5107514	731.4951013	685.489622
29	: 2	38	68	10	743.4951013	729.4794512	683.4739719
29	: 3	38	66	10	741.4794512	727.4638012	681.4583219
29	: 4	38	64	10	739.4638012	725.4481511	679.4426718
30	: 0	39	74	10	761.5420515	747.5264014	701.5209221
30	: 1	39	72	10	759.5264014	745.5107514	699.5052721
30	: 2	39	70	10	757.5107514	743.4951013	697.489622
30	: 3	39	68	10	755.4951013	741.4794512	695.4739719
30	: 4	39	66	10	753.4794512	739.4638012	693.4583219
30	: 5	39	64	10	751.4638012	737.4481511	691.4426718
30	: 6	39	62	10	749.4481511	735.4325011	689.4270217
30	: 7	39	60	10	747.4325011	733.416851	687.4113717
30	: 8	39	58	10	745.416851	731.4012009	685.3957216
31	: 0	40	76	10	775.5577016	761.5420515	715.5365722
31	: 1	40	74	10	773.5420515	759.5264014	713.5209221
31	: 2	40	72	10	771.5264014	757.5107514	711.5052721
31	: 3	40	70	10	769.5107514	755.4951013	709.489622
31	: 4	40	68	10	767.4951013	753.4794512	707.4739719
31	: 5	40	66	10	765.4794512	751.4638012	705.4583219
31	: 6	40	64	10	763.4638012	749.4481511	703.4426718
31	: 7	40	62	10	761.4481511	747.4325011	701.4270217
31	: 8	40	60	10	759.4325011	745.416851	699.4113717

Table S17: Mass list dataset 1 and 2 - MGDG negative ion mode (part 3/6)

Name	dbs	C	H	O	mass(form[+C2H3O2] name[C2H3O2])	mass(form[+HCOO] name[HCOO])	mass(form[-H] name[-H])	
32	:	0	41	78	10	789.5733516	775.5577016	729.5522223
32	:	1	41	76	10	787.5577016	773.5420515	727.5365722
32	:	2	41	74	10	785.5420515	771.5264014	725.5209221
32	:	3	41	72	10	783.5264014	769.5107514	723.5052721
32	:	4	41	70	10	781.5107514	767.4951013	721.489622
32	:	5	41	68	10	779.4951013	765.4794512	719.4739719
32	:	6	41	66	10	777.4794512	763.4638012	717.4583219
32	:	7	41	64	10	775.4638012	761.4481511	715.4426718
32	:	8	41	62	10	773.4481511	759.4325011	713.4270217
33	:	0	42	80	10	803.5890017	789.5733516	743.5678723
33	:	1	42	78	10	801.5733516	787.5577016	741.5522223
33	:	2	42	76	10	799.5577016	785.5420515	739.5365722
33	:	3	42	74	10	797.5420515	783.5264014	737.5209221
33	:	4	42	72	10	795.5264014	781.5107514	735.5052721
33	:	5	42	70	10	793.5107514	779.4951013	733.489622
33	:	6	42	68	10	791.4951013	777.4794512	731.4739719
33	:	7	42	66	10	789.4794512	775.4638012	729.4583219
33	:	8	42	64	10	787.4638012	773.4481511	727.4426718
34	:	0	43	82	10	817.6046518	803.5890017	757.5835224
34	:	1	43	80	10	815.5890017	801.5733516	755.5678723
34	:	2	43	78	10	813.5733516	799.5577016	753.5522223
34	:	3	43	76	10	811.5577016	797.5420515	751.5365722
34	:	4	43	74	10	809.5420515	795.5264014	749.5209221
34	:	5	43	72	10	807.5264014	793.5107514	747.5052721
34	:	6	43	70	10	805.5107514	791.4951013	745.489622
34	:	7	43	68	10	803.4951013	789.4794512	743.4739719
34	:	8	43	66	10	801.4794512	787.4638012	741.4583219
35	:	0	44	84	10	831.6203018	817.6046518	771.5991725
35	:	1	44	82	10	829.6046518	815.5890017	769.5835224
35	:	2	44	80	10	827.5890017	813.5733516	767.5678723
35	:	3	44	78	10	825.5733516	811.5577016	765.5522223
35	:	4	44	76	10	823.5577016	809.5420515	763.5365722
35	:	5	44	74	10	821.5420515	807.5264014	761.5209221
35	:	6	44	72	10	819.5264014	805.5107514	759.5052721
35	:	7	44	70	10	817.5107514	803.4951013	757.489622
35	:	8	44	68	10	815.4951013	801.4794512	755.4739719
36	:	0	45	86	10	845.6359519	831.6203018	785.6148225
36	:	1	45	84	10	843.6203018	829.6046518	783.5991725
36	:	2	45	82	10	841.6046518	827.5890017	781.5835224
36	:	3	45	80	10	839.5890017	825.5733516	779.5678723
36	:	4	45	78	10	837.5733516	823.5577016	777.5522223
36	:	5	45	76	10	835.5577016	821.5420515	775.5365722
36	:	6	45	74	10	833.5420515	819.5264014	773.5209221
36	:	7	45	72	10	831.5264014	817.5107514	771.5052721
36	:	8	45	70	10	829.5107514	815.4951013	769.489622

Table S18: Mass list dataset 1 and 2 - MGDG negative ion mode (part 4/6)

Name	dbs	C	H	O	mass(form[+C2H3O2] name[C2H3O2])	mass(form[+HCOO] name[HCOO])	mass(form[-H] name[-H])	
37	:	0	46	88	10	859.651602	845.6359519	799.6304726
37	:	1	46	86	10	857.6359519	843.6203018	797.6148225
37	:	2	46	84	10	855.6203018	841.6046518	795.5991725
37	:	3	46	82	10	853.6046518	839.5890017	793.5835224
37	:	4	46	80	10	851.5890017	837.5733516	791.5678723
37	:	5	46	78	10	849.5733516	835.5577016	789.5522223
37	:	6	46	76	10	847.5577016	833.5420515	787.5365722
37	:	7	46	74	10	845.5420515	831.5264014	785.5209221
37	:	8	46	72	10	843.5264014	829.5107514	783.5052721
38	:	0	47	90	10	873.667252	859.651602	813.6461226
38	:	1	47	88	10	871.651602	857.6359519	811.6304726
38	:	2	47	86	10	869.6359519	855.6203018	809.6148225
38	:	3	47	84	10	867.6203018	853.6046518	807.5991725
38	:	4	47	82	10	865.6046518	851.5890017	805.5835224
38	:	5	47	80	10	863.5890017	849.5733516	803.5678723
38	:	6	47	78	10	861.5733516	847.5577016	801.5522223
38	:	7	47	76	10	859.5577016	845.5420515	799.5365722
38	:	8	47	74	10	857.5420515	843.5264014	797.5209221
38	:	9	47	72	10	855.5264014	841.5107514	795.5052721
38	:	10	47	70	10	853.5107514	839.4951013	793.489622
39	:	0	48	92	10	887.6829021	873.667252	827.6617727
39	:	1	48	90	10	885.667252	871.651602	825.6461226
39	:	2	48	88	10	883.651602	869.6359519	823.6304726
39	:	3	48	86	10	881.6359519	867.6203018	821.6148225
39	:	4	48	84	10	879.6203018	865.6046518	819.5991725
39	:	5	48	82	10	877.6046518	863.5890017	817.5835224
39	:	6	48	80	10	875.5890017	861.5733516	815.5678723
39	:	7	48	78	10	873.5733516	859.5577016	813.5522223
39	:	8	48	76	10	871.5577016	857.5420515	811.5365722
39	:	9	48	74	10	869.5420515	855.5264014	809.5209221
39	:	10	48	72	10	867.5264014	853.5107514	807.5052721
40	:	0	49	94	10	901.6985521	887.6829021	841.6774228
40	:	1	49	92	10	899.6829021	885.667252	839.6617727
40	:	2	49	90	10	897.667252	883.651602	837.6461226
40	:	3	49	88	10	895.651602	881.6359519	835.6304726
40	:	4	49	86	10	893.6359519	879.6203018	833.6148225
40	:	5	49	84	10	891.6203018	877.6046518	831.5991725
40	:	6	49	82	10	889.6046518	875.5890017	829.5835224
40	:	7	49	80	10	887.5890017	873.5733516	827.5678723
40	:	8	49	78	10	885.5733516	871.5577016	825.5522223
40	:	9	49	76	10	883.5577016	869.5420515	823.5365722
40	:	10	49	74	10	881.5420515	867.5264014	821.5209221
40	:	11	49	72	10	879.5264014	865.5107514	819.5052721
40	:	12	49	70	10	877.5107514	863.4951013	817.489622

Table S19: Mass list dataset 1 and 2 - MGDG negative ion mode (part 5/6)

Name	dbs	C	H	O	mass(form[+C2H3O2] name[C2H3O2])	mass(form[+HCOO] name[HCOO])	mass(form[-H] name[-H])
41	: 0	50	96	10	915.7142022	901.6985521	855.6930728
41	: 1	50	94	10	913.6985521	899.6829021	853.6774228
41	: 2	50	92	10	911.6829021	897.667252	851.6617727
41	: 3	50	90	10	909.667252	895.651602	849.6461226
41	: 4	50	88	10	907.651602	893.6359519	847.6304726
41	: 5	50	86	10	905.6359519	891.6203018	845.6148225
41	: 6	50	84	10	903.6203018	889.6046518	843.5991725
41	: 7	50	82	10	901.6046518	887.5890017	841.5835224
41	: 8	50	80	10	899.5890017	885.5733516	839.5678723
41	: 9	50	78	10	897.5733516	883.5577016	837.5522223
41	: 10	50	76	10	895.5577016	881.5420515	835.5365722
41	: 11	50	74	10	893.5420515	879.5264014	833.5209221
41	: 12	50	72	10	891.5264014	877.5107514	831.5052721
42	: 0	51	98	10	929.7298523	915.7142022	869.7087229
42	: 1	51	96	10	927.7142022	913.6985521	867.6930728
42	: 2	51	94	10	925.6985521	911.6829021	865.6774228
42	: 3	51	92	10	923.6829021	909.667252	863.6617727
42	: 4	51	90	10	921.667252	907.651602	861.6461226
42	: 5	51	88	10	919.651602	905.6359519	859.6304726
42	: 6	51	86	10	917.6359519	903.6203018	857.6148225
42	: 7	51	84	10	915.6203018	901.6046518	855.5991725
42	: 8	51	82	10	913.6046518	899.5890017	853.5835224
42	: 9	51	80	10	911.5890017	897.5733516	851.5678723
42	: 10	51	78	10	909.5733516	895.5577016	849.5522223
42	: 11	51	76	10	907.5577016	893.5420515	847.5365722
42	: 12	51	74	10	905.5420515	891.5264014	845.5209221
43	: 0	52	100	10	943.7455023	929.7298523	883.724373
43	: 1	52	98	10	941.7298523	927.7142022	881.7087229
43	: 2	52	96	10	939.7142022	925.6985521	879.6930728
43	: 3	52	94	10	937.6985521	923.6829021	877.6774228
43	: 4	52	92	10	935.6829021	921.667252	875.6617727
43	: 5	52	90	10	933.667252	919.651602	873.6461226
43	: 6	52	88	10	931.651602	917.6359519	871.6304726
43	: 7	52	86	10	929.6359519	915.6203018	869.6148225
43	: 8	52	84	10	927.6203018	913.6046518	867.5991725
43	: 9	52	82	10	925.6046518	911.5890017	865.5835224
43	: 10	52	80	10	923.5890017	909.5733516	863.5678723
43	: 11	52	78	10	921.5733516	907.5577016	861.5522223
43	: 12	52	76	10	919.5577016	905.5420515	859.5365722
44	: 0	53	102	10	957.7611524	943.7455023	897.740023
44	: 1	53	100	10	955.7455023	941.7298523	895.724373
44	: 2	53	98	10	953.7298523	939.7142022	893.7087229
44	: 3	53	96	10	951.7142022	937.6985521	891.6930728
44	: 4	53	94	10	949.6985521	935.6829021	889.6774228
44	: 5	53	92	10	947.6829021	933.667252	887.6617727
44	: 6	53	90	10	945.667252	931.651602	885.6461226
44	: 7	53	88	10	943.651602	929.6359519	883.6304726
44	: 8	53	86	10	941.6359519	927.6203018	881.6148225
44	: 9	53	84	10	939.6203018	925.6046518	879.5991725
44	: 10	53	82	10	937.6046518	923.5890017	877.5835224
44	: 11	53	80	10	935.5890017	921.5733516	875.5678723
44	: 12	53	78	10	933.5733516	919.5577016	873.5522223

Table S20: Mass list dataset 1 and 2 - MGDG negative ion mode (part 6/6)

Name	dbs	C	H	O	mass(form[+C2H3O2] name[C2H3O2])	mass(form[+HCOO] name[HCOO])	mass(form[-H] name[-H])
45	: 0	54	104	10	971.7768025	957.7611524	911.7556731
45	: 1	54	102	10	969.7611524	955.7455023	909.740023
45	: 2	54	100	10	967.7455023	953.7298523	907.724373
45	: 3	54	98	10	965.7298523	951.7142022	905.7087229
45	: 4	54	96	10	963.7142022	949.6985521	903.6930728
45	: 5	54	94	10	961.6985521	947.6829021	901.6774228
45	: 6	54	92	10	959.6829021	945.667252	899.6617727
45	: 7	54	90	10	957.667252	943.651602	897.6461226
45	: 8	54	88	10	955.651602	941.6359519	895.6304726
45	: 9	54	86	10	953.6359519	939.6203018	893.6148225
45	: 10	54	84	10	951.6203018	937.6046518	891.5991725
45	: 11	54	82	10	949.6046518	935.5890017	889.5835224
45	: 12	54	80	10	947.5890017	933.5733516	887.5678723
46	: 0	55	106	10	985.7924525	971.7768025	925.7713232
46	: 1	55	104	10	983.7768025	969.7611524	923.7556731
46	: 2	55	102	10	981.7611524	967.7455023	921.740023
46	: 3	55	100	10	979.7455023	965.7298523	919.724373
46	: 4	55	98	10	977.7298523	963.7142022	917.7087229
46	: 5	55	96	10	975.7142022	961.6985521	915.6930728
46	: 6	55	94	10	973.6985521	959.6829021	913.6774228
46	: 7	55	92	10	971.6829021	957.667252	911.6617727
46	: 8	55	90	10	969.667252	955.651602	909.6461226
46	: 9	55	88	10	967.651602	953.6359519	907.6304726
46	: 10	55	86	10	965.6359519	951.6203018	905.6148225
46	: 11	55	84	10	963.6203018	949.6046518	903.5991725
46	: 12	55	82	10	961.6046518	947.5890017	901.5835224
47	: 0	56	108	10	999.8081026	985.7924525	939.7869732
47	: 1	56	106	10	997.7924525	983.7768025	937.7713232
47	: 2	56	104	10	995.7768025	981.7611524	935.7556731
47	: 3	56	102	10	993.7611524	979.7455023	933.740023
47	: 4	56	100	10	991.7455023	977.7298523	931.724373
47	: 5	56	98	10	989.7298523	975.7142022	929.7087229
47	: 6	56	96	10	987.7142022	973.6985521	927.6930728
47	: 7	56	94	10	985.6985521	971.6829021	925.6774228
47	: 8	56	92	10	983.6829021	969.667252	923.6617727
47	: 9	56	90	10	981.667252	967.651602	921.6461226
47	: 10	56	88	10	979.651602	965.6359519	919.6304726
47	: 11	56	86	10	977.6359519	963.6203018	917.6148225
47	: 12	56	84	10	975.6203018	961.6046518	915.5991725
48	: 0	57	110	10	1013.823753	999.8081026	953.8026233
48	: 1	57	108	10	1011.808103	997.7924525	951.7869732
48	: 2	57	106	10	1009.792453	995.7768025	949.7713232
48	: 3	57	104	10	1007.776802	993.7611524	947.7556731
48	: 4	57	102	10	1005.761152	991.7455023	945.740023
48	: 5	57	100	10	1003.745502	989.7298523	943.724373
48	: 6	57	98	10	1001.729852	987.7142022	941.7087229
48	: 7	57	96	10	999.7142022	985.6985521	939.6930728
48	: 8	57	94	10	997.6985521	983.6829021	937.6774228
48	: 9	57	92	10	995.6829021	981.667252	935.6617727
48	: 10	57	90	10	993.667252	979.651602	933.6461226
48	: 11	57	88	10	991.651602	977.6359519	931.6304726
48	: 12	57	86	10	989.6359519	975.6203018	929.6148225

### Dataset 3

Table S21: Mass list dataset 3 - PC negative ion mode

Name	dbs	C	H	O	P	N	mass(form[+HCOO] name[HCOO])	mass(form[-CH3] name[-CH3])	oxidation-state	
34	:	0	42	84	8	1	1	806.5916675	746.5705335	;OH;2OH;3OH;O;2O
34	:	1	42	82	8	1	1	804.5760145	744.5548843	;OH;2OH;3OH;O;2O
34	:	2	42	80	8	1	1	802.5603621	742.5392396	;OH;2OH;3OH;O;2O

### Dataset 4

Table S22: Mass list dataset 4 - TG positive ion mode (part 1/2)

Name	dbs	C	H	O	mass(form[+NH4] name[NH4])	oxidation-state	
44	:	3	47	84	6	762.6617131	OH
44	:	4	47	82	6	760.6460631	OH
46	:	2	49	90	6	792.7086633	
46	:	3	49	88	6	790.6930133	;OH
46	:	4	49	86	6	788.6773632	;OH;2OH
46	:	5	49	84	6	786.6617131	
48	:	0	51	98	6	824.7712636	;OH
48	:	1	51	96	6	822.7556135	;OH
48	:	2	51	94	6	820.7399634	;OH
48	:	3	51	92	6	818.7243134	;OH
48	:	4	51	90	6	816.7086633	
48	:	5	51	88	6	814.6930133	OH
50	:	0	53	102	6	852.8025637	
50	:	1	53	100	6	850.7869136	;OH;2OH;3OH
50	:	2	53	98	6	848.7712636	;OH;2OH
50	:	3	53	96	6	846.7556135	;OH;2OH
50	:	4	53	94	6	844.7399634	;OH;2OH
50	:	5	53	92	6	842.7243134	;OH
50	:	6	53	90	6	840.7086633	;OH
50	:	7	53	88	6	838.6930133	;OH
52	:	0	55	106	6	880.8338638	
52	:	1	55	104	6	878.8182138	
52	:	2	55	102	6	876.8025637	;OH;2OH;3OH;4OH
52	:	3	55	100	6	874.7869136	;OH;2OH;3OH
52	:	4	55	98	6	872.7712636	;OH;2OH;3OH;4OH
52	:	5	55	96	6	870.7556135	;OH;2OH;3OH;4OH
52	:	6	55	94	6	868.7399634	;OH;2OH;3OH;4OH
52	:	7	55	92	6	866.7243134	;OH;2OH

Table S23: Mass list dataset 4 - TG positive ion mode (part 2/2)

Name	dbs	C	H	O	mass(form[+NH4])	name[NH4])	oxidation-state
54	:	0	57	110	6	908.8651640	
54	:	1	57	108	6	906.8495139	;OH
54	:	2	57	106	6	904.8338638	;OH;2OH
54	:	3	57	104	6	902.8182138	;OH;2OH
54	:	4	57	102	6	900.8025637	;OH;2OH
54	:	5	57	100	6	898.7869136	;OH;2OH;3OH;4OH
54	:	6	57	98	6	896.7712636	;OH;2OH;3OH;4OH
54	:	7	57	96	6	894.7556135	;OH;2OH;3OH;4OH
54	:	8	57	94	6	892.7399634	;OH;2OH;3OH;4OH
54	:	9	57	92	6	890.7243134	;OH;2OH;3OH
56	:	0	59	114	6	936.8964641	;OH
56	:	1	59	112	6	934.8808140	
56	:	2	59	110	6	932.8651640	;OH
56	:	3	59	108	6	930.8495139	;OH
56	:	4	59	106	6	928.8338638	;OH
56	:	5	59	104	6	926.8182138	;OH;2OH
56	:	6	59	102	6	924.8025637	;OH
56	:	7	59	100	6	922.7869136	;OH
56	:	8	59	98	6	920.7712636	;OH
58	:	0	61	118	6	964.9277642	;OH
58	:	1	61	116	6	962.9121141	
58	:	2	61	114	6	960.8964641	;OH
58	:	3	61	112	6	958.8808140	;OH
58	:	4	61	110	6	956.8651640	;OH;4OH
58	:	5	61	108	6	954.8495139	;OH
58	:	6	61	106	6	952.8338638	;OH;2OH
58	:	7	61	104	6	950.8182138	OH
60	:	0	63	122	6	992.9590643	OH
60	:	1	63	120	6	990.9434143	
60	:	2	63	118	6	988.9277642	;OH
60	:	3	63	116	6	986.9121141	;OH
60	:	4	63	114	6	984.8964641	;OH
60	:	5	63	112	6	982.8808140	;OH
60	:	6	63	110	6	980.8651640	;OH
60	:	7	63	108	6	978.8495139	;OH
60	:	9	63	104	6	974.8182138	OH
62	:	1	65	124	6	1018.9747144	
62	:	2	65	122	6	1016.9590643	
62	:	3	65	120	6	1014.9434143	;OH;2OH
62	:	4	65	118	6	1012.9277642	;OH
62	:	5	65	116	6	1010.9121141	;OH
62	:	6	65	114	6	1008.8964641	
64	:	2	67	126	6	1044.9903645	
64	:	3	67	124	6	1042.9747144	
64	:	4	67	122	6	1040.9590643	;OH
64	:	5	67	120	6	1038.9434143	
64	:	6	67	118	6	1036.9277642	
66	:	3	69	128	6	1071.0060145	
66	:	4	69	126	6	1068.9903645	;2OH;3OH
66	:	5	69	124	6	1066.9747144	

Table S24: Mass list dataset 4 - DG positive ion mode

Name	dbs	C	H	O	mass(form[+NH4])	name[NH4])	oxidation-state
30	:	3	33	58	5	552.4633477	4OH
30	:	5	33	54	5	548.4320475	3OH
32	:	0	35	68	5	586.5415980	
32	:	1	35	66	5	584.5259479	;OH
32	:	2	35	64	5	582.5102979	;OH
34	:	0	37	72	5	614.5728981	;2OH
34	:	1	37	70	5	612.5572481	;OH;2OH;3OH
34	:	2	37	68	5	610.5415980	;OH;2OH;3OH
34	:	3	37	66	5	608.5259479	;OH;2OH;3OH
34	:	4	37	64	5	606.5102979	;OH;2OH
34	:	5	37	62	5	604.4946478	3OH
36	:	0	39	76	5	642.6041983	
36	:	1	39	74	5	640.5885482	;2OH
36	:	2	39	72	5	638.5728981	;OH;2OH;3OH;4OH
36	:	3	39	70	5	636.5572481	;OH;2OH;3OH;4OH
36	:	4	39	68	5	634.5415980	;OH;2OH;3OH;4OH
36	:	5	39	66	5	632.5259479	;OH;2OH;3OH
36	:	6	39	64	5	630.5102979	;OH;2OH
36	:	8	39	60	5	626.4789977	;3OH
38	:	0	41	80	5	670.6354984	
38	:	1	41	78	5	668.6198483	;2OH
38	:	2	41	76	5	666.6041983	;2OH
38	:	3	41	74	5	664.5885482	;OH
38	:	4	41	72	5	662.5728981	;OH
38	:	5	41	70	5	660.5572481	OH
40	:	0	43	84	5	698.6667985	;OH
40	:	1	43	82	5	696.6511484	2OH
40	:	2	43	80	5	694.6354984	;OH
40	:	3	43	78	5	692.6198483	
40	:	4	43	76	5	690.6041983	;4OH
42	:	0	45	88	5	726.6980986	
42	:	1	45	86	5	724.6824486	
42	:	2	45	84	5	722.6667985	;4OH
42	:	3	45	82	5	720.6511484	;2OH
42	:	4	45	80	5	718.6354984	2OH
42	:	5	45	78	5	716.6198483	3OH
44	:	0	47	92	5	754.7293988	
44	:	1	47	90	5	752.7137487	
44	:	2	47	88	5	750.6980986	
44	:	3	47	86	5	748.6824486	
44	:	4	47	84	5	746.6667985	2OH
44	:	5	47	82	5	744.6511484	;OH;2OH
44	:	6	47	80	5	742.6354984	

Table S25: Mass list dataset 4 - PC positive ion mode

Name	dbs	C	H	O	N	P	mass(form[+H])	name[H])	oxidation-state
34	:	4	42	76	8	1	1	754.5392285	
36	:	3	44	82	8	1	1	784.5861787	;3OH
36	:	4	44	80	8	1	1	782.5705286	;3OH;4OH
36	:	6	44	76	8	1	1	778.5392285	OH

Table S26: Mass list dataset 4 - DGDG positive ion mode

Name	dbs	C	H	O	mass(form[+NH4]	name[NH4])	oxidation-state
32	:	0	47	88	15	910.6472449	
32	:	1	47	86	15	908.6315948	
32	:	2	47	84	15	906.6159447	
34	:	0	49	92	15	938.6785450	
34	:	1	49	90	15	936.6628949	;2OH
34	:	2	49	88	15	934.6472449	;2OH
34	:	3	49	86	15	932.6315948	;OH
34	:	4	49	84	15	930.6159447	
36	:	2	51	92	15	962.6785450	;OH;2OH
36	:	3	51	90	15	960.6628949	;2OH
36	:	4	51	88	15	958.6472449	;OH;2OH
36	:	5	51	86	15	956.6315948	;OH
36	:	6	51	84	15	954.6159447	
38	:	2	53	96	15	990.7098451	
38	:	3	53	94	15	988.6941950	
38	:	4	53	92	15	986.6785450	
38	:	5	53	90	15	984.6628949	
42	:	2	57	104	15	1046.7724454	
42	:	3	57	102	15	1044.7567953	
44	:	2	59	108	15	1074.8037455	

Table S27: Mass list dataset 4 - PE positive ion mode

Name	dbs	C	H	O	N	P	mass(form[+H]	name[H])	oxidation-state
34	:	1	39	76	8	1	718.5392285		
34	:	2	39	74	8	1	716.5235784		
34	:	3	39	72	8	1	714.5079284		
36	:	2	41	78	8	1	744.5548786		
36	:	3	41	76	8	1	742.5392285		
36	:	4	41	74	8	1	740.5235784	;OH	
36	:	5	41	72	8	1	738.5079284	;OH	
36	:	6	41	70	8	1	736.4922783		

Table S28: Mass list dataset 4 - PI positive ion mode

Name	dbs	C	H	O	P	mass(form[+NH4]	name[NH4])	oxidation-state
34	:	2	43	79	13	1	852.5607518	
34	:	3	43	77	13	1	850.5451018	
36	:	2	45	83	13	1	880.5920519	
36	:	3	45	81	13	1	878.5764019	
36	:	4	45	79	13	1	876.5607518	

Table S29: Mass list dataset 4 - LPC positive ion mode

Name	dbs	C	H	O	N	P	mass(form[+H]	name[H])	oxidation-state
14	:	0	22	46	7	1	468.3095629		
16	:	0	24	50	7	1	496.3408631	;OH	
16	:	1	24	48	7	1	494.3252130		
18	:	0	26	54	7	1	524.3721632		
18	:	1	26	52	7	1	522.3565131	;OH	
18	:	2	26	50	7	1	520.3408631	;OH;2OH	
18	:	3	26	48	7	1	518.3252130	;OH	
20	:	0	28	58	7	1	552.4034633		
20	:	1	28	56	7	1	550.3878133		
20	:	2	28	54	7	1	548.3721632		
22	:	0	30	62	7	1	580.4347634		
22	:	1	30	60	7	1	578.4191134		

Table S30: Mass list dataset 4 - LPE positive ion mode

Name	dbs	C	H	O	N	P	mass(form[+H] name[H])	oxidation-state
16	:	0	21	44	7	1	1	454.2939129
18	:	0	23	48	7	1	1	482.3252130
18	:	1	23	46	7	1	1	480.3095629
18	:	2	23	44	7	1	1	478.2939129
18	:	3	23	42	7	1	1	476.2782628
20	:	0	25	52	7	1	1	510.3565131
20	:	2	25	48	7	1	1	506.3252130

Table S31: Mass list dataset 4 - PG positive ion mode

Name	dbs	C	H	O	P	mass(form[+NH4] name[NH4])	oxidation-state
34	:	1	40	77	10	1	766.5603579
34	:	2	40	75	10	1	764.5447078
34	:	3	40	73	10	1	762.5290578
36	:	2	42	79	10	1	792.5760079
36	:	3	42	77	10	1	790.5603579
36	:	4	42	75	10	1	788.5447078
36	:	5	42	73	10	1	786.5290578

Table S32: Mass list dataset 4 - LPG positive ion mode

Name	dbs	C	H	O	P	mass(form[+H] name[H])	oxidation-state
16	:	0	22	45	9	1	485.2884931
18	:	2	24	45	9	1	509.2884931

Table S33: Mass list dataset 4 - MGDG positive ion mode

Name	dbs	C	H	O	mass(form[+H] name[H])	oxidation-state	
34	:	2	43	78	10	772.5944214	
36	:	2	45	82	10	800.6257216	
36	:	3	45	80	10	798.6100715	;OH;2OH
36	:	4	45	78	10	796.5944214	;OH;2OH
36	:	5	45	76	10	794.5787714	;OH;2OH
36	:	6	45	74	10	792.5631213	;OH

Table S34: Mass list dataset 4 - MG positive ion mode

Name	dbs	C	H	O	mass(form[+H] name[H])	oxidation-state	
16	:	0	19	38	4	331.2853833	
18	:	0	21	42	4	359.3166834	
18	:	1	21	40	4	357.3010334	
18	:	2	21	38	4	355.2853833	
18	:	3	21	36	4	353.2697332	;OH
18	:	4	21	34	4	351.2540832	

### 5.3 Novel LDA Features

64

The screenshot shows the Lipid Data Analyzer 2.7.0 interface. The title bar indicates the software version and license. The main window has several tabs: Quantitation, Batch Quantitation, Statistical Analysis, Display Results (selected), Settings, Help, and About. Below the tabs, there are two text input fields for file paths and two buttons: 'Open Chrom' and 'Open Result'. A 'Start Display' button is also present.

The 'Results' section displays a table with the following data:

Name	Area
34:0[OH]_11.82_H...	22786.693
34:0[2OH]_10.91_...	702826.1
34:0[2OH]_10.89_...	99534.484
34:0[2OH]_13.38_...	143015.53
34:0[2OH]_14.49_...	8922.554
34:0[3OH]_13.46_...	966.26373
34:0[3OH]_13.50_...	42596.117
34:1[OH]_11.19_H...	5665776.5
34:1[OH]_12.16_H...	3997338.8
34:1[OH]_13.80_H...	163092.84
34:1[OH]_11.53_...	471331.66
34:1[OH]_12.16_...	316392.78
34:1[2OH]_10.83_...	78610.41
34:1[3OH]_10.36_...	58871.93
34:1[3OH]_10.23_...	13609.236
34:2[OH]_11.66_H...	4091021.0
34:2[OH]_11.66_...	349845.06
34:2[OH]_12.51_...	11614.232
34:2[3OH]_10.69_...	124906.82
34:2[3OH]_10.09_...	16484.133
34:2[3OH]_10.82_...	12192.495
34:2[3OH]_13.21_...	3341.4148

Below the table, there are control options for the results display:

- m/z: 1.5 + m/z: 2.5 [Da] Update
- Start: [ ] Stop: [ ] [min]
- Lock m/z range [ ] - [ ]
- Show MSn  Show 2D-View
- Show MSn Only
- Chain Evidence Only

Figure S1: Screenshot of LDA; new view modes

Lipid Data Analyzer 2.7.0\_nightly QTOF settings noIntensity © 2019 - Jürgen Hartler, Andreas Ziegl, Gerhard G Thallinger - GNU GPL v3 license

Quantitation Batch Quantitation Statistical Analysis Display Results Settings Help About

prData\OxPC\_Standard\LD\OrbiTrap\_velos\_pro\_HCD(no Intensity ms2PrecursorTolerance=0.1)\180816\_oxPC\_10ng.chrom  
OxPC\_Standard\LD\OrbiTrap\_velos\_pro\_HCD(no Intensity ms2PrecursorTolerance=0.1)\180816\_oxPC\_10ng\_negative.xlsx

Open Chrom Start Display  
Open Result

Results  
oxPC

Name	Area
34:0[2OH]_10.91_...	702826.1
34:0[2OH]_10.89_...	99534.484
34:1[OH]_11.19_H...	5665776.5
34:1[OH]_12.16_H...	3997338.8
34:1[OH]_13.80_H...	163092.84
34:1[OH]_11.53_-...	471331.66
34:1[OH]_12.16_-...	316392.78
34:1[2OH]_10.83_...	78610.41
34:1[3OH]_10.36_...	58871.93
34:2[OH]_11.66_H...	4091021.0
34:2[OH]_11.66_-...	349845.06
34:2[3OH]_10.69_...	124906.82
34:2[3OH]_13.21_...	3341.4148

- m/z: 1.5 + m/z: 2.5 [Da] Update  
Start: [ ] Stop: [ ] [min]  
 Look m/z range [ ] - [ ]  
 Show MSn  Show 2D-View  
 Show MSn Only  
 Chain Evidence Only

Figure S2: Screenshot of LDA; new view modes: example

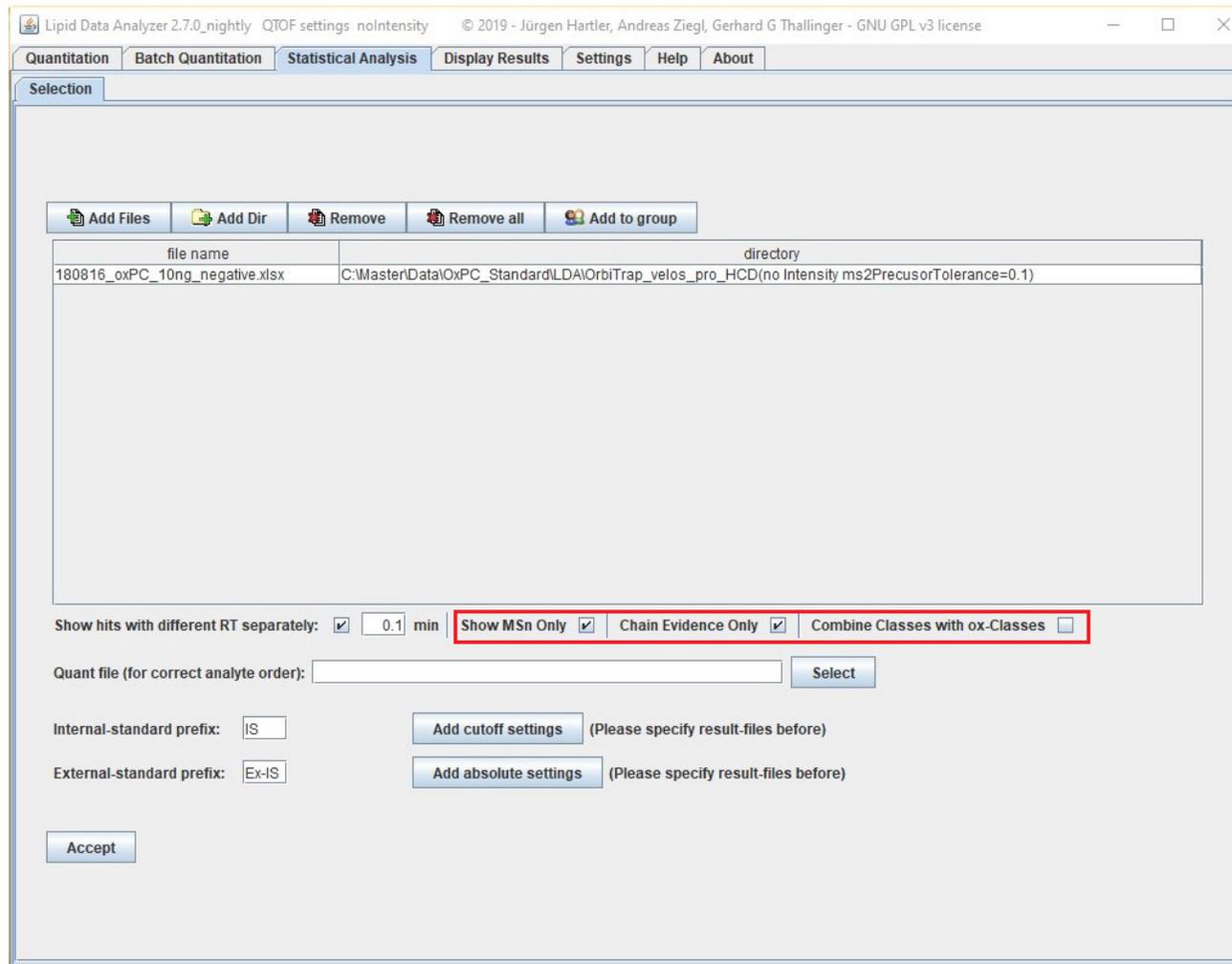


Figure S3: Screenshot of LDA; new options for statistics module

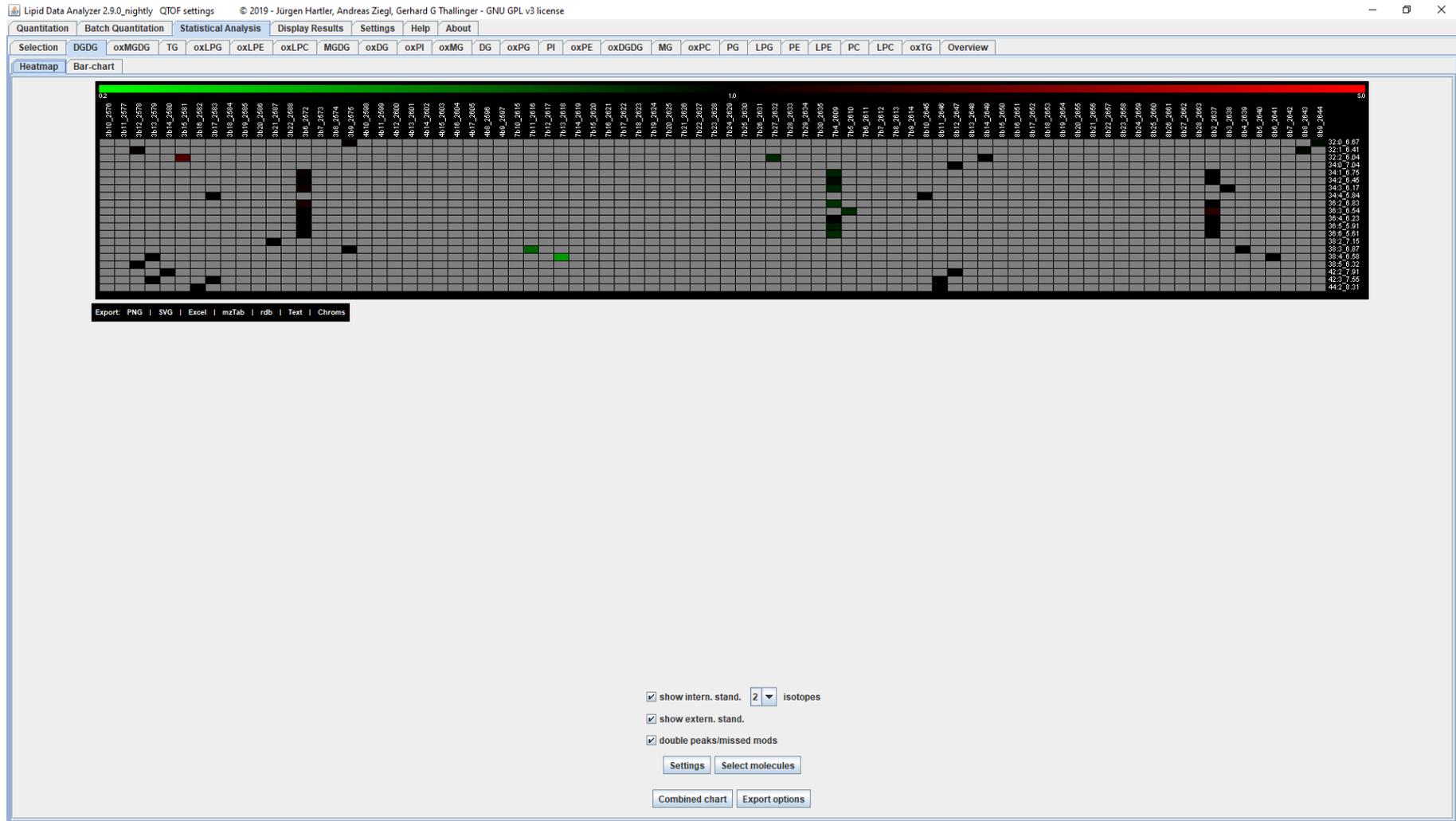


Figure S4: Screenshot of LDA; statistics module: classes and ox-classes separate

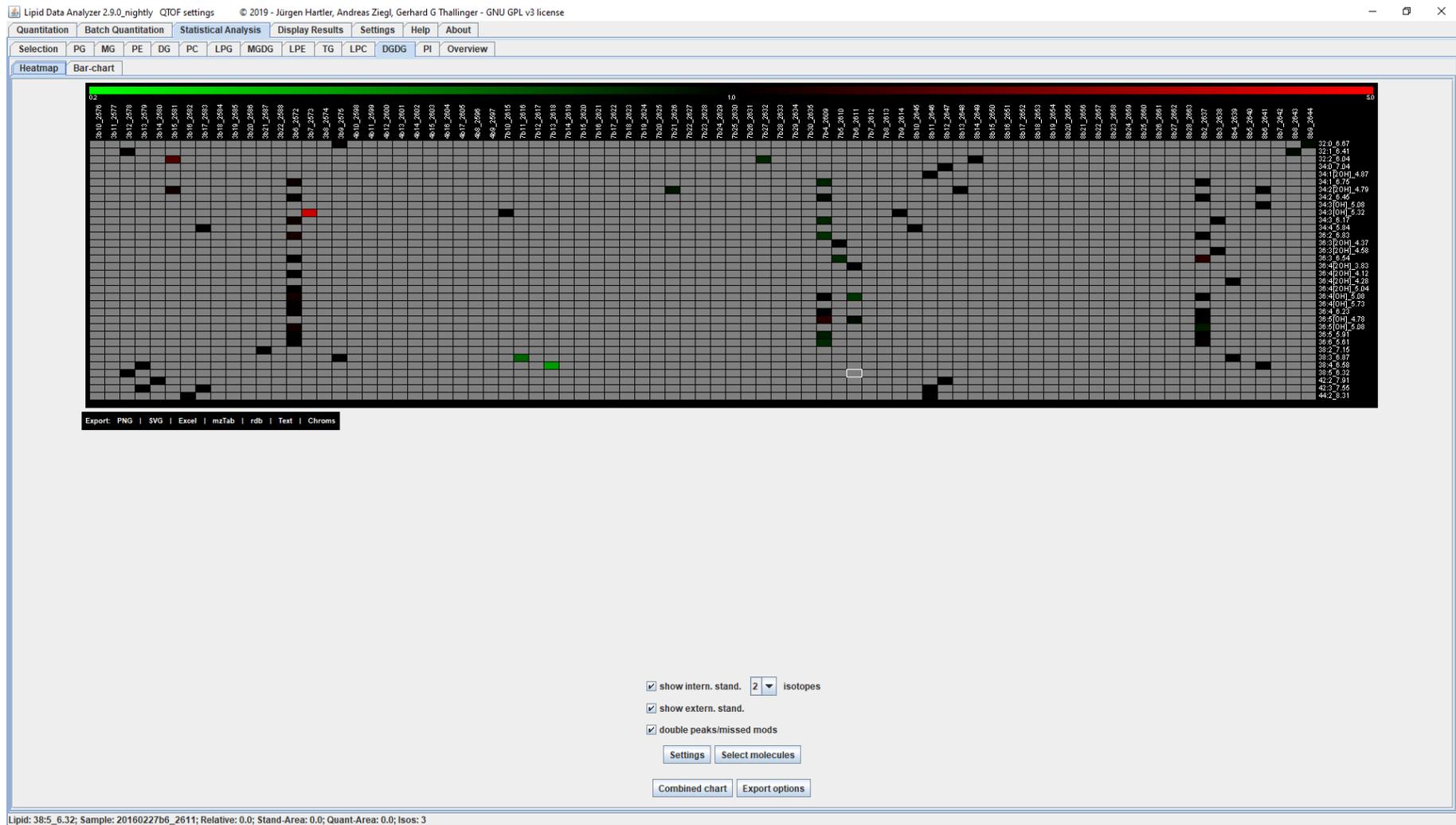


Figure S5: Screenshot of LDA; statistics module: classes and ox-classes combined

## 5.4 Lipid Data Analyzer Results

### 5.4.1 Dataset 1

Table S35: MS<sup>2</sup> confirmed lipids in dataset 1 (positive ion mode) by Lipid Data Analyzer

	Lipid Species		Lipid Species		Lipid Species		Lipid Species
1	DG(34:1)	29	PC(35:3)	57	PE(36:4)	85	TG(53:3)
2	DG(34:2)	30	PC(36:1)	58	PE(36:5)	86	TG(53:4)
3	DG(36:2)	31	PC(36:2)	59	PE(38:3)	87	TG(54:0)
4	DG(36:3)	32	PC(36:3)	60	PE(38:4)	88	TG(54:1)
5	DG(36:4)	33	PC(36:4)	61	PE(38:5)	89	TG(54:2)
6	DG(38:5)	34	PC(36:5)	62	PE(38:6)	90	TG(54:3)
7	DG(38:6)	35	PC(37:2)	63	PE(38:7)	91	TG(54:4)
8	LPC(16:0)	36	PC(37:4)	64	PE(40:5)	92	TG(54:5)
9	LPC(18:0)	37	PC(38:2)	65	PE(40:6)	93	TG(54:6)
10	LPC(18:1)	38	PC(38:3)	66	PE(40:7)	94	TG(54:7)
11	LPC(18:2)	39	PC(38:4)	67	TG(48:1)	95	TG(55:3)
12	LPC(20:4)	40	PC(38:5)	68	TG(48:2)	96	TG(56:3)
13	LPC(22:6)	41	PC(38:6)	69	TG(48:3)	97	TG(56:4)
14	LPE(18:0)	42	PC(38:7)	70	TG(50:0)	98	TG(56:5)
15	P-PE(36:4)	43	PC(39:4)	71	TG(50:1)	99	TG(56:6)
16	P-PE(38:4)	44	PC(39:6)	72	TG(50:2)	100	TG(56:7)
17	PC(32:0)	45	PC(40:4)	73	TG(50:3)	101	TG(56:8)
18	PC(32:1)	46	PC(40:5)	75	TG(50:5)	102	TG(56:9)
19	PC(32:2)	47	PC(40:6)	76	TG(51:2)	103	TG(58:10)
20	PC(33:1)	48	PC(40:7)	77	TG(51:3)	104	TG(58:5)
21	PC(33:2)	49	PC(40:8)	78	TG(52:1)	105	TG(58:6)
22	PC(34:0)	50	PC(42:10)	79	TG(52:2)	106	TG(58:7)
23	PC(34:1)	51	PE(24:0)	80	TG(52:3)	107	TG(58:8)
24	PC(34:2)	52	PE(34:1)	81	TG(52:4)	108	TG(58:9)
25	PC(34:3)	53	PE(34:2)	82	TG(52:5)	109	TG(62:14)
26	PC(34:4)	54	PE(36:1)	83	TG(52:6)		
27	PC(35:1)	55	PE(36:2)	84	TG(53:2)		
28	PC(35:2)	56	PE(36:3)	74	TG(50:4)		

Table S36: MS<sup>2</sup> confirmed lipids in dataset 1 (negative ion mode) by Lipid Data Analyzer

	Lipid Species		Lipid Species		Lipid Species		Lipid Species
1	Cer(16:0)	29	P-PE(42:6)	57	PC(38:5)	85	PE(39:4)
2	Cer(22:0)	30	PC(32:0)	58	PC(38:6)	86	PE(39:6)
3	Cer(23:0)	31	PC(32:1)	59	PC(38:7)	87	PE(40:4)
4	Cer(24:0)	32	PC(32:2)	60	PC(39:4)	88	PE(40:5)
5	Cer(24:1)	33	PC(33:2)	61	PC(39:6)	89	PE(40:6)
6	LPC(16:0)	34	PC(34:0)	62	PC(40:4)	90	PE(40:7)
7	LPC(16:1)	35	PC(34:1)	63	PC(40:5)	91	PE(40:8)
8	LPC(17:0)	36	PC(34:2)	64	PC(40:6)	92	PG(34:2)
9	LPC(18:0)	37	PC(34:3)	65	PC(40:7)	93	PG(36:3)
10	LPC(18:1)	38	PC(34:4)	66	PC(40:8)	94	PG(36:4)
11	LPC(18:2)	39	PC(35:2)	67	PC(42:10)	95	PG(38:6)
12	LPC(20:3)	40	PC(35:3)	68	PE(24:0)	96	PG(40:7)
13	LPC(20:4)	41	PC(35:4)	69	PE(34:1)	97	PG(40:8)
14	LPC(22:5)	42	PC(36:0)	70	PE(34:2)	98	PG(42:10)
15	LPC(22:6)	43	PC(36:1)	71	PE(34:3)	99	PG(44:11)
16	LPE(16:0)	44	PC(36:2)	72	PE(36:1)	100	PG(44:12)
17	LPE(18:0)	45	PC(36:3)	73	PE(36:2)	101	PI(34:2)
18	LPE(18:1)	46	PC(36:4)	74	PE(36:3)	102	PI(36:3)
19	LPE(18:2)	47	PC(36:5)	75	PE(36:4)	103	PI(36:4)
20	LPE(20:4)	48	PC(36:6)	76	PE(36:5)	104	PI(37:4)
21	LPE(22:6)	49	PC(37:1)	77	PE(37:4)	105	PI(38:3)
22	P-PE(36:4)	50	PC(37:2)	78	PE(38:1)	106	PI(38:4)
23	P-PE(38:4)	51	PC(37:4)	79	PE(38:2)	107	PI(38:5)
24	P-PE(38:5)	52	PC(37:6)	80	PE(38:3)	108	PI(38:6)
25	P-PE(38:6)	53	PC(38:1)	81	PE(38:4)	109	PI(40:6)
26	P-PE(40:4)	54	PC(38:2)	82	PE(38:5)	110	PS(38:4)
27	P-PE(40:5)	55	PC(38:3)	83	PE(38:6)	111	PS(38:6)
28	P-PE(40:6)	56	PC(38:4)	84	PE(38:7)	112	PS(40:6)

## 5.4.2 Dataset 2

Table S37: MS<sup>2</sup> confirmed lipids in dataset 2 (positive ion mode) by Lipid Data Analyzer

	Lipid Species		Lipid Species		Lipid Species		Lipid Species
1	DG(32:0)	15	LPC(16:0)	29	PE(36:3)	43	TG(52:4)
2	DG(34:0)	16	LPC(18:2)	30	PE(36:5)	44	TG(52:5)
3	DG(34:2)	17	LPC(18:3)	31	PG(34:3)	45	TG(52:6)
4	DG(34:3)	18	MGDG(34:3)	32	PG(34:4)	46	TG(54:2)
5	DG(34:4)	19	MGDG(36:5)	33	PG(36:7)	47	TG(54:3)
6	DG(36:0)	20	MGDG(36:6)	34	PI(34:2)	48	TG(54:4)
7	DG(36:4)	21	MGMG(18:3)	35	PI(34:3)	49	TG(54:5)
8	DG(36:5)	22	PC(34:2)	36	TG(34:0)	50	TG(54:6)
9	DG(36:6)	23	PC(34:3)	37	TG(50:1)	51	TG(54:7)
10	DG(44:6)	24	PC(36:4)	38	TG(50:2)	52	TG(54:8)
11	DGDG(34:2)	25	PC(36:5)	39	TG(50:3)	53	TG(54:9)
12	DGDG(34:3)	26	PC(36:6)	40	TG(52:1)	54	TriGDG(36:6)
13	DGDG(36:6)	27	PE(34:2)	41	TG(52:2)		
14	DGMG(18:3)	28	PE(34:3)	42	TG(52:3)		

Table S38: MS<sup>2</sup> confirmed lipids in dataset 2 (negative ion mode) by Lipid Data Analyzer

	Lipid Species		Lipid Species		Lipid Species		Lipid Species
1	DGDG(34:1)	16	LPE(18:2)	31	PC(36:4)	46	PI(34:3)
2	DGDG(34:2)	17	LPE(18:3)	32	PC(36:5)	47	PI(36:4)
3	DGDG(34:3)	18	MGDG(34:3)	33	PC(36:6)	48	PI(36:5)
4	DGDG(34:4)	19	MGDG(34:4)	34	PE(31:1)	49	PI(36:6)
5	DGDG(35:3)	20	MGDG(35:6)	35	PE(32:2)	50	PS(31:1)
6	DGDG(35:6)	21	MGDG(36:6)	36	PE(34:2)	51	PS(33:3)
7	DGDG(36:3)	22	MGMG(18:3)	37	PE(34:3)	52	PS(36:4)
8	DGDG(36:4)	23	P-PE(36:6)	38	PE(36:3)	53	PS(36:5)
9	DGDG(36:6)	24	P-PE(38:7)	39	PE(36:4)	54	PS(40:3)
10	DGMG(16:0)	25	P-PE(38:8)	40	PE(36:5)	55	SQDG(32:0)
11	DGMG(18:3)	26	PC(33:4)	41	PE(36:6)	56	SQDG(34:3)
12	LPC(16:0)	27	PC(34:1)	42	PG(36:4)	57	SQDG(36:3)
13	LPC(18:2)	28	PC(34:2)	43	PG(36:5)	58	SQDG(36:6)
14	LPC(18:3)	29	PC(34:3)	44	PG(36:6)	59	SQMG(18:3)
15	LPE(16:0)	30	PC(36:3)	45	PI(34:2)		

### 5.4.3 Dataset 4

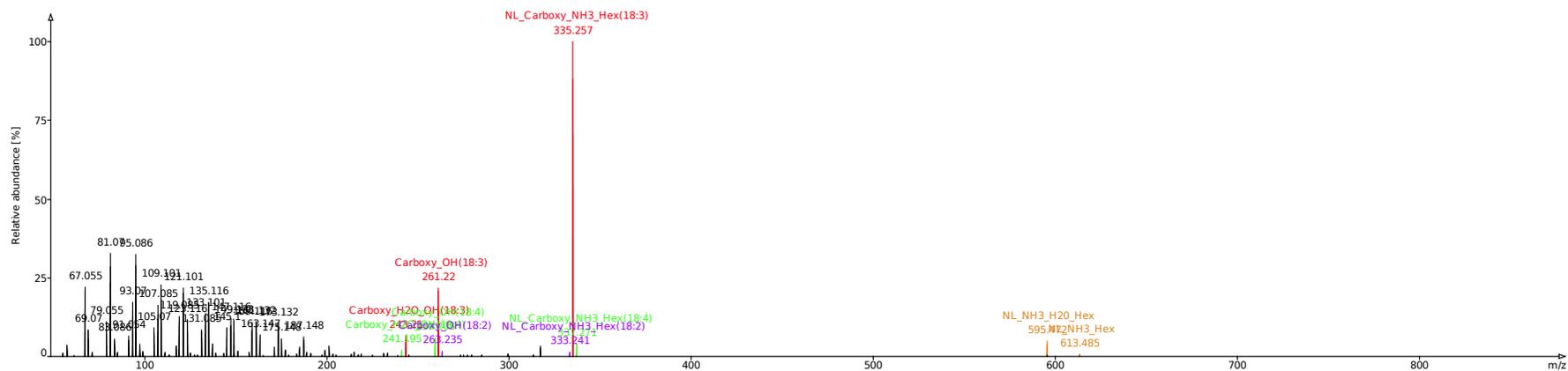
Table S39: MS<sup>2</sup> confirmed lipids in dataset 4 (positive ion mode) by Lipid Data Analyzer (part 1/2)

	Lipid Species	Lipid Species	Lipid Species
1	DG(32:0)	43	DG(42:1)
2	DG(32:1)	44	DG(42:2)
3	DG(32:2)	45	DG(42:3)
4	DG(34:0)	46	DG(44:0)
5	DG(34:1)	47	DG(44:2)
6	oxDG(34:1[2OH])	48	oxDG(44:4[2OH])
7	oxDG(34:1[OH])	49	DGDG(32:0)
8	DG(34:2)	50	DGDG(32:1)
9	oxDG(34:2[2OH])	51	DGDG(32:2)
10	oxDG(34:2[OH])	52	DGDG(34:0)
11	DG(34:3)	53	DGDG(34:1)
12	oxDG(34:3[2OH])	54	oxDGDG(34:1[2OH])
13	oxDG(34:3[3OH])	55	DGDG(34:2)
14	oxDG(34:3[OH])	56	oxDGDG(34:2[2OH])
15	DG(34:4)	57	oxDGDG(34:2[OH])
16	oxDG(34:4[2OH])	58	DGDG(34:3)
17	oxDG(34:4[OH])	59	oxDGDG(34:3[OH])
18	DG(36:0)	60	DGDG(34:4)
19	DG(36:1)	61	DGDG(36:2)
20	oxDG(36:1[2OH])	62	DGDG(36:3)
21	DG(36:2)	63	oxDGDG(36:3[2OH])
22	oxDG(36:2[2OH])	64	oxDGDG(36:3[OH])
23	oxDG(36:2[4OH])	65	DGDG(36:4)
24	oxDG(36:2[OH])	66	oxDGDG(36:4[2OH])
25	DG(36:3)	67	oxDGDG(36:4[3OH])
26	oxDG(36:3[2OH])	68	oxDGDG(36:4[OH])
27	oxDG(36:3[OH])	69	DGDG(36:5)
28	DG(36:4)	70	oxDGDG(36:5[OH])
29	oxDG(36:4[2OH])	71	DGDG(36:6)
30	oxDG(36:4[OH])	72	oxDGDG(36:6[OH])
31	DG(36:5)	73	DGDG(38:2)
32	oxDG(36:5[OH])	74	DGDG(38:3)
33	DG(36:6)	75	DGDG(38:4)
34	oxDG(36:6[2OH])	76	DGDG(38:5)
35	oxDG(36:6[OH])	77	DGDG(42:2)
36	DG(38:0)	78	DGDG(42:3)
37	DG(38:2)	79	DGDG(44:2)
38	DG(38:4)	80	MGDG(34:2)
39	DG(40:0)	81	oxMGDG(34:2[OH])
40	DG(40:2)	82	MGDG(36:2)
41	DG(40:3)	83	MGDG(36:3)
42	DG(42:0)	84	oxMGDG(36:3[2OH])
		85	MGDG(36:4)
		86	oxMGDG(36:4[2OH])
		87	oxMGDG(36:4[OH])
		88	MGDG(36:5)
		89	oxMGDG(36:5[OH])
		90	MGDG(36:6)
		91	oxMGDG(36:6[OH])
		92	PC(32:0)
		93	PC(32:1)
		94	PC(34:1)
		95	PC(34:2)
		96	oxPC(34:2[2OH])
		97	PC(34:3)
		98	oxPC(34:3[OH])
		99	PC(36:1)
		100	PC(36:2)
		101	PC(36:3)
		102	oxPC(36:3[OH])
		103	PC(36:4)
		104	oxPC(36:4[2OH])
		105	oxPC(36:4[OH])
		106	PC(36:5)
		107	oxPC(36:5[2OH])
		108	oxPC(36:5[OH])
		109	PC(36:6)
		110	PC(38:3)
		111	PC(38:4)
		112	PE(34:2)
		113	PE(34:3)
		114	PE(36:2)
		115	PE(36:3)
		116	PE(36:4)
		117	PE(36:5)
		118	PE(36:6)
		119	PI(34:1)
		120	PI(34:2)
		121	PI(34:3)
		122	PI(36:2)
		123	PI(36:3)
		124	PI(36:4)
		125	PI(36:5)
		126	oxTG(44:4[OH])

Table S40: MS<sup>2</sup> confirmed lipids in dataset 4 (positive ion mode) by Lipid Data Analyzer (part 2/2)

Lipid Species		Lipid Species		Lipid Species	
127	TG(46:2)	171	oxTG(52:5[2OH])	215	TG(56:4)
128	TG(46:3)	172	oxTG(52:5[3OH])	216	oxTG(56:4[OH])
129	oxTG(46:3[OH])	173	oxTG(52:5[OH])	217	TG(56:5)
130	TG(46:4)	174	TG(52:6)	218	oxTG(56:5[OH])
131	oxTG(46:4[2OH])	175	oxTG(52:6[2OH])	219	TG(56:6)
132	oxTG(46:4[OH])	176	oxTG(52:6[3OH])	220	oxTG(56:6[OH])
133	TG(46:5)	177	oxTG(52:6[4OH])	221	TG(56:7)
134	TG(48:0)	178	oxTG(52:6[OH])	222	TG(56:8)
135	oxTG(48:0[OH])	179	TG(52:7)	223	TG(58:0)
136	TG(48:1)	180	oxTG(52:7[OH])	224	TG(58:1)
137	oxTG(48:1[OH])	181	TG(54:0)	225	TG(58:2)
138	TG(48:2)	182	TG(54:1)	226	oxTG(58:2[OH])
139	oxTG(48:2[OH])	183	oxTG(54:1[OH])	227	TG(58:3)
140	TG(48:3)	184	TG(54:2)	228	oxTG(58:3[OH])
141	TG(48:4)	185	TG(54:3)	229	oxTG(58:4[OH])
142	oxTG(48:5[OH])	186	oxTG(54:3[OH])	230	TG(58:5)
143	TG(50:0)	187	TG(54:4)	231	oxTG(58:5[OH])
144	TG(50:1)	188	oxTG(54:4[2OH])	232	TG(58:6)
145	oxTG(50:1[2OH])	189	oxTG(54:4[OH])	233	oxTG(58:6[OH])
146	oxTG(50:1[OH])	190	TG(54:5)	234	oxTG(58:7[OH])
147	TG(50:2)	191	oxTG(54:5[2OH])	235	TG(60:1)
148	oxTG(50:2[2OH])	192	oxTG(54:5[3OH])	236	TG(60:2)
149	oxTG(50:2[OH])	193	oxTG(54:5[OH])	237	TG(60:3)
150	TG(50:3)	194	TG(54:6)	238	TG(60:4)
151	oxTG(50:3[2OH])	195	oxTG(54:6[2OH])	239	oxTG(60:4[OH])
152	oxTG(50:3[OH])	196	oxTG(54:6[3OH])	240	TG(60:5)
153	TG(50:4)	197	oxTG(54:6[4OH])	241	oxTG(60:5[OH])
154	oxTG(50:4[OH])	198	oxTG(54:6[OH])	242	TG(60:6)
155	TG(50:5)	199	TG(54:7)	243	oxTG(60:6[OH])
156	oxTG(50:5[OH])	200	oxTG(54:7[2OH])	244	TG(62:1)
157	TG(50:6)	201	oxTG(54:7[3OH])	245	TG(62:2)
158	oxTG(50:7[OH])	202	oxTG(54:7[OH])	246	TG(62:3)
159	TG(52:0)	203	TG(54:8)	247	TG(62:4)
160	TG(52:1)	204	oxTG(54:8[2OH])	248	TG(62:5)
161	TG(52:2)	205	oxTG(54:8[3OH])	249	oxTG(62:5[OH])
162	oxTG(52:2[OH])	206	oxTG(54:8[OH])	250	TG(62:6)
163	TG(52:3)	207	oxTG(54:9[2OH])	251	TG(64:2)
164	oxTG(52:3[2OH])	208	oxTG(54:9[3OH])	252	TG(64:3)
165	oxTG(52:3[OH])	209	oxTG(54:9[OH])	253	TG(64:4)
166	TG(52:4)	210	TG(56:0)	254	TG(64:5)
167	oxTG(52:4[2OH])	211	TG(56:1)	255	TG(64:6)
168	oxTG(52:4[3OH])	212	TG(56:2)	256	TG(66:3)
169	oxTG(52:4[OH])	213	TG(56:3)	257	TG(66:4)
170	TG(52:5)	214	oxTG(56:3[OH])	258	TG(66:5)

### 5.4.4 Dataset 2: Selected Spectra



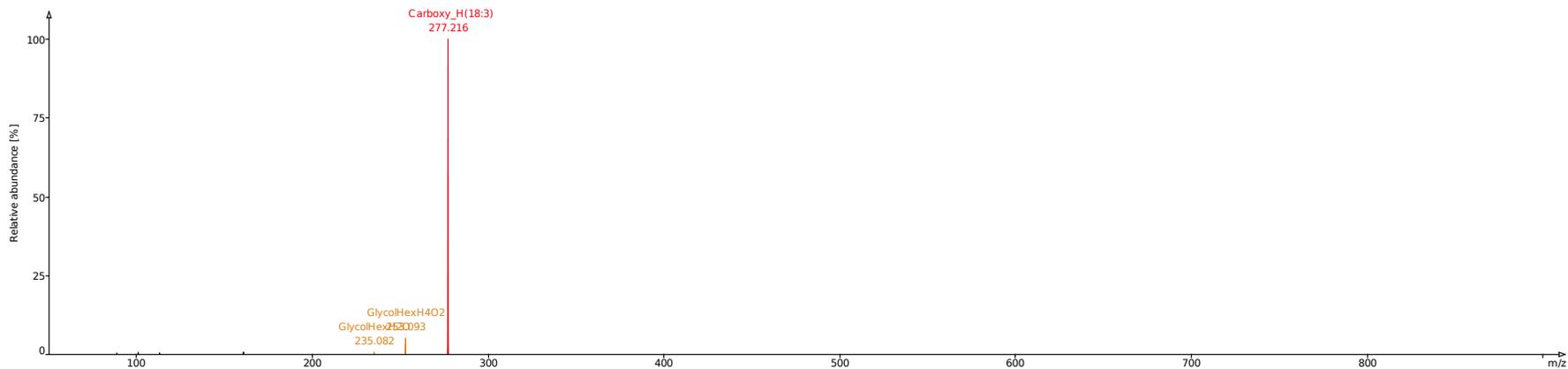


Figure S8: MS<sup>2</sup> spectrum of MGDG(36:6)+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 2 file *20130729\_NP-ve\_003.raw* (negative ion mode)

75

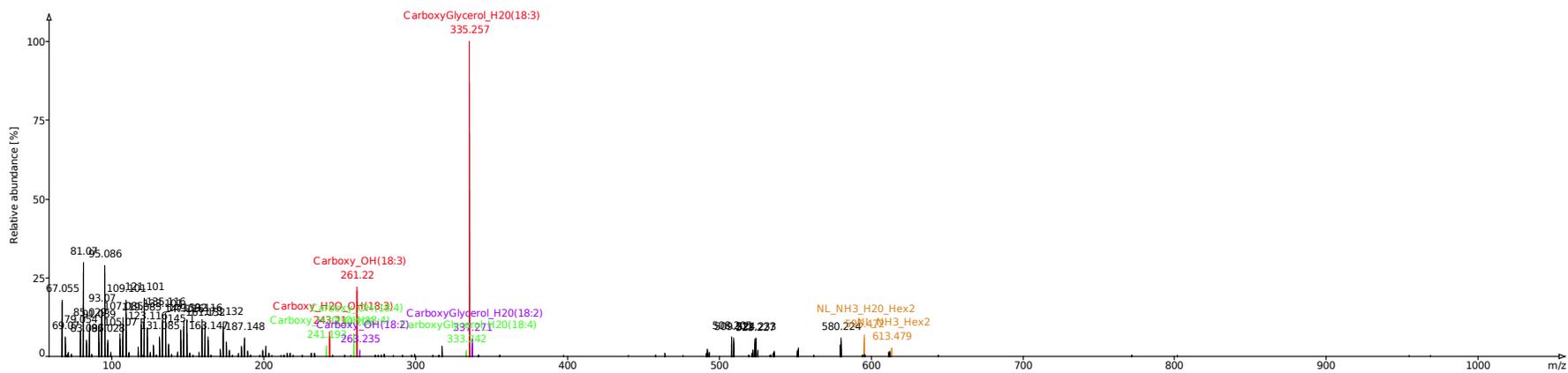


Figure S9: MS<sup>2</sup> spectrum of DGDG(36:6)+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 2 file *20130726\_NP+ve\_004.raw* (positive ion mode)

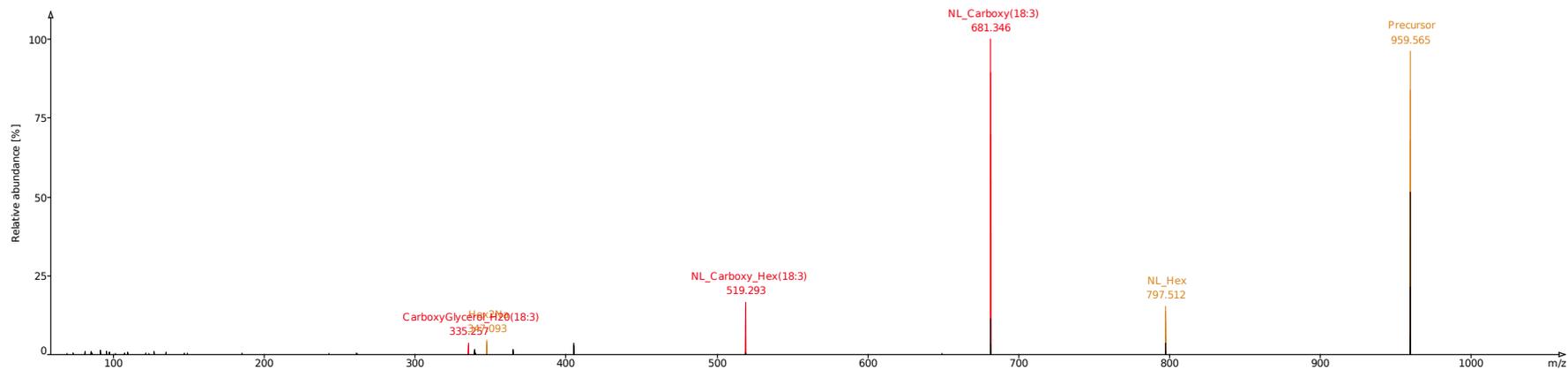


Figure S10: MS<sup>2</sup> spectrum of DGDG(36:6)+Na<sup>+</sup> in positive ion mode; reported by LDA in dataset 2 file *20130726\_NP+ve\_005.raw* (positive ion mode)

76

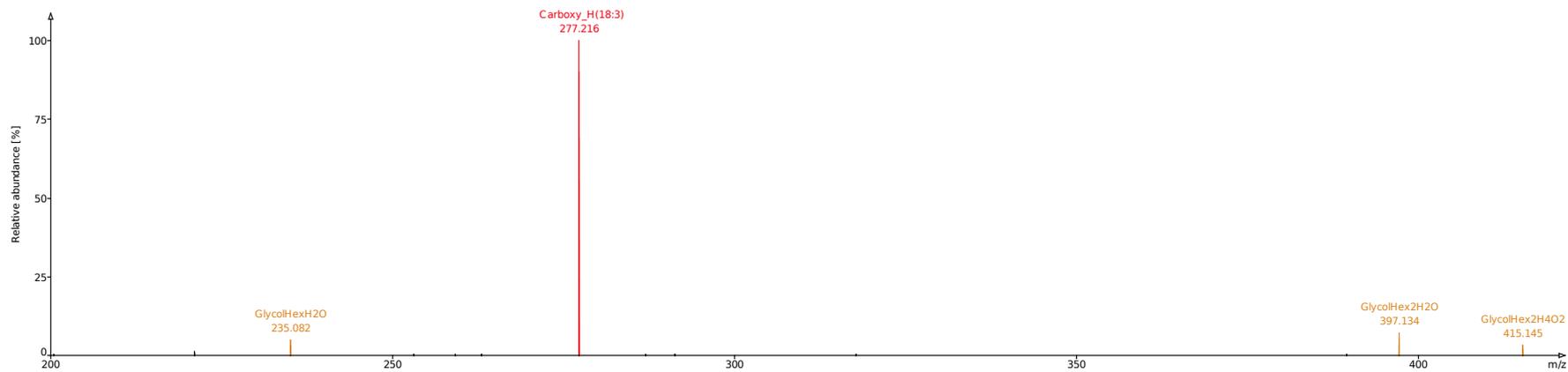


Figure S11: MS<sup>2</sup> spectrum of DGDG(36:6)-H<sup>-</sup>; reported by LDA in dataset 2 file *20130729\_NP-ve\_007.raw* (negative ion mode)

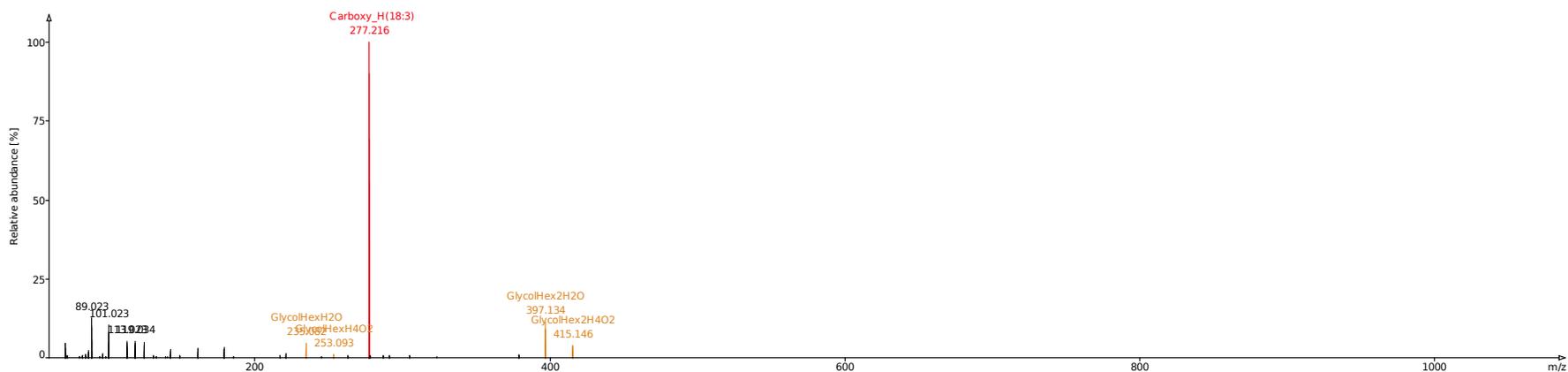


Figure S12: MS<sup>2</sup> spectrum of DGDG(36:6)+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 2 file *20130729\_NP-ve\_003.raw* (negative ion mode)

77

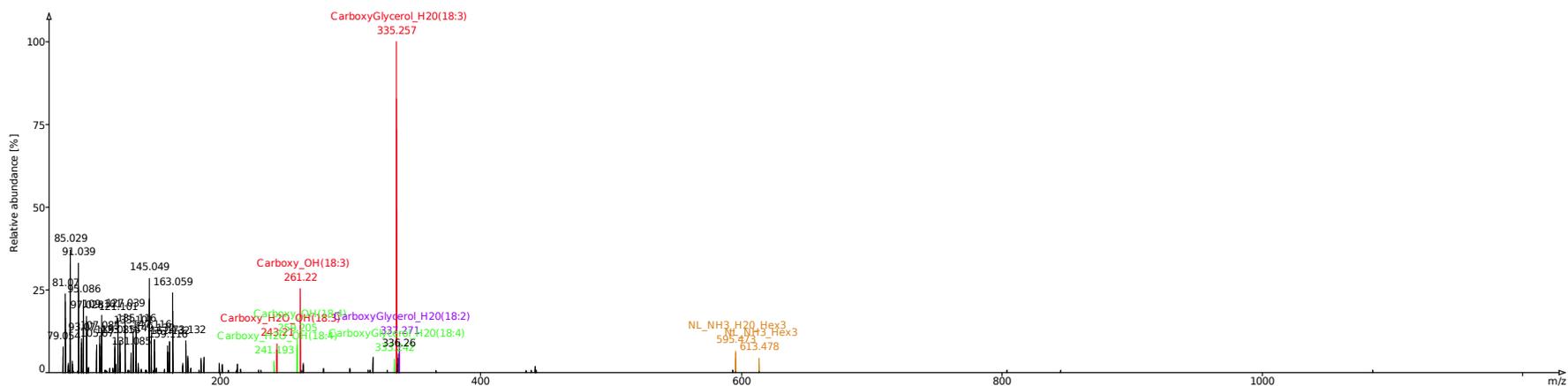
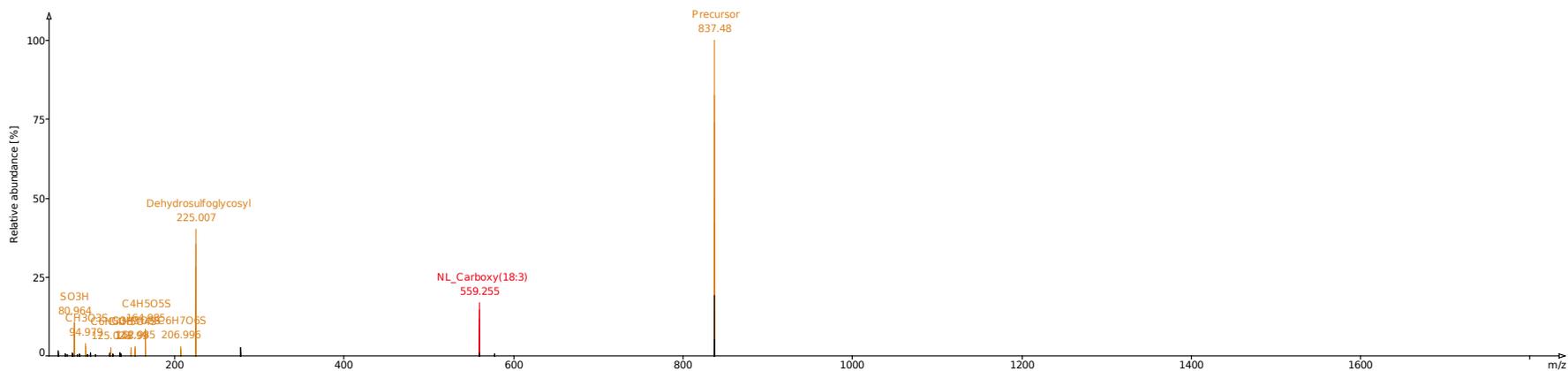


Figure S13: MS<sup>2</sup> spectrum of TriGDG(36:6)+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 2 file *20130726\_NP+ve\_005.raw* (positive ion mode)



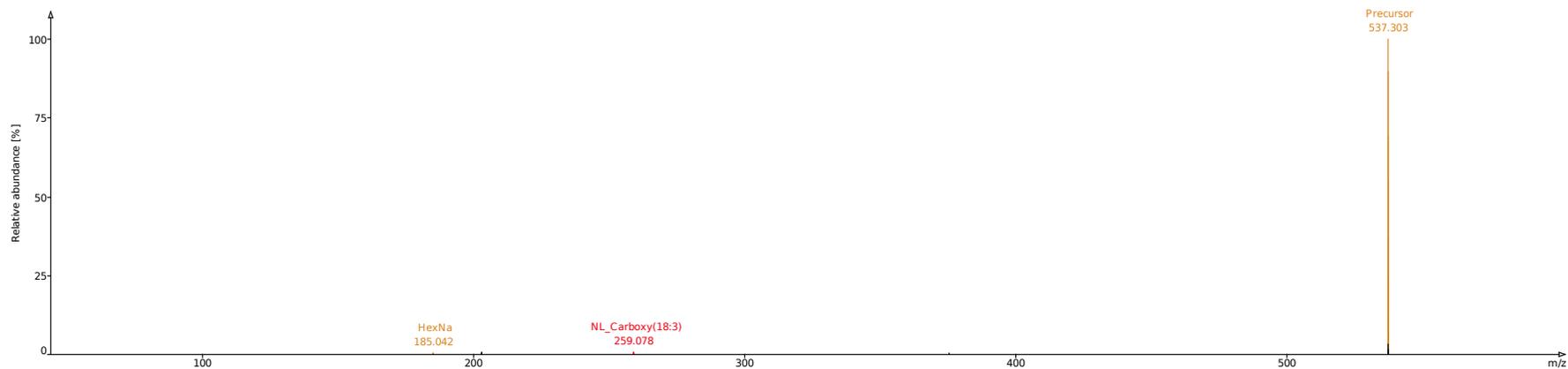


Figure S16: MS<sup>2</sup> spectrum of MGMG(18:3)+Na<sup>+</sup>; reported by LDA in dataset 2 file *20130726\_NP+ve\_009.raw* (positive ion mode)

79

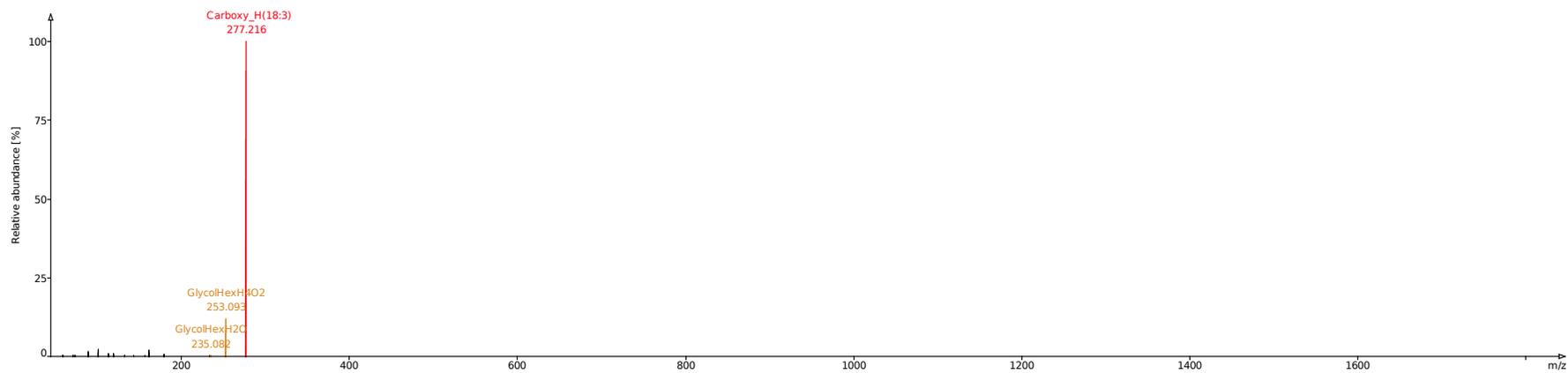


Figure S17: MS<sup>2</sup> spectrum of MGMG(18:3)+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 2 file *20130729\_NP-ve\_003.raw* (negative ion mode)

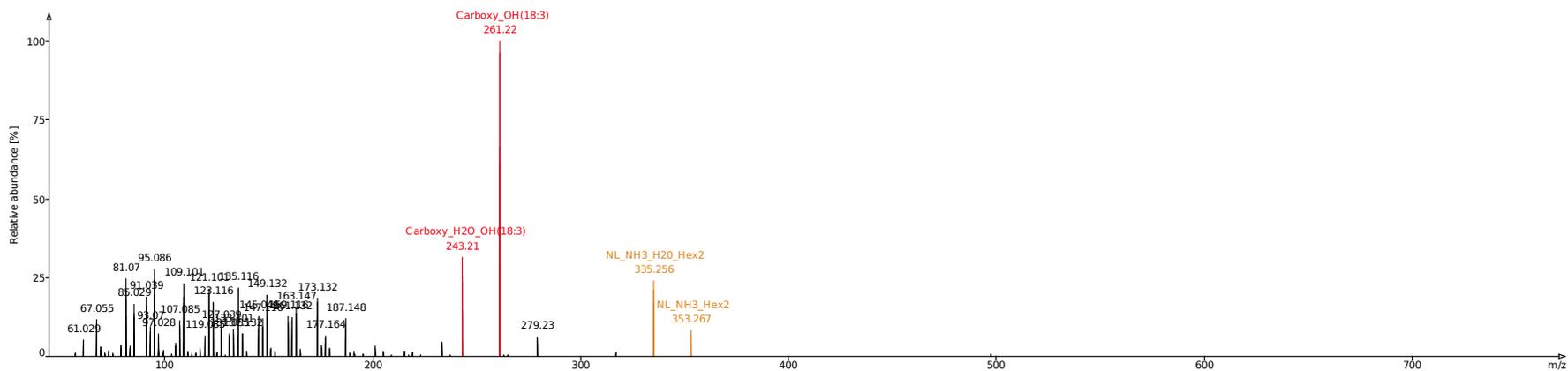


Figure S18: MS<sup>2</sup> spectrum of DGMG(18:3)+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 2 file *20130726\_NP+ve\_009.raw* (positive ion mode)

08

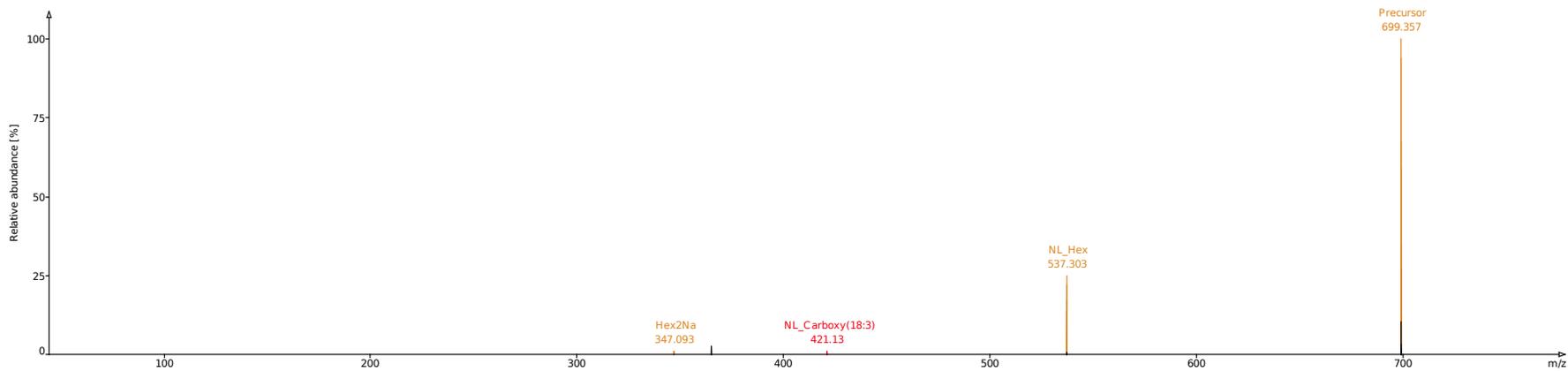


Figure S19: MS<sup>2</sup> spectrum of DGMG(18:3)+Na<sup>+</sup>; reported by LDA in dataset 2 file *20130726\_NP+ve\_009.raw* (positive ion mode)

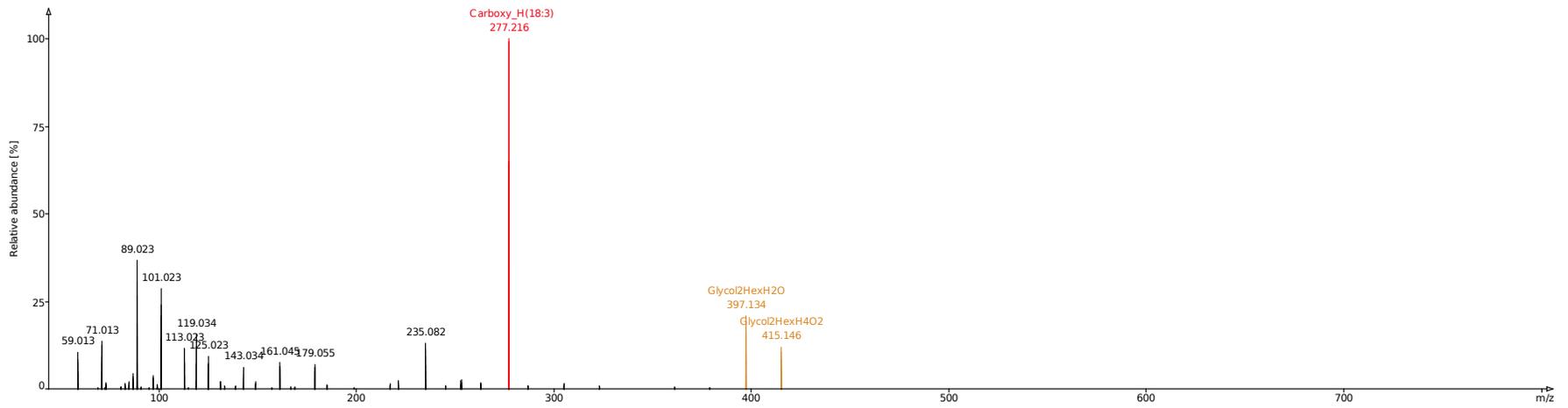


Figure S20: MS<sup>2</sup> spectrum of DGMG(18:3)+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 2 file *20130729\_NP-ve\_003.raw* (negative ion mode)

81

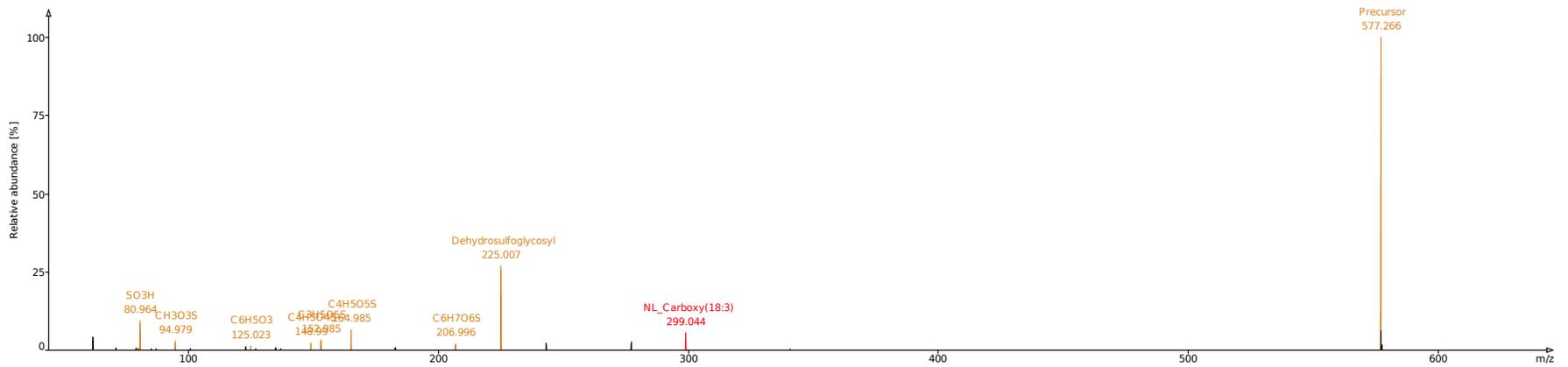


Figure S21: MS<sup>2</sup> spectrum of SQMG(18:3)-H<sup>-</sup>; reported by LDA in dataset 2 file *20130729\_NP-ve\_015.raw* (negative ion mode)

### 5.4.5 Dataset 3: Selected Spectra

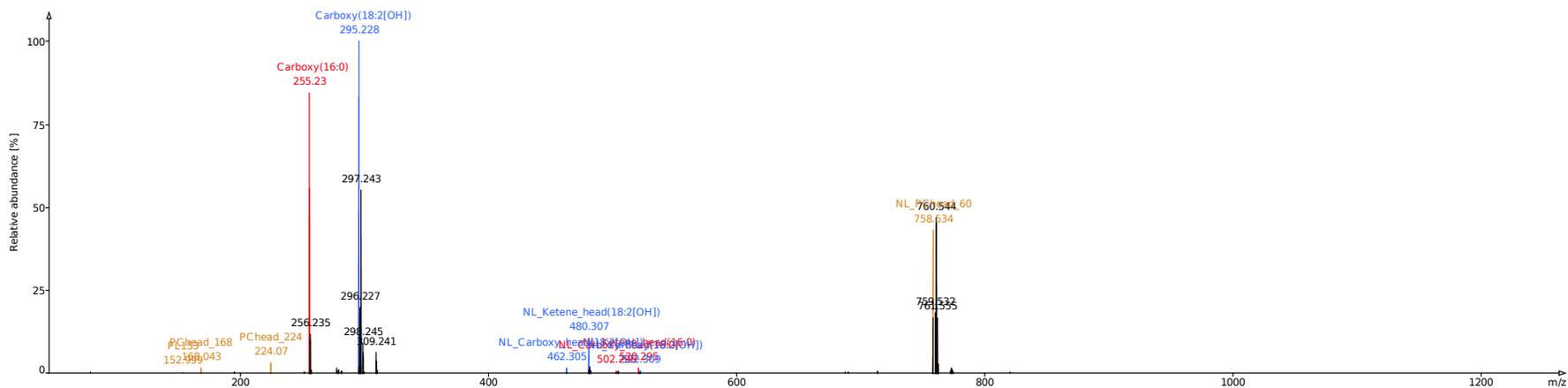


Figure S22: MS<sup>2</sup> spectrum of oxPC(34:2[OH])+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 3 file *180816\_oxPC\_10ng.mzXML* (negative ion mode)

82

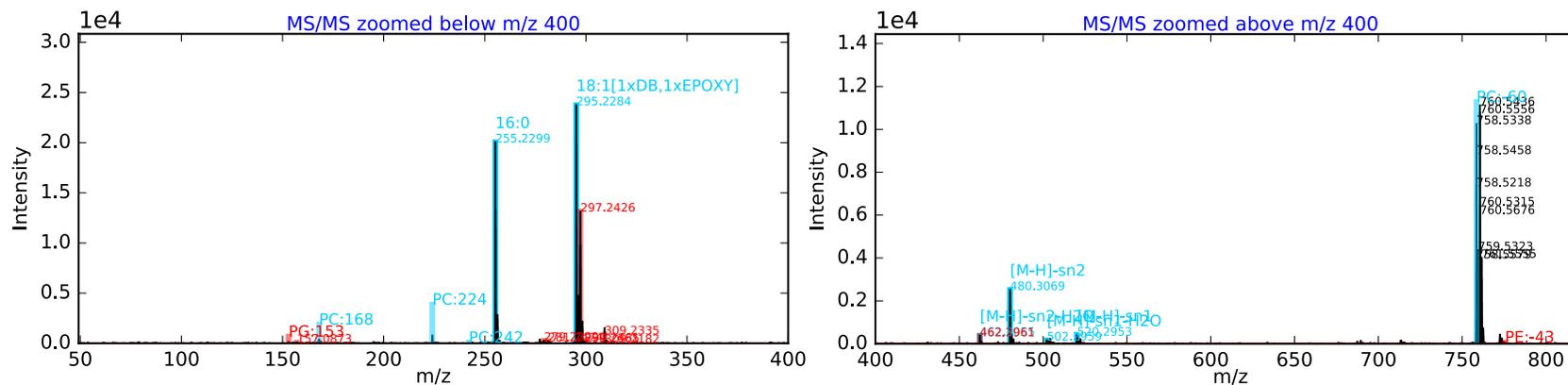


Figure S23: MS<sup>2</sup> spectrum of oxPC(34:1[O])+HCO<sub>2</sub><sup>-</sup>; reported by LPPtiger in dataset 3 file *180816\_oxPC\_10ng.mzXML* (negative ion mode)

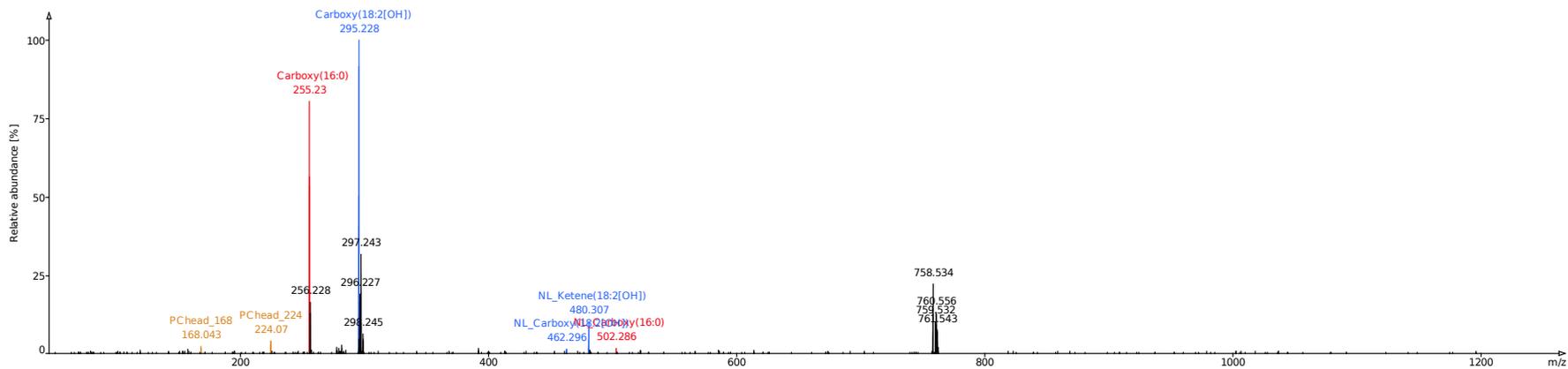


Figure S24: MS<sup>2</sup> spectrum of oxPC(34:2[OH])-CH<sub>3</sub><sup>-</sup>; reported by LDA in dataset 3 file *180816-oxPC\_10ng.mzXML* (negative ion mode)

83

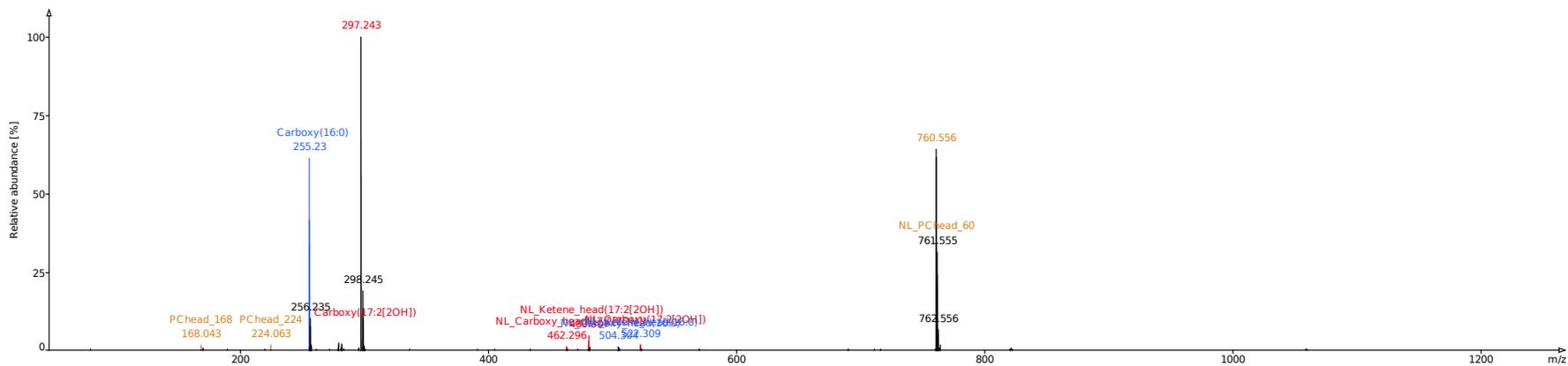


Figure S25: MS<sup>2</sup> spectrum of oxPC(33:2[2OH])+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 3 file *180816-oxPC\_10ng.mzXML* (negative ion mode)

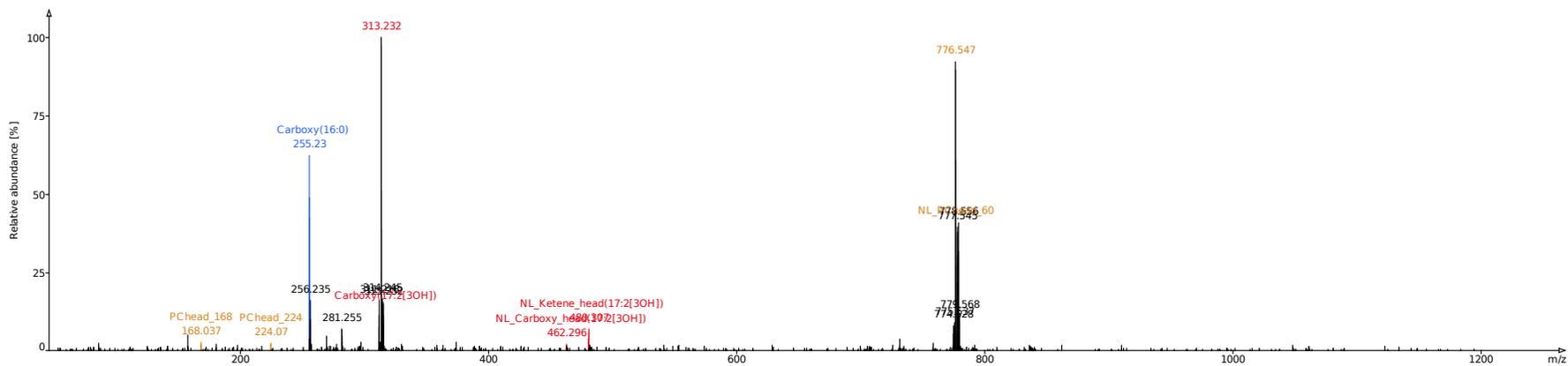


Figure S26: MS<sup>2</sup> spectrum of oxPC(33:2[3OH])+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 3 file *180816-oxPC\_10ng.mzXML* (negative ion mode)

84

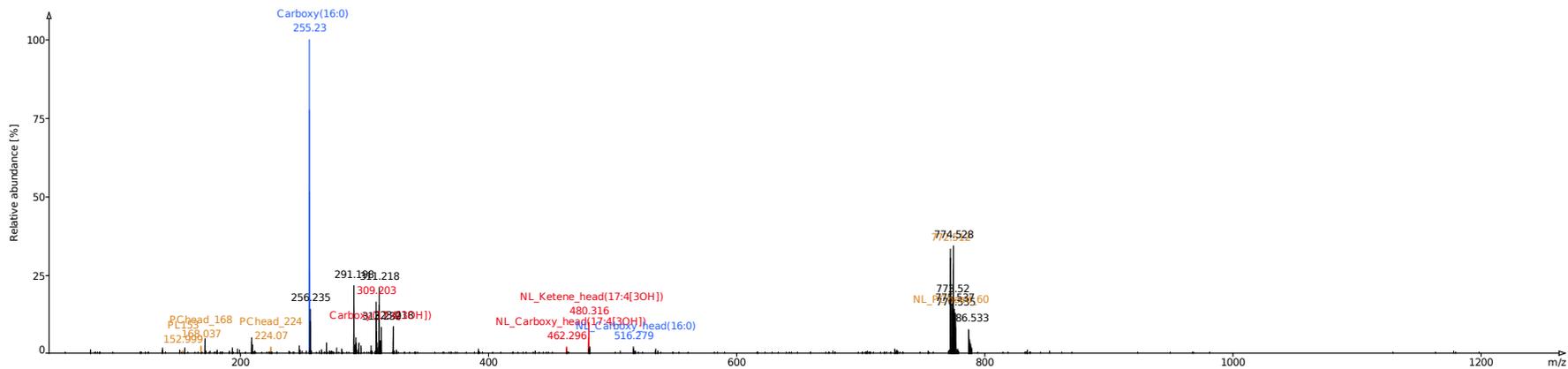


Figure S27: MS<sup>2</sup> spectrum of oxPC(33:4[3OH])+HCO<sub>2</sub><sup>-</sup>; reported by LDA in dataset 3 file *180816-oxPC\_10ng.mzXML* (negative ion mode)

## 5.4.6 Dataset 4: Selected Spectra

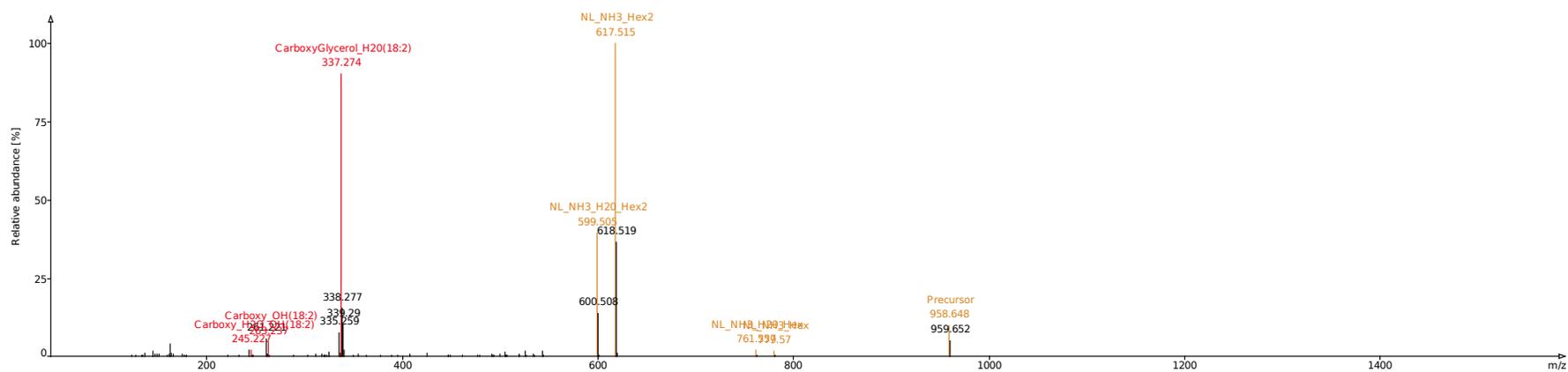


Figure S28: MS<sup>2</sup> spectrum of DGDG(36:4)+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 4 file *20160223b6\_2572.mzXML* (positive ion mode)

85

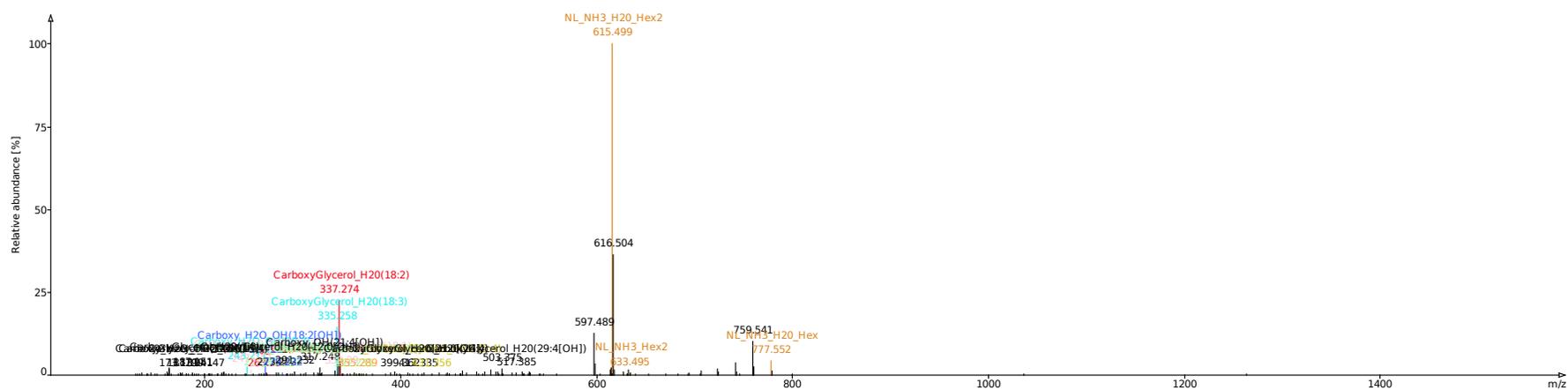


Figure S29: MS<sup>2</sup> spectrum of oxDGDG(36:4[OH])+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 4 file *20160223b6\_2572.mzXML* (positive ion mode)

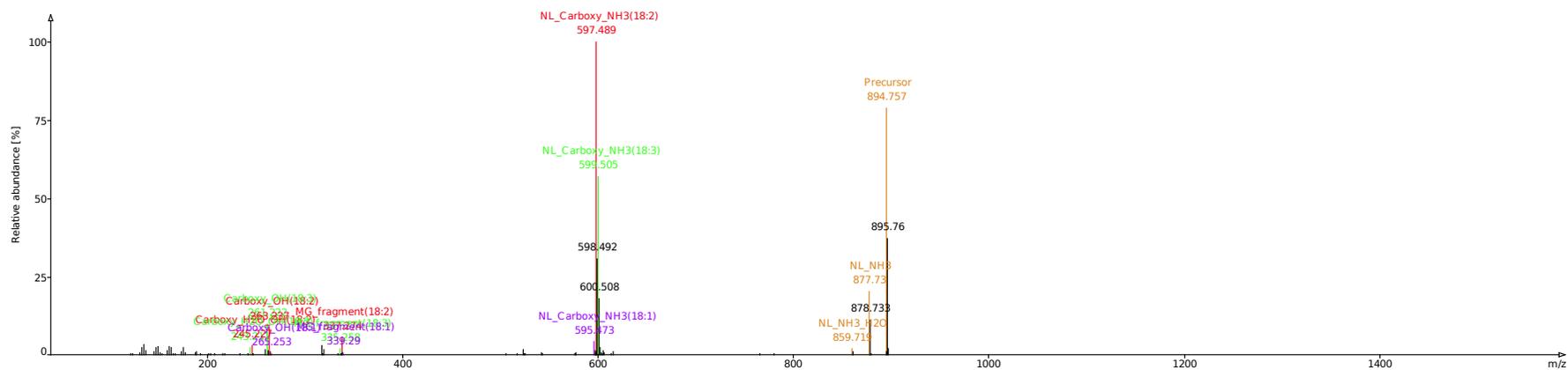
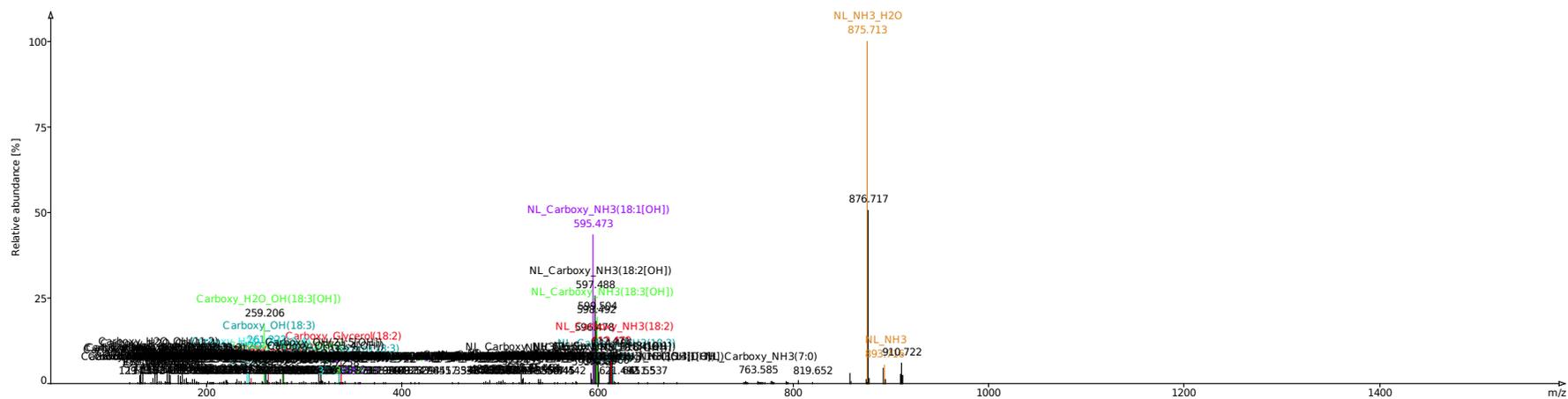


Figure S30: MS<sup>2</sup> spectrum of TG(54:7)+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 4 file *20160223b6\_2572.mzXML* (positive ion mode)



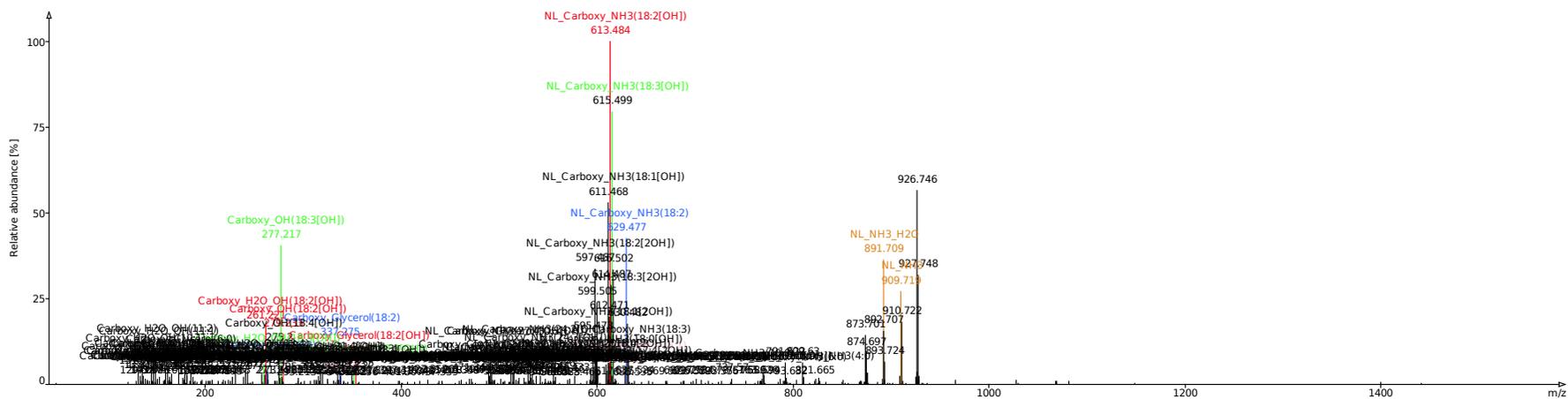


Figure S32: MS<sup>2</sup> spectrum of oxTG(54:7[2OH])+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 4 file *20160223b6\_2572.mzXML* (positive ion mode)

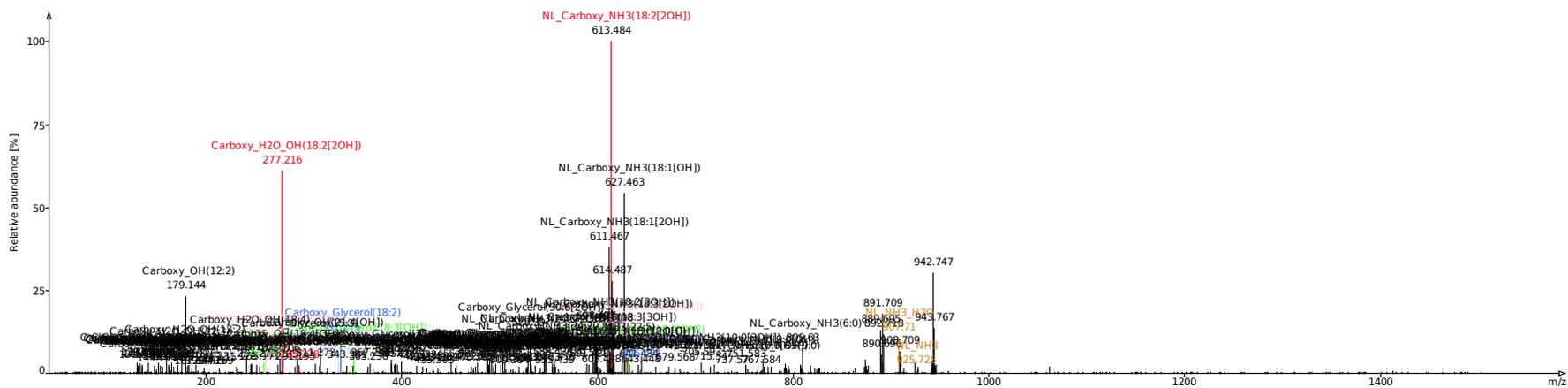


Figure S33: MS<sup>2</sup> spectrum of oxTG(54:7[3OH])+NH<sub>4</sub><sup>+</sup>; reported by LDA in dataset 4 file *20160223b6\_2572.mzXML* (positive ion mode)

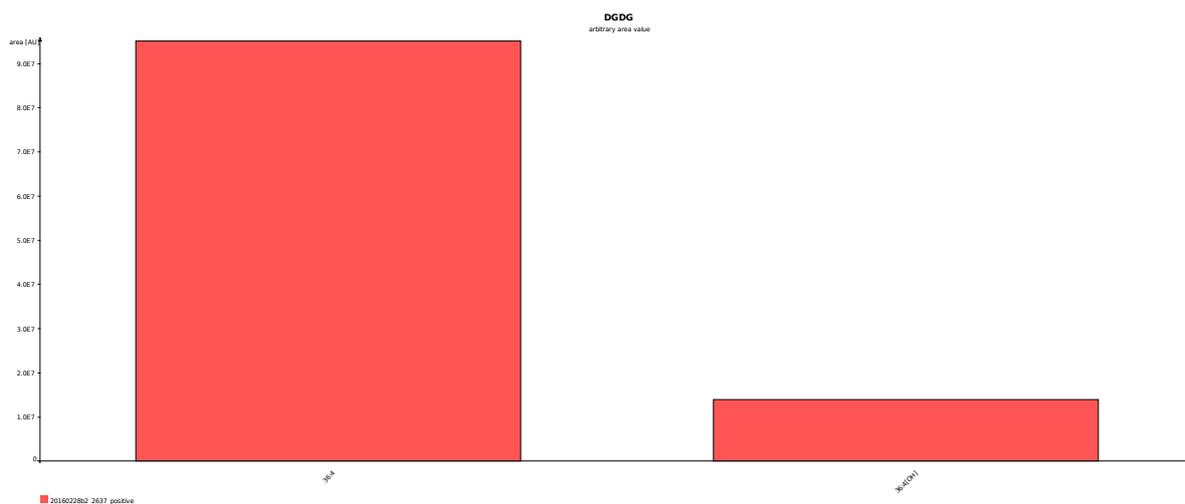


Figure S34: Abundances of DGDG(36:4) and oxDGDG(36:4[OH]); reported by LDA in dataset 4 file *20160223b6\_2572.mzXML* (positive ion mode)

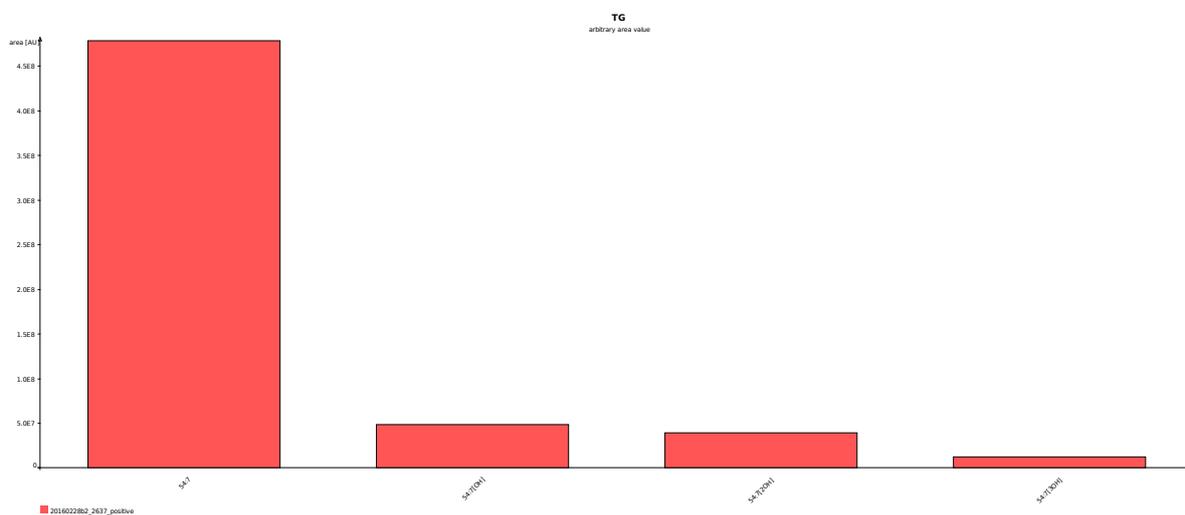


Figure S35: Abundances of TG(54:7), oxTG(54:7[OH]), oxTG(54:7[2OH]) and oxTG(54:7[3OH]); reported by LDA in dataset 4 file *20160223b6\_2572.mzXML* (positive ion mode)

## 5.5 Galactolipids - Comparison

The following tables show a direct comparison between the galactolipids (molecular species) found by Lipid Data Analyzer and LipidMatch, whereby cells with gray content refer to molecular species that were identified by LipidMatch but not by LipidMatch Flow.

Table S41: Identified galactolipids in dataset 2: LDA vs. LipidMatch (part 1/3)

Lipid Species	Adduct	Lipid Data Analyzer	LipidMatch
MGDG(28:3)	[M+Na] <sup>+</sup>	-	MGDG(10:0_18:3)
MGDG(34:2)	[M+HCO <sub>2</sub> ] <sup>-</sup>	MGDG(16:0_18:2)	MGDG(16:0_18:2)
MGDG(34:3)	[M+NH <sub>4</sub> ] <sup>+</sup>	MGDG(16:0_18:3)	-
	[M+Na] <sup>+</sup>	-	MGDG(16:0_18:3)
	[M+HCO <sub>2</sub> ] <sup>-</sup>	MGDG(16:0_18:3)	MGDG(16:0_18:3)
		-	MGDG(16:1_18:2)
MGDG(34:4)	[M+HCO <sub>2</sub> ] <sup>-</sup>	MGDG(16:1_18:3)	MGDG(16:1_18:3)
MGDG(35:6)	[M+HCO <sub>2</sub> ] <sup>-</sup>	MGDG(17:3_18:3)	-
MGDG(36:4)	[M+NH <sub>4</sub> ] <sup>+</sup>	MGDG(18:2_18:2)	MGDG(18:2_18:2)
		MGDG(18:1_18:3)	MGDG(18:1_18:3)
MGDG(36:5)	[M+NH <sub>4</sub> ] <sup>+</sup>	MGDG(18:1_18:4)	-
		MGDG(18:2_18:3)	-
MGDG(36:6)	[M+NH <sub>4</sub> ] <sup>+</sup>	MGDG(18:3_18:3)	MGDG(18:3_18:3)
		MGDG(18:2_18:4)	MGDG(18:2_18:4)
	[M+Na] <sup>+</sup>	MGDG(18:3_18:3)	MGDG(18:3_18:3)
	[M+HCO <sub>2</sub> ] <sup>-</sup>	MGDG(18:3_18:3)	MGDG(18:3_18:3)
		-	MGDG(18:2_18:4)
MGDG(36:7)	[M+Na] <sup>+</sup>	-	MGDG(18:3_18:4)

Table S42: Identified galactolipids in dataset 2: LDA vs. LipidMatch (part 2/3)

Lipid Species	Adduct	Lipid Data Analyzer	LipidMatch
DGDG(34:1)	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(16:0_18:1)	DGDG(16:0_18:1)
DGDG(34:2)	[M+NH <sub>4</sub> ] <sup>+</sup>	DGDG(16:0_18:2)	-
	[M+Na] <sup>+</sup>	DGDG(16:0/18:2)	not implemented
	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(16:0_18:2)	DGDG(16:0_18:2)
DGDG(34:3)	[M+NH <sub>4</sub> ] <sup>+</sup>	DGDG(16:0_18:3)	DGDG(16:0_18:3)
		-	DGDG(16:1_18:2)
	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(16:0_18:3)	DGDG(16:0_18:3)
		-	DGDG(16:1_18:2)
DGDG(34:4)	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(16:1_18:3)	DGDG(16:1_18:3)
DGDG(35:3)	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(17:0_18:3)	DGDG(17:0_18:3)
		DGDG(17:1_18:2)	DGDG(17:1_18:2)
DGDG(35:6)	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(17:3_18:3)	-
DGDG(36:3)	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(18:1_18:2)	DGDG(18:1_18:2)
		DGDG(18:0_18:3)	DGDG(18:0_18:3)
		-	DGDG(16:0_20:3)
DGDG(36:4)	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(18:2_18:2)	DGDG(18:2_18:2)
		DGDG(18:1_18:3)	DGDG(18:1_18:3)
DGDG(36:6)	[M+NH <sub>4</sub> ] <sup>+</sup>	DGDG(18:3_18:3)	DGDG(18:3_18:3)
		DGDG(18:2_18:4)	-
			DGDG(14:1_22:5)
	[M+Na] <sup>+</sup>	DGDG(18:3_18:3)	not implemented
	[M-H] <sup>-</sup>	DGDG(18:3_18:3)	not implemented
	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGDG(18:3_18:3)	DGDG(18:3_18:3)

Table S43: Identified galactolipids in dataset 2: LDA vs. LipidMatch (part 3/3)

Lipid Species	Adduct	Lipid Data Analyzer	LipidMatch
TriGDG(36:6)	[M+NH <sub>4</sub> ] <sup>+</sup>	TriGDG(18:3/18:3) TriGDG(18:2_18:4)	not implemented
SQDG(32:0)	[M-H] <sup>-</sup>	SQDG(16:0/16:0)	SQDG(16:0_16:0)
SQDG(34:3)	[M-H] <sup>-</sup>	SQDG(16:0/18:3)	SQDG(16:0_18:3)
SQDG(36:3)	[M-H] <sup>-</sup>	SQDG(18:0/18:3) -	SQDG(18:0_18:3) SQDG(16:0_20:3)
SQDG(36:6)	[M-H] <sup>-</sup>	SQDG(18:3/18:3) -	SQDG(18:3_18:3) SQDG(16:1_20:5)
MGMG(18:3)	[M+Na] <sup>+</sup> [M+NH <sub>4</sub> ] <sup>+</sup> [M+HCO <sub>2</sub> ] <sup>-</sup>	MGMG(18:3) MGMG(18:3) MGMG(18:3)	not implemented
DGMG(16:0)	[M+HCO <sub>2</sub> ] <sup>-</sup>	DGMG(18:3)	not implemented
DGMG(18:3)	[M+Na] <sup>+</sup> [M+NH <sub>4</sub> ] <sup>+</sup> [M+HCO <sub>2</sub> ] <sup>-</sup>	DGMG(18:3) DGMG(18:3) DGMG(18:3)	not implemented
SQMG(18:3)	[M-H] <sup>-</sup>	SQMG(18:3)	not implemented

### 5.5.1 LipidMatch-only Identifications

Table S44: LipidMatch-only identifications; a comment column provides further information for each lipid

Molecular Species Adduct	Files	Comment
[MGDG(10:0_18:3)+Na] <sup>+</sup>	20130726_NP+ve_xxx.raw xxx = 007-013	LDA did not identify this lipid species, because of missing <i>HexNa</i> fragment LipidMatch identification made by tiny abundances; also identified by LipidMatch Flow.
[MGDG(16:0_18:3)+Na] <sup>+</sup>	20130726_NP+ve_xxx.raw xxx = 003-016	LDA identified lipid molecular species, but not with this adduct Missing <i>HexNa</i> fragment; also identified by LipidMatch Flow.
[MGDG(16:1_18:2)+HCO <sub>2</sub> ] <sup>-</sup>	20130729_NP-ve_xxx.raw xxx = 004-007, 009-012, 015	LDA identified lipid species, but not this lipid molecular species (missing chain fragments) LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.
[MGDG(18:2_18:4)+HCO <sub>2</sub> ] <sup>-</sup>	20130729_NP-ve_xxx.raw xxx = 004-013, 015	LDA identified lipid species, but not this lipid molecular species (missing chain fragments) LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.
[MGDG(18:3_18:4)+Na] <sup>+</sup>	20130726_NP+ve_xxx.raw xxx = 005-008, 010, 012, 013	LDA did not identify this lipid species, because of missing <i>HexNa</i> fragment LipidMatch identification made by tiny abundances; also identified by LipidMatch Flow.
[DGDG(16:1_18:2)+NH <sub>4</sub> ] <sup>+</sup>	20130726_NP+ve_xxx.raw xxx = 003-005, 013	LDA identified lipid species, but not this lipid molecular species (missing chain fragments) LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.
[DGDG(16:1_18:2)+HCO <sub>2</sub> ] <sup>-</sup>	20130729_NP-ve_xxx.raw xxx= 003, 004, 006, 008, 013, 014, 016	LDA identified lipid species, but not this lipid molecular species (missing chain fragments) LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.
[DGDG(16:0_20:3)+HCO <sub>2</sub> ] <sup>-</sup>	20130729_NP-ve_xxx.raw xxx = 005-008, 009- 012, 014, 016	LDA identified lipid species, but not this lipid molecular species (missing chain fragments); LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.
[DGDG(14:1_22:5)+NH <sub>4</sub> ] <sup>+</sup>	20130726_NP+ve_xxx.raw xxx = 007, 013	LDA identified lipid species, but not this lipid molecular species (missing chain fragments) LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.
[SQDG(16:0_20:3)-H] <sup>-</sup>	20130729_NP-ve_xxx.raw xxx = 008, 009	LDA identified lipid species, but not this lipid molecular species (missing chain fragments) LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.
[SQDG(16:1_20:5)-H] <sup>-</sup>	20130729_NP-ve.016.raw	LDA identified lipid species, but not this lipid molecular species (missing chain fragments) LipidMatch identification made by tiny abundances; not identified by LipidMatch Flow.

## Selected Spectra

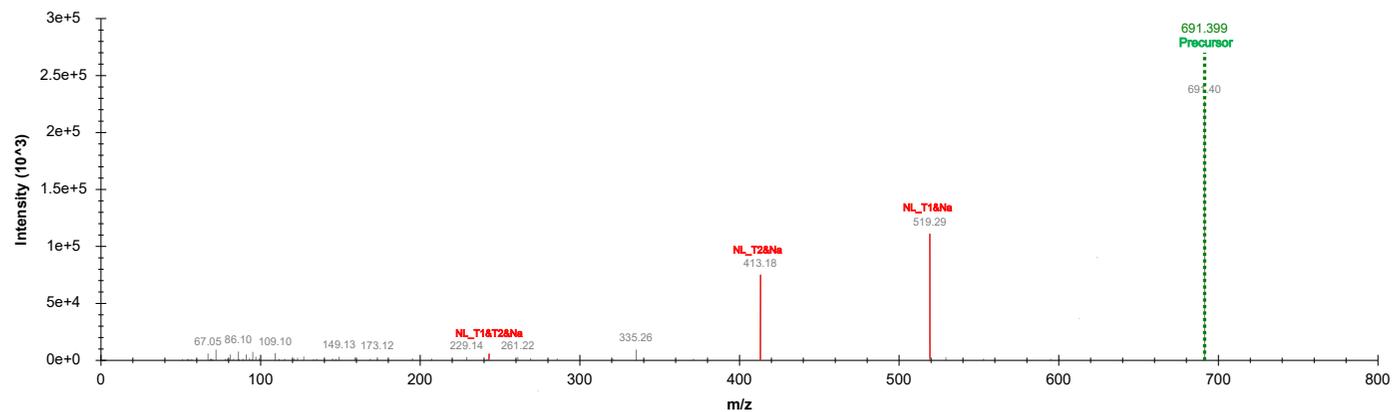


Figure S36: MS<sup>2</sup> spectrum of MGDG(10:0.18:3)+Na<sup>+</sup>; reported by LM and LMF in dataset 2 file *20130726\_NP+ve.007.raw* (negative ion mode)

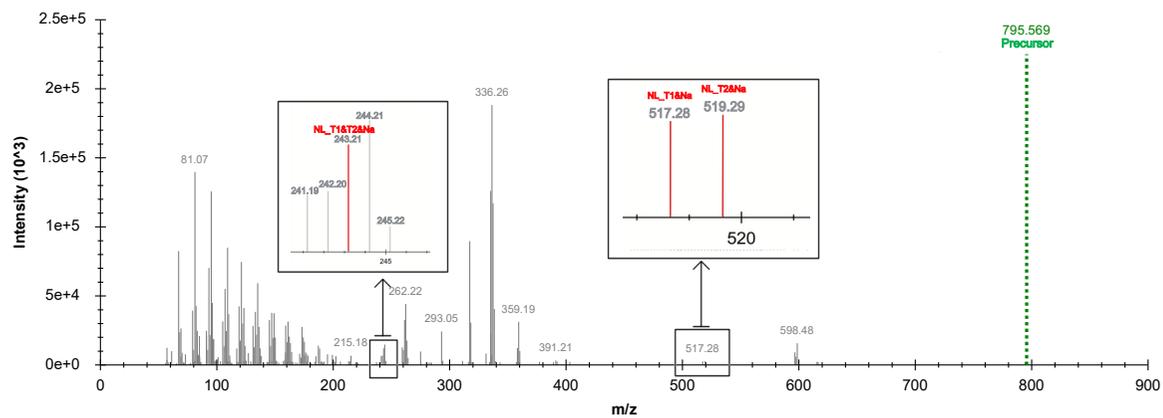


Figure S37: MS<sup>2</sup> spectrum of MGDG(18:3.18:4)+Na<sup>+</sup>; reported by LM and LMF in dataset 2 file *20130726\_NP+ve.007.raw* (negative ion mode)

## 5.6 Noisy Spectra

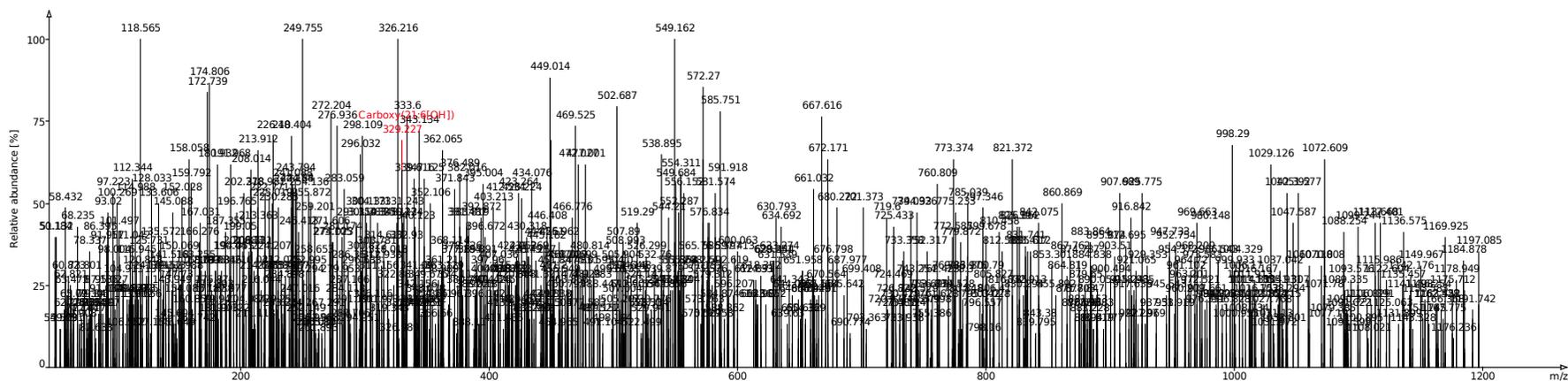


Figure S38: Very noisy spectrum: false positive identification; reported by LDA in dataset 3 file *180816\_oxPC\_10ng.mzXML* (negative ion mode)

93

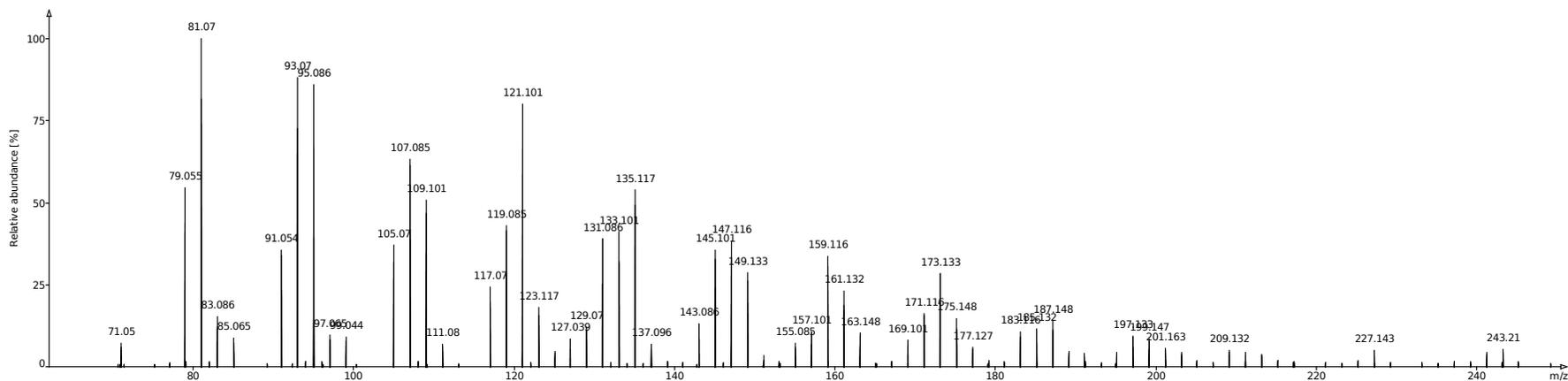


Figure S39: Zoomed low mass cluster observed in  $\text{NH}_4^+$  adducts of various lipid classes; dataset 2 file *20130726\_NP+ve\_009.raw* (positive ion mode)