Lipidomics of Biological Samples for the Assessment of Physiological and Pathological Processes

by

Adriana Zardini Buzatto

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
University of Alberta

ABSTRACT

The *omic* sciences are the interdisciplinary study of all the biological molecules within an organism. The combination of different *omics* approaches is a powerful way to unravel physiological and pathological processes. In the past decades, we have seen major developments and breakthroughs in the genomics, transcriptomics and proteomics fields, but there is still a need for further research into the small molecules that control our metabolism. Lipids are a large class of hydrophobic molecules involved in energy storage, signaling, modulation of gene expression and membranes. Lipidomics focuses on comprehensive analysis of lipids and their interactions with other molecules. Although the potential of lipidomics to study physiological and pathological processes is undeniable, reliable methodologies for the comprehensive assessment of the lipidome in biological samples are still needed.

Many biological samples are available in small volumes or limited amounts and display very complex lipid compositions. Nano-scale liquid chromatography allied to mass spectrometry (nanoLC-MS) offers extremely high sensitivity, although it is known to be technically more challenging than the conventional liquid chromatography approach. The first part of this work describes the development and optimization of a nanoLC-MS method for routine analysis of the lipidome of small volumes of biological samples (1.0 to 2.5 µL) with high sensitivity. The nanoLC method, mass spectrometry conditions and sample preparation by liquid-liquid extraction of the lipidome were fully optimized, including an evaluation of contamination sources. The method can be easily adapted to other types of biological samples where only limited volumes are available. A pilot study for biomarker discovery of spinal cord injury using the developed nanoLC-MS method applied to blood serum, cerebrospinal fluid and intraparenchymal microdialysate samples

is presented in Chapter III. Furthermore, an application of the method after minor modifications for extremely diluted exosome samples is described in Chapter IV.

The second part of this work focuses on the application of comprehensive, untargeted lipidomics of biological samples by liquid chromatography coupled to mass spectrometry to study the physiological and pathological effects of different conditions upon the lipidome. In Chapter V, the mechanism of action of a novel vaccine against respiratory syncytial virus, the leading cause of acute lower respiratory infections in young children, is investigated in a rat model by a combination of untargeted lipidomics and chemical isotope labeling-based metabolomics. Metabolic pathways involving the lung surfactant layer, synthesis and regulation of amino acids and unsaturated fatty acids were modulated by immunization and viral challenge. Chapter VI describes the use of untargeted lipidomics of serum to identify novel biomarkers of Parkinson's disease and predict the progression to dementia up to three years before noticeable symptoms. The results provide valuable new information to study the relationship between neurological diseases and the lipidome. Chapter VII investigates the lipidome alterations in different tissues upon treatment with dexamethasone, a glucocorticoid highly prescribed by the medical community but often associated with severe adverse effects. Finally, Chapter VIII shows an application of untargeted, comprehensive lipidomics to investigate alterations in the lipid metabolism related to cystic fibrosis. The study emphasizes the importance of odd-chain fatty acids, previously seen as irrelevant in humans due to their inherent low concentrations, for pathological processes.

This work illustrates that lipidomic profiling can provide a comprehensive understanding of metabolic alterations caused by physiological and pathological processes in the organism. The results described herein confirm the potential of high-quality untargeted lipidomics studies for physiological and pathogenesis analyses, as well as for biomarker discovery.

PREFACE

A version of Chapter II was recently accepted for publication at Analytica Chimica Acta as: Adriana Zardini Buzatto, Brian Kwon and Liang Li, Development of a NanoLC-MS Workflow for High-Sensitivity Global Lipidomic Analysis (Aug/2020). Part of the project was also presented as a poster at the 65th American Society for Mass Spectrometry (ASMS) Conference on Mass Spectrometry and Allied Topics (June 3-8, 2017, Indianapolis, Indiana, USA) as: Adriana Zardini Buzatto, Jaspaul Tatlay, Brian Kwon and Liang Li; Development of a high-sensitivity method for lipidomics of small volumes of biological fluids by nanoLC-MS. This work is part of a consortium established between University of British Columbia (Dr. Brian Kwon and Dr. Leonard Foster), University of Alberta (Dr. Liang Li) and the Rick Hansen Institute to study biomarkers of spinal cord injury through lipidomics, metabolomics, genomics and proteomics. Blood serum and cerebrospinal fluid samples from pigs were obtained through a collaboration with Dr. Brian Kwon (University of British Columbia, Vancouver, BC, Canada). I was responsible for the method development for the comprehensive, untargeted lipidomics of small volumes of biological samples using nanoLC-MS, including experimental design and execution, data processing and interpretation, and manuscript preparation. Dr. Liang Li contributed to the experimental design, supervised the project and edited the manuscript. Dr. Kwon provided the pig samples and edited the manuscript.

Chapter III was part of the same consortium to study biomarkers of spinal cord injury through a collaboration with Dr. Brian Kwon. Versions of this chapter were presented as an oral presentation at the 101st Canadian Chemistry Conference and Exhibition (CSC, May 27 - 31, 2018, Edmonton, AB, Canada) and as a poster at the 66th American Society for Mass Spectrometry (ASMS) Conference on Mass Spectrometry and Allied Topics (June 2-9, 2018, San Diego,

California, USA) as: Adriana Zardini Buzatto, Brian Kwon and Liang Li; Lipidomics of spinal cord injury: high sensitivity analysis of small volumes of biological fluids by nanoLC-MS. I was responsible for the experimental design and execution, data processing, data interpretation and writing. Dr. Liang Li supervised the project and designed the experiments. Dr. Brian Kwon designed the animal trial and supervised the project.

A version of Chapter IV is currently being prepared for publication. The work was a collaboration with Dr. David M. Lubman (University of Michigan Medical Center, Ann Arbor, MI, USA). Serum samples and exosome isolates were obtained by Dr. Jianhui Zhu and Jie Zhang at the University of Michigan Medical Center. I was responsible for the experimental design and execution, data processing, interpretation and writing. Dr. Liang Li designed the experiments and supervised the project. Dr. David Lubman and his group collected the samples, designed the experiments and supervised the project.

Versions of Chapter V were published as: (1) Indranil Sarkar, Adriana Zardini Buzatto, Ravendra Garg, Liang Li and Sylvia van Drunen Little-van den Hurk, Metabolomic and immunological profiling of respiratory syncytial virus infection after intranasal immunization with a subunit vaccine candidate, Journal of Proteome Research 2019, 18(3), 1145-1161; and (2) Adriana Zardini Buzatto, Indranil Sarkar, Sylvia van Drunen Little-van den Hurk and Liang Li, Comprehensive lipidomic and metabolomic analysis for studying metabolic changes in lung tissue induced by a vaccine against respiratory syncytial virus, ACS Infectious Diseases 2020.[1,2] The project was performed through a collaboration with Dr. Sylvia van Drunen Little-van den Hurk's research group (University of Saskatchewan, Saskatoon, SK, Canada), which developed the novel vaccine candidate, handled the animals, collected the samples and performed the immunology experiments reported in Sarkar *et al*.[1] I was responsible for the method optimization for tissue

samples, experimental design and execution, data processing, data interpretation, and manuscript preparation and editing. Dr. Indranil Sarkar (University of Saskatchewan, Saskatoon, SK, Canada) handled the animals, collected the samples and prepared and edited the manuscripts. Dr. Sylvia van Drunen Little-van den Hurk and Dr. Liang Li designed the experiments, edited the manuscripts and supervised the project.

Chapter VI was a collaboration with Dr. Richard Camicioli (Neuroscience and Mental Health Institute and Department of Medicine - Neurology) and Dr. Roger A. Dixon (Neuroscience and Mental Health Institute and Department of Psychology, University of Alberta, Edmonton, AB, Canada) to study lipid alterations induced by Parkinson's disease and transition to dementia. A version of the work was presented as a poster at the 67th American Society for Mass Spectrometry (ASMS) Conference on Mass Spectrometry and Allied Topics (June 2-6, 2019, Atlanta, Georgia, USA) as: Adriana Zardini Buzatto, Barinder Bajwa, Jaspaul Tatlay, Roger A. Dixon, Richard Camicioli and Liang Li; Lipidomics of Parkinson's Disease: towards more accurate diagnosis methods through *omics* technologies. I was responsible for the lipidomics method optimization, experimental design and execution, data processing, interpretation and writing. Barinder Bajwa (University of Alberta, Edmonton, AB, Canada) helped with the sample extraction steps. Dr. Liang Li designed the experiments, supervised the project and edited the text. Dr. Richard Camicioli and Dr. Roger A. Dixon recruited and assessed the patients and healthy controls, collected the samples, supervised the work and edited the text. A version of this chapter is currently being prepared for publication.

Chapter VII was a collaboration with Dr. Anas M. Abdel Rahman and Dr. Majed Dasouki (King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia). I was responsible for the experimental design, sample analysis, data processing, interpretation and writing. Dr. Anas

M. Abdel Rahman and Dr. Majed Dasouki's teams handled the animals and collected the samples. Dr. Liang Li designed the experiments, supervised the work and edited the text. Xiaohang Wang (University of Alberta, Edmonton, AB, Canada) helped with the sample homogenization and further worked on metabolomics experiments using the aqueous layer obtained after liquid-liquid extraction.[3] A version of this chapter is currently being prepared for publication by Dr Anas M. Abdel Rahman's team.

Chapter VIII was also developed through a collaboration with Dr. Anas M. Abdel Rahman and Dr. Majed Dasouki (King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia). A version of this chapter was conditionally accepted for publication as Adriana Zardini Buzatto, Mai Abdel Jabar, Imran Nizami, Majed Dasouki, Liang Li and Anas Abdel Rahman, Lipidome Alterations Induced by Cystic Fibrosis, CFTR Mutation, and Lung Function, Journal of Proteome Research (Sep / 2020). I was responsible for the experimental design, sample analysis, data processing, interpretation and manuscript preparation. Dr. Anas M. Abdel Rahman and Dr. Majed Dasouki's teams recruited the patients and collected the samples. Dr. Anas M. Abdel Rahman also edited the manuscript and supervised the experiments. Dr. Liang Li designed the experiments, supervised the work and edited the manuscript.

Other conference presentations for projects not described in this thesis, but performed at University of Alberta as a PhD student and candidate include: (1) Nan Wang, Xian Luo, Shuang Zhao, Yiman Wu, Zhendong Li, Wei Han, Jaspaul Tatlay, Yunong Li, Kevin Hooton, Dorothea Mung, Adriana Zardini Buzatto, Xiaohang Wang, Aiko Barsch, Ulrike Schweiger Hufnagel and Liang Li, Enabling high-confidence human endogenous metabolite identification via high-resolution MS/MS retention-time library (poster), 64th American Society for Mass Spectrometry (ASMS) Conference on Mass Spectrometry and Allied Topics (June 5-9, 2016, San Antonio,

Texas, USA); (2) Barinder Bajwa, Adriana Zardini Buzatto, Eric Smith and Liang Li; Lipidomics of Alzheimer's disease and cerebral amyloid angiopathy: identification of potential biomarkers in human plasma by UHPLC-MS; 67th American Society for Mass Spectrometry (ASMS) Conference on Mass Spectrometry and Allied Topics (June 2-6, 2019, Atlanta, Georgia, USA); and (3) Adriana Zardini Buzatto, Shuang Zhao, Ulrike Schweiger Hufnagel, Aiko Barsch and Liang Li; Integration of metabolomic and lipidomic workflows for studying biological samples; 67th American Society for Mass Spectrometry (ASMS) Conference on Mass Spectrometry and Allied Topics (June 2-6, 2019, Atlanta, Georgia, USA).

Supporting figures for Chapters II, V, VI and VIII can be found in the Appendixes A, B, C and D, respectively. Supporting tables are available with Dr. Liang Li.

ACKNOWLEDGMENTS

This work could not be completed without the guidance and supervision Dr. Liang Li. His enthusiasm and knowledge for research encouraged me to keep pursuing excellence in scientific research. I would like to express my most sincere gratitude to Dr. Li. I would also like to thank my supervisory committee, Dr. Sara Styler and Dr. John Vederas. Their comments and advice were extremely valuable and will stay with me in the future.

I am also thankful to all the collaborators that made the seven projects described in this work possible, as well as other projects that I was involved in during my PhD trajectory. I particularly thank Dr. Aiko Barsch, Dr. Ulrike Schweiger Hufnagel, Dr. Brian Kwon, Sara Assadian, Dr. Indranil Sarkar, Dr. Sylvia van Drunen Little-van den Hurk, Dr. Anas M. Abdel Rahman, Dr. Richard Camicioli, Dr. Roger A. Dixon, Dr. David Lubman and Jianhui Zhu.

The help and support of my past and current colleagues at the Li group was essential for an enjoyable and enriching experience for the last five years. Their comments and questions during group meetings, as well as enlightening discussions and help in the lab, were deeply appreciated. Special thanks to Jaspaul Tatlay for my initial training, Dr. Randy Whittal for his assistance with nanoLC maintenance and all the staff at the Department of Chemistry of University of Alberta. The funding, facilities and staff offered by University of Alberta were essential for the completion of this work.

Last but not least, my special appreciation to my parents, Urbano Celso Buzatto and Sonia Zardini Buzatto, and my family for their unconditional love and support. This work could not be finished without their encouragement and efforts. To my husband, Hugo da Silva Pessoa, who accepted the challenge of moving to a new country to follow my dream, thank you. This work is also yours.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
ACP	Acyl-carrier-protein
AMP	Adenosine monophosphate
ANOVA	Analysis of Variance
ASAH	N-acyl sphingosine amidohydrolase 1
ATP	Adenosine triphosphate
AUC	Area under the curve
BEH	Bridged ethylene hybrid
BMP	Bis(monoacylglycero)phosphate
BPC	Base-peak chromatogram
CAR	Acyl Carnitine
CCNA	Canadian Consortium on Neurodegeneration in Aging
CDR	Clinical Dementia Rating
CE	Cholesteryl ester
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane regulator
CI	Confidence interval
CIHR	Canadian Institutes of Health Research
CIRS	Cumulative Illness Rating Scale
CL	Cardiolipin
COPD	Chronic obstructive pulmonary disease
COVID	Coronavirus disease
COX	Cyclooxygenase
CRC	Canada Research Chairs
CSF	Cerebrospinal fluid
CSV	Comma-separated value
DA	Discriminant analysis
DCM	Dichloromethane

Dex	Dexamethasone
DG	Diacylglycerol
DGC	Acyl-sn-glycero-3-O-carboxy-(hydroxymethyl)-choline
DGD	Digalactosyldiacylgylcerol
DGDMS	Diacylglyceryltrimethylhomoserine
DGKQ	Diacylglycerol kinase
DGT	Diacylglyceryltrimethyl-homoserine
DHA	Docosahexaenoic acid
DNA	Deoxyribonucleic acid
DOPA	L-3,4-dihydroxyphenylalanine (Levodopa)
DPA	Docosapentaenoic acid
DPPC	Dipalmitoyl-phosphatidylcholine
DRS	Dementia Rating Scale
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray
EV	Extracellular vesicle
FA	Fatty acyl
FAB	Frontal Assessment Battery
FAHFA	Fatty Acid ester of Hydroxyl Fatty Acid
FC	Fold-change
FDR	False-discovery rate
FEV1	Forced expiratory volume for the first second of the forced breath
FVC	Forced vital capacity
$GABA_A$	γ-aminobutyric acid
GBA	Glucosylceramidase beta
GC	Gas chromatography
ICH-GCP	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use - Good Clinical Practice
GDG	Monogalactosyldiacylgylcerols
GDS	Geriatric Depression Scale
GL	Glycerolipid

GMG	Sulfoquinovosylmonoacyl-glycerol
GOLD	Global Initiative for Chronic Obstructive Lung Disease
GP	Glycerophospholipid
GPI	Glycosylphosphatidylinositol
GR	Glucocorticoid receptors
HEK	Human embryonic kidney
HETE	Hydroxyeicosatetraenoic acid
HMDB	Human Metabolome Database
HPA	Hypothalamic-pituitary-adrenal
HPLC	High-performance liquid chromatography
HY	Hoehn and Yahr Rating Scale
ID	Identification
IFN	Interferon
IL	Interleukin
ILCNC	International Lipid Classification and Nomenclature Committee
IMS	Ion mobility mass spectrometry
IPA	2-propanol or isopropanol
IRT	Immunoreactive trypsinogen
KEGG	Kyoto Encyclopedia of Genes and Genomes
KFSHRC	King Faisal Specialist Hospital and Research Center
LC	Liquid chromatography
LDL	Low-density lipoprotein
LLE	Liquid-liquid extraction
LPA	Lysophosphatidic acid
LPC	Lysophosphatidylcholine
LPE	Lysophosphatidylethanolamine
LPG	Lysophosphatidylglycerol
LPI	Lysophosphatidylinositol
LPS	Lysophosphatidylserine
LSL	Lysoglycosphingolipid
MD	Intraparenchymal microdialysis fluid

MG	Monoacylglycerol
MGDG	Monogalactosyldiacylglycerol
MIPC	Ceramide phosphoinositol
MMSE	Mini-Mental Status Examination
MPA	Mobile phase A
MPB	Mobile phase B
MRI	Magnetic resonance imaging
MS	Mass spectrometry
MTBE	Methyl t-butyl ether
MVB	Multivesicular bodies
NA	Not available
NAA	N-acyl amine
NADP	Nicotinamide adenine dinucleotide phosphate
NAE	N-acyl ethanolamine
NAT	N-acyl taurine
NBD	Nucleotide-binding domain
NHS	National Health Service
NMDA	N-methyl-D-aspartate
NMR	Nuclear magnetic resonance
OCFA	Odd-chain fatty acyl
OOB	Out-of-bag
OPLS	Orthogonal partial least square
PA	Phosphatidic acid
PBS	Phosphate Buffered Saline
PC	Phosphatidylcholine
PCA	Principal component analysis
PCCA	Propionyl-CoA carboxylase subunit alpha
PCCB	Propionyl-CoA carboxylase subunit beta
PCEP	Poly[di(sodiumcarboxylatoethylphenoxy)]-phosphazene
PD	Parkinson's disease
PDD	Parkinson's disease dementia

PDND	Parkinson's disease - no dementia
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PIP	Phosphatidylinositol-phosphate
PKA	Protein kinase A
PLS	Partial least-squares
PP	Polypropylene
PS	Phosphatidylserine
PTFE	Polytetrafluoroethylene
PUFA	Polyunsaturated fatty acyls
QC	Quality control
QTOF	Quadrupole Time-of-Flight
RNA	Ribonucleic acid
ROC	Receiver Operating Characteristic
RPLC	Reversed-phase liquid chromatography
RSD	Relative standard deviation
RSV	Respiratory syncytial virus
SD	Standard deviation
SCI	Spinal cord injury
SEC	Size-exclusion chromatography
SM	Sphingomyelin
SMPDB	Small Molecule Pathway Database
SP	Sphingolipids
SPB	Sphingoid base or Sphingoid base-1-phosphate
SQDG	Sulfoquinovosyldiacylglycerols
ST	Sterol
TCA	Tricarboxylic acid
TG	Triacylglycerol
TNF	Tumor necrosis factor
TOF	Time-of-Flight

TXB	Thromboxane
UC	Ultracentrifugation
UHPLC	Ultra-high-performance liquid chromatography
UPDRS	Unified Parkinson's Disease Rating Scale
UPLC	Ultra-Performance Liquid Chromatography
UV	Ultraviolet
VLDL	Very low-density lipoprotein
VIDO-InterVac	Vaccine and Infectious Disease Organization-International Vaccine Centre
WE	Wax ester
WHO	World Health Organization

I

Chapter I: Introduction

1.1. Omic sciences for the study of biological samples

Systems biology is a combination of analytical and biochemical approaches to study the relationship between different parts of a biological system. It is an interdisciplinary field that moves away from the traditional scientific method of focusing on small sections or parts of an organism; instead, the goal is a holistic approach to study the complex interactions that are part of all physiological and pathogenic processes, from the molecular level to the entire organism. This complex field of study requires the integration between different areas, including bioanalytical chemistry, biochemistry, biology, bioinformatics and statistics.

The *omic* sciences, *i.e.,* genomics, transcriptomics, proteomics, metabolomics, lipidomics and glycomics, are particularly useful for the study of systems biology. The combination of different *omic* approaches is powerful way to unravel physiological and pathological processes within an organism. In the past three decades, we have seen major developments and breakthroughs in the genomics, transcriptomics and proteomics fields, but there is still a need for further research into the small molecules that control our metabolism. Metabolomics is the comprehensive study of all metabolites and their relationships, including nucleic acids, amino acids, sugars, small peptides and lipids. Lipidomics, a branch of metabolomics first named by Han and Gross in 2003, is an integral part of systems biology and the main focus of this work.[4–8]

Lipidomics is the characterization of all lipid species in a biological sample, including their interactions, biological properties and dynamics. Lipids exhibit a variety of functions in an

organism, including cell and organelle membranes, signalling, inter-cell communication, energy storage, protein trafficking and folding, amongst others. Their concentrations are affected by genetics, age, sex, diet, physical exercise, circadian rhythms, immune status and diseases. The potential of lipidomics for the medical, pharmaceutical and nutritional fields is extensive, including the identification of pathological and physiological processes, understanding pathogenesis and disease progression, identification of novel diagnostic biomarkers, prediction and evaluation of treatment efficacy and searching for new therapeutic targets. However, the lipidome of biological samples is highly complex and diverse; hence, the lipidomics field has been constantly changing and evolving for the past 30 years. A typical lipidomics experiment usually involves the sample collection and storage, preparation (homogenization, extraction), analysis (chromatography, mass spectrometry, nuclear magnetic resonance), data processing (alignment, mass correction, filtering), identification, normalization, statistics and evaluation of biological implications. Much emphasis has been placed upon the biochemistry of lipids, but there is still a need to further develop better analytical approaches to study such complex mixtures of biological molecules.[4,9–14]

1.2. Lipids and their implications in health and disease

Although this work is mostly focused on the analysis of lipids, their characteristics and relationships are a vital part of lipidomics. Hence, the next sections provide an overview of lipid classes and their main biochemical reactions, followed by a review of methods for lipid analysis. Detailed discussions for selected lipid species and subclasses are provided with each chapter.

1.2.1. Lipids: overview

Lipids (from the Greek *lipos*, "fat") are roughly defined as hydrophobic biological molecules that are soluble in organic solvents but have low solubility in water. Examples of lipids that are easily recognizable include fats, oils, vitamins and hormones. However, this broad definition does not do justice to the enormous importance of lipids in metabolism: signalling, immune response, energy production and storage, cell structure, protein folding, trafficking, maintenance of homeostasis and others. Lipid bilayers are vital to compartmentalize living matter in cells and organelles, allowing the concentrations on each side of the membrane to be dramatically different, an essential characteristic for most processes in our body. More than half of the dry weight of the human brain is composed by lipids.[9,12,15] Unlike proteins, nucleic acids and polysaccharides, lipids exhibit a much greater structural variety. They are mostly composed by a combination of different hydrophobic hydrocarbon chains and a polar headgroup, although not all lipids follow the same rules. [9,15] Hence, they can be divided into classes that share structural similarities.

1.2.2. Lipid classification

A comprehensive classification system for lipids was proposed by Fahy *et al*. under the leadership of the International Lipid Classification and Nomenclature Committee.[16,17] Lipids were divided into eight major classes, each containing a number of subclasses: fatty acyls, glycerolipids, sphingolipids, sterols, glycerophospholipids, saccharolipids, polyketides and prenol lipids. Examples of molecular structures that can be found in each of the eight main categories are shown in Figure I-1. [4,16–18]

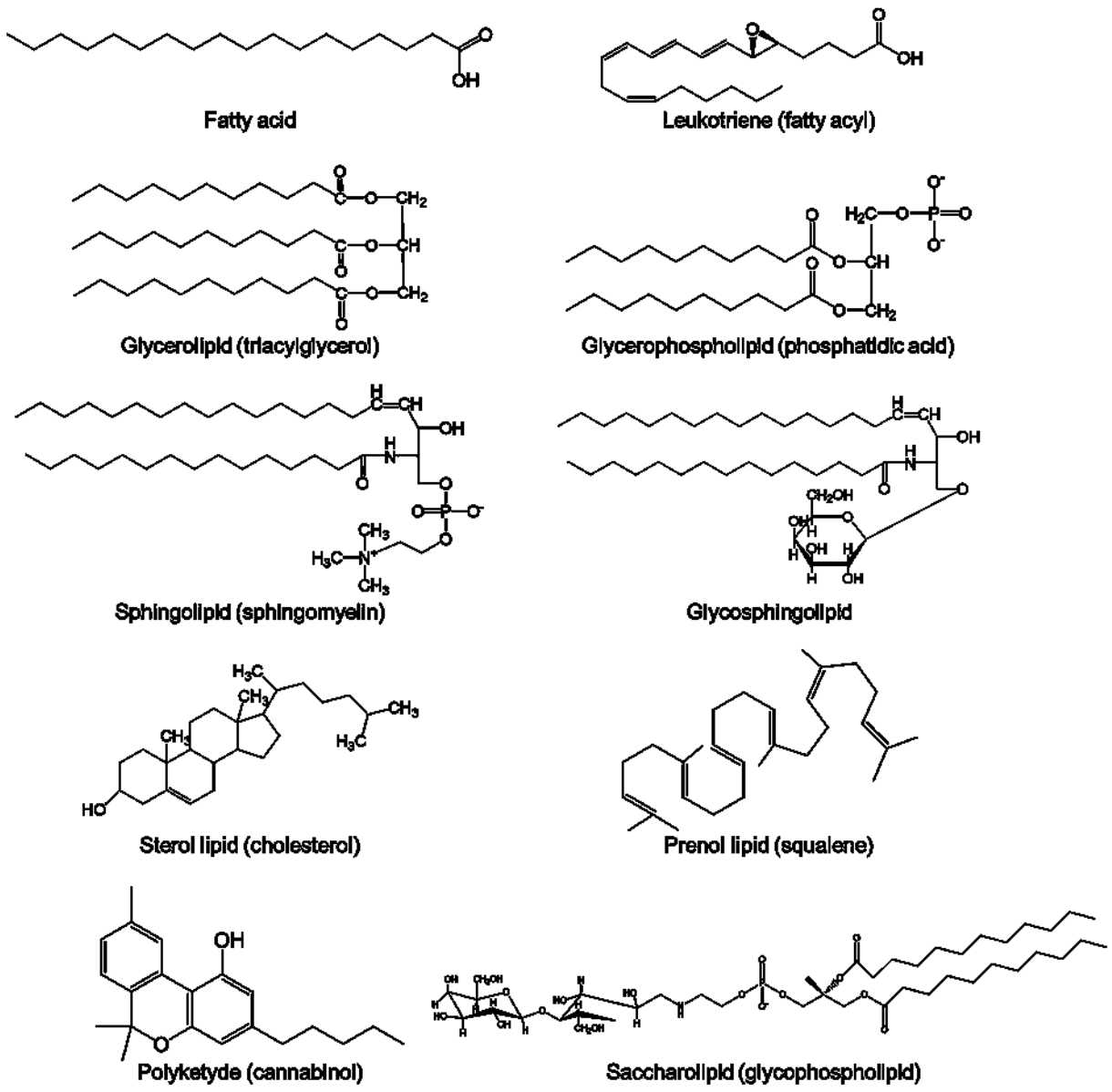

Figure I-1. Examples of structures found in each lipid class.

Fatty acids (FA) are usually composed by a hydrocarbon chain and a terminal carboxyl group (Figure I-1). The polar headgroup is typically deprotonated under physiological conditions, although the pKa of long-chain fatty acids increases with the number of carbons in the hydrophobic tail. They are the building blocks of most lipid classes, being commonly found esterified to other functional groups. The class is divided into 14 subclasses characterized by modifications to the

core structure, each one containing subgroups, *viz.* fatty acids and conjugates, octadecanoids, eicosanoids, docosanoids, fatty alcohols, fatty aldehydes, fatty esters, fatty amides, fatty nitriles, fatty ethers, hydrocarbons, oxygenated hydrocarbons, fatty acyl glycosides and others.[4,16–18] Eicosanoids (prostaglandins, thromboxanes, leukotrienes, lipoxins and endocannabinoids) are particularly important for metabolism as signalling molecules derived from the oxidation of arachidonic acid or other polyunsaturated fatty acids. They act in inflammatory processes, allergy, immune response, regulation of pregnancy and sexual function, pain, cell growth, blood pressure control, blood flow and others. The most common fatty acid residues have 14 to 22 carbons, but the hydrophobic tails can range from 2 to 36 carbons. Longer fatty acids are possible but rare. Most long-chain fatty acids found in nature contain one or more double bonds that are usually cis-oriented, which bends the lipid structure, contributing to membrane fluidity. Saturated fatty acids and trans-oriented unsaturated lipids are correlated with cardiovascular diseases due to their increased tendency to tightly aggregate in aqueous media, depositing on the walls of arteries. [4,9,15]

Glycerolipids are composed by a glycerol unit esterified to one, two or three fatty acyls (mono-, di- or triacylglycerols, respectively). The class also includes glyceroglycans, which contain a sugar residue attached to the glycerol backbone. Triacylglycerols (TG) are the most common glycerolipid and constitute the major energy source of our organism upon oxidation (Figure I-1). They are stored in the adipose tissue until extra energy is required, which also provides thermal insulation.[15]

Glycerophospholipids are essential components of cell membranes due to their characteristic structure composed by one or two hydrophobic fatty acyl residues (“hydrophobic tail”) esterified to a glycerol-3-phosphate backbone that may contain a polar headgroup (choline, inositol, ethanolamine or serine). The fatty acyl residue esterified to the *sn*-1 position of the

glycerol backbone is usually saturated, while the *sn*-2 group tends to be mono- or polyunsaturated. Species missing one of the fatty acyl residues due to the action of specific enzymes are known as lysoglycerophospholipids. Phosphatidic acid (PA, Figure I-1), also known as acylglycerophosphate, is the simplest of the glycerophospholipids, being the main source for the biosynthesis of other subclasses. The neutral acylglycerophosphocholines (also known as phosphatidylcholines, PC) and ethanolamines (PE) are the major components of cell membranes. The class also includes cardiolipins (CL, diphosphatidylglycerols) and plasmalogens (alkenyl-ether glycerophosphates, denoted by the prefix "O-" or suffix "e" in the nomenclature of lipid species).[4,15–17,19,20]

Sphingolipids are neutral derivatives of sphingoid bases such as sphingosine, a C_{18}-amino alcohol synthesized from serine and palmitoyl-CoA. They are typically enriched in neural tissue and can also be found in the outer layer of membranes, where they may form microdomains known as lipid rafts and act in signal transduction, cell recognition, immune response and stabilization. Ceramides (Cer), formed by an N-acylated sphingoid base, are one of the most abundant types of sphingolipids, but they can be modified to contain additional headgroups, *e.g.* sphingomyelins (SM, ceramides containing a phosphocholine or phosphoethanolamine unit - Figure I-1), cerebrosides (HexCer, hexosylceramides containing a sugar residue – glucose, galactose, N-acetylglucosamine and others, Figure I-1), sulfatides (Sulf, anionic sulfated sugar residue) and gangliosides (oligosaccharide group containing a sialic acid unit). Sphingomyelin and ceramides can regulate the activity of protein kinases and phosphatases, while gangliosides act as specific receptors for hormones and other molecules on cell surfaces.[4,15,20]

Sterol lipids are composed by a rigid structure derived from cyclopentanoperhydrophenanthrene, made of three 6-membered and one 5-membered fused,

nonplanar rings (Figure I-1). The class is divided into 6 subclasses, *viz.* sterols (cholesterol and derivatives, cholesteryl esters, ergosterols, stigmasterols, etc.), steroids (estrogens, androgens, glucocorticoids, mineralocorticoids, progestogins), secosteroids (vitamin D and derivatives), bile acids, steroid conjugates (glucoronides, sulfates, glycine and taurine conjugates), and others.[16–18,20] Cholesterol is the most commonly found sterol in animals and a major component of membranes (Figure I-1). Although it has been previously deemed as a nutritional villain, the rigid structure of cholesterol grants an essential ordering effect upon membranes, ensuring higher stability and control of permeability towards other molecules. Cholesterol also acts as a precursor of most hormones, including glucocorticoids (control of inflammatory reaction, response to stress and regulation of carbohydrate, protein and lipid metabolism, *e.g.* cortisol), mineralocorticoids (control of salt, water and ionic balance, *e.g.* aldosterone), sexual hormones (control of sexual function and phenotype, menstrual cycle and pregnancy, *e.g.* testosterone, estradiol and progesterone), bile acids (absorption of dietary lipids in the intestine) and the various forms of vitamin D. Cholesterol can also be esterified with fatty acyls to form the highly hydrophobic cholesteryl esters.[4,15]

The combination of two or more isoprene units (2-methyl-1,3-butadiene) compose the prenol lipids, also known as terpenes, which include carotenoids, quinones, hydroquinones, vitamin A, vitamin E, vitamin K2, ubiquinone (coenzyme Q) and others. Squalene and lanosterol, two C_{30}-triterpenes, are the precursors of cholesterol and other steroids (Figure I-1).[4,15]

Polyketides and saccharolipids are usually not included in untargeted lipidomics studies due to their different characteristics (higher polarity and solubility in water) when compared to other lipid classes. Polyketides are lipids synthesized from propionyl-coenzyme A (CoA) and methylmalonyl-CoA (Figure I-1). They include important medications, such as tetracyclines, anti-

cancer agents and anti-parasitic drugs, although some are potent toxins, *e.g.* aflatoxin B. Last, saccharolipids are glycerophospholipids that contain a sugar moiety attached to the polar headgroup (Figure I-1).[4,15]

1.2.3. Lipid membranes

Lipid bilayers form the boundaries of cells and organelles, acting as barriers to polar molecules. Membranes not only compartmentalize biological matter, but also control many biological functions, regulate the transport of molecules and ions, control intracellular communication and modulate gene expression. Although lipids are the main component of membranes, they may contain all types of macromolecules, except for nucleic acids. Lipid membranes are disordered, fluid bilayers kept together by a delicate balance between weak intermolecular forces. The amphiphilic characteristic of most lipids allows their natural aggregation in aqueous medium due to the hydrophobic effect, forming the bilayers that constitute most structures of our organisms. When hydrophobic molecules are put in an aqueous medium, the hydrogen bond network of water molecules need to be reorganized around them, which decreases the entropy of the system. Hence, lipids tend to aggregate to minimize the contact between their hydrocarbon tails and water, preserving the entropy. The self-assembled arrays of lipids, proteins and other molecules are mostly governed by hydrophobic dipole interactions between the hydrocarbon tails. The polar headgroups of lipids are oriented towards the surface (aqueous media), where they may form ionic and electrostatic interactions, whereas the hydrophobic tails aggregate to minimizes the contact with water. The hydrophobic core acts as a barrier to ions and polar molecules but can interact with hydrophobic compounds. Polar molecules and ions are transported through the bilayers by the action of protein channels or carriers.[4,9,15]

The two sides of biological membranes are not equal, granting transport of molecules in specific directions and allowing interaction with signalling molecules, hormones and reactions on the outer layer, typically enriched with sphingolipids, phosphatidylcholine and cholesterol. Glycoproteins and glycolipids are usually oriented with the carbohydrate group facing the exterior side of the membrane, where they can act as anchors or signalling receptors. Furthermore, enzymes that cleave the polar headgroup or remove fatty acyl residues can modify the shape and function of membranes. Phospholipases, *i.e*., enzymes that catalyze the cleavage of a fatty acyl residue of glycerophospholipids, are usually found outside of cells and cannot cross the membranes; hence, lipids on the external leaflet are more susceptible to hydrolysis. Flipases and translocases promote specific lipids to change membrane sides by facilitated diffusion (no energy consumption) or active transport with adenosine triphosphate (ATP) hydrolysis.[9,15]

The structure of membranes is based on a subtle equilibrium between ionic forces, hydrogen bonds, dipole interactions and weak dispersion forces. Glycerophospholipids are particularly prone to form bilayers due to their pair of fatty acyl residues and polar headgroups. However, all polar lipids naturally aggregate in water. Proteins and other lipids can induce changes in the membrane structure and permeability. For example, the stiff steroid core of cholesterol requires ordering of the fatty acyl groups of the glycerophospholipids that compose most membranes, which increases their thickness and reduces the fluidity. The hydrophilic hydroxyl group of cholesterol tends to stay in the aqueous interface, while the steroid core requires conformational ordering of fatty acyl residues for stronger interactions, resulting in a bilayer that is thicker and tighter, but still fluid. Lipid rafts are membrane regions with elevated proportions of cholesterol and sphingolipids that result in higher ordering and organization. Rafts are involved in inter-cell communication, trafficking, cell surface adhesion, signalling and motility. [49,15]

1.2.4. Lipid biosynthesis and metabolism

Most lipids in the human body come from diet, but selected classes are biosynthesized within the cells. The detailed mechanisms of biosynthesis and metabolism of lipids are highly diverse and beyond the scope of this work. However, the basic principles are required to understand the implications of lipidomic analysis of biologic samples. The following sections briefly describe the mechanisms of lipid synthesis and metabolization. Enzymatic modifications of lipids can further alter their function in the organism.[4]

1.2.4.1.Digestion and transport

Most lipids acquired through diet are the highly hydrophobic triacylglycerols. Lipases secreted by the pancreas catalyze the hydrolysis of fatty acyl residues from triacylglycerols and other lipids in the lipid-water interface inside the small intestine. Bile acids are cholesterol derivatives synthesized by the liver and stored in the gallbladder that act as “detergents” when secreted in the small intestine during digestion. The bile acids emulsify the dietary lipids into micelles, allowing their transport across the intestinal walls. The dietary lipid products are then reconverted to triacylglycerols and packed with dietary cholesterol into lipoprotein particles known as chylomicrons, which are carried by the bloodstream to other organs and tissues. Chylomicrons transport dietary triacylglycerols and digestion products to muscle and adipose tissue, as well as dietary cholesterol to the liver, through the lymphatic and circulatory systems. Endogenous triacylglycerols and cholesterol synthesized in the liver are transported by similar particles known as very low-density lipoproteins (VLDL). Both types of particles can adhere to binding sites on the walls of capillaries in skeletal muscle and adipocytes, where they are

hydrolyzed by lipoprotein lipase. The resulting mono- and diacylglycerols and fatty acids are then taken up by the tissues and oxidized to produce energy or modified for synthesis of other lipids.[15,21]

1.2.4.2. Energy production

Fatty acids are released from triacylglycerols stored in adipose tissue by hormone-sensitive lipases in the cytoplasm, being further transported through the bloodstream to other organs as albumin complexes. Hormone-sensitive triacylglycerol lipase is regulated by phosphorylation in response to levels of cyclic adenosine monophosphate (cAMP) levels , which are in turn controlled by glucagon, epinephrine and norepinephrine. Alternatively, insulin decreases cAMP levels, inactivating hormone-sensitive lipase when there is no demand for energy. The ratio between glucagon and insulin determines the rate of fatty acid metabolism. The glycerol released by the action of lipases upon triacylglycerols enters the glycolysis pathway, while the fatty acid molecules undergo β-oxidation in the mitochondria or peroxisome to produce energy. Peroxisomal oxidation in animals shortens the long-chain fatty acids, which are then further degraded in the mitochondria. However, oxidation of fatty acids in plants fully happen in the peroxisome and glyoxissome. [15,21]

The β-oxidation of fatty acids is a progressive oxidation of 2-carbon units, starting at the carbon atom in the β position to the carboxyl headgroup. First, fatty acids are activated in the cytosol by an ATP-dependent reaction catalyzed by acyl-CoA synthetases to form a thioester bond with coenzyme A. However, acyl-CoA species cannot cross the mitochondrial membranes. The fatty acids are then transported into the mitochondria in the form of acyl-carnitines by the action of carnitine acyltransferases. Once inside the mitochondria, the acyl-carnitine is reconverted to acyl-CoA. The β-oxidation of fatty acids occurs in a series of four enzyme-dependent reactions that includes two dehydrogenation steps with FAD and NAD^+ as hydrogen receptors. Each round

of β-oxidation results in a fatty acyl-CoA unit shorter in two carbon atoms, an acetyl-CoA molecule and ATP. The complete oxidation of fatty acids with even number of carbons results in acetyl-CoA, whereas odd-numbered fatty acids lead to acetyl and propionyl-CoA. The generated acetyl-CoA molecules are oxidized by the citric acid cycle and oxidative phosphorylation to produce ATP. The complete oxidation of palmitic acid, a C_{16} fatty acid, can generate up to 129 molecules of ATP. [15,21]

1.2.4.3.Lipogenesis and metabolism

Lipids are synthesized by two main pathways. The condensation of acetyl-CoA and malonyl-CoA generates fatty acyls, glycerolipids and glycerophospholipids, while the condensation of branched-chain five-carbon pyrophosphates with a carbocation intermediate leads to prenol lipids and sterols. The biosynthesis of fatty acids is the opposite process of β-oxidation and occurs in the cytosol of adipocytes and hepatocytes. Acetyl-CoA molecules are created in the mitochondria by oxidative decarboxylation of pyruvate or oxidation of fatty acids. Then, the molecules are transported to the cytosol by the ATP-driven tricarboxylate transport system in the form of citrate. Fatty acid synthesis starts with the activation of acetyl-CoA as malonyl-CoA by the action of biotin-dependent acetyl-CoA carboxylase. The acyl groups of acetyl-CoA are then anchored to acyl-carrier proteins (ACP) through a thioester bond, followed by a series of ATP-dependent reactions that elongates the acyl group by two carbons via condensation with malonyl-ACP. The cycle is sequentially repeated to elongate the fatty acid, usually culminating in palmitic acid (C_{16}). Elongases and desaturases in the mitochondria and endoplasmic reticulum further modify the fatty acid to convert it into long-chain and unsaturated molecules. [15,19,21] Polyunsaturated fatty acids, fatty acid esters and amides are substrates to oxygenase and

cyclooxygenase enzymes, which can introduce an oxygen atom in the hydrocarbon tail, an essential mechanism for the synthesis of prostaglandins and thromboxanes. Lipoxygenases, on the other hand, produce hydroperoxides such as leukotrienes, important mediators in inflammatory response and allergies. Non-steroidal anti-inflammatory drugs can inhibit the action of oxygenase enzymes to reduce inflammatory processes and control immune response. Polyunsaturated fatty acids can also undergo spontaneous radical-induced autoxidation, a process usually associated with oxidative stress. [19]

Triacylglycerols are obtained by further enzymatic reactions in the endoplasmic reticulum, mitochondria or peroxisome, where fatty acyl-CoAs are condensed with glycerol-3-phosphate. The resulting lysophosphatidic acid can be converted into a phosphatidic acid by acylation, which is then further modified to di- or triacylglycerols by the action of phosphatidic acid phosphatases and acyl transferases. The obtained phosphatidic acids and diacylglycerols are also the precursors of other glycerophospholipids, *viz*. phosphatidylcholines, ethanolamines, inositols and glycerols, as well as cardiolipins. Phospholipases and lipid kinases are two families of enzymes that modify glycerophospholipids to alter their function in the organism. The enzymes are involved is almost all pathophysiological processes, including inflammation, oxidative stress and infections. Sphingolipids are mostly synthesized from N-acyl-sphingosine in the endoplasmic reticulum, which is obtained from condensation of palmitoyl-CoA and serine. Cholesterol is synthesized in the liver by the condensation of six isoprene units, resulting in squalene. The 30-carbon linear lipid is further cyclized with the aid of oxygen and oxidized to lanosterol, which is finally converted into cholesterol. The molecule can be modified to a bile acid or esterified with a fatty acyl residue to form cholesteryl esters. [15,19,21]

1.2.5. Cholesterol: villain or hero?

Cholesterol is an essential component of membranes and a precursor of vital molecules for the organism. It can be obtained through diet or produced and metabolized in the liver. Cholesterol is transported to peripheral tissues by the blood stream as low-density lipoproteins (LDL), *i.e.,* packages that contain cholesteryl esters (esterified cholesterol) and triacylglycerols wrapped in a layer of glycerophospholipids and proteins. LDL receptors on the surface of cells enable the permeation of cholesterol and triacylglycerols through receptor-mediated endocytosis. The lipids are then employed for membrane regulation, as hormones and for other vital functions. High-density lipoproteins (HDL), in turn, transport cholesterol from peripheral tissue and organs back to the liver for excretion or modifications. Cholesterol is either oxidized to bile salts or modified to hormones and other vital molecules in the liver. Once oxidized, bile salts are transported to the gallbladder and intestines, where the cholesterol derivatives are reabsorbed or excreted.[9,15]

Cells can regulate the levels of cholesterol, LDL and other lipids through surface receptors. The amount of cholesterol in the blood stream depends on the partition between LDL and HDL. When the levels of cholesterol in the cells are high, the transcription of LDL-receptors is suppressed, as well as the biosynthesis of cholesterol. The excess is transported to the liver for processing via HDL particles. If the process is not well-regulated, cholesterol will accumulate in the blood stream, where it can be oxidized and further processed by macrophages, resulting in foam cells. The fat-laden macrophages accumulate on the walls of blood vessels, forming plaques (atherosclerosis) that may block coronary arteries, causing heart attacks and strokes.[9,15]

1.2.6. Lipids and diseases

The previous sections illustrate the vast importance of lipids for the metabolism. Consequently, the hydrophobic molecules are also involved in most pathogenic and physiological processes in our organism. The relationship between high-fat diets and heart conditions is well recognized, but lipid dysregulation is also related to immune response against pathogens[2,22–29], neurological diseases[30–39], respiratory conditions[2,40–44], oxidative stress[45–51] and inflammatory processes[41,52–58]. Yet, the regulation of lipids in the organism is not yet well characterized. Although general relationships between the major lipid classes are known, the pathways involving the minor subclasses and lipid species remain to be determined.[59] Hence, the study of the lipid composition of different biological samples, as well as their relationships and alterations in response to physiological and pathological processes, is a highly attractive field for research and the main objective of this work.

1.3. Analytical methods for lipidomics

Over the past two decades, many technical advances have been made regarding methods for the analysis of biological molecules. Although much focus has been put on genomics, proteomics and metabolomics, there is still a need for reliable analytical approaches that include sample preparation, analysis, data processing, normalization and identification of lipids. Lipidomic profiling can be achieved through targeted, untargeted or shotgun approaches. The experimental design selected for an application depends on its objectives, expected outcomes and nature of the samples. Targeted lipidomics focuses in one or a few selected lipid species or subclasses, while untargeted techniques attempt to profile the full composition of a sample that may contain thousands of different lipid species. Targeted lipidomics is often used when the molecules involved

in a particular biological process are previously known or suspected. It provides high selectivity and sensitivity, being appropriate for applications that require absolute quantification, such as biomarker validation. The low complexity of results and simple data processing when compared to untargeted and shotgun approaches is attractive, although the conclusions and implications that can be obtained are limited. Untargeted lipidomics, on the other hand, can detect thousands of molecules in one unique analytical procedure, generating extremely complex datasets with a high amount of information. It is usually employed for biomarker discovery and pathway analysis, along with relative quantification. Shotgun lipidomics, *i.e.*, the direct injection of organic extracts into a mass spectrometer without prior separation, is used to identify patterns of changes in response to physiological or pathological processes. Intra-source separation of lipid subclasses can be achieved by changing the ionization polarity during ESI and the composition of the sample media. Shotgun lipidomics usually provides reliable results for lipids with moderate concentrations, which can be analyzed within minutes. The approach is a quick and efficient method to obtain a fingerprint of the lipid composition of a sample, but suffers from limited dynamic range, strong ion suppression, aggregation effects and overlap of isomeric and isobaric lipids.[10,11,60]

This work focuses on untargeted lipidomics of biological samples to study pathological and physiological processes. Untargeted applications are usually the starting point for the investigation of unknown biological processes, as a great number of molecules are analyzed in a single experiment. It allows the relative quantification and identification of hundreds or thousands of compounds, in addition to the capability to find unexpected molecules and trends. Usually, untargeted lipidomics does not require absolute quantifications, but only relative comparisons between two or more conditions (*e.g.*, healthy *versus* diseased). However, the complexity of the

obtained data requires specialized techniques for sample preparation, analysis and data processing.[10]

A typical lipidomics experiment starts with extraction of the lipid fraction of a biological sample, followed by separation, analysis, detection, data processing, statistics and data reporting. Within each step, an array of different options is available; hence, the experiments must be carefully planned and adapted to the application to ensure reliable and meaningful outcomes. The sample preparation includes the addition of internal standards, homogenization, isolation and purification of lipids. The separation of the lipidome can be achieved through liquid (LC) and gas chromatography (GC), although direct infusion without a prior separation technique is also employed for shotgun lipidomics. The analysis is usually performed with mass spectrometry (MS) allied to electrospray (ESI), nano-ESI or electron impact ionization, but nuclear magnetic resonance (NMR) has also been described. Last, the processing of the enormous amount of data generated by untargeted lipidomics requires mass recalibration (for MS-based applications), peak picking, alignment, identification, normalization and statistical analysis, followed by bioinformatics and pathway analysis. The structural characterization and identification of lipids, as well as the determination of absolute or relative concentrations, remain the biggest challenges for lipidomics. Derivatization methods for specific lipid classes have been previously proposed to improve analytical separation, detection and identification, but there is still no universal derivatization approach that allows comprehensive, untargeted lipidomics due to the variety of functional groups that can be found in lipids. The following sections describe the most common techniques employed for each step, with special emphasis on the comprehensive, untargeted methods employed in this work.[12,59,61]

1.3.1. Sample preparation and storage

Lipids are highly prone to oxidation, peroxidation, hydrolysis and other modifications, particularly when extracted and exposed to normal atmospheric conditions. Hydrolytic enzymes, proteases and phosphatases present in biological fluids and tissues can modify the relative abundance of lipid species after sample collection, *e.g.* an artificial increase in the levels of lysoglycerophospholipids or oxidized glycerophospholipids. Hence, the preservation of the lipid composition of samples is essential to any preparation methodology. It is highly recommended that raw biological samples and lipid extracts are immediately snap-frozen after collection and stored under low temperature and protected from light to minimize degradation and enzymatic activity (-20°C for short-term and -80°C for long-term storage). The use of techniques that require long-term exposure of lipid samples to room temperature, light, oxygen, multiple freeze-thaw cycles, heat or acidic/basic conditions is highly undesirable. Although antioxidants are sometimes suggested to preserve lipid samples, they may not be compatible with the analytical techniques or cause extra modifications and contamination in the samples. Hence, the sample preparation steps must be performed quickly and accurately. Lipid extracts can be further purged with inert gas (nitrogen, argon) for storage to prevent oxidation. It is not advisable to store dried or concentrated lipid extracts under normal atmospheric conditions due to the possibility of degradation.[4,12]

The preparation of samples may require a mechanical homogenization step for tissue or cells, such as sonication or a bead-beater. The hydrophobic nature of most lipids in aqueous biological samples or homogenates allow the use of liquid-liquid extraction with organic solvents to isolate the lipidome. The use of non-polar solvents disrupts membranes, lipid droplets and lipoprotein particles, allowing the isolation of the lipidome from the more polar compounds. Alternatively, solid-phase extraction methods, microwave-assisted extraction and supercritical

fluids have also been used for lipid isolation, particularly for targeted analysis. Ideally, extraction procedures should provide quantitative lipid isolates without major degradation, contamination or modification of the sample components. The most commonly used methods employ chloroform, methanol and water or salt solutions to promote a biphasic system, *e.g.* the popular Folch and Bligh & Dyer methods.[62–64] The organic layer is then separated, dried under nitrogen flow or using a SpeedVac (room temperature), and resuspended in an appropriate medium for analysis. The chloroform-based methods usually allow high recoveries of low to mid-polarity lipids, such as triacylglycerols, glycerophospholipids and sphingolipids, but have a lower performance for the most polar lipids, *e.g.* small-chain fatty acids.[4,6,14]

It is worth emphasizing that the use of non-polar organic solvents for extraction requires a cautious evaluation of glassware and plastic ware. Chloroform, dichloromethane, methanol and isopropanol, commonly used for lipidomics, can extract contaminants and plasticizers from vessels, caps and pipet tips. Although glassware is usually less prone to contamination, it has a high associated cost for single-use vessels and may promote selective retention of lipids in non-silanized glass walls. Disposable vessels with polytetrafluoroethylene (PTFE) or Teflon lining and fittings are usually indicated for lipidomics. Quality control through blank extractions (water instead of sample) to verify contamination, pooled samples and standards with each batch of sample extractions is essential to ensure reliable results. Chapter II includes an evaluation of contamination sources for liquid-liquid extraction of lipids and clear examples of the effects of employing unsuitable vessels for lipidomics. [4,6,14]

The Folch liquid-liquid extraction method was originally developed for the isolation of lipids from tissue samples over 50 years ago, although it is still recognized as one of the most reliable preparation methods for lipidomics.[62,63] It provides a high coverage of lipid classes through

a combination of chloroform and methanol in a 2:1 ratio, followed by a clean-up step with water or saline solution to a final proportion of 8:4:3 chloroform/ methanol/ water. The low polarity of chloroform promotes the selective extraction of hydrophobic compounds, while the combination with methanol controls the partition coefficient of intermediate-polarity compounds, disrupts hydrogen bonding between polar lipids and proteins, and promotes protein precipitation. The water wash step separates the undesirable polar compounds, such as proteins, sugars and amino acids. The adjustment of solvent ratios can increase the recovery of particular lipid classes. Nowadays, many researchers opt to substitute chloroform by dichloromethane due to lower carcinogenic risk, less restrictions to acquisition and reduced cost, while maintaining similar chemical and physical characteristics. The Folch method was employed for this work with dichloromethane and minor modifications for each application, as described in the Experimental section of all chapters.[4,12,62,63,65,66]

A limitation to chloroform or dichloromethane biphasic extraction methods for biological samples is that the organic layer is located under a solid protein disk. The separation of the organic fraction requires pipetting the solvent through the upper aqueous layer and the protein disk, a procedure that can be tricky if performed by unexperienced analysts, particularly for samples with a high protein content. An extraction method employing methyl-*tert*-butyl ether (MTBE), methanol and water was reported by Matyash *et al.* in 2008 and has become increasingly popular in the last decade.[67] MTBE promotes the organic phase to the upper layer of the biphasic system, facilitating the separation. The method provides similar recoveries and coverage of lipid subclasses to the Folch extraction, as well as comparable reproducibility, depending on the experience of the analyst. However, the procedure requires a long incubation period during which lipids can be

easily degraded or modified. A comparison between the Folch and the MTBE method is discussed in Chapter II. [12,67–70]

Quality control of sample extraction steps can ensure the reliability and stability of lipidomics. Usually, a pooled mixture composed by small aliquots of all samples is extracted along with each batch of samples. Alternatively, a sample with similar composition can be used if the available volumes of biological fluids are too small for pooling. Reference standards are also highly recommended to control retention time and mass shifts. Blank extracts (extraction of water or solvent instead of sample) are useful to control contamination and background noise, particularly for methods that employ organic solvents. Although quality control steps are rarely discussed or reported, they are essential for a correct evaluation of lipidomics results. Alterations in lipid concentrations cannot be attributed to physiological or pathological processes if the possibility of analytical interference is not controlled.[14]

1.3.2. Separation

Although the lipidome can be investigated without the use of separation techniques (shotgun lipidomics, *i.e.* direct injections of the organic extract), prior chromatography can reduce the complexity of the sample and eliminate contaminants, improving the quality of the obtained data. Separation approaches minimize aggregation effects, reduce ion suppression and increase the dynamic range for detection. Chromatography is a powerful technique that further allows the detection of lipids with very small concentrations, which is hardly ever achieved for shotgun lipidomics due to the complexity and high ion suppression of extracts. High-performance liquid chromatography is the most common approach for untargeted lipidomics, owing to high resolution, versatility and possibility of automation. Nanoscale-liquid chromatography (nanoLC) can further

increase the sensitivity and allow untargeted lipidomics of droplets of biological fluids, as discussed in Chapter II. Reversed-phase conditions are employed to separate lipids by hydrophobicity, while hydrophilic-interaction chromatography (HILIC) promotes separation based on the polar headgroups. Gas chromatography has also been used for the analysis of small fatty acids, triacylglycerols and sterols, but most lipids require lengthy derivatization procedures to increase their volatility and resistance to high temperatures.[134,12]

Reversed-phase chromatography is the most common approach for lipid separation. The mobile phases must have a high content of strong organic solvents to ensure elution of the hydrophobic lipids, but chloroform and dichloromethane are not often indicated due to the possibility of corrosion and damage to metal surfaces and polymeric fittings in the instruments, as well as to columns. The strong organic solvents may also cause precipitation of buffers, additives and sample components inside columns and instruments. The analysis of lipids demand different and sometimes extreme analytical conditions, but the composition of mobile phases, injection volumes and handling of LC instruments require training and attention as overloading, peak distortions, co-elution, aggregation and damage to expensive instruments and columns are often seen and may impact research results.[59]

The separation of lipids allows an additional criterion to increase the accuracy of identifications, as lipids with similar structures will elute within the same retention time range. Hence, lipid standards may be used to determine expected retentions to further confirm the identification of endogenous species. This principle was used in this work as a filtering step for isomeric and isobaric identifications of lipids, as described in Chapter II.[10,12]

1.3.3. Analysis-

The complexity of the lipidome, associated with concentrations that can vary in 6-8 orders of magnitude, requires the use of high sensitivity, high resolution techniques with extensive dynamic ranges that can also provide structural information for identification. Hence, mass spectrometry (MS) is the ideal choice for untargeted lipidomics, although nuclear magnetic resonance (NMR) has also been described for the study of the lipidome. NMR is a non-destructive and fast technique, but it has limited sensitivity and can be challenging for highly complex samples. MS-based techniques offer the required broad coverage, high sensitivity and extensive dynamic range, in addition to the possibility of structural information, imaging and coupling to liquid and gas chromatography. Imaging techniques, *e.g*. MALDI, further provide the special localization of lipids at the cost of lower analytical coverage and reduced quantitation performance. Although triple quadrupole (QqQ) mass spectrometers are still constantly used due to the variety of scan modes, high sensitivity for selected molecules and lower cost, high-resolution instruments capable of detecting a wide range of masses with higher mass accuracy are typically employed for comprehensive lipidomics, including quadrupole ion trap, quadrupole time-of-flight (QToF) and orbitrap. The MS^n capability of ion trap instruments offers better identifications, but the low mass cut-off and sensitivity decay for multiple *tandem*-MS experiments hinder their widespread application for lipidomics. Orbitrap-MS, on the other hand, has excellent resolutions and mass accuracy, but the high cost and slower spectra acquisition rate required for full resolution are the main disadvantages. The combination between quadrupole and time-of-flight mass spectrometry provides excellent resolution, high mass accuracy and extensive mass range allied to *tandem*-MS capability, being highly appropriate for untargeted lipidomics. Recently, the use of ion mobility MS (IMS) for lipidomics has been on the rise as it adds an additional dimension of

separation for isomeric and isobaric compounds (drift time). However, IMS does not solve the classical issues of lipidomic analysis, *e.g.* ion suppression and overlap between lipids with very similar structures and masses. [6,10–12,14]

The combination of high-resolution mass spectrometry with electrospray ionization (ESI) and liquid chromatography (LC) is the current gold standard for lipidomics. However, the inherent amphipathic characteristic of most lipids allied to the complexity of biological samples grants intense ion suppression for ESI-MS. The ionization efficiency of a lipid depends upon the polar headgroup, the number of carbons in the hydrophobic tail and the presence of unsaturation or other functional groups. The intensity of lipid peaks will also be impacted the detector response and ion transmission, further complicating the quantitation of species. Hence, normalization approaches that correct ion suppression and detector response are essential for meaningful comparisons. Nevertheless, many lipid species are not easily protonated or deprotonated ($[M+H]^+$ and $[M-H]^-$ ions). Charged or zwitterionic lipids usually have higher ionization efficiencies, but non-polar lipids require derivatizations or adducts for MS analysis. The use of additives and modifiers in the mobile phase or sample medium is required to ensure adduct formation, enhancing the sensitivity of lipidomic analysis, as discussed in Chapter II. However, acidic mobile phases, commonly employed for LC-ESI-MS under positive ionization, are not advisable for lipidomics due to the possibility of cleavage of vinyl-ether bonds of plasmalogens and hydrolysis of lipid species.[4,12,13]

1.3.4. Data processing

The processing of the complex results obtained from untargeted lipidomics is an essential step to obtain high-quality data. The raw data must be corrected for mass errors, aligned and normalized to ensure that the biological and biochemical differences are kept, while analytical

effects that may affect the results are minimized. The correspondence determination, *i.e.*, alignment of signals across multiple samples that correspond to the same lipid, may be complicated by retention time drifts, mass variations and sample degradation over time. Quality control is essential to ensure reliable and meaningful results, as well as differentiation between signals from analytes and random noise. However, there is no consensus in the scientific community regarding the best tools for processing the highly complex lipidomics datasets. A few software packages have been previously developed for metabolomics and may be adapted for lipidomics data (MZmine, XCMS, MetaboScape). Specialized software is also available (Lipid Data Analyzer, LipidSearch, MSDial LipidBlast, LipidXplorer), but their use still requires adaptations to each analytical platform and application.[71–74] This work employed MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA) for data alignment, correction, filtering and MS/MS identification, a software developed for metabolomics that can be further adapted for lipidomics. Although the software is a powerful tool for data alignment, it does not fulfill all needs for lipidomics. The normalization, putative identification, standardization of nomenclature and data checking steps, further discussed in the next sections, still mostly depend on manual work. [4,10,11,13]

1.3.5. Identification and nomenclature

Lipids can have a variety of different structures that can be extremely different, but also very similar. Many isomeric lipids differ only in the position of an unsaturation or a functional group. However, lipid structures reflect their class or subclass characteristics and can generate specific fragments when *tandem*-MS with collision-induced dissociation (CID) is employed. The characteristic product ions and neutral losses obtained by CID-MS/MS of lipids usually include the headgroup, fatty acyl residues and other functional groups, depending on the structure of the

lipid and acquisition parameters. Targeted lipidomics can be very effective for positive MS/MS identification of lipids as a small number of species are selected for fragmentation. However, comprehensive, untargeted lipidomic approaches are limited by the number of ions that can be fragmented within an injection. Data-dependent analysis is based on the selection of a predetermined number of precursor ions from a survey scan, followed by CID and acquisition of the obtained fragments. The method allows the acquisition of thousands of MS/MS spectra and has the potential for identifying a high number of unknown lipids. Precursor ions are sequentially screened during the chromatographic elution of each peak. Although chromatographic separation reduces the complexity of the data, tens of lipids can still be detected within the small retention time range of a peak; hence, the most intense ions are usually selected for fragmentation, whereas the less intense species are not identified by *tandem*-MS. The obtained fragmentation spectra are then compared to spectral databases for positive identifications. Usually, the precursor ions are first scanned within a set mass-to-charge (m/z) error. Then, a similarity score (MS/MS score, between 0 and 1000) is calculated to assess the resemblance between the measured and the reference spectra. A high MS/MS score is related to better identifications. An isotope pattern score, known as mSigma, can also be used to verify the similarity between the measured and the expected isotope pattern for each ion as an extra layer of identification. Low mSigma values represent better isotope pattern matches. The three parameters, *viz.* m/z error, MS/MS score and mSigma, were employed to select MS/MS identifications for all applications shown in this work. Unfortunately, most MS/MS approaches cannot determine the position of unsaturation, which requires target methods with cleavage or derivatization of double bonds. [6,12,14]

A major limitation for *tandem*-MS identification of lipids is the relatively small number of entries in libraries and databases that are currently available. The high complexity of lipid extracts

containing thousands of different molecules with variable concentrations also restrains the number of precursors that can be fragmented in an analytical experiment. Hence, many species cannot be identified by *tandem*-MS. Putative identification by mass-match is a less accurate but still valid option to determine lipid species. It is worth emphasizing that lipids can have multiple isomers and isobars even when accurate masses are defined, meaning that multiple putative identifications are possible. Also, the composition of aliphatic chains and positions of unsaturation and branching cannot be specified. The term "annotation" has been mentioned as more appropriate for lipids identified by exact mass-match; however, this work adopts the more common "putative identification" to prevent confusion. The level of putative identifications is typically restricted to the molecular lipid (*i.e.*, the headgroup or subclass, total number of carbons in fatty acyl residues and total number of unsaturation) rather than the structural identification (positions and stereochemical configuration of fatty acyl residues, unsaturation and other functional groups, such as branching, cyclization and hydroxyls).[12,13,59]

One of the biggest challenges for untargeted lipidomics is the lack of a consensus for nomenclature of lipid species, classes and subclasses. Although several lipid databases can be found in the literature and online, most of them develops their own nomenclature rules, which further complicates the reporting and discussion of lipidomic results. Several initiatives have attempted to standardize nomenclatures and techniques for lipidomics, including LipidMaps (https://www.lipidmaps.org), the Lipidomics Standards Initiative (https://lipidomics-standards-initiative.org), the European Lipidomics Initiative, the Lipid Mass Consortium (USA) and the Lipid Bank (Japan).[10,12,18,20,75,76] The LipidMaps database is a great resource for researchers focusing on lipidomics that was adopted for all chapters of this work. The database is a compendium of 44,701 biologically relevant lipids (July/2020) and computationally generated

molecules that can be used for putative mass-match searches and general information regarding lipid classes, subclasses and species. The lipid nomenclature and subclass distribution of LipidMaps is often employed for reporting lipidomics data. The compendium partnered with the Lipidomics Standards Initiative on 2018, which led to an update on the nomenclature employed for the mass search tool on 2019. However, the database has not yet been fully modified (July/2020), leading to confusion regarding lipid names. Recently, a tool for nomenclature standardization was introduced to LipidMaps to convert lipid names from different sources into a standardized version, although further improvements are still required. For example, a search for the saturated fatty acid with 16 carbons FA(16:0) returns the common name Palmitic Acid; however, the Lipidomics Standards Initiative suggests a standardized nomenclature of FA 16:0.[20,76] The identity of this simple example could easily be confirmed, but differences in the nomenclature of the thousands of complex lipids that can be detected in a lipidomics experiment, containing a multitude of modifications, can be highly confusing, particularly for inexperienced personnel.[12,20,61,75,76]

Unfortunately, although this problem has been constantly reported in the lipidomics community, there is still no consensus regarding the processing, identification and reporting of lipidomics data. This can be observed within the different chapters of this work, developed between 2016 and 2020. Although the same databases were employed for all chapters (LipidMaps[16–18,75,77], MSDial LipidBlast[73,78], Human Metabolome Database[79–81] and MassBank of North America - MoNA[82,83]) and an effort was made to uniformize the nomenclature, there are differences that were incorporated as the databases were updated. The nomenclature adopted for all chapters is based on the abbreviated lipid subclass, followed by the number of carbons in fatty acyl residues, number of unsaturation and other functional groups.[20] For example, the 16-carbon

saturated fatty acid commonly known as palmitic acid is named FA 16:0 with minor differences amongst chapters, while the same species with one double bond is FA 16:1. A phosphatidylcholine with two palmitate residues is named PC 16:0_16:0 if the MS/MS fragmentation spectra was acquired and matched to a database; or PC 32:0 for putative identification, as the identity of individual fatty acyl residues cannot be confirmed without fragmentation.[76]

Even when an accurate identification of the detected lipids is achieved, another challenge remains: lipid pathways and networks are yet to be fully explored, as discussed in Chapter V. The integration of lipidomics with other *omic* sciences may help in this regard, as well as the amplification of high-quality lipidomic studies.

1.3.6. Ion suppression and normalization approaches

LC-MS-based lipidomics approaches have ion abundances that are affected by several factors, including biological differences, sample preparation, the strong ion suppression effect that is often observed for complex lipid mixtures, different ionization efficiencies, mass-dependent ion transmission within the MS instrument and detector response. Amphipathic molecules tend to stay in the air/water interface during electrospray ionization as the polar headgroup is solvated, but the hydrophobic tails are directed to the air to preserve the entropy of the system. Molecules that are on the surface of droplets are preferentially ionized, resulting in higher intensities, whereas other molecules may be suppressed. The polar headgroup, fatty acid chain length, modifications and degree of unsaturation may affect the ionization efficiency of different lipids. Furthermore, the instrument response usually decreases with increasing fatty acyl chain length. Longer fatty acyl residues have more hydrophobic intermolecular interactions, protecting the polar headgroup, which can decrease the ionization efficiency. The effect is even more pronounced for high

concentrations as hydrophobic lipids tend to aggregate. Appropriate normalization of the obtained data with multiple internal standards, added to the sample as soon as possible during the preparation step, is essential for reliable and meaningful results. Targeted or untargeted lipidomics should not be performed without a valid internal normalization approach to remove experimental bias while keeping the biological differences. The relative intensities or peak areas of lipid species are not directly related to their concentrations or abundances in biological samples, but can be compared if appropriate normalization procedures are applied.[4,12–14,84,85]

External calibration methods cannot account for matrix effects; hence, they are not indicated for complex lipid extracts. Internal standards are expected to suffer similar effects as endogenous lipids with analogous characteristics and concentrations, hence allowing the correction of differences that may affect the ion count. Isotope-labeled internal standards are recommended, although lipids containing odd-chain fatty acyl residues have been previously employed as internal standards due to low natural abundancy in biological samples. Unfortunately, that approach resulted in under-reporting the importance of odd-chain fatty acids, which can be detected by high-sensitivity methods and are associated with pathogenic processes, as discussed in more details in Chapter VIII. Furthermore, although they have reduced concentrations, odd-chain fatty acids are naturally found in biological samples, reducing the efficacy of the normalization.[12]

The absolute quantification of lipids using mass spectrometry requires the use of multiple stable isotopologues for each lipid subclass to account for different ionization efficiencies, ion suppression, ion transmission and detector response. Ideally, one isotope-labeled standard should be used for each quantified molecule to ensure similar structures and properties. A stable isotopologue added at the beginning of sample preparation will suffer similar effects as the

molecule of interest, ensuring adequate normalization and quantification. Although that approach can be achieved for targeted analysis of a small number of lipids, an untargeted lipidomics experiment can provide the detection of thousands of different compounds. The acquisition of so many standards would be not only extremely expensive, but also generate further issues as to the addition of so many exogenous compounds to an already very complex biological sample. Nevertheless, the commercial availability of isotope-labeled lipid standards is still limited. Reliable normalization for relative quantification approaches can be achieved with one stable isotopologue standard for each polar lipid subclass if low concentrations can be maintained. Relative quantification, *i.e*., the comparison between the concentrations or intensities of species between two or more conditions without the need to determine their exact values, is commonly used for LC-MS biomarker discovery. The approach for lipid normalization using one representative isotope-labeled internal standard for each class or subclass is recommended and widely employed in the untargeted lipidomics field. Since the ionization efficiency is highly dependent upon the polar headgroup, individual lipid species within the same subclass have similar instrument response under low concentrations. Hence, all lipids that belong to the same subclass are normalized by one unique isotope-labeled internal standard that is structurally similar to the analytes and elutes within the same chromatographic retention time range. A commercially available mixture of 14 deuterated standards belonging to different lipid subclasses was used for normalization in this work (Splash Lipidomix Mass Spec Standard, Avanti Polar Lipids, Alabaster, AL, USA). After identification of the detected lipids and determination of their subclasses, each lipid was matched to one of the deuterated internal standards, according to subclass similarity and retention time. The intensity of the endogenous lipid was divided by the intensity of the matched deuterated internal standard to correct for differences that may happen during sample preparation

and analysis. It is worth emphasizing that the adopted strategy is widely recognized by the lipidomics community for relative comparisons between samples. Absolute quantification requires a higher number of isotope-labeled internal standards to normalize for different response factors and ensure accurate results, but it is not necessary for comparisons between similar samples if stable instrumental conditions and low sample concentrations can be maintained. [10,12,14]

1.3.7. Statistics applied to lipidomics

Given the complexity and high dimensionality of lipidomics data, statistical models are essential to find trends and identify lipids that are significantly altered between the studied conditions. A deep discussion of statistical models is beyond the scope of this work, but the basic principles are required for the correct evaluation of lipidomics data. The massive complexity of terabytes of multidimensional lipidomics and metabolomics data is prone to small data processing and statistical faults that are hardly detected, leading to false patterns and conclusions. We are looking for a few small changes within thousands of random molecules, which can be highly challenging. The reproducibility and validity of *omics* data has been questioned due to a lack of reliable statistics and validation. Understanding the statistical models to avoid false performance-enhancing practices is essential for high quality results.[86]

Most statistical models employed for lipidomics require a normal distribution and are heavily affected by data scaling. The application of statistical models and tests in skewed data yield invalid results with no real meaning. Hence, auto-scaling (*i.e.*, mean-centering with unit variance) and normalization to the median value, by the total ion count (summed intensities of all lipids employed for statistics) or quantile (ranked normalization to uniformize two distributions) is often required. Although normalization by internal standards has been previously discussed, the

extra normalization step before statistical analysis removes systematic variations that are not related to biological differences and may not be corrected by the internal standards, *e.g.* samples with different dilutions (*e.g.*, saliva from a person that drinks 2 L of water *versus* a person who drinks 500 mL of water per day) and batch effects. Total intensity and median normalization assume that the change in intensity due to sample dilution or batch effects is uniform across all peaks; hence, a fixed scaling factor can be used for correction. Normalization by total intensity is highly affected by outliers or very intense features with high variability, while median normalization is less dependent upon such factors. Quantile normalization is based on forcing the distributions of compared datasets to be the same. Briefly, the values of each feature are ranked within each sample. Then, the average of each feature is determined across all samples and used to replace the values of features with the same rank or quantile. Finally, the transformed values are replaced in the original order, resulting in the correction of technical variability and batch effects. Unfortunately, there is no "recipe" for the most appropriate handling of lipidomics data. Different normalization and scaling approaches will affect the obtained results and should be considered along with the data structure and question being evaluated.[11,85,87]

Principal Component Analysis (PCA) is useful to verify trends in the data, as well as clustering of experimental or injection replicates and quality control experiments. The statistical model reduces the dimensionality of the data by using orthogonal components, known as principal components. The first principal component explains as much of the original variance as possible, with decreasing portions for further components. Caution must be taken regarding data normalization and scaling, as PCA requires normal distribution and unit variance. If no data scaling is applied, the separation will be heavily affected by the most abundant lipids as they will have the highest absolute variation, causing misleading conclusions. The simple, non-supervised model is

not expected to result in a full separation between different groups (*e.g.* diseased *versus* healthy) in the score plot. PCA is a dimensionality reduction tool that may not be able to unravel the high complexity of the lipidome, particularly if the studied conditions only cause small changes for a few lipids in a sea of thousands of molecules. PCA can also be affected by noise, outliers and confounding factors. However, it is an excellent initial statistical tool to evaluate the performance of the chosen methods, data quality and initial trends. The clustering of quality control and sample replicates in PCA score plots are indications of the technical reproducibility and suitability of the employed methods.[11]

Supervised statistical models take the class of each sample into consideration for the separation. Partial Least Squares – Discriminant Analysis (PLS-DA) is a popular statistical model for classification of complex datasets. A data matrix containing intensities for each detected feature is correlated to a second matrix containing the classification of each sample (*e.g.*, healthy or diseased) through a linear regression. Caution must be taken regarding overfitting for datasets with a small number of samples and thousands of variables. In overfitted models, the observed separation may be due to modelling of random noise rather than meaningful variables. Thus, although an overfitted model is still a valid tool to evaluate the linear regression between the employed variables and the sample categories for the studied dataset, its performance may not be replicated for independent sample cohorts and result in false positives. The model must be validated by cross-validation and permutation tests to ensure reliable performance for biomarker discovery approaches. The parameters R^2 and Q^2 are employed to evaluate a PLS-DA model through cross-validation. R^2 is defined as the correlation coefficient between the predicted and real sample categories, *i.e.* high R^2 values are related to correct sample classification. Q^2 is a measure of the prediction ability of the developed model, *i.e.*, the average prediction error of a number of

cross-validations where one or a portion of the samples are retained to construct the model and further used to evaluate class prediction. There is no consensus for acceptable R^2 and Q^2 values, but high R^2 with low Q^2 indicate model overfitting. Permutation tests are performed by randomly assigning classes to different samples and calculating the performance of the resulting models to evaluate the significance of the original PLS-DA model. R^2, Q^2 and p values for permutation tests are provided for all PLS-DA models employed in this work. [88,89]

Random forest is a classification method that employs several randomized decision trees to predict the classification of samples. A random forest is a collection of a number of randomized regression trees. The combination of the decision from all trees gives the sample classification. It has a reliable performance for datasets in which the number of variables far exceeds the number of samples, as usually is the case for exploratory metabolomics and lipidomics. Each decision tree is prone to overfitting, but the combination of a high number of them results in a robust classification model. [90,91]

Univariate statistical methods also represent an important way of selecting significantly altered molecules, although the correlation between different compounds is disregarded. Analysis of Variance (ANOVA), t-tests and Volcano plots (fold change – FC *versus* p for Student's t-test) are often employed for *omics* applications. However, the raw p values may not be suitable to evaluate the significance of lipids in complex datasets. There is much confusion in the literature regarding the interpretation of p values. When it was first introduced in the 1920s by the statistician Ronald Fisher, it was meant as an additional method to verify if it was worth further investigating a variable.[92] The main idea was to provide a numeric measure to check if an observation was random by calculating the probability that results as extreme (or more) as the found values could be obtained if there was no difference between the compared groups. It was intended as a tool to

select targets for further experiments rather than as a unique evidence for scientific conclusions. Nowadays, it is often seen as an indication that the mean values for two compared conditions are "statistically" different, but that is an incorrect affirmation. The p value represents the area below the probability density curve of the t-test, *i.e.*, the probability of obtaining the observed or a more extreme result if the null hypothesis (equal means for two different conditions, H_0: $\mu_1 = \mu_2$) is true. The significance level (α) chosen to reject the null hypothesis is an arbitrary value that is historically chosen as 0.05. A correct interpretation of $p = 0.05$ would be that the probability of observing a difference between the means of the tested conditions as large or larger than the observed value by chance alone is of 1 in 20.[93] A low p value doesn't mean that the null hypothesis of equal means is true. The p value by itself does not support conclusions about the probabilities of the tested hypothesis and carries no information about the magnitude of the difference between the conditions. A small p value may have no practical significance if the numeric difference between the two compared groups is very small. However, it can be a useful tool for rejecting the null hypothesis of equal means if interpreted correctly and in combination with other meaningful parameters, such as fold-change. The combination of the p value with other methods, such as volcano plots and multivariate statistics, can increase the reliability of the conclusions. Furthermore, a p value can only be correctly interpreted if authors report the corresponding fold-changes, sample size, data exclusions, manipulations, scaling and modifications. Nevertheless, complex datasets with multiple hypothesis testing have increasing probability of erroneous conclusions, as the chance of a false-positive result increases when more than one variable is evaluated. A variety of multiple testing correction approaches can be used to reduce false positives when multiple features are assessed together. For this work, p values were adjusted for false-discovery rate, where the null hypothesis is rejected if the k^{th} ascending-ordered p value is larger

than $\alpha\, k/n$, where α is the significance level (usually 0.05), k is the rank of the p value in ascending order and n is the total number of features.[94–98]

The statistical treatment of *omics* data is still a highly debated field. It is unlikely that researchers will reach conclusions on the most appropriate way to handle the complex datasets in the foreseeable future. Yet, the application of statistical models is essential to evaluate differences between the studied groups. This work includes comprehensive statistical analysis for each chapter, performed according to the current principles and requirements for reliable evaluations to the best of our knowledge.

1.4. Limitations and challenges of current lipidomic approaches

Lipidomics is a growing field that is constantly changing. Although new technologies, methods and applications are published almost daily, there are still many questions and problems that must be addressed. The previous sections have showcased the common approaches to study lipids and some of the issues found in the literature.

First, there is still a need for analytical methods that allow a global profiling of lipids with reasonable sample amounts. The combination of lipidomics with other *omic* technologies for systems biology is hardly achieved if high volumes of precious biological fluids (>50 µL) or high masses of tissues (>350 mg) are required for each application. Second, the quality control of lipidomics data is essential to ensure reliable results, but usually not discussed or even mentioned in the literature. Sample preparation methods that require lipids to be exposed for long periods of time may seem attractive at first, but the results must be carefully evaluated. Furthermore, the precision of sample preparation and analysis methods is not always discussed, but essential for claiming that differences in lipid intensities or concentrations are related to a biological process

instead of technical flaws. Third, the normalization of lipidomics data is a vital step to obtain quantitative results, whether relative or absolute comparisons are being made. The use of internal standards for normalization, as well as data scaling and corrections, cannot be overlooked. Fourth, the statistical analysis and bioinformatics of lipids require careful planning and execution. The high complexity of the obtained datasets can easily hide small errors that will affect the biological implications and conclusions. The potential of lipidomics is undeniable but it can only be achieved with careful research and analytical development.

1.5. Objectives

Lipids are a tightly regulated but highly diverse class of biomolecules involved in virtually all physiological and pathological processes in an organism. Besides their function in regulating metabolism, lipids are also involved in a variety of diseases, *e.g.* atherosclerosis, diabetes, obesity and cancer.[13,14] Hence, lipidomics is a highly attractive field that offers numerous possibilities, but the associated challenges described in the previous sections still require extensive research and the development of high-quality analytical approaches. This work aimed to unravel the relationship between different diseases and lipid metabolism through the use of comprehensive, untargeted lipidomics for relative quantification.

The first objective consisted in the development of appropriate analytical methods for the sample preparation, chromatographic separation, mass spectrometry detection, data processing, identification and normalization of lipids. Chapter II describes the method development and optimization for comprehensive, untargeted lipidomics of small volumes of biological samples (1.0 – 2.5 µL) using nanoLC-MS. Then, the developed techniques were applied to study a variety of biological samples in Chapters III and IV (serum, cerebrospinal fluid, intraparenchymal

microdialysate and exosome). Although the main focus of chapters II, III and IV was the analytical performance, the biochemistry of lipids was also briefly discussed in each chapter.

The second objective was the application of developed UHPLC-MS methods to different types of biological samples. Chapters V to VIII describe a variety of applications of comprehensive, untargeted lipidomics to study pathological processes, including a novel vaccine candidate for respiratory syncytial virus (Chapter V, which also comprises metabolomics via chemical isotope labeling), biomarkers for Parkinson's disease diagnosis and prediction of dementia onset before noticeable symptoms (Chapter VI), effects of glucocorticoids upon different tissues (Chapter VII), and biomarkers of cystic fibrosis (Chapter VIII). The main biochemical implications of the analytical findings are discussed for each chapter.

This work aimed to showcase the importance of lipids in pathological and pathogenic processes, as well as the benefits of optimizing analytical and bioinformatics methods. The combination of lipidomics with other *omic* applications may provide the missing links to understand the many aspects of physiology and pathology that remain unexplained by the scientific community.

II

Chapter II: Development of a NanoLC-MS Workflow for High-Sensitivity Global Lipidomic Analysis[1,2,3]

1. Introduction

Lipidomics has become an essential part of systems biology, providing valuable information on physiological and pathological processes involving lipids.[99] The investigation of the lipid composition of biological samples, as well as their biochemical interactions with metabolites, proteins and other lipids, is essential for biomarker research.[13,14,61] Despite great advances in lipid analysis methods in the last two decades, accurate quantification and rapid identification of lipid species in a comprehensive and high-throughput manner are still needed.[14,61,99]The complexity of biological samples and immense structural diversity of lipids summed to strong ion suppression requires the development of adequate methods for comprehensive lipidomics. Lipid analysis is further complicated by the low available amounts of precious biological samples, the limited capability of data processing routines, variations in lipid nomenclature and limited commercial availability of isotopic standards for data normalization and quantification. Chapter I provides a broad review of current techniques and challenges for lipidomics.[12,13,61,100,101]

[1] A version of this chapter was accepted for publication at Analytica Chimica Acta on Aug / 2020.

[2] Supporting figures for this chapter are available at Appendix A. Supporting tables are available with Dr. Liang Li.

[3] Pig blood and cerebrospinal fluid samples employed for this chapter were collected through a collaboration with of Dr. Brian Kwon (University of British Columbia, Vancouver, Canada).

The combination of lipidomics with other *omics* technologies can provide a comprehensive panel of the biochemical processes involved in physiological and pathological conditions; however, the small volumes of biological samples available usually restrain the use of different analytical techniques. Nanoflow liquid chromatography (nanoLC) offers extremely high sensitivity for small volumes of diluted samples, low organic solvent consumption and reduced waste. It can be easily coupled to electrospray ionization-mass spectrometry (ESI-MS), making it ideal for multiple *omics* applications in small amounts of diluted samples.[100] NanoLC has been previously employed for the selective analysis of intact phosphatidylcholine molecules in soybean, bovine brain and bovine liver, as well as for lipids in tissue, plasma and human cerebrospinal fluid (CSF).[102–107] Danne-Rasche *et al.* described a method for the sensitive analysis of lipids in *Saccharomyces cerevisiae* extract by nanoLC-MS, employing a long 110 min gradient focusing on phospholipids. The work achieved an increase in sensitivity of 2 to 3 orders of magnitude when compared to a narrow-bore HPLC system, showcasing the advantage of nanoLC applications [108]. Yet, a method for sensitive analysis of diluted biological fluids with comprehensive lipid class coverage within a more reasonable analysis time is still required. This work describes the development of a high-sensitivity nanoLC-MS method for untargeted lipidomics of small volumes of biological samples. As part of the nanoLC-MS lipidomics workflow, we further optimized a modified Folch liquid-liquid extraction protocol based on dichloromethane and methanol for the comprehensive analysis of lipids in 1.0 – 2.5 μL of biological fluids. The evaluation of the method for lipidomics of serum and cerebrospinal fluid samples is also discussed.

2. Experimental

2.2.1 Chemicals and reagents

The chemicals and reagents employed for Chapter II to VIII were purchased from Sigma-Aldrich (St. Louis, MO, USA; ammonium formate and formic acid), Avanti Polar Lipids (Alabaster, AL, USA; Splash Lipidomix Mass Spec Standard), Fisher Scientific (Waltham, MA, USA; analytical grade dichloromethane and LC-MS grade methanol, 2-propanol, acetonitrile and water) and Honeywell (Charlotte, NC, USA; LC-MS grade methanol, 2-propanol, acetonitrile and water).

2.2.2 Instrumentation

The nanoLC-MS experiments were performed on a Waters nanoAcquity UPLC system (Waters Corporation, Milford, MA, USA) employing Thermo Acclaim PepMap C18 nanoViper trap (20 mm × 75 µm × 3 µm) and analytical columns (150 mm × 75 µm × 2 µm; Thermo Fisher Scientific, Waltham, MA, USA). The nanoLC instrument was coupled to an Impact II quadrupole time-of-flight (QTOF) mass spectrometer through a CaptiveSpray NanoBooster electrospray ionization (ESI) source (Bruker Daltonics, Billerica, MA, USA) with acetonitrile-enriched nitrogen gas.

The UHPLC experiments were performed on a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific, Waltham, MA, USA) employing a Waters Acquity BEH C18 column (5 cm × 2.1 mm × 1.7 µm; Waters Corporation, Milford, MA, USA) and coupled to the same mass spectrometer through a conventional ESI source.

2.2.3 Analysis methods

The initial nanoLC method for lipidomics of biological fluids was selected after a literature review, namely: MPA - 20 mM NH_4COOH, 5 mM formic acid (FA) in 45:45:10 methanol/ acetonitrile/ water (v/v/v); MPB - 20 mM NH_4COOH, 5 mM formic acid in 2-propanol; trapping for 1 min at 7 µL/min (95% MPA); flowrate of 300 nL/min; column temperature of 45°C; 38 min gradient (0 min – 5% MPB; 2 min – 5% MPB; 12 min – 30% MPB; 24 min – 90% MPB; 28 min – 1% MPB; 38 min – 1% MPB) followed by 20 min of re-equilibrium at 95% MPA; and 2.0 µL injection. The initial nanoLC conditions were optimized for mobile phase composition, gradient, temperature, trapping time, trapping mobile phase composition, trapping flowrate, and injection volume. Initial MS conditions were also fully optimized, namely: capillary voltage of 1300 V, dry gas flow rate of 3.0 L/min, ion source temperature of 200°C, spectra acquisition rate of 2 Hz, and NanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar. Optimal conditions were selected based on peak intensity, visual chromatographic separation and the number of detected features. The relative standard deviation for peak intensities between replicates was also evaluated as a measure of technical reproducibility. Only features detected in all experimental replicate analyses were considered for method optimization.

After the optimization, the sensitivity of the nanoLC method was compared to similar conditions applied to UHPLC-MS, *i.e.,* MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water; MPB - 10 mM NH_4COOH in 95:5 2-propanol/ water; 250 µL/min; 40 °C; 22 min gradient (0 min – 5% MPB; 1.8 min – 5% MPB; 8.5 min – 30% MPB; 18 min – 95% MPB; 22 min – 95% MPB) followed by 10 min of re-equilibrium (0 min – 95% MPB; 3 min – 95% MPB; 4 min – 5% MPB; 10 min – 95% MPB); and 4 µL injection. The UHPLC method was

previously optimized (data not shown), but MS conditions were identical to the method optimized for nanoLC analysis.

2.2.4 Sample preparation [4]

The classical Folch method for liquid-liquid extraction of lipids by dichloromethane and methanol was optimized for handling 1.0 – 10.0 µL of blood serum. All optimization experiments were performed with serum samples obtained from Yucatan miniature pigs unless otherwise stated. Initially, 1.0 µL of serum was mixed with 55.4 µL of methanol, 110.8 µL of dichloromethane and 33.4 µL of water. After resting for 10 min at room temperature, the mixture was centrifuged for 10 min (12,000 rpm, 4°C) and the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min. The dried extract was resuspended immediately in 10 µL of 6:4 MPA/MPB and diluted with 90 µL of 9:1 MPA/MPB. A mixture of 14 deuterated lipids in methanol (Splash Lipidomix Mass Spec Standard, Avanti Polar Lipids) was further added to the serum samples in equal volumes before extraction for internal standardization, namely: lysophosphatidylcholine – LPC d7-18:1, lysophosphatidylethanolamine – LPE d7-18:1, monoacylglycerol – MG d7-18:1, d7-cholesterol, phosphatidylglycerol – PG d7-15:0/18:1 (sodium salt), phosphatidylinositol – PI d7-15:0/18:1 (ammonium salt), phosphatidylcholine – PC d7-15:0/18:1, phosphatidylethanolamine – PE d7-15:0/18:1, phosphatidylserine – PS d7-15:0/18:1 (sodium salt), phosphatidic acid – PA d7-15:0/18:1 (sodium salt), sphingomyelin – SM d9-d18:1/18:1, diacylglycerol – DG d7-15:0/18:1, triacylglycerol – TG d7-15:0/18:1/15:0, and cholesteryl ester – CE d7-18:1. Table II-1 provides

[4] Serum and CSF samples from Yucatan miniature pigs were obtained through a collaboration with Dr. Brian Kwon (University of British Columbia, Vancouver, BC, Canada).

details on molecular weight, formula and concentration of the internal standard mixture. The sample extraction method was optimized for: initial sample volume, equilibration temperature and time; drying time; composition of solvents for resuspension and dilution of the extract; and composition and proportions of extraction solvents. We also evaluated contamination sources for the extraction method, including polypropylene (PP) microcentrifuge tubes and autosampler vial inserts (PP and glass).

Table II-1. Composition of the deuterated lipid standard mixture (14 standards in methanol solution) employed for internal standardization of lipidomics of biological samples (Splash Lipidomix Mass Spec Standard, Avanti Polar Lipids).

Standard	Formula	Molecular weight	Concentration (µg/mL)
MG d7-18:1	$C_{21}H_{33}D_7O_4$	363.59	1.8
CE d7-18:1	$C_{45}H_{71}D_7O_2$	658.16	329.1
Cholesterol-d7	$C_{27}H_{39}D_7O$	393.71	98.4
LPE d7-18:1	$C_{23}H_{39}D_7NO_7P$	486.64	4.9
LPC d7-18:1	$C_{26}H_{45}D_7NO_7P$	528.72	23.8
DG d7-15:0/18:1	$C_{36}H_{61}D_7O_5$	587.98	8.8
PA d7-15:0/18:1	$C_{36}H_{61}D_7NaO_8P$	689.94	6.9
PE d7-15:0/18:1	$C_{38}H_{67}D_7NO_8P$	711.03	5.3
SM d9-d18:1/18:1	$C_{41}H_{72}D_9N_2O_6P$	738.12	29.6
PC d7-15:0/18:1	$C_{41}H_{73}D_7NO_8P$	753.11	150.6
PS d7-15:0/18:1	$C_{39}H_{66}$ $D_7NNaO_{10}P$	777.02	3.9
PG d7-15:0/18:1	$C_{39}H_{67}D_7NaO_{10}P$	764.02	26.7
TG d7-15:0/18:1/15:0	$C_{51}H_{89}D_7O_6$	812.37	52.8
PI d7-15:0/18:1	$C_{42}H_{75}D_7NO_{13}P$	847.13	8.5

2.2.5 Data processing

The detected features for each evaluated condition during the nanoLC-MS optimization were picked by the software Bruker Compass Data Analysis 3.1 (Bruker Daltonics, Billerica, MA, USA) and aligned on Profile Analysis 2.1 (Bruker Daltonics, Billerica, MA, USA), using S/N threshold of 3, correlation coefficient threshold of 0.7, minimum compound length of 10 spectra, smoothing width of 5 spectra, and retention time and mass tolerances for alignment automatically selected by the software. The data processing routine was fully optimized, focusing on maximizing the number of detected features. However, the combination of two software took an unreasonable amount of time for processing high sensitivity data (extracting and aligning >10,000 features). Hence, after the nanoLC-MS and extraction method optimization, the data processing routine was adapted for Bruker MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA).

2.2.6 MS/MS identification of detected features

MS/MS identification of lipids for the optimized method was performed with the following LC-MS/MS libraries: MS-Dial LipidBlast (https://fiehnlab.ucdavis.edu/projects/LipidBlast)[73,78], the Human Metabolome Database (HMDB, https://hmdb.ca)[80,81,109] and MassBank of North America (MoNA, a repository encompassing spectra from MassBank, HMDB, LipidBlast, the Fiehn's lab FAHFA lipid library[82], the Riken Center's oxidized phospholipid library[83] and others, https://mona.fiehnlab.ucdavis.edu). The libraries were combined to MetaboScape 4.0 for spectral match using precursor m/z tolerance of 5.0 mDa and (1) MS/MS score threshold of 500 and mSigma (isotope pattern match) threshold of 150; or (2) MS/MS score threshold of 100 and mSigma threshold of 50. Internal standard features were found by retention time and accurate m/z values. Lipid subclasses and categories followed the International Lipid Classification and

Nomenclature Committee (ILCNC), the LipidMaps database (2019-2020 update) and the Lipidomics Standards Initiative (Table II-2). [16–18,20,75,76]

Table II-2. Abbreviation employed for lipid subclasses in Chapters II, III, IV, VII and VIII.

Abbreviation	Common name	Lipid Subclass	Lipid Class
Acer	Acylceramide	Ceramides	Sphingolipids
AcylGlcADG	Glucosyldiacylglycerol	Glycosyldiradylglycerols	Glycerolipids
BMP	Bis[monoacylglycero]phosphate / Hemibismonoacylglycerophosphate / lysobisphosphatidic acid	Monoacylglycerophospho-monoradylglycerols	Glycerophospholipids
Car	Fatty acylcarnitine	Fatty esters	Fatty Acyls
CE	Cholesteryl ester	Cholesteryl esters	Sterol lipids
Cer	Ceramide	Ceramides	Sphingolipids
CL	Cardiolipin	Glycerophosphoglycerophosphoglycerols	Glycerophospholipids
CoA	Fatty acyl CoEnzyme A	Fatty esters	Fatty Acyls
DG	Diacylglycerol (Diglyceride)	Diradylglycerols	Glycerolipids
DGC	Acyl-sn-glycero-3-O-carboxy-(hydroxymethyl)-choline	Other Glycerolipids	Glycerolipids
DGD	Digalactosyldiacylgylcerol	Glycosyldiacylglycerols	Glycerolipids
DGT	Diacylglyceryltrimethyl-homoserine	Other Glycerolipids	Glycerolipids
FA	Fatty acid / Fatty ester / Wax ester / Fatty acid derivative	Fatty acids and conjugates	Fatty acyls
FC	Free cholesterol	Cholesterol	Sterol lipids
GMG	Sulfoquinovosylmonoacyl-glycerol	Glycosylmonoradylglycerols	Glycerolipids
HexCer	Hexosylceramides / Glucosylceramide / Lactosylceramide	Neutral glycosphingolipids	Sphingolipids
LPA	Lysophosphatidic acid / Cyclic Lysophosphatidic acid	Glycerophosphates	Glycerophospholipids
LPC	Lysophosphatidylcholine	Glycerophosphocholines	Glycerophospholipids
LPE	Lysophosphatidylethanolamine	Glycerophosphoethanolamines	Glycerophospholipids
LPG	Lysophosphatidylglycerol	Glycerophosphoglycerols	Glycerophospholipids
LPI	Lysophosphatidylinositol	Glycerophosphoinositols	Glycerophospholipids
LPS	Lysophosphatidylserine	Glycerophosphoserines	Glycerophospholipids

LSL	Lysosphingomyelin / lysoglycosphingolipid	Sphingoid bases	Sphingolipids
MG	Monoacylglycerol (Monoglyceride)	Monoradylglycerols	Glycerolipids
MIPC	Ceramide phosphoinositol	Phosphosphingolipids	Sphingolipids
NAA	N-acylamine / N-acylamide / N-acyl-taurine	Fatty amides	Fatty Acyls
PA	Phosphatidic acid	Glycerophosphates	Glycerophospholipids
PC	Phosphatidylcholines	Glycerophosphocholines	Glycerophospholipids
PE	Phosphatidylethanolamine / Diacylglycerophosphoethanolamine	Glycerophosphoethanolamines	Glycerophospholipids
PE-Cer	Ceramide phosphoethanolamine	Phosphosphingolipids	Sphingolipids
PEtOH	Phosphatidylethanol	Other Glycerophospholipids	Glycerophospholipids
PG	Phosphatidylgylcerol	Glycerophosphoglycerols	Glycerophospholipids
PI	Phosphatidylinositol / Diacylglycerophosphoinositolglycan /	Glycerophosphoinositols	Glycerophospholipids
PIP	Phosphatidylinositol-phosphate	Glycerophosphoinositol phosphates	Glycerophospholipids
PS	Phosphatidylserines	Glycerophosphoserines	Glycerophospholipids
SM	Sphingomyelin / Ceramide-1-phosphate	Phosphosphingolipids	Sphingolipids
SPB	Sphingoid base / Sphingoid base-1-phosphate	Sphingolipids	Sphingolipids
ST	Sterol / Cholesterol and Derivatives	Sterols	Sterol Lipids
Sulf	Sulfoglycosphingolipid (sulfatide)	Acidic glycosphingolipids	Sphingolipids
SulfDG	Sulfoquinovosyldiacylglycerols	Glycosyldiradylglycerols	Glycerolipids
TG	Triacylglycerol (Triglyceride)	Triradylglycerols	Glycerolipids

2.2.7 Putative identification of lipids

Features that were not identified by MS/MS were inputted on the LipidMaps database for putative identification based on mass-match (m/z error smaller than 5.0 mDa). Lipids can have multiple isomers and isobars within the mass tolerance of 5.0 mDa; hence, all possible isomeric and isobaric lipids for the same feature were filtered and ranked by a 5-tier system to ensure more accurate identifications. All isomeric and isobaric possibilities that passed initial filters for retention time and adduct detection were kept for each mass-matched feature, but they were ranked according to the characteristics of the employed method and biological fluids. A similar approach was employed for Chapters III to VIII with minor modifications.

First, we applied a retention time filter to exclude identification possibilities that could not elute in the detected retention time, *e.g*. a triacylglycerol is expected to have high retention by the reverse-phase chromatography conditions employed for this work; hence, identifications for triacylglycerols with low retention times were excluded. We employed lipid standards and high-confidence MS/MS identifications to determine the expected retention time range for each lipid subclass, sub-divided into 4 fatty acyl length groups (sum of the number of carbons for all fatty acyl residues): less than 14 carbons; 14 to 20 carbons; 21 to 33 carbons; and 34 carbons or more. The expected retention time ranges were then used to filter out identification possibilities that could not elute in the detected times.

Second, we employed a filter for the identified adducts. Most lipid classes are not easily ionized by electrospray as the common $[M+H]^+$ and $[M-H]^-$ ions; instead, they are detected as adducts that depend upon the modifiers employed for mobile phase and sample medium. The mobile phases and chromatographic conditions used for the optimized conditions allow different adducts to be detected, including $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$, $[M-H_2O+H]^+$, $[M-H]^-$,

$[M+HCOO]^-$ and $[M-CH_3]^-$. However, different lipid subclasses can form different ions and adducts, *e.g.,* simple fatty acids are often detected as $[M-H]^-$ but triacylglycerols cannot be ionized by the loss of a proton, being more commonly detected as the $[M+NH_4]^+$ adduct. Therefore, we also employed lipid standards to determine the possible adducts for each lipid subclass. While identification possibilities for adducts that could not be detected were excluded (*e.g.*, triacylglycerols as $[M-2H]^{2-}$), all the possible matches were scored according to their expected ionization. The most likely adduct possibilities for each feature were given a score of one, while the less likely possibilities were given higher scores.

Third, the identifications for the same feature were scored by their mass errors. While all identifications had a maximum mass error of 5.0 mDa, errors smaller than 5.0 ppm were given a score of one, while errors between 5.0 and 7.5 ppm received two points; 7.5 to 10.0 ppm received three; 10.0 to 15.0 ppm received four; and errors higher than 15.0 ppm received five points.

Fourth, the isomeric and isobaric identifications were scored according to an odd or even number of carbons in fatty acyl groups. Even-numbered fatty acyls are more commonly found in mammals due to the acetyl-CoA biosynthesis route (Chapter I), although odd-chain fatty acyl groups can be detected in smaller amounts. Therefore, identification possibilities for isomers or isobars with even-chain fatty acyls received a score of one, while odd-chain fatty acyls were scored two points.

Fifth, the presence of functional groups other than the expected for each lipid subclass (oxidation, dehydration, cyclization, etc.) was scored two points, while the absence of extra groups received one point due to being more common in biological fluids. Although modifications are typical for lipids, they require the action of specific enzymes or processes, leading to lower abundance when compared to unmodified species.

The identification possibilities that passed the retention time and adduct filter were ranked from the lowest to the highest total score. The most likely identification, *i.e.* lowest score, was chosen to determine the lipid subclass for normalization (Table II-2, p.48), although the remaining possibilities were kept for further reference. The positions of double bonds and the stereospecific configuration of glycerol derivatives were not determined in this study.

3. Results and Discussion

The initial sample extraction and nanoLC method resulted in low performance and robustness, with constant instrument issues (blockages, leaking), insufficient separation, peak tailing and fronting, intense background (contamination), and low sensitivity (Figure II-1A). The subsequent method optimization was separated into two phases. First, we optimized the initial trapping conditions, mobile phase composition, mobile phase flow rate, gradient and column temperature using the lipid extract from a pig serum. The MS acquisition method was also optimized for capillary voltage, dry gas flowrate, dry temperature, spectra acquisition rate and CaptiveSpray NanoBooster gas pressure. We finalized the first phase of optimization by evaluating the background intensity and contamination from plastic and glassware (sample extraction and sample injection vessels). The chromatograms presented substantial alterations after the evaluation of contamination sources. Furthermore, we introduced a mixture of 14 deuterated standards belonging to different lipid subclasses to assess the separation and sensitivity of the evaluated conditions, as well as for internal standardization (Table II-1). For the second phase of the optimization, we re-evaluated the separation gradient, trapping and MS acquisition method with a pool of 8 pig serum samples and the deuterated standard mix. The sample extraction method was also optimized. The method was finalized with a fine adjustment of the separation gradient and

injection volume. Detailed descriptions of the individual steps for the development of an optimized workflow are described below. Supporting figures for this chapter, *i.e.* chromatograms obtained for each step of optimization, are available in Appendix A (p. 404).

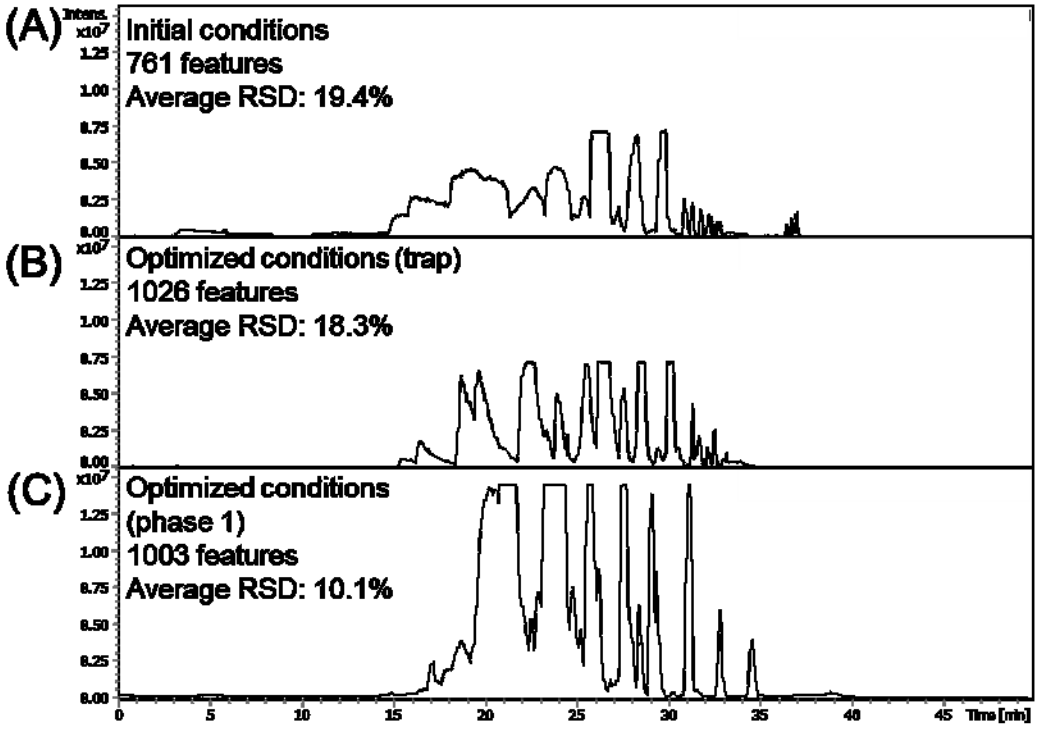

Figure II-1. Base peak chromatograms (BPC) obtained with a serum sample from a pig for the first phase of nanoLC method optimization. (A) initial nanoLC-MS method, namely: MPA – 20 mM NH_4COOH, 5 mM formic acid in 45:45:10 methanol/ acetonitrile /water; MPB - 20 mM NH_4COOH, 5 mM formic acid in 2-propanol; trapping for 1 min at 7 µL/min (95% MPA); 300 nL/min; 45ºC; 38 min gradient (0 min – 5% MPB, 2min – 5% MPB, 12 min – 30% MPB, 24 min – 90% MPB, 28 min – 1% MPB, 38 min – 1% MPB); 20 min equilibrium (95% MPA); 2 µL injection; electrospray ion source capillary voltage of 1300 V; nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar; dry nitrogen gas flow rate of 3 L/min; ion source temperature of 200ºC; and spectra acquisition rate of 2 Hz. (B) Initial optimization of trapping flow rate, time and

mobile phase composition (7.0 µL/min, 1.5 min, 95% MPA). (C) Optimized nanoLC-MS method for the first phase, namely: MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water; MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water; trapping for 1.5 min at 7.0 µL/min (100%A); 45ºC; 300 nL/min; 50 min gradient (0 min – 30% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 30 min – 90% MPB, 31 min – 30% MPB, 50 min – 30% MPB); 20 min equilibrium (95% MPA); 2 µL injection; electrospray ion source capillary voltage of 1300 V; nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.20 bar; dry nitrogen gas flow rate of 4 L/min; ion source temperature of 220ºC; and spectra acquisition rate of 1 Hz. Average RSD: average relative standard deviation between injection replicates for all detected features, employed as a measure of reproducibility.

2.3.1 nanoLC separation method (initial optimization)

The narrow nanoLC capillary columns have limited flowrates (nanoliter range) and are easily overloaded by small amounts of sample (less than 1.0 µL). The internal volume of the employed analytical column (15 cm × 75 µm) is calculated as 0.663 µL without considering the volume occupied by the stationary phase; hence, a 1.0 µL injection would fill the entire column, causing distorted peaks and bad chromatography. Furthermore, the nanoliter-range flowrates cause excessive long dead times. A trapping step with a smaller C18 capillary column with larger particles (2 cm × 75 µm, 3 µm particles) was employed before the injection of samples in the C18 analytical column (2 µm particles). The small length and larger particle size of the trap column allows higher flowrates for loading the sample without risking damage to the delicate capillary due to excessive pressure. The C18 trap column, a waste line and the analytical column were connected to a valve (Figure II-2A). During the trapping step, the sample was injected into the trap

column, followed by the mobile phase at high flowrate (microliter-range) directed to the waste line (Figure II-2B). The goal was to eliminate non-retained, polar compounds and the excess of solvent, creating a narrow sample plug inside the trap column. After trapping, the flow rate was reduced and re-directed to the analytical column (Figure II-2C). The clean-up step provided by trapping the sample in a small C18 column is vital for reducing the long dead time of nanoliter-range flowrates and ion suppression, as well as preventing the overload of the capillary analytical column, which can greatly increase the sensitivity of the analysis. The reversed-phase character of the trap column requires hydrophilic conditions, *i.e.*, high proportions of weak mobile phase (MPA), for maximum retention of lipids. Low flowrates and longer trapping times can improve the sample clean-up, peak shapes and intensity for the most hydrophobic compounds. However, a narrow injection band and adequate peak shapes for small, less hydrophobic lipids can only be obtained for the opposite conditions. We first performed an initial optimization of the trapping parameters to ensure the detection of a reasonable number of features for further method development. The initial trapping flow rate, time and mobile phase composition were evaluated between 3.0 and 10.0 µL/min for 0.75 to 1.5 min, with 90 to 99% of MPA (Appendix A, Figure A - 1, Figure A - 2 and Figure A - 3, respectively). The best conditions were selected based on the number of detected features, total peak intensity and average relative standard deviation (RSD) for injection replicates, *i.e.,* trapping at 7.0 µL/min for 1.50 min at 95% of MPA. The initial optimization of the trapping conditions increased the number of detected features from 761 to 1026 (34.8%, Figure II-1B). The overall reproducibility of peak intensities was also improved, as showed by the average RSD for injections replicates (from 19.4% to 18.3%).

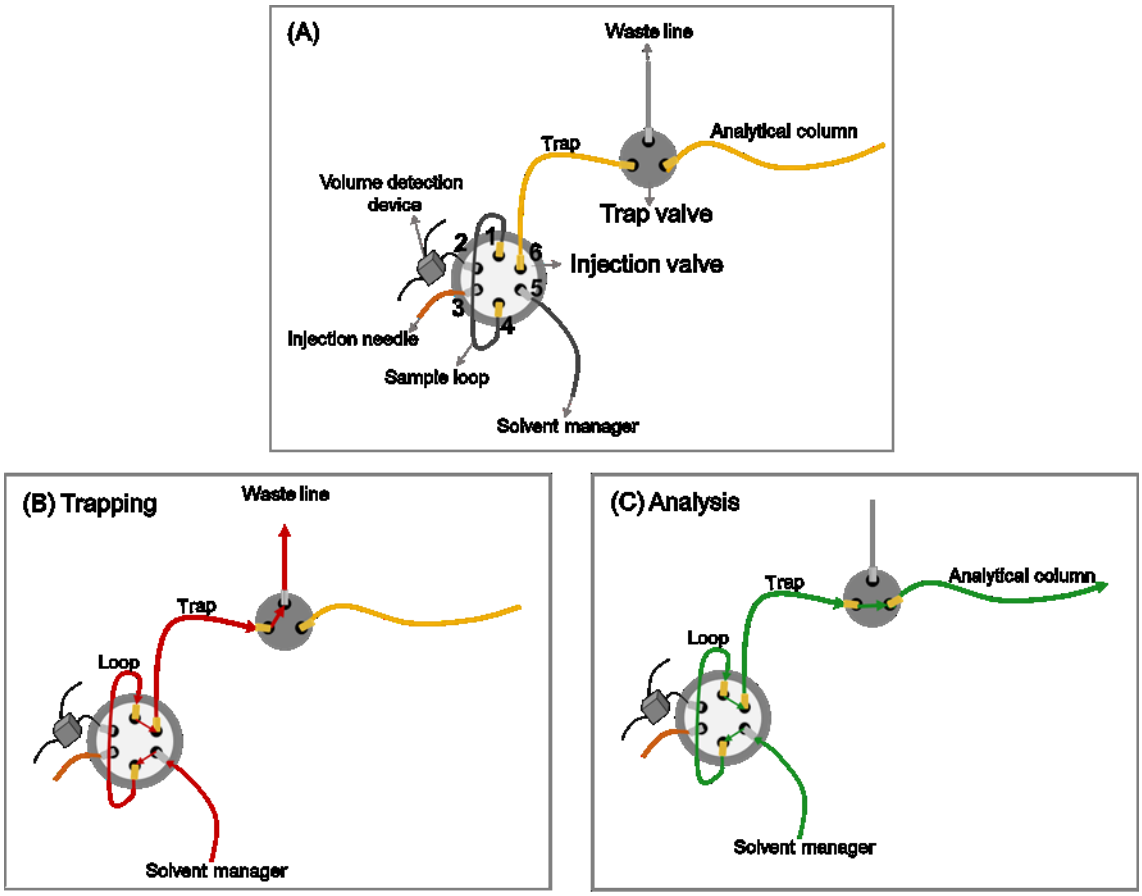

Figure II-2. Scheme for the injection and trap valves for the nanoLC instrument. (A) Fittings connected to the 6-port injection valve and 3-port trap valve. The injection valve is connected to the sample loop (ports 1 and 4), volume detection device (port 2), sample needle (port 3), solvent manager (port 5) and the trap column (port 6), as indicated in the figure. The sample needle introduces the sample into the loop. The mobile phase carriers the sample from the loop to the trap column. The volume detection device uses air gaps to measure the sample volume. (B) Mobile phase path during the trapping step, indicated in red. The mobile phase flows from the solvent manager through the sample loop to push the sample into the trap column, with the flow directed to the waste line. (C) Mobile phase path during the analytical separation, indicated in green. The mobile phase flows through the sample loop and the trap column, pushing the trapped analytes into the analytical column for separation and detection.

Next, we optimized the analytical mobile phase composition (Table II-3, Figure II-3). The hydrophobic character of most lipids requires the use of strong organic solvents for elution under reversed-phase chromatography, even for the weak mobile phase (MPA). The combination of the hydrogen-bonding character of methanol, π interactions of acetonitrile and high polarity of water can provide the high selectivity required for MPA, while 2-propanol in the strong mobile phase B (MPB) ensures elution of the most hydrophobic lipids, such as cholesteryl esters and triacylglycerols.[110,111] Therefore, the initial MPA contained a combination of methanol, acetonitrile and water (45:45:10 v/v/v), whereas MPB was composed of pure 2-propanol. The solvent composition was chosen after a literature review; however, the high viscosity of MPB caused dangerously high pressures (close to the column limit of 600 bar), which led to constant instrumental issues (leaking in zero-volume fittings, damage to the delicate sprayer tip, blockages and capillary fissures). We evaluated different proportions of solvents in MPA and MPB, searching for better separation and robustness (Table II-3). The viscosity of MPB was controlled through the addition of 5 to 10% of water, but higher percentages caused an undesirable increase in retention times of low polarity lipids and a decrease in the total number of detected features (Test ID Initial *versus* Test ID 1, Figure II-3).

Table II-3. Mobile phase compositions evaluated for method optimization. The optimized composition was test ID 7.

Test ID	MPA			MPB			Detected features
	Ammonium formate	Formic acid	Solvents (methanol/ acetonitrile/ H_2O)	Ammonium formate	Formic acid	Solvents (2-propanol/ H_2O)	
Initial	20 mM	5 mM	47.5 : 47.5 : 5	20 mM	5 mM	100 : 0	1026
1	20 mM	0.1% (26.5 mM)	47.5 : 47.5 : 5	20 mM	0.1%	90 : 10	559
2	20 mM	5 mM	47.5 : 47.5 : 5	-	0.1%	90 : 10	673
3	-	0.1%	47.5 : 47.5 : 5	-	0.1%	90: 10	229
4	10 mM	-	0 : 60 : 40	20 mM	0.1%	90 : 10	759
5	10 mM	0.1%	47.5 : 47.5 : 5	10 mM	0.1%	90 : 10	376
6	10 mM	-	47.5 : 47.5 : 5	10 mM	-	90 : 10	441
7	10 mM	-	50 : 40 : 10	10 mM	-	95 : 5	1106

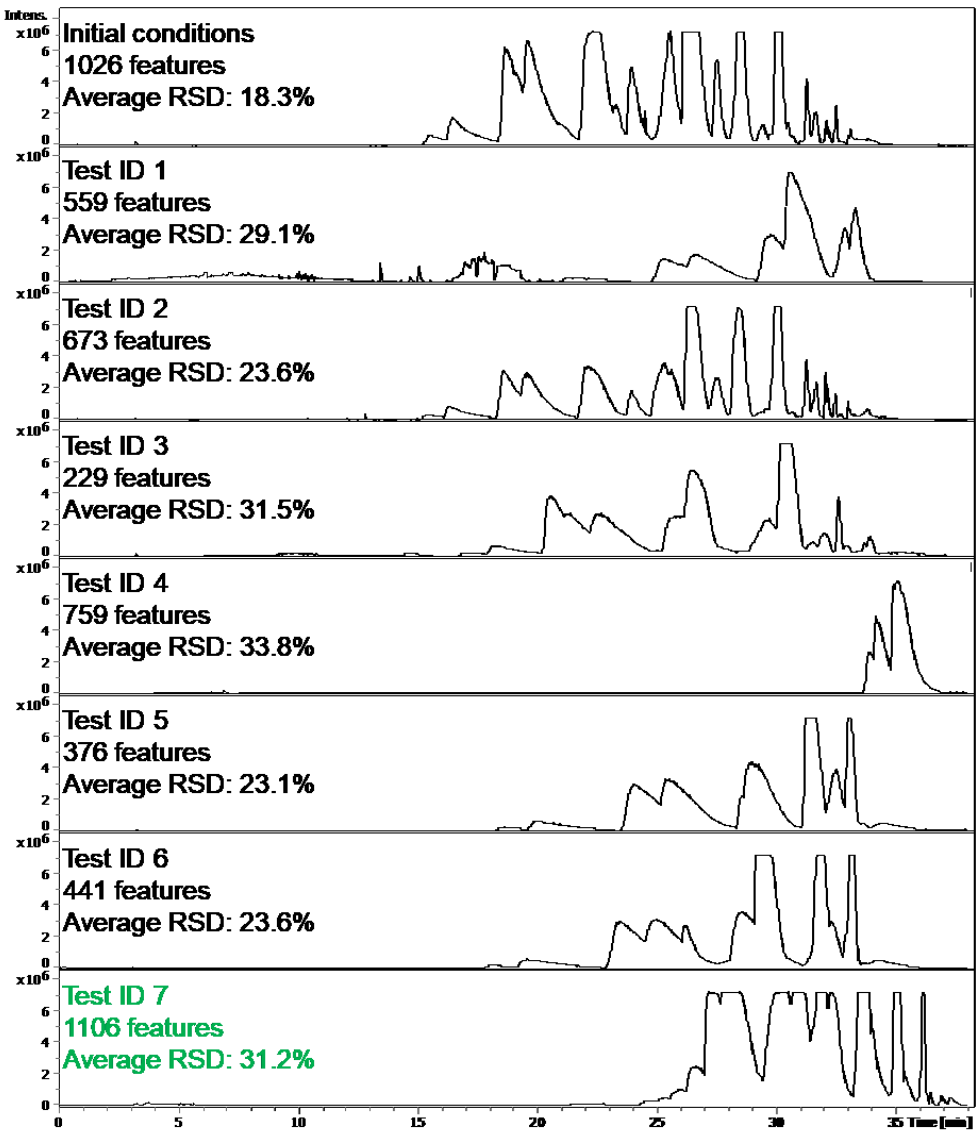

Figure II-3. BPC obtained for the optimization of mobile phase composition using a serum sample from a pig. The compositions of the tested mobile phases are described in Table II-3. NanoLC method: trapping for 1.50 min at 7 µL/min (95% MPA); 300 nL/min; 45ºC, 38 min gradient (0 min – 5% MPB, 2 min – 5% MPB, 12 min – 30% MPB, 24 min – 90% MPB, 28 min – 1% MPB, 38 min – 1% MPB); 20 min equilibrium (95% MPA); 2 µL injection. Initial MS conditions: electrospray ion source capillary voltage of 1300 V, dry gas flow rate of 3.0 L/min, source temperature of 200ºC, spectra acquisition rate of 2 Hz and nanoBooster acetonitrile-enriched

nitrogen gas pressure of 0.10 bar. The initial sample preparation method is described in the Experimental Section.

Along with the mobile phase solvent optimization, we also evaluated the additive composition and concentrations. Most lipid species are not easily ionized under soft ionization techniques such as ESI. While charged or zwitterionic molecules can be easily detected by ESI under positive or negative ionization ($[M+H]^+$ or $[M-H]^-$ ions), neutral or non-polar lipids are not promptly ionized and require the use of additives to form adducts, *e.g.,* acylglycerols and cholesteryl esters.[13,100,112] Different mobile phase additives have been employed in an attempt to promote ionization and, consequently, improve the sensitivity of lipid analysis, *viz.*, ammonium formate, ammonium acetate and formic acid. Ammonium formate has been previously described as an optimal modifier for most lipids under positive ESI conditions, whereas formic acid is commonly added to promote positive ionization, although acidic conditions may cause the degradation of lipids through cleavage of vinyl-ether bonds of plasmalogens and hydrolysis.[68,113] Ammonium acetate has been cited as a better modifier for negative mode but the use of the same mobile phase composition and methodology for both polarities simplifies the analysis routine, avoids long equilibrium runs and decreases the amount of waste.[113]

For this work, the initial mobile phases A and B contained a combination of 20 mM of ammonium formate and 5 mM of formic acid.[12,114–119] However, ammonium formate has low solubility in 2-propanol and acetonitrile, which may result in the precipitation of salt crystals over time, clogging the narrow nanoLC column and the glass ion source sprayer tip. Moreover, the acidic initial conditions often result in peak tailing for compounds that contain deprotonated phosphate groups due to hydrogen bonding with free silanol groups in the C18 stationary phase,

as well as complexation of phosphatidyl and carbonyl groups by metal surfaces in the sample path (tubing, injection needle, etc.). In contrast, neutral conditions can promote better peak shapes for phospholipids while avoiding acid degradation of susceptible lipids.[120]

Different mobile phase compositions were evaluated, focusing on improved sensitivity and simpler, robust preparation (Table II-3, Figure II-3). The concentration of formic acid was first increased to 0.1%, but the number of detected features decreased to about half of the initial value (Test ID 1). The use of formic acid as the only additive in MPB (Test ID 2) and both MPA and MPB (Test ID 3) were also evaluated, but the later provided strong peak tailing due to the acidic pH and the smallest number of detected features (Figure II-3). Therefore, formic acid didn't offer enough advantages to be included in the mobile phase composition. Although a smaller concentration of ammonium formate decreased the number of detected features (Test IDs 1 *versus* 5), the method robustness increased due to less precipitation. A compromise was achieved for 10 mM ammonium formate in both MPA and MPB. A higher percentage of water in MPA further improved peak shapes and the number of detected features, probably due to an improvement of trapping for lipids with lower hydrophobicity (Test ID 7, Figure II-3). The combination of MPA – 10 mM ammonium formate in 50:40:10 methanol/ acetonitrile/ water (v/v/v) and MPB - 10 mM ammonium formate in 95:5 2-propanol/ water (v/v) delivered the best peak shapes and a high number of detected features. The same mobile phase composition was employed for the nanoLC-MS and UHPLC-MS applications described in Chapters III to VIII.

The nanoliter-range flowrate of nanoLC separations leads to extremely high sensitivities due to reduced sample dilution during analysis and increased ESI ionization efficiency. However, it can also cause long dead times and peak broadening. The flow rate was varied between 300 and 400 nL/min. The smallest value provided a higher number of detected features, but poor separation

and peak shapes (Appendix A - Figure A - 4). In contrast, the highest value was chosen due to smaller analysis time and relative standard deviations (*i.e.*, better reproducibility for peak intensities), but it also slightly decreased the sensitivity and increased the pressure.

Next, 26 different gradients were assessed (Figure A - 5, Figure A - 6 and Figure A - 7). The best conditions were achieved with 0 min – 30% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 30 min – 90% MPB, 31 min – 30% MPB, 50 min – 30% MPB. The optimized gradient, combined with the previously optimized flowrate of 400 nL/min, led to dangerously high pressures; therefore, the flow rate was reduced to 300 nL/min (Figure A - 8).

The separation temperature was further evaluated between 40 and 50°C (Figure A - 9). An increase in temperature promotes the diffusion of the solute in the column, but also a faster partition between the mobile and the stationary phase. Likewise, it can provide more reasonable pressures for higher percentages of MPB, but it can also cause degradation of lipids. Therefore, the intermediate temperature of 45°C was chosen. Finally, the trap mobile phase composition (70 to 100% of MPA, Figure A - 10), flowrate (4.0 to 10.0 µL/min, Figure A - 11) and time (1.0 to 2.0 min, Figure A - 12) were re-evaluated, with the best separation and highest number of detected features found for 7.0 µL/min for 1.5 min at 100% of MPA.

2.3.2 Mass spectrometry acquisition method (initial optimization)

The analysis of biological samples by liquid chromatography coupled to electrospray quadrupole time-of-flight mass spectrometry (ESI-QqToF-MS) is well established due to high resolutions, ample m/z coverage and high sensitivities. For this work, we employed the Captive Spray NanoBooster ESI ion source, designed specifically for nanoLC-MS coupling. The NanoBooster technology allows the enrichment of nitrogen gas that flows around the emitter tip

with volatile solvents, which can improve ionization efficiency through charge state control. The use of acetonitrile as an enrichment solvent has been previously evaluated for peptide analysis, providing an overall increase in the average charge state and signal intensity.[121] However, a detailed optimization of detection parameters is required for optimal performance.

The positive mode ESI-QqToF parameters were optimized concomitantly, *i.e.*, capillary voltage (1200 to 1400 V, Figure A - 13), acetonitrile-enriched nitrogen gas pressure (NanoBooster, 0.10 to 0.30 bar, Figure A - 14), nitrogen gas flow rate (2.0 to 4.0 L/min, Figure A - 15), ion source temperature (180 to 220 °C, Figure A - 16) and spectra acquisition rate (1 to 3 Hz, Figure A - 17). The reduced flow rate of nanoLC can easily handle smaller acquisition rates, which provides higher sensitivities. Therefore, the rate of 1 Hz was selected for further experiments. Higher capillary voltages resulted in more detected features, but we also observed higher baseline noise and in-source fragmentation of lipid standards; therefore, the best compromise was chosen at 1300 V. We found a higher number of detected features and total intensity for ion source temperature of 220°C, nitrogen gas flow rate of 4 L/min and NanoBooster gas pressure of 0.20 bar (Figure A - 18).

The nanoLC-MS method optimization for Phase 1 resulted in an increase of 31.8% for detected features, allied to higher intensity and lower average relative standard deviation between injection replicates (Figure II-1C).

2.3.3 Evaluation of contamination sources

The intense chromatograms obtained for blank extractions (extraction of water instead of the sample) prompted an examination of sources of contamination. We first employed 1.5 mL polypropylene (PP) microcentrifuge tubes from Fisher Scientific (Waltham, MA, USA) for sample

extraction, but the blank extracts were unreasonably intense (Figure II-4A). The application of single-use plastic ware for extractions with organic solvents is not ideal, although more practical and affordable than glassware. We hypothesized that different brands and sizes of tubes could provide different backgrounds for lipid extraction. Hence, we evaluated two brands of similar PP microcentrifuge tubes with different sizes (Fisher Scientific and Rose Scientific Ltd - Edmonton, AB, Canada, Figure II-4B) and two types of vial inserts for sample extraction and injection (PP autosampler inserts from Canadian Life Science, Peterborough, ON, Canada; and glass autosampler inserts from Agilent Technologies, Santa Clara, CA, USA, Figure II-5A). Overall, we observed that the contamination level from plastic ware varies with the brand (Figure II-4B, Figure A - 19) and the size of tubes (Figure II-4B, Figure A - 20), as well as with different batch numbers for the same manufacturer. The extraction in PP autosampler inserts provided a cleaner blank extraction when compared to the microcentrifuge tubes (Figure II-5, Figure A - 21). Even though the glass inserts provided low contamination (Figure II-5), the elevated costs associated with single-use glassware cannot be justified by the small differences found between the PP and glass inserts (*circa* 7 times more expensive for glass, Figure A - 22). We also examined the effect of pre-washing the plastic tubes with solvents but found no improvement (Figure A - 23, Figure A - 24 and Figure A - 25). Hence, the further steps of method optimization were performed with extractions and injections in 250 µL PP autosampler inserts (Figure II-5B). After extraction, samples were stored in 250 µL PP autosampler inserts placed in amber vials capped with PTFE-lined caps at 4°C and injected within a maximum of 48h. However, we recommend the use of glassware for long-term storage of lipid samples and solutions in strong organic solvents.

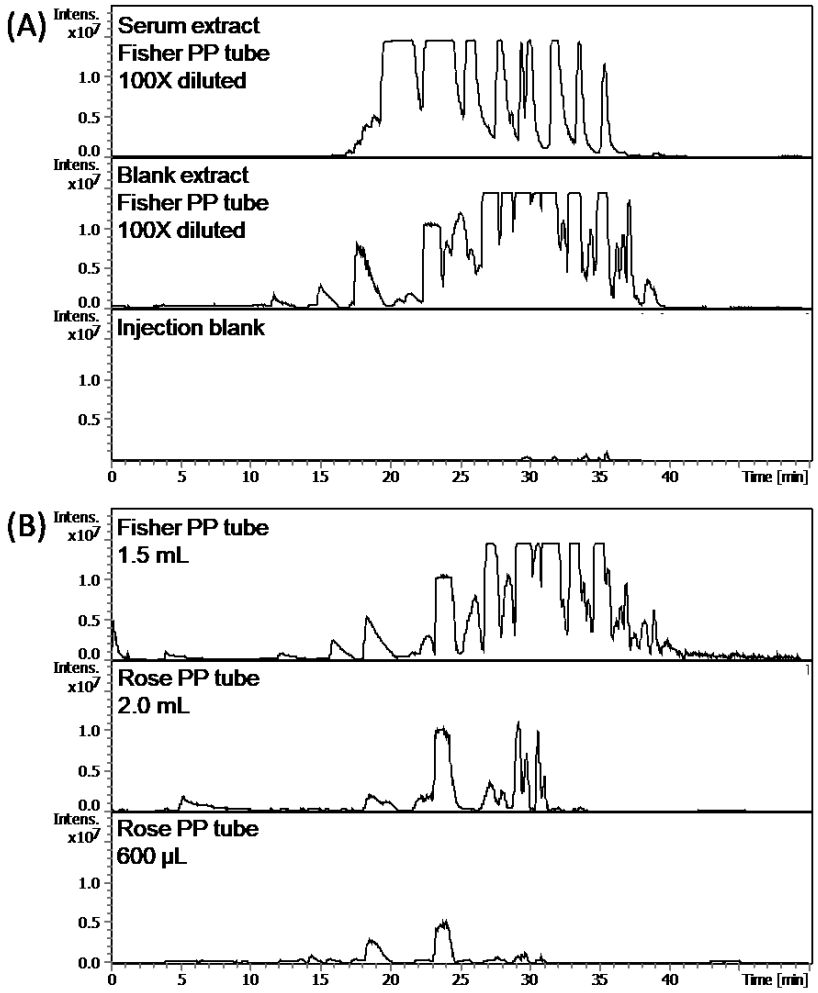

Figure II-4. Evaluation of contamination from extractions with dichloromethane and methanol in polypropylene (PP) microcentrifuge tubes from different brands. After extraction with each type of tube, the resuspended extract was transferred to a PP autosampler insert for injection in all cases. (A) A serum sample from a pig compared to a blank extract (extraction of water instead of the sample) and an injection blank (pure mobile phase); both extractions were performed in 1.5 mL PP microcentrifuge tubes from Fisher Scientific acquired between 2016 and 2017. (B) Extraction blanks obtained with 1.5 mL PP microcentrifuge tubes from Fisher Scientific compared to 2.0 mL and 600 µL PP microcentrifuge tubes from Rose Scientific, acquired between 2016 and 2017.

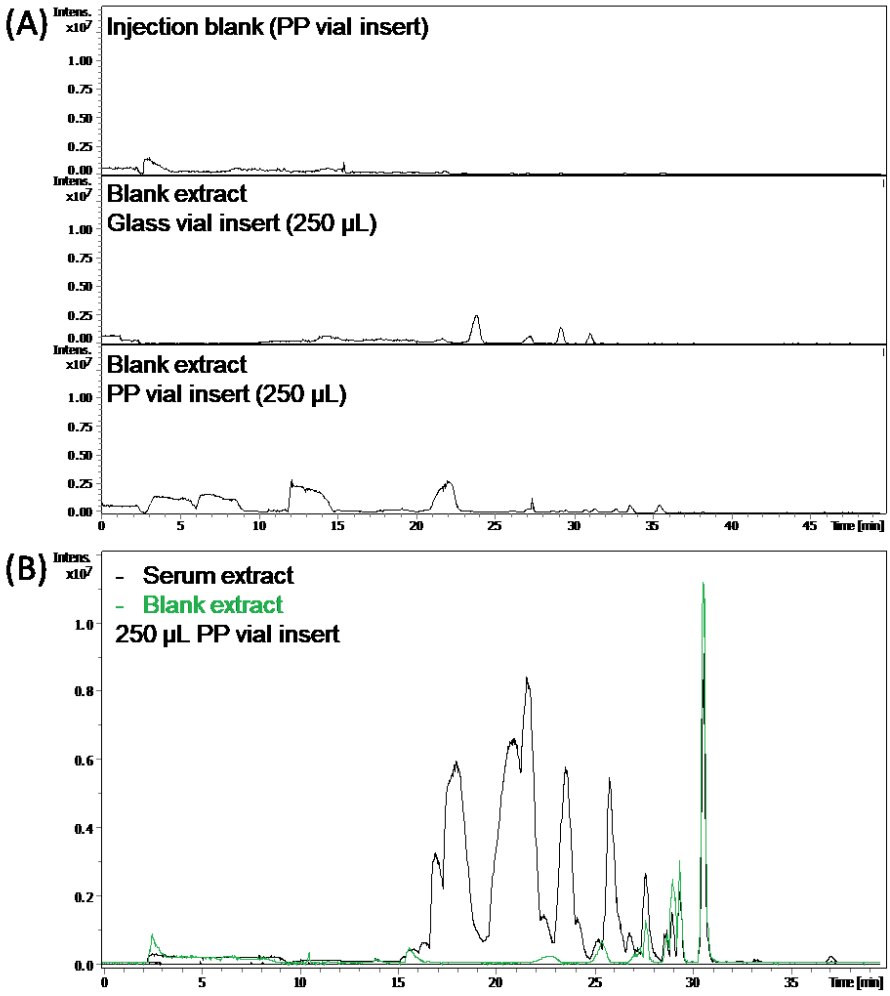

Figure II-5. Evaluation of contamination from extractions with dichloromethane and methanol in PP and glass autosampler inserts. After extraction with each type of tube, the resuspended extract was transferred to a PP autosampler insert for injection in all cases. (A) Chromatograms obtained for extraction blanks performed in PP and glass autosampler inserts compared to a blank injection. (B) Chromatograms obtained for a serum sample from a pig compared to a blank extract, with both extractions and injections performed in PP autosampler inserts.

2.3.4 Re-evaluation of the nanoLC separation method

We observed massive differences for the chromatograms acquired with the optimized plastic ware (Figure A - 26). Furthermore, we chose to re-evaluate the separation method using 14 deuterated lipid standards belonging to different lipid subclasses (Table II-1, Figure A - 27). The standard mixture was used not only for optimization of the separation between lipid subclasses but also as internal standards, *i.e.*, added to each sample before extraction to correct for ion suppression and small differences that may happen during analysis. The gradient and trapping parameters were re-evaluated first (0 to 7.0 µL/min; 0 to 1.5 min , Figure A - 28 to Figure A - 39). The best conditions were found for trapping for 1.25 min at 5.0 µL/min, followed by a 50 min gradient separation (0 min – 0% MPB, 10 min – 30% MPB, 20 min – 70% MPB, 32 min – 95% MPB, 50 min – 95% MPB) and 20 min of re-equilibrium at 100% MPA.

2.3.5 Fine-tuning of the MS acquisition method

The MS acquisition method was fine-tuned for NanoBooster acetonitrile-enriched nitrogen gas pressure (0.10 to 0.30 bar, Figure A - 40 and Figure A - 41), capillary voltage (1200 to 1450 V, Figure A - 42 and Figure A - 43), ion source temperature (190 to 240°C, Figure A - 44 and Figure A - 45) and nitrogen dry gas flow rate (2.5 to 5.0 L/min, Figure A - 46 and Figure A - 47). Each parameter was first re-evaluated with the lipid deuterated standard mix (Table II-1) and then fine-tuned with a pool of serum from 8 pigs. The optimized conditions were: NanoBooster acetonitrile-enriched nitrogen gas pressure of 0.15 bar; capillary voltage of 1375 V; ion source temperature of 190 °C; and nitrogen gas flow rate of 2.5 L/min (Figure A - 48). The number of detected features increased by 44.3%, and we also found higher total intensity (summed intensity for all detected features). When compared to the initial nanoLC-MS method, the number of

detected features increased from 761 to 3532 (Figure II-6), with reduced contamination and better separation of lipid subclasses.

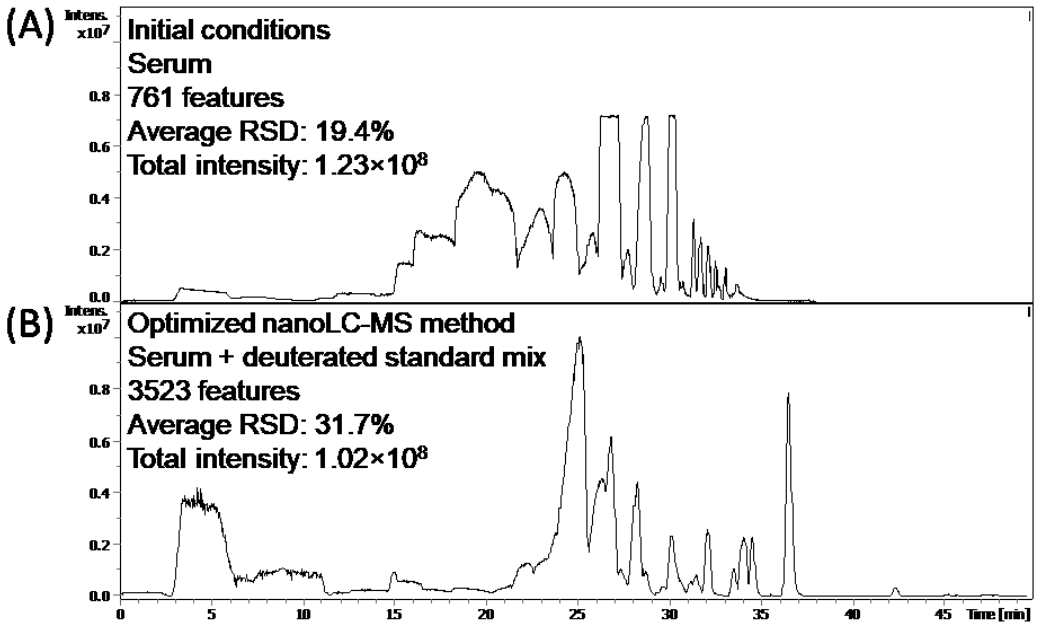

Figure II-6. Base peak chromatograms obtained for (A) a serum sample from a pig with the initial nanoLC-MS and sample extraction methods, before evaluation of contamination sources; and (B) a pool of serum samples from 8 pigs and a mixture of 14 deuterated lipids as internal standards (Table II-1) with the optimized nanoLC-MS method.

2.3.6 Extraction of lipids from serum samples

The Folch method is one of the most popular liquid-liquid extraction procedures for lipids, originally designed for 1 g or more of homogenized tissue. The classical method employs 20 parts of a mixture of chloroform and methanol 2:1 (v/v) for each part of the aqueous sample, followed by a clean-up step with water or a salt solution to achieve a total ratio of 8:4:3 chloroform/ methanol/ water.[62,63] For this work, chloroform was substituted by dichloromethane due to its

lower carcinogenicity risk, lower restrictions to the acquisition and reduced cost.[65,122] The classical method was further modified to handle small volumes of samples. For the initial extraction method, 1.0 µL of serum sample was mixed with 55.4 µL of methanol, 110.8 µL of dichloromethane and 33.4 µL of water. After resting for 10 min at room temperature to reach equilibrium, the mixture was centrifuged for 10 min (12,000 rpm, 4°C) and the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min. The dried extracted was resuspended immediately in 10 µL of 6:4 MPA/MPB and diluted in 90 µL of 9:1 MPA/MPB (100× dilution). It is worth emphasizing that lipids can be easily degraded by oxidation when exposed to oxygen, peroxides, light and heat. Extraction procedures must be optimized in order to avoid excessive degradation, particularly for unsaturated species. Fast methods that minimize the direct contact between sample components and air, acidic conditions, high temperature and light are preferable.[4]

We first optimized the volume of serum for extraction between 1.0 and 5.0 µL (Appendix A - Figure A - 49). The volumes of dichloromethane, methanol and water were not altered for this test, *i.e.*, the proportions between sample and solvents were evaluated. After the extraction, all tested conditions were diluted 100× to the extracted volume of serum, *e.g.*, the 2.5 µL sample aliquot was diluted to a total volume of 250.0 µL after the drying step, whereas the 1.0 µL sample aliquot was diluted to 100.0 µL. The extraction of 2.5 µL of serum provided an increase of 88.7% in detected features when compared to the initial volume of 1.0 µL.

Next, we focused on the composition of solvents for resuspension of the dried extract by remaking it in 40 to 80% of MPA/MPB (v/v), but the best compromise between solubilization of the small lysophosphatidylglycerols and the long-chain, strongly hydrophobic triacylglycerols was achieved by 50% MPA/MPB (v/v, Figure A - 50), which provided the highest total intensity and number of detected features. The dilution solvents were also evaluated, with the best peak shapes

and total intensity for 100% MPA (Figure A - 51). The weak mobile phase composition for dilution provides better conditions for the trapping step, ensuring that the less hydrophobic lipids will be retained while the sample is introduced in the C18 trap column.

The extraction procedure was further optimized for equilibrium time (0 to 15 min, Figure A - 52) and equilibrium temperature (room, 4°C, or -20°C, Figure A - 53). The equilibration between the organic and aqueous phase of the biphasic liquid-liquid extraction before centrifuging is required to achieve the desired partition of lipids and other hydrophilic molecules. The highest number of detected features and lowest average relative standard deviation (relative standard deviation for peak intensities between extraction replicates for each detected feature, averaged for all detected features, taken a measure of reproducibility) were found for 10 min of equilibrium at room temperature. The SpeedVac drying time (15 to 45 min, Figure A - 54) was also evaluated between 15 and 45 min. The organic phase was mainly composed of highly volatile dichloromethane, which was visibly dried within 15 min. However, any moisture or traces of solvents that may remain in the extract can influence lipid resuspension, as well as oxidation processes. We found that the initial drying time of 30 min provided the highest number of detected features and total intensity (Figure A - 55).

Next, we focused on the extraction solvents (Figure II-7, Figure A - 56). The use of methyl t-butyl ether (MTBE) for the extraction of lipids has been previously reported as an alternative to dichloromethane or chloroform.[67] The reproducibility of LLE methods based on dichloromethane or chloroform has been questioned due to the difficulty for selectively removing the bottom organic layer, particularly for samples with high protein content, which form a thick pellet disk between the layers.[67,70,123] The introduction of a pipette tip through the solid pellet to achieve the bottom organic layer, allied to dripping losses of the organic solvents, may result in low

reproducibility if the procedure is not carefully executed by experienced personal. The extraction with MTBE promotes the organic phase to the upper layer, increasing reproducibility while keeping similar recoveries and lipid class coverage.[67,70] The methodology was first described by Matyash *et al.* for 200 µL of serum, 1.5 mL of methanol and 5.0 mL of MTBE, followed by 1 hour incubation period at room temperature in a shaker.[67] However, the sample volume demanded by the procedure is excessively high for studies with limited amounts of biological samples. We adapted the original method for 2.5 µL or serum (2.5 µL of the mixture of deuterated lipids as internal standards - Table II-1, 16.2 µL of methanol, 62.5 µL of MTBE and 15.6 µL of water), but observed slightly higher average relative standard deviations for peak intensities (RSD, *i.e.* lower reproducibility), lower total peak intensity and a longer extraction time when compared to our modified Folch method (1 h incubation in a shaker and 1h30min drying time for MTBE; 30 min drying time and no incubation for the modified or original Folch method, Appendix A - Figure A - 56). A slight decrease in the number of detected features of 4.5% was also observed. The substitution of dichloromethane by MTBE did not achieve the expected reproducibility improvement for extractions replicates, with an average relative standard deviation for peak intensities of 23.3% for MTBE, while the modified and original Folch methods resulted in 20.0 and 21.6%, respectively. We attributed our observations to an increased period for evaporation of MTBE, which has a higher boiling point (55°C *versus* 39.8°C for dichloromethane) and higher solubility of water, summed to a 1-hour incubation period at room temperature. We hypothesize that the small volume of the organic solvents may have favoured the exposure of lipids to sources of oxidation and degradation, *e.g.* air, light, plastic/glassware and sample components, during the long extraction (about 3h for MTBE *versus* 1h for the modified Folch method), mitigating the expected gain in reproducibility. Nevertheless, the longer incubation and drying time reduced the

throughput of sample preparation. When executed carefully, the dichloromethane-based extraction offered a lower relative standard deviation between experimental replicates. The combination of methanol, dichloromethane and water for the extraction of small volumes of serum provided the best sensitivity within a reasonable experimental timeframe.

The optimized extraction procedure was also compared to the results obtained with the original Folch method (Figure II-7, Figure A - 56). The method optimized herein for small volumes of biological fluids resulted in the detection of 15.8% more features with higher reproducibility. The small solvent volumes employed to match the original proportions of the Folch method to 2.5 µL of the sample (2.5 µL of the mixture of deuterated lipids as internal standards - Table II-1, 14.2 µL of methanol, 33.3 µL of dichloromethane and 8.2 µL of water) hamper the pipetting step for the organic layer, particularly for samples with high protein content.

Finally, we also evaluated the benefits of repeating the extraction of the aqueous phase for a second time before the drying step for all three sample preparation methods (Figure II-7). The aqueous phase and protein pellet obtained after the first extraction were re-extracted with the organic solvents, following the same procedure. The organic phases were then combined for drying. The number of detected features and the reproducibility were reduced for the modified Folch method with two combined extractions, probably due to degradation of lipids caused by the longer period required to complete both extractions allied to extra pipetting steps with potential dripping losses (Figure II-7). Even though the MTBE method provided slightly more detected features with the combined two extractions (3791 detected features) when compared to the modified Folch method with one unique extraction (3650 detected features), the time required for sample preparation was unreasonable (about 4.5h for MTBE with two combined extractions *versus* 1h for modified Folch with one extraction).

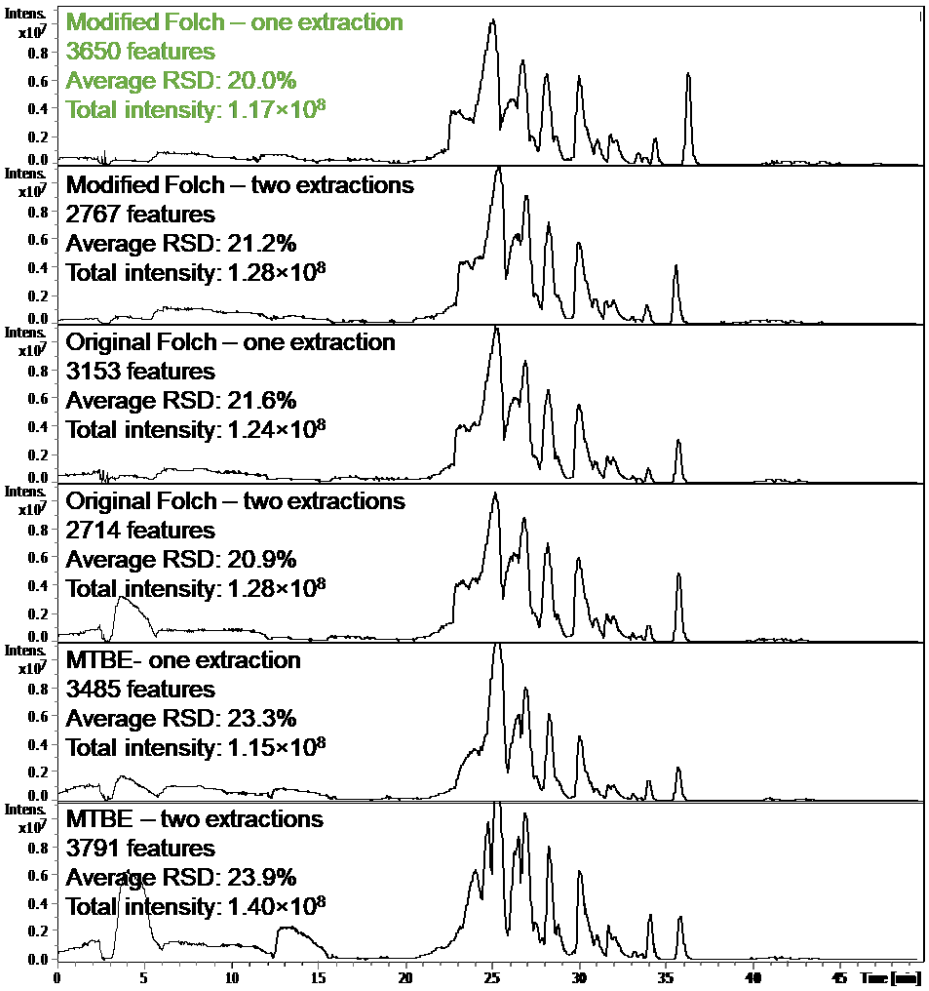

Figure II-7. Evaluation of repeating the extraction procedure for the aqueous phase using the original and modified Folch methods, as well as the MTBE method described by Matyash *et al.*, but adapted for 2.5 µL of serum.[62,63,67,70] Modified Folch method: 2.5 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 53.3 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 31.9 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 10.0 µL of 1:1 MPA/MPB and diluted with 90.0 µL of 9:1 MPA/MPB. Folch method: the

same procedure was followed, but with 2.5 µL of serum, 2.5 µL of the mixture of internal standards, 14.2 µL of methanol, 33.3 µL of dichloromethane and 8.2 µL of water. Adapted MTBE method: 2.5 µL of serum, 2.5 µL of the mixture of internal standards, 16.2 µL of methanol and 62.5 µL of MTBE were incubated in a shaker for 1 h at room temperature, followed by vortex with 15.6 µL of water; the remaining steps were identical to the modified Folch method, except for drying time of 1h30min. For two extractions, the exact same procedures were followed, but remaining aqueous phase and protein pellet were re-extracted with methanol and dichloromethane; or methanol and MTBE. The organic phases were combined for drying.

Finally, the sample dilution was evaluated between 10 and 40× (Figure A - 57), with best results obtained with 10× (extraction of 2.5 µL of sample with final dilution of the dried extract to 25.0 µL), which allowed the detection of 3602 features for positive ionization and 11 internal standards from different lipid classes (LPC, LPE, SM, PC, PE, PS, PG, PI, DG, TG and CholE, Table II-1 and Table II-2). However, we were able to detect over 3000 features at 40× dilution (extraction of 2.5 µL of serum sample with final dilution of the dried extract to 100.0 µL), *i.e.*, the method can be easily adapted to smaller sample volumes without substantial sensitivity losses for applications with more restricted sample amounts. The application of the optimized method to 1.0 µL of microdialysate fluid, as well as 2.5 µL of serum and cerebrospinal fluid, are presented in Chapter III.

A chromatogram obtained with the optimized sample extraction procedure is presented in Figure II-8B. The optimization of the sample extraction method led to an increase of 81.8% in detected features (from 1981 to 3602 features, Figure II-8). For the optimized method, an aliquot of 2.5 µL of serum was vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipids (internal

standards, Table II-1) and 53.2 µL of methanol. Next, 110.8 µL of dichloromethane was added, followed by vortex for 20 s. The cleanup step was performed with 32.0 µL of water and 10 s vortex. The biphasic mixture rested at room temperature for 10 min and was centrifuged for 10 min at 12,000 rpm and 4°C. An aliquot of 105.3 µL of the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min. The dried extract was resuspended with 2.5 µL of 1:1 MPA/MPB, vortexed for 60 s and diluted with 22.5 µL of MPA. After preparation, the resuspended extracts were kept at 4°C for a maximum of 48 h; or at -20°C under a nitrogen atmosphere for up to 21 days. It is worth noticing that unsaturated lipids can be easily oxidized when exposed to air or light; therefore, the dried extract must be either immediately resuspended and analyzed, or properly stored in a non-oxidant atmosphere at low temperature.

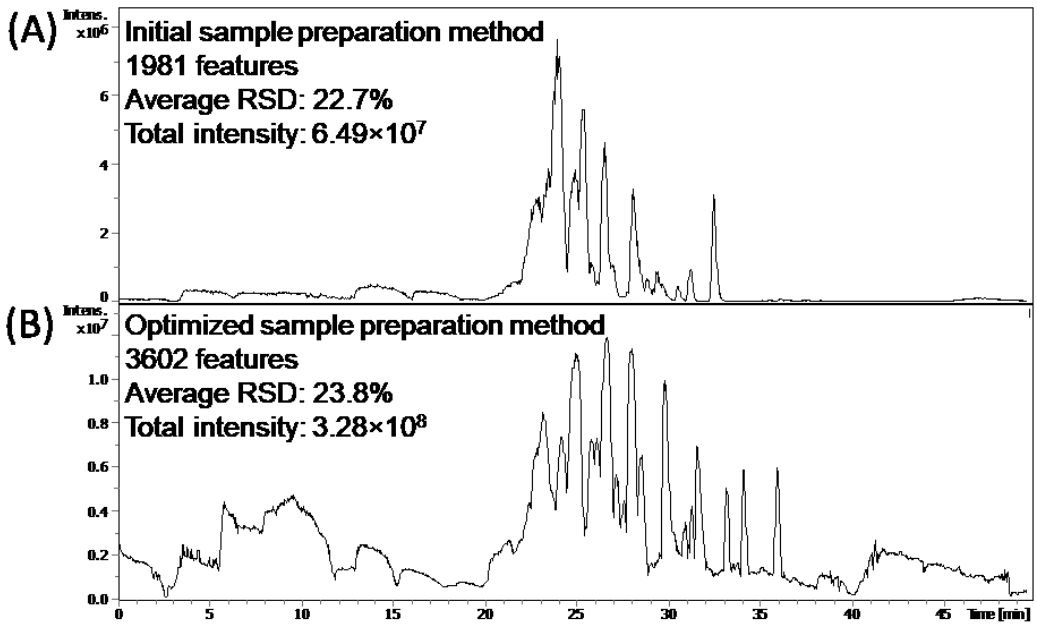

Figure II-8. Base peak chromatograms obtained for a pool of 8 serum samples from pigs using (A) the initial sample preparation method; and (B) the optimized sample preparation method.

2.3.7 Injection volume

The small internal volume of the nanoLC analytical capillary column usually restricts the amount of sample that can be injected without overloading to less than 1.0 µL. However, the initial trapping step adopted and optimized herein allows the injection of larger volumes, since the most hydrophilic components and solvent excess are directed to a waste line, whereas the more hydrophobic compounds are retained by the C18 trap column as a narrower sample band. The optimization of the injection volume guarantees optimal sensitivity while avoiding overloading the column, which could lead to preferential retention of hydrophobic compounds and distorted peaks. The injection volume was evaluated between 1.0 and 4.0 µL (sample loop size of 5.0 µL), and we detected the highest number of features with 4.0 µL injection (Figure A - 58).

2.3.8 *Data processing*

Different parameters were evaluated for peak picking and alignment of the chromatograms using MetaboScape 4.0 (Bruker Daltonics), namely: internal re-calibration masses (background ions detected for all injections between 13.5 and 18.5 min for positive ionization, and 18.2 and 20.2 min for negative ionization), minimum intensity (3000 cts), correlation coefficient (0.8), retention time tolerance (60 s), m/z tolerance (5.0 mDa) and minimum compound length for alignment (8 spectra). The characteristic low flowrate of nanoLC can result in higher variation of retention times and broader chromatographic peaks, which requires higher retention time tolerances for alignment. However, the high mass accuracy achieved after internal re-calibration allows for narrow m/z tolerances.

Statistical models used to evaluate the data cannot handle missing values; hence, we employed recursive extraction for features that were found for at least 10% of injections with the

alignment parameters, *i.e.* they were searched again in the raw chromatograms with a less strict minimum peak length of 4 spectra and no minimum intensity threshold. The aligned data was filtered by detection in at least 80% of injections within each group. A combination of positive and negative ion modes is required for full profiling of the lipid composition of biological fluids; therefore, features detected for positive and negative ionization were merged into a unique list with tolerances of 10 ppm and 30 s.

2.3.9 Final nanoLC-MS method

The last method optimization step consisted of fine-tuning the gradient and re-equilibrium of the trap and analytical columns. We previously employed a 20 min equilibrium run between injections, but the total analysis time of 70 min (50 min analytical run, 20 min equilibrium) reduced the sample throughput. To decrease the equilibrium time while still ensuring good separation, we took advantage of the long dead time of nanoLC separations by starting the equilibration while still acquiring the sample injection. Although the dead time after an injection is below 3 min due to trapping, it increases to about 8 min with the optimized analytical nanoLC flow of 300 nL/min. We used the 8 min dead time to start the re-equilibrium of the columns at 42 min (100% MPA), whereas data acquisition for the sample was finished at 50 min. The re-equilibrium run employed between sample injections was further reduced to 5 min at 100% MPA, resulting in a total analysis time of 56.25 min, including the trapping step. Although we acknowledge that 56.25 min per sample is usually not seen as high throughput, it still represents a great improvement when compared to usual lipidomics nanoLC methods, which can take 1.5 to 3 hours per sample injection.[108,124,125] A comparison between the initial method for a serum sample from a pig and the optimized chromatograms for a pool of serum samples from 8 pigs and a pool of serum from 100

healthy humans is presented in Figure II-9. The extracted ion chromatograms for the internal standards (Table II-1) are shown in Appendix A - Figure A - 59, while the blank extract for the optimized method is presented in Figure A - 60.

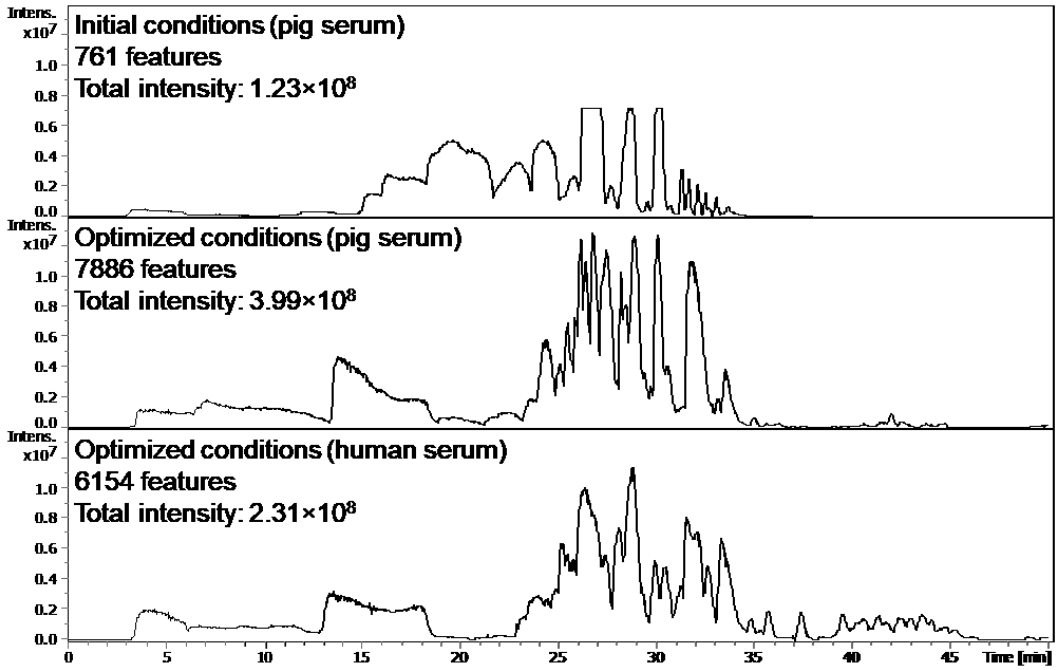

Figure II-9. Comparison between the initial nanoLC-MS and sample preparation method (top), the optimized methods for a pool of serum from 8 pigs (mid) and a pool of serum from 100 healthy humans (bottom chromatogram). Optimized nanoLC method: trapping at 5.0 µL/min for 1.25 min (100% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 50 min gradient (0 min – 0% MPB, 10 min – 30% MPB, 20 min – 70% MPB, 25 min – 80% MPB, 30 min – 95% MPB, 40 min – 95% MPB, 42 min – 0% MPB, 50 min – 0% MPB), 5 min equilibrium (100% MPA); 4 µL injection. Optimized MS method: electrospray ion source capillary voltage of 1375 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.15 bar, dry nitrogen gas flow rate of 2.5 L/min, ion source temperature of 190ºC. Optimized sample preparation method: 2.5 µL of a

pool of blood serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 53.3 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 31.9 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 2.5 µL of 1:1 MPA/MPB and diluted with 22.5 µL of 9:1 MPA/MPB.

2.3.10 Evaluation of the optimized method

Although the positive ion mode of ESI is more commonly employed, some polar classes of lipids are better ionized under negative ionization, *e.g.*, PI, PS and PG (Table II-2). Hence, a combination of positive and negative ionization is required for the comprehensive profiling of a biological sample. We merged the positive and negative aligned feature list with a tolerance of 10 ppm and 30 s. The optimized method provided highly sensitive detection of 9900 to 12200 features by employing only 1.0 to 2.5 µL of serum samples from healthy humans (Figure A - 61). Similar conditions applied to a UHPLC-MS system allowed the detection of less than a third of the features, proving that the nanoLC setup is appropriate for highly sensitive analysis of biological fluids (Figure II-10, Figure A - 62 and Figure A - 63). The number of detected features with the optimized nanoLC-MS conditions could not be achieved with UHPLC-MS, even without sample dilution, *i.e.*, extraction of 25.0 µL of sample and final dilution of the dried extract to 25.0 µL. The method was also applied to 2.5 µL of a pool of cerebrospinal fluid samples from 8 pigs, with the detection of over 13000 features (Figure A - 64).

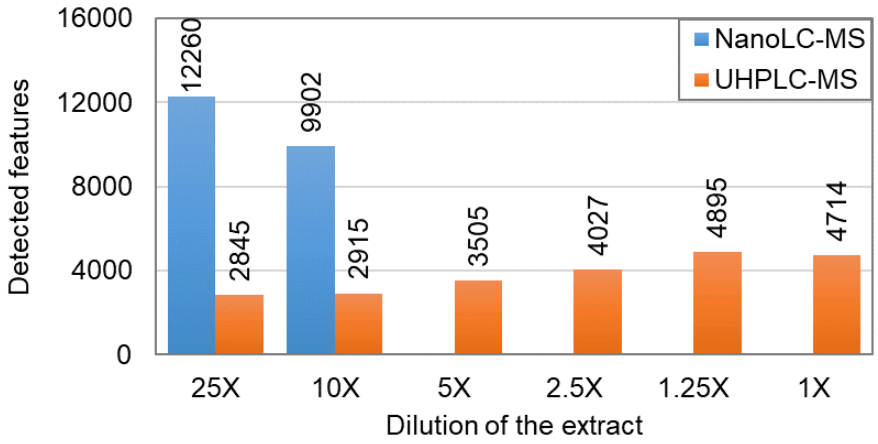

Figure II-10. Comparison between detected features for the optimized nanoLC-MS method and similar conditions applied to a previously optimized UHPLC-MS method.

The detected features for 2.5 µL of a pool of samples from 100 healthy humans (10× dilution) were employed for MS/MS identification. A total of 211 features were matched to known lipids found in the MSDial Lipid Blast, Human Metabolome Database (HMDB) and MassBank of North America LC-MS/MS libraries.[73,78,82,83,109] The remaining 9654 features were inputted on the LipidMaps database for accurate mass search with a tolerance of 5.0 mDa and 5611 were putatively matched to biologically relevant lipids (Supp. Table 1) after application of the developed 5-tier filtering and scoring approach to select isomeric or isobaric identification possibilities, as described in the Experimental section (*2.2.7. Putative identification of lipids*, p. 50).[77] The identification of 5842 lipids by MS/MS and accurate mass-match resulted in high coverage of 36 different lipid subclasses belonging to six categories (sphingolipids, glycerolipids, glycerophospholipids, fatty acyls, sterol lipids and others, Table II-2 and Figure II-11). Our future work will focus on improving the MS/MS fragmentation to extend the number of high-confidence identifications.

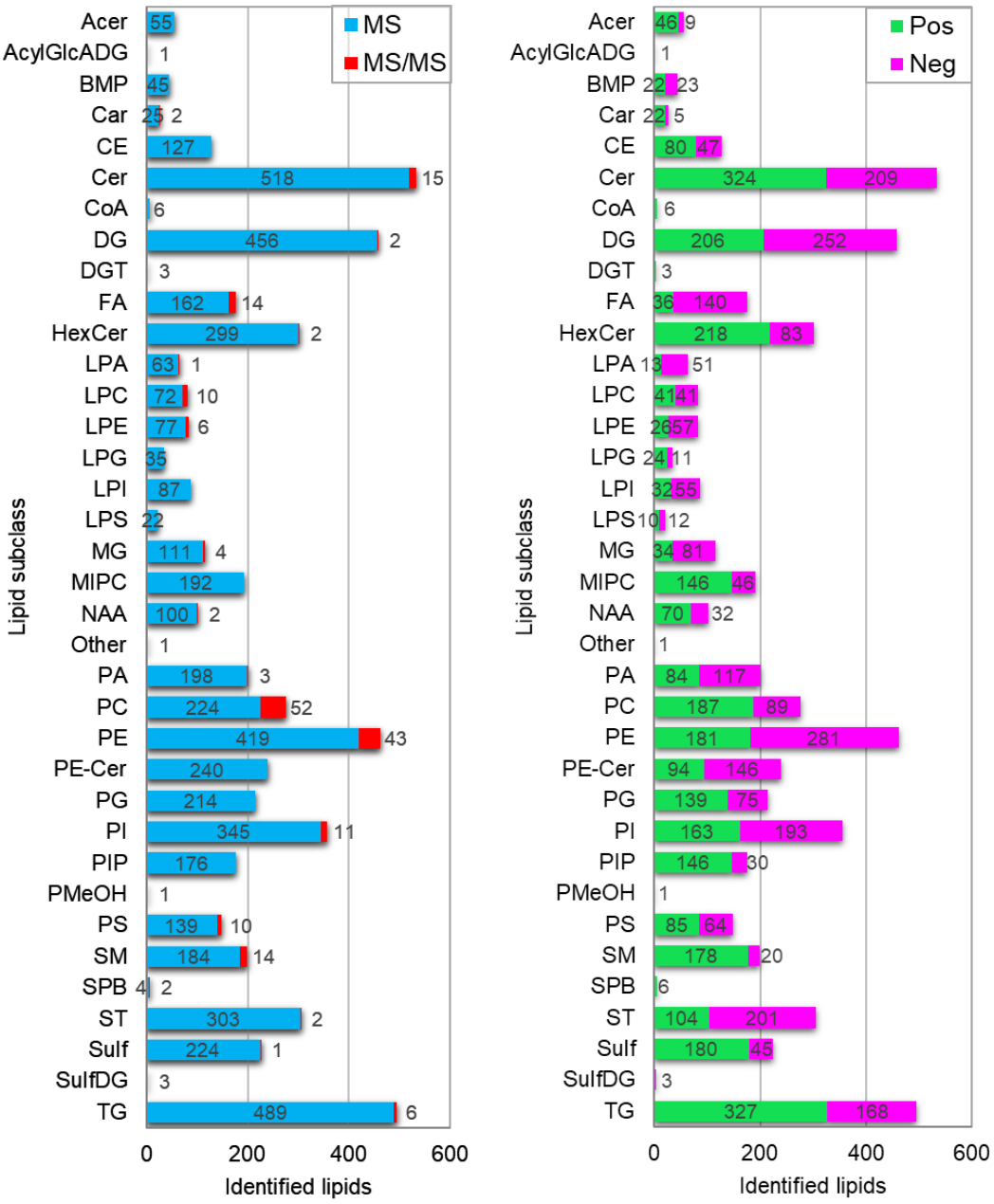

Figure II-11. Subclasses for the 5842 lipids identified for a pool of serum from 100 healthy humans with the optimized nanoLC-MS and sample preparation methods, divided into the level of identification (MS/MS - blue, or accurate mass match, MS - red) and ionization polarity (positive – green, or negative - magenta). Putative identifications were filtered and ranked according to the

5-tier system described in the Experimental section (*2.2.7. Putative identification of lipids*, p. 50). Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

Although a high coverage of lipids was achieved, the complexation of lipids that contain free phosphatidyl and carbonyl groups with metal surfaces in the sample path (injection needle, tubing, ion source, etc.) caused noticeable peak broadening and tailing, particularly for the more polar lysophospholipids and phosphatidic acids. The reversible complexation of the electron-donating phosphatidyl group by metal ions, *e.g.*, Fe(III), is a well-known issue in the analysis of lipids by LC-MS. [120,126] Since the use of chelators in the mobile phase is not advisable for nanoLC-MS due to the possible damage to the instruments and columns, phosphoric acid has been previously described as a possible additive for mobile phase and sample resuspension to improve peak shapes. However, it was not employed herein due to the possibility of corrosion of the delicate nanoESI ion source and emitter tip, as well as clogging and ion suppression.[108,127,128] As an alternative, ammonium bicarbonate buffers have also been shown to mitigate peak tailing, but they can easily decompose into ammonia, CO_2 and water, creating air bubbles in the system. [129] Our goal was to develop a robust method suitable for comprehensive lipidomic analysis, but the use of such additives would increase the need for constant maintenance. Nevertheless, we have optimized a robust, high coverage method for the analysis of very small amounts of diluted biological samples (1.0 to 2.5 µL). We are currently working to improve the MS/MS fragmentation and identification of the detected features.

4. Conclusions

This work described the development and optimization of a nanoLC-MS method for the untargeted, comprehensive lipidomics of small volumes of biological samples. The extremely high sensitivity of nanoLC is well known; however, the low robustness and constant instrumental issues of most methods prevent its wide application to lipidomics. The method development described herein focused on robustness, sensitivity and reproducibility, providing appropriate conditions for routine applications with low amounts of diluted samples. We also described the optimization of a lipid extraction procedure for low volumes of samples (1.0 to 2.5 µL), including an examination of contamination sources. The identification of 5842 lipids from 36 different subclasses was achieved with only 2.5 µL of blood serum. While improved sensitivity by nanoLC is well recognized, the work described herein demonstrates the possibility of achieving reproducible conditions for the analysis of biological samples with comprehensive lipidomic coverage. The sample preparation can be completed within 1h for a maximum of two samples or 1h30min for groups of up to 10 samples, whereas the nanoLC-MS analysis takes 56.25 min per injection. The application of the developed method to different types of diluted samples for biomarker discovery and biological research is described in Chapters II and III. Further applications have been performed but are currently under data processing.

5. Acknowledgments

We thank Jaspaul Tatlay for training in the initial operation of the nanoLC-MS instruments, Dr. Randy Whittal (University of Alberta, Edmonton, AB, Canada) for sharing his extensive knowledge in nanoLC maintenance, and Dr. Ulrike Schweiger Hufnagel and Dr. Aiko Barsch (Bruker Daltonik GmbH, Bremen, Germany) for their assistance in MetaboScape data analysis.

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research, the Canada Research Chairs program, Brain Canada, MITACS, Canada Foundation for Innovation, Genome Canada and Alberta Innovates.

III

Chapter III: Lipidomics of Spinal Cord Injury: Pilot Study[5]

3.1 Introduction

The spinal cord is composed by nervous tissue enclosed by the vertebral column that extends from the brainstem to the lumbar region. The tubular structure acts as a communication pathway between the brain and the body. Nerve signals from the motor cortex travel through the spinal cord to reach the peripheral nervous system, allowing movement control, sensory function and reflexes. The human spinal cord is composed by a peripheral layer of white matter, containing nerve fibers or myelinated axons, three internal layers of grey matter made of nerve cells, and a central canal of cerebrospinal fluid (CSF) protected by meninges, *viz.* dura mater, arachnoid mater and pia mater. The subarachnoid space between the arachnoid and the pia matter also contains CSF, which provides mechanical and immunological protection, as well as regulation of cerebral blood flow, homeostasis and waste management for the neural tissue.[130,131]

A spinal cord injury (SCI) is characterized by a partial or complete rupture of the spinal cord that may cause loss of sensory function and muscle control below the injury site. The position and extend of the injury determines which parts of the body are affected, *i.e.*, a lesion on the top cervical segments may cause total paralysis of the body, whereas lesions in the lumbar region are related to loss of movement and sensation in the lower limbs. A SCI may also affect autonomic

[5] A collaboration between Dr. Brian Kwon (University of British Columbia, Canada), the Rick Hansen Institute (Vancouver, Canada) and Dr. Liang Li was established for this study. Animal handling and sample collection was performed under the supervision of Dr. Brian Kwon.

functions, such as control of respiratory movements, heart rate, blood pressure and body temperature, being potentially fatal without immediate medical care. [131–133]

The primary injury caused by the mechanical trauma upon the delicate spinal cord may cause immediate loss of function by contusion, compression, maceration, acute stretch, shearing and laceration of the nervous tissue.[134] Within a few minutes, a cascade of biological processes will lead to secondary injuries, causing further cell death and loss of neuronal tissue, including hemorrhage, ischemia, necrosis, inflammation and swelling.[135] While no effective, specific treatment is currently available for SCI, the mortality risk increases with the injury level and severity, but is mostly related to immediate, high-quality medical assessment to ensure immobilization, adequate spinal cord perfusion, management of excessive inflammation and life support.[136] Glucocorticoids (*e.g.*, methylprednisolone) have been historically used to control inflammation and as neuroprotective agents, although their efficacy for SCI, allied to severe adverse effects, is still highly controversial.[135,137,138] Surgical interventions may be also required to align the column and decompress the affected area. Unfortunately, the spinal cord has a poor remyelination potential, limited plasticity and low possibility of axon regrowth, leading to a small chance of full spontaneous recovery once the neuronal tissue is lost. SCI patients may show different levels of neurological recovery that is usually more pronounced within the first 6 months after the injury, although improvements may occur for up to 5 years. However, the prognosis for recovery is uncertain and affected by many factors, such as the injury level and availability of prompt specialized medical care. The most severe patients are completely dependent on caregivers and assistive technology. [132,133,135,139–141]

The devastating physical, social and economical consequences for patients and their families emphasize the need for new therapies and biomarkers for better management of the

condition. The patients' health and welfare are highly dependent on the initial injury assessment and treatment; yet, no biomarkers or laboratory-based tests have been approved by the medical and scientific community for a quick and accurate SCI diagnosis and prediction of neurological recovery. The injury is usually diagnosed through a combination of imaging techniques (plain X-ray, computerized tomography - CT scans, magnetic resonance imaging - MRI) and clinical examination (sensation to pinprick and touch, muscle strength). X-rays cannot show the SCI but may be useful to find fractures of the cervical spine. CT scans and MRIs are highly recommended to diagnose and assess the injury but depend on the availability of expensive instrumentation and trained personal, often requiring transfers of patients with possible unstable spines. Biomarkers could provide a safe and accurate measure of the extent and severity of the primary injury, as well as prediction of neurological recovery. Although there is profuse ongoing research on biomarker candidates, their usefulness for SCI assessment has not been established by the scientific and medical community. A few proteomic and inflammatory biomarker candidates have been previously suggested, but large population studies and clinical validation are yet to be performed.[135,136,138,142–144]

The currently available treatments for SCI are limited by the absence of specific drugs, but new clinical trials are compromised by the lack of accurate statistical data, unpredictable neurological recovery and unsuitable animal models. The inaccurate diagnosis and unpredictable spontaneous neurological recovery hamper the statistical evaluation of clinical trials, requiring large numbers of patients that are not easily recruited. Nevertheless, the biomechanical variability of human SCI requires the use of animal models for pre-clinical studies to reduce the heterogeneity, which are often performed with rodents. Over the past few decades, many promising therapies were successful in rodent model studies, but their efficacy was not translated to human clinical

trials. [134,135,138] Rodents are inexpensive models that require only basic housing facilities, easily available to researchers. Hence, rats are commonly used to study human conditions but differences in size, anatomy, metabolism and recovery potential impacts the translation between the animal model and human trials. Ideally, animal models should resemble humans as much as possible, but handling costs and complications must also be considered. Although primates are highly similar to humans, the high operational costs and large specialized facilities required to handle the animals are inaccessible for most research groups. [131–133,141,142] Recently, Yucatan miniature pigs (20 – 25 kg) have been proposed as a plausible alternative to primates due to comparable spinal cord size, similar physiology and lower associated costs, but further studies are required to prove the possibility of translation to human trials.[145]

Biomarkers that correctly reflect the extent of damage to the spinal cord and predict neurological recovery are required to improve SCI diagnosis and assessment, as well as the outcome of new clinical trials.[142] Lipids are a diverse class of biomolecules involved in an assortment of processes.[15] Inflammation, a hallmark for secondary injuries after a spinal cord trauma, is closely related to lipid signalling and oxidation; hence, the application of lipidomics to study SCI is promissing.[146,147] Nevertheless, there is still a need of developing highly sensitive analytical tools for profiling lipids with high coverage in limited amounts of samples. NanoLC-MS offers high sensitivity for reduced volumes of diluted samples, being ideal for handling small amounts of biological fluids. This work describes a pilot study for the evaluation of the lipidic composition of fluids obtained from Yucatan miniature pigs as animal model candidates after a controlled spinal cord injury. This work is part of a consortium between the University of Alberta (Edmonton, AB, Canada), the University of British Columbia (Vancouver, BC, Canada) and the Rick Hansen Institute (Vancouver, BC, Canada) to study biomarkers of spinal cord injury through

lipidomics, metabolomics, genomics and proteomics of spinal cord injuries in humans and mini-pigs as animal models.[134,145,148] We employed a previously developed nanoLC-MS method (Chapter II) for untargeted, comprehensive lipidomics of low volumes of blood serum, CSF and intraparenchymal microdialysate fluid (MD) obtained from the interstitial region around the spinal cord after a controlled injury. Although serum sampling is less invasive, CSF may provide a better representation of the pathology as it flows through the spinal cord structure.[147,149] The parenchymal microdialysate obtained directly from the injury site may offer further essential information; however, the small volumes usually available (< 10 μL) hampers its investigation by traditional techniques.[134,142] We compared the three types of biological fluids collected at different time points after the controlled injury to investigate the usefulness of untargeted lipidomics by nanoLC-MS to study the condition. The main objective of this work was not biomarker discovery, but the evaluation of the developed nanoLC-MS method (Chapter II) to assess the lipidome composition of different types of samples for future, large-scale studies. The use of reduced volumes of biological fluids may allow the application of other *omic* technologies, *e.g*. metabolomics, proteomics and genomics, to further characterize the composition of the samples, which amplifies the possibilities for finding biomarker candidates and molecules that may help unravelling the biological processes associated with SCI.

3.2 Experimental

3.2.1 Animal model

Blood serum, cerebrospinal fluid (CSF) and intraparenchymal microdialysate fluid (MD) were collected from 4 Yucatan miniature pigs (20-25 kg) that suffered a controlled spinal cord injury. The animal model and controlled injury were previously described elsewhere.[134,145] Briefly,

a T10/T11 laminectomy was performed to expose the spinal cord. An impact weigh of 50 g was dropped directly onto the exposed spinal cord from a defined height (20 cm for serum and CSF; 50 cm for MD) to generate the lesion. Immediately after the contusion, a compression weight of 150 g was placed on top of the impactor for 5 min (blood serum and CSF) or 60 min (MD) to simulate sustained compression.[134,145] Blood serum and CSF samples were collected through subcutaneous access ports for two of the animals at 0h (pre-injury), 24h, 72h and 168h after the lesion (Table III-1). The remaining animals were kept under anesthesia for collection of MD at 4.75, 6.75, 8.75 and 10.75 h post-injury, followed by euthanasia. Four microdialysate probes were placed in the interstitial region around the spinal cord (about 15 and 35 mm cranial and caudal to the impact site). A perfusate isotonic solution with composition similar to CSF (147 mM NaCl, 2.7 mM KCl, 1.2 mM $CaCl_2$, 0.85 mM $MgCl_2$) was flowed through the probe at a rate of 0.5 µL/min to a total sample volume of 7 µL. All collected samples were immediately frozen and kept under -80°C until analysis. Samples were concomitantly collected for other related applications.[134,145,148,150]

Table III-1. Sample set employed for the pilot study on spinal cord injury (N = 2).

Fluid	Time point 1	Time point 2	Time point 3	Time point 4
Serum	0h	24h	72h	168h
Cerebrospinal fluid (CSF)	0h	24h	72h	168h
Intraparenchymal microdialysate (MD)	4.75h	6.75h	8.75h	10.75h

3.2.2 Sample extraction and nanoLC-MS analysis

A modified Folch method for liquid-liquid extraction of lipids by dichloromethane and methanol was previously optimized (Chapter II). Briefly, samples were thawed for 1h at 4°C and vortexed for 1 min. For serum and CSF, aliquots of 2.5 µL of biological fluid (serum or CSF) and 2.5 µL of an internal standard mixture composed by 14 deuterated lipids (Splash Lipidomix Mass Spec Standard, Avanti Polar Lipids, Table II-1, p. 45) were mixed with 53.2 µL of methanol, 110.8 µL of dichloromethane and 32.0 µL of water, with vortex for 10 to 20 s between additions. After resting for 10 min at room temperature, the mixture was centrifuged for 10 min (12,000 rpm, 4°C) and the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min. The dried extracted was resuspended immediately in 2.4 µL of 1:1 mobile phase A (MPA) / mobile phase B (MPB), vortexed for 30 s, and then diluted in 22.6 µL of MPA. For MD, each sample had a total volume of approximately 7 µL. Hence, a similar procedure was followed, but due to the reduced sample volumes, aliquots of 1.0 µL of sample and 1.0 µL of internal standard mixture (Table II-1, p. 45) were extracted to allow for further analysis procedures. All samples were extracted and stored in 250 µL polypropylene inserts (Canadian Life Science, Peterborough, ON, Canada) placed inside amber injection vials capped with PTFE-lined caps (Agilent Technologies, Santa Clara, CA, USA). A detailed discussion on sources of contamination and extraction vessels can be found in Chapter II (*2.3.3. Evaluation of contamination sources*, p. 63). The extracts were stored at 4°C for a maximum of 4 days before injection. Since the main goal of the experiment was to reduce the required sample volume, we opted for employing a pooled mixture of serum from 100 healthy individuals for quality control. However, features that were detected only for the QC injections were excluded before statistical analysis as they were not found for the samples.

The extracts were analyzed by nanoLC-ESI-QToF-MS with the optimized method described in Chapter II. Briefly, the method conditions included: MPA - 10 mM ammonium formate in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB - 10 mM ammonium formate in 95:5 isopropanol/ water; trapping for 1.25 min at 5 µL/min (100% MPA); 300 nL/min; 45°C; 50 min gradient (0 min – 0% MPB; 10 min – 30% MPB; 20 min – 70% MPB; 32 min – 95% MPB; 50 min – 95% MPB) followed by a 20 min run for re-equilibrium (0 min – 90% MPB; 1 min – 100% MPA; 20 min – 100% MPA); and 4.0 µL injection. The instrumentation and reagents were described in Chapter II (*2.2.1. Chemicals and reagents* and *2.2.2. Instrumentation*, p. 42). All experiments were performed with extraction duplicates. Positive and negative ionization were acquired in separate, consecutive injections. Quality control (QC) was performed through injections of two aliquots of extracted pooled serum sample (a mixture of serum obtained from 100 healthy individuals), *viz.* three QC injections before and three injections after all sample extracts, plus one QC injection after every 8 sample injections.

3.2.3 Data processing

The detected features were re-calibrated, selected and aligned by the software MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA). Alignment parameters included a minimum peak intensity (*i.e.*, peak height) threshold of 2000 cts, minimum peak length of 22 spectra, mass recalibration with background contaminants detected in all injections, retention time tolerance of 60s and m/z tolerance of 5.0 mDa. Missing values were substituted by recursive extraction of the raw data, during which the minimum intensity threshold was removed and a minimum peak length of 11 spectra was applied. All experiments were performed in duplicates and only features detected in more than 80% of injections for at least one of the time points were

considered for statistical analysis. Positive and negative ionization results were merged with a retention time tolerance of 30 s and m/z tolerance of 10 ppm. The remaining missing values for features detected in at least 50% of injections in each time group were replaced by the minimum within-group intensity. However, missing values for features not detected in at least 50% of injections within a time group were substituted by the minimum global intensity detected for all injections.

3.2.4 Identification: tandem-MS and accurate mass

The QC samples were further analyzed by CID-QToF *tandem*-MS (MS/MS) for positive identification. The collision energy for fragmentation was stepped between 15 (150 m/z) and 45 eV (1000-1500 m/z). Precursor ions were selected for fragmentation by the software oToF Control (Bruker Daltonics) in auto-MS/MS mode, *i.e.*, the most intense ions are fragmented by CID in each MS-MS/MS cycle. However, the less intense ions cannot be fragmented as the number of precursors chosen for CID is limited by the number of scans that can be performed within the time range of a chromatographic peak, which may contain tens of species. In order to fragment a higher number of precursor ions, the mass range for precursor selection was reduced to 100 Da intervals, *i.e.,* one injection with precursors limited to m/z 150-250, a second injection for m/z 250-350, etc. All the collected MS/MS spectra were matched to the aligned peak list with retention time tolerance of 15 to 60 s and m/z tolerance of 5.0 mDa. The MS/MS fragmentation spectra were searched on the MSDial LipidBlast[73,78], Human Metabolome Database[79,81,109,151] and MassBank of North America LC-MS/MS libraries[82,83] (m/z tolerance of 5.0 mDa) in two identification tiers: tier 1 contained molecule identified with MS/MS score higher than 500, while tier 2 included molecules with mSigma (isotope pattern match) smaller than 50 and MS/MS score between 100

and 500. The identification procedure employed for this work followed the guidelines of the Lipidomics Standards Initiative (https://lipidomics-standards-initiative.org).[20,76] The position of unsaturation and glycerol derivatives were not determined.

Features that could not be identified by MS/MS match were inputted into the LipidMaps database for accurate mass identification with mass-to-charge (m/z) tolerance of 5.0 mDa (https://www.lipidmaps.org).[17,75] Lipids are a broad class of molecules that can have many isomers and isobars. Hence, each detected feature, characterized by a combination of m/z and retention time, can have multiple mass-based putative identifications. We employed a six-tier filtering and scoring approach to determine the most likely putative identification to determine the lipid subclass. The five-tier approach described in Chapter II (*2.2.7. Putative identification of lipids*, p. 50) was further improved with a sixth scoring tier to consider the sensitivity of the method for each lipid subclass and its expected concentrations in biological fluids. Isomeric possibilities that are not easily detected by nanoLC-MS or not commonly found in biological fluids were given higher scores, whereas the most likely detection possibilities received lower scores. After summing the scores for each step, the isomeric or isobaric identification for each feature with the lowest total score was chosen to determine the lipid subclass. All the possibilities that passed the retention time and adduct filtering were still kept for further reference, ordered from the lowest to the highest total score.

3.2.5 Normalization and statistical analysis

The identified lipids were matched to one of the 14 deuterated internal standards (Table II-1, p. 45) according to lipid subclass and retention time similarity. The peak intensity (*i.e.* peak height) of each identified lipid was divided by the intensity of the matched internal standard for

normalization of ion suppression and small differences that can occur during sample handling. The structural similarity between the deuterated standards and the matched lipids, eluting within the same retention time windows, ensure that ion suppression effects are mitigated for reliable relative comparisons. The normalization procedure employed herein is widely used by the lipidomics community and recommended by the Lipidomics Standards Initiative. Further information on normalization of lipids can be found in Chapter I (*1.3.1.6. Ion suppression and normalization approaches*, p. 29).[12,152]

The normalized intensity ratios (intensity of feature / intensity of internal standard) were used for uni- and multivariate statistical analyses on MetaboAnalyst 4.0 (https://www.metaboanalyst.ca).[153] The web-based platform allows processing of a maximum of 5000 features. Hence, the identified and normalized lipids were filtered by relative standard deviation (RSD) for quality control replicates to remove the compounds with high variation, ensuring that a maximum of 5000 were kept for statistics. The filtered lipids were further normalized by quantile to ensure uniform distribution and auto-scaled for statistical evaluation, including Principal Component Analysis (PCA), Partial-Least Squares – Discriminant Analysis (PLS-DA), Analysis of Variance (ANOVA) and Volcano plot analysis (fold change, FC, *versus* p value for Student's t-test with unequal variances adjusted for false-discovery rate, p). Lipids were considered as significantly altered for $p < 0.05$ and $FC \geq 1.5$ or ≤ 0.67. A discussion on statistical models and the validity of p values is available in Chapter I (*1.3.1.7. Statistics applied to lipidomics*, p. 32).

3.3 Results

3.3.1. Serum

Twelve internal standards were detected in positive ionization with a maximum m/z error of 2.2 mDa or 5.1 ppm. For negative ionization, eleven standards were found with a maximum m/z error of 3.2 mDa or 3.8 ppm. The deuterated cholesterol and PG internal standards were not found for positive ionization, while PI, cholesterol and cholesteryl ester were not detected in negative ionization. We did not expect to detect the cholesterol species in negative mode due to their inherent structure. Although the concentrations of the commercial mixture employed for this work are advertised as similar to blood serum, some of the standards had low concentrations that resulted in intensities that were much inferior than the values found for similar lipids (Table II-1, p. 45). When allied to elution in retention time ranges with strong ion suppression from many other sample components, their intensities became too low for accurate detection, particularly for PI. A similar effect was observed for all projects described in this work. Although the Splash Lipidomix commercial mixture from Avanti Polar Lipids is highly useful, the relative concentrations between the standards may not be suitable for all applications (Table II-1, p. 45).

The nanoLC-MS analysis of pig serum samples resulted in the identification of 16,638 lipids, including 82 in tier 1 (m/z error ≤5.0 mDa and MS/MS score ≥500), 178 in tier 2 (m/z error ≤5.0 mDa, MS/MS score between 100 and 500, and mSigma ≤50) and 16,350 in tier 3 (putative mass-match with m/z error ≤5.0 mDa), in addition to 28 internal standard features (different adducts for positive and negative ionization, Supp. Table 2). The most abundant lipid category was glycerophospholipids with 37.5% of the identified species (6223 lipids), followed by sphingolipids with 30.4% (5043 lipids, Figure III-1A and B). Serum had a high number of phosphatidylcholines, phosphatidylethanolamines and triacylglycerols when compared to CSF and MD (Figure III-1C).

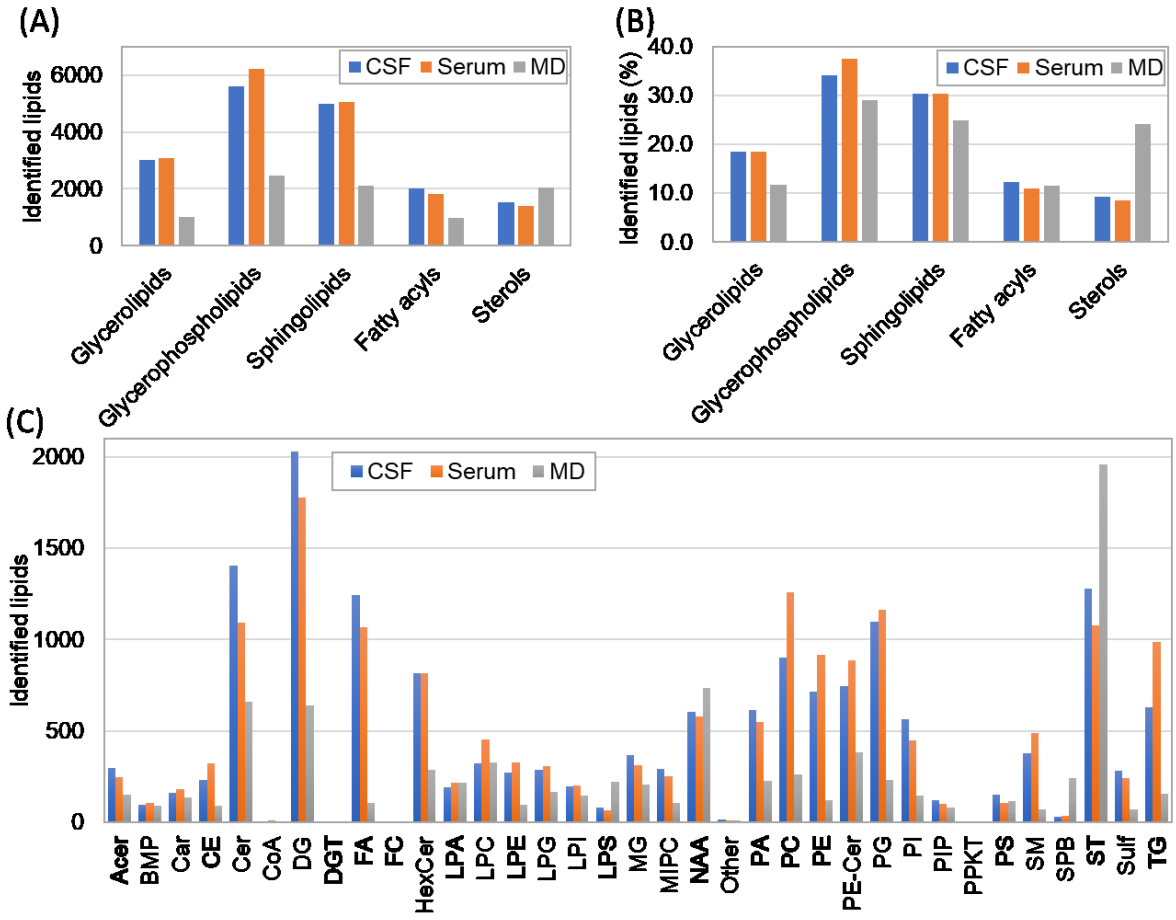

Figure III-1. Lipid species identified for serum, CSF and MD samples from Yucatan miniature pigs (20-25 kg) after a controlled spinal cord injury. Lipids are divided into (A) categories with number of absolute identifications; (B) categories with percentual values to the total number of identified species for each fluid; and (C) lipid subclasses with absolute number of identifications.

The identified species were filtered by RSD to remove lipids with higher variation and match the limit of 5000 features imposed by MetaboAnalyst 4.0.[153] The 4722 lipids with RSD ≤20% were auto-scaled and normalized by quantile for statistical analysis. The PCA score plot displayed tightly clustered QC injections, showcasing the reproducibility and suitability of the employed methods (Figure III-2A). The PCA score plot without the QC replicates (Figure III-2B)

showed full separation between the samples collected 0h and 24h post-injury, while the samples collected after 72 and 168h were clustered together. Unfortunately, the small number of samples and high complexity of the dataset resulted in an overfitted PLS-DA model, as indicated by the low Q^2 value (Figure III-2C). Although the PLS-DA model for this pilot study was overfitted and did not pass the permutation test (p of 0.10), the separation found between the different time points and the excellent R^2 value indicate that the development of large-scale studies to assess the development of SCI using the nanoLC method are reasonable.

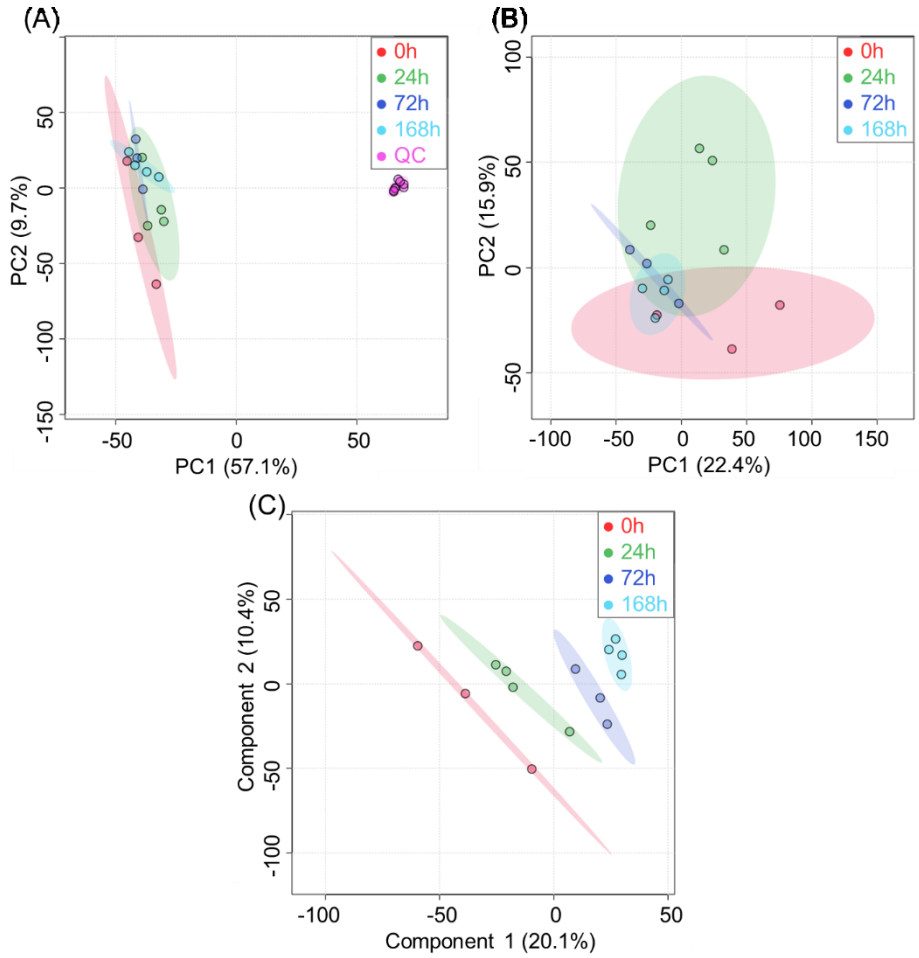

Figure III-2. Statistical analysis for lipidomics of serum samples from miniature pigs (N = 2) collected 0, 24, 72 and 168 h after a controlled spinal cord injury. (A) PCA score plot with 8 QC

replicates (pooled serum from 100 healthy humans); (B) PCA score plots without QCs; (C) PLS-DA score plot with 2 components (R^2 of 0.9893, Q^2 of 0.5498 and p of 0.10).

The small number of biological replicates (N = 2) decreases the reliability of statistical models to find significantly altered lipids that may be related to biological effects after the controlled spinal cord injury. However, this project was meant as a pilot study to ensure adequate sample handling and analysis for future research. Hence, we performed univariate statistical analysis using ANOVA and Volcano plots. The ANOVA for the four time points resulted in 224 lipids with $p < 0.05$, including 67 ceramides (5.8% of the ceramides employed for statistics), 48 PCs (9.6%), 25 TGs (6.2%) and 20 SMs (Supp. Table 3). The samples collected at 0h post-injury were further compared to 24h, 72h and 168h through Volcano plot analysis (Figure III-3). Overall, there were only mild changes in the lipidome of serum samples collected up to seven days after the injury. First, the lipidome of serum at 0h post-injury was compared to 24h, resulting in 21 significantly altered lipids (FC ≤ 0.67 or ≥ 1.5 and $p < 0.05$, Figure III-3, Supp. Table 3). The significantly affected compounds included 5 sphingomyelins (all with higher intensities for 24 h), 5 ceramides and 4 phosphatidylcholines (all with lower intensities for 24h). Also, 13 of the significantly altered lipids were oxidized, with 76.9% of them displaying higher normalized intensities for the 24h group. The most intense significant fold-changes were found for the putatively identified PC 40:5 (or the isomers/isobars PC O-40:6;O, CerP 48:6;O4, PE 43:5 or PC 38:2, with FC 0h / 24h of 9.4 and p of 0.02) and PG 40:3 (or the isomers/isobars PG O-40:4;O, PI O-38:3 and DG 50:12;O2, with FC 0h / 24 h of 4.7 and p of 0.01). The MS/MS identified PC 22:0_18:2 was also significantly altered with FC 0h / 24h of 2.2 and p of 0.04.

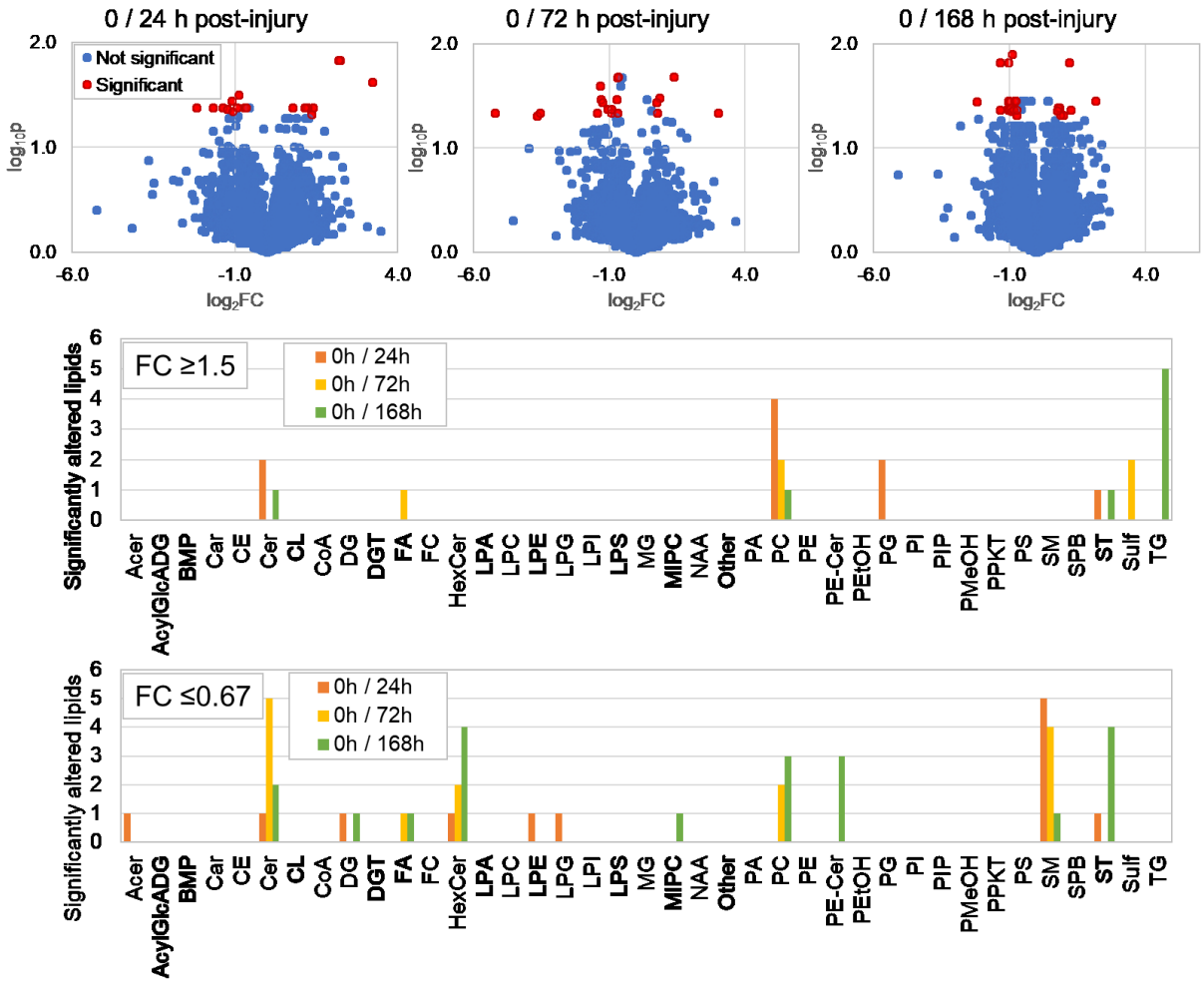

Figure III-3. Volcano plot analysis for serum obtained from miniature pigs (N = 2) that suffered a controlled spinal cord injury, collected at 0h, 24h, 72h and 168h post-injury. Lipids were considered significantly altered for FC ≤0.67 or ≥1.5 and p <0.05.

Second, we compared the lipidome of serum collected at 0h with 72h post-injury, resulting in 19 significantly altered lipids (Figure III-3, Supp. Table 3). All the 7 significantly affected ceramides had higher intensities for the later time point, as well as all 4 sphingomyelins. We also noticed that 71.4% of the affected oxidized lipids (10 compounds) were elevated for 72h. The most

intense significant fold-changes were found for the putatively identified Cer 47:3;O (FC 0h / 72h of 0.03 and p of 0.046) and HexCer 39:2;O4 (FC 0h / 72h of 0.08 and p of 0.049).

Last, we compared the lipidome at 0h with 168h post-injury and 28 lipids were significantly altered (Figure III-3, Supp. Table 3). Most of the 11 significantly affected ceramides (90.9%) were elevated at 168h, but all the triacylglycerols had reduced intensities (5), indicating alterations in the energy metabolism within 7 days after the injury. Most oxidized lipids were once again elevated for the later point (18 species, 78.3%). The most intense significant fold-changes were found for PE-Cer 40:3;O5 (or the isomers/isobars SM 36:3;O3, PA O-44:7, SM 38:3;O5 and DG 45:7;O, with FC 0h/168h of 0.21 and p of 0.04) and ST 22:0;O7;Hex (FC 0h / 168h of 4.5 and p of 0.04).

No lipids were affected for all three comparisons (0 / 24h, 0 / 72h and 0 / 168h post-injury), but two lipids were significantly altered on both the 0h / 24h and 0h / 72h comparisons, namely PC 40:5 (or the isomers/isobars PC O-40:6;O , CerP 48:6;O4, PE 43:5 and PC 38:2, with FC 0h / 24h of 9.4 and FC 0h / 72h of 8.1) and SM 46:2;O2 (or the isomers/isobars PE-Cer 49:2;O2, CerP 51:3;O2 and TG 49:0;O3, with FC 0h / 24h of 0.64 and FC 0h / 72h of 0.54). One lipid was significantly affected for the 0h / 72h and 0h / 168h comparisons, *i.e.* HexCer 38:5;O3 (or the isomers/isobars DG 44:5;O2, TG 44:4;O, TG O-44:5;O2 and CE 24:5;O, with FC 0h / 72h of 0.40 and FC 0h / 168 h of 0.50). Furthermore, two lipids were significantly affected for both the 0h / 24h and 0h / 168h comparisons, but not for the 0h / 72h comparison, *i.e.*, SM 36:1;O4 (or the isomers/isobars PE-Cer 39:1;O4, CerP 41:2;O4, CE 24:6;O and PE 36:1, with FC 0h / 24h of 0.42 and FC 0h / 168h of 0.40) and Cer 37:2;O2 (FC 0h / 24h of 2.37 and 0h / 168 h of 1.73).

3.3.1. Cerebrospinal fluid (CSF)

Twelve internal standards were detected in positive ionization with m/z error between 0.3 and 5.0 mDa (0.04 and 5.8 ppm, respectively). For negative ionization, 12 deuterated standards were also detected with m/z error between 0.2 and 3.0 mDa (0.3 and 3.4 ppm, respectively). The cholesterol and cholesteryl ester standards were not found for negative ionization, whereas the cholesterol and phosphatidylinositol standards were not detected in positive mode due to a combination of ionization efficiency, ion suppression and low concentrations.

The lipidomics of CSF samples resulted in the identification of 16,432 lipids (61.1% of the detected features), including 84 in tier 1, 110 in tier 2 and 16,241 in tier 3. The samples displayed higher numbers of diacylglycerols (2030), total ceramides (3556) and fatty acids (1244) when compared to serum and MD (Figure III-1C, Supp. Table 4). However, the overall composition of CSF lipidome was very similar to serum samples (Figure III-1A and B).

The 4745 lipids with RSD ≤24% were normalized by quantile and auto-scaled for statistical analysis. The PCA score plot resulted in tightly clustered QC samples (Figure III-4A) and partial separation between the four time points ((Figure III-4B). The samples collected at 24h were separated from the other time points on the first and second principal components, indicating changes in the metabolism of the animal within the first day after the injury. The remaining time points were separated only in the second principal component, showing less intense changes, but the last time point at 7 days after injury was fully separated. The low number of biological replicates (N = 2) and high complexity of the dataset also resulted in a heavily overfitted PLS-DA model. However, the model shows that the different time points can be separated with high correlation and indicate that further, large-scale studies may provide more information for evaluation of spinal cord injuries.

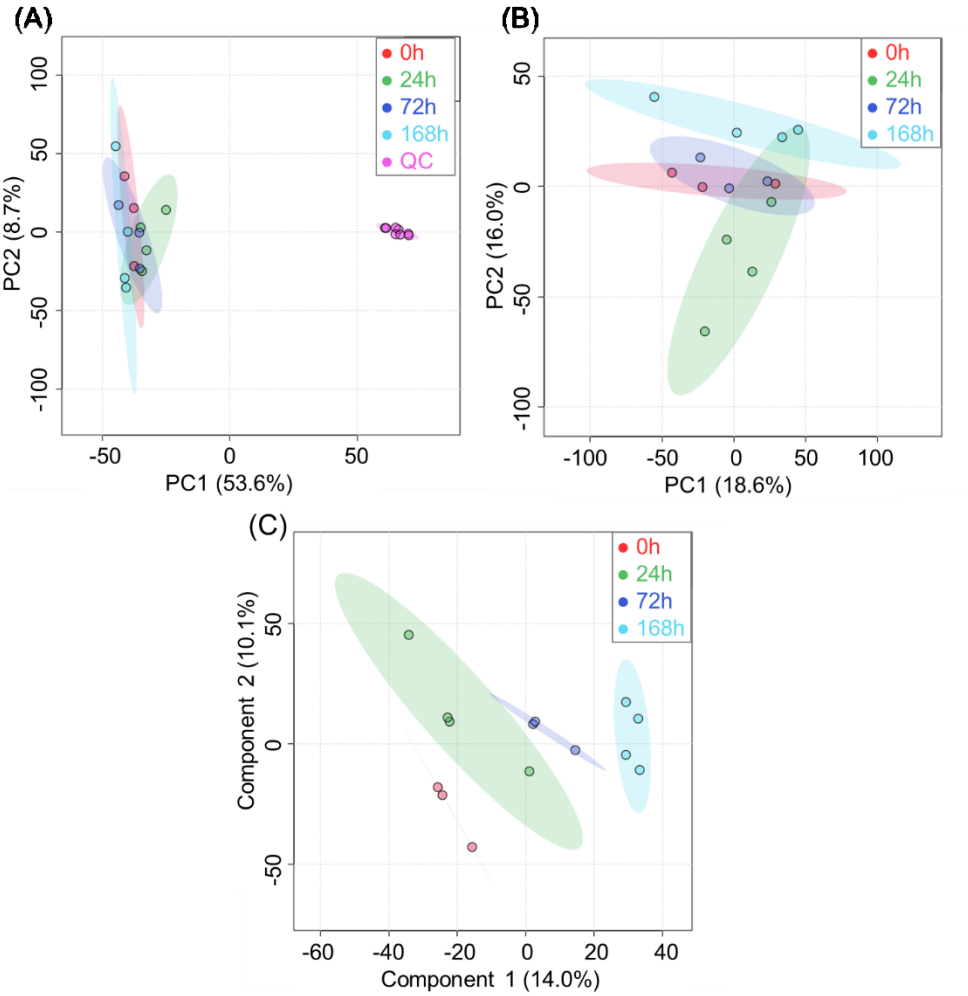

Figure III-4. Multivariate statistical analysis for CSF samples obtained from pigs 0, 24, 72 and 168h after a controlled spinal cord injury (N = 2). (A) PCA score plot with 8 QC replicates (pooled serum from 100 healthy humans); (B) PCA score plot without QC replicates; (C) PLS-DA score plot with 2 components (R^2 of 0.9721, Q^2 of 0.2468 and p of 0.16 for 1000 permutations).

The ANOVA resulted in 71 significant compounds ($p < 0.05$), including 14 phosphatidylcholines (PC), 13 phosphatidylglycerols (PG) and 8 diacylglycerols (Supp. Table 5) The samples collected at 0h post-injury were further compared to 24, 72 and 168h by Volcano plot analysis (Figure III-5, Supp. Table 5). First, the samples collected at 0h were compared to 24h

post-injury, but only three lipids were significantly altered, *i.e.* DG 52:8;O2 (or PG O-42:0;O, with FC 0/24h of 0.38 and p of 0.04), ST 18:1;O7;T (FC 0/24h of 2.4 and p of 0.04) and ST 23:0;O (FC 0/24h of 2.4 and p of 0.04).

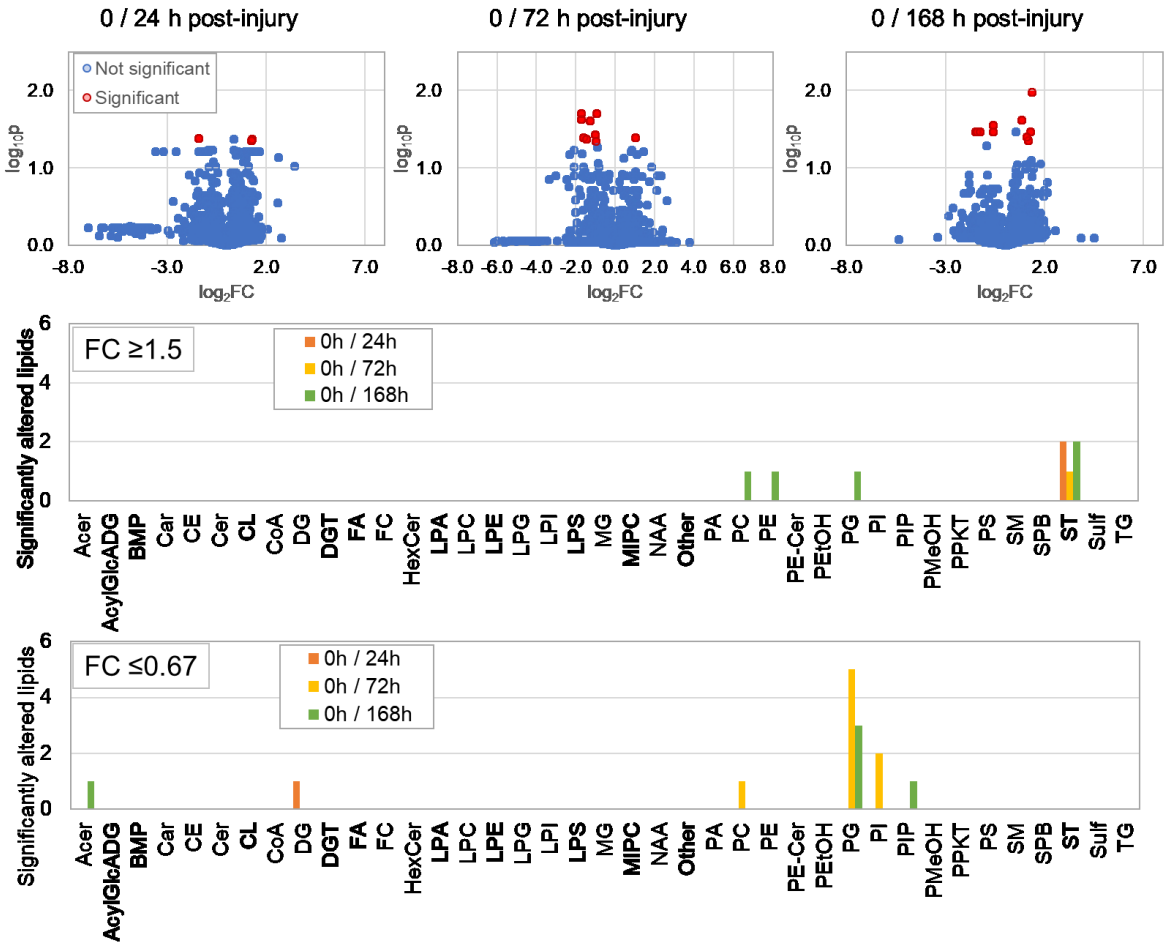

Figure III-5. Volcano plot analysis for CSF samples collected from miniature pigs 0h, 24h, 72h and 168h after a controlled spinal cord injury (N = 2).

Second, the samples collected at 0h were compared to 72h post-injury (Figure III-5, Supp. Table 5). Nine lipids were significantly altered, including 5 phosphatidylglycerols (PG), two phosphatidylinositols (PI), one sterol (ST) and one phosphatidylcholine (PC). All the eight

significant glycerophospholipids were elevated for 72h. The highest significant fold-changes were found for PC 58:11 (or the isomers/isobars PC O-58:12;O, PI-Cer 54:2;O6, PS 61:10 and PE 61:11, with FC 0/72h of 0.30 and p of 0.02) and PI 49:4;O (FC 0/72h of 0.30 and p of 0.02).

Third, the CSF lipidome at 0h was compared to 168h post-injury and ten lipids were significantly affected, including 4 phosphatidylglycerols (PG) and two sterols (ST, Figure III-5, Supp. Table 5). The highest significant fold-changes were found for ACer 55:6;O4 (or the isomers/isobars ACer 53:3;O4, Cer 53:4;O5 and ACer 55:5;O5, with FC 0/168h of 0.36 and p of 0.03) and PG 40:5 (or the isomers/isobars PG O-40:6;O, PI O-38:5 and DG 50:14;O2, with FC 0/168h of 0.38 and p of 0.03). Furthermore, one lipid was significantly affected for both the 0/24h and 0/168h comparisons, namely ST 18:1;O7;T (FC 0/24h of 2.43 and FC 0/168h of 2.18), and one for the 0/72h and 0/168h comparisons, *viz.* ST 20:2;O4;T (or the isomers/isobars NAT 20:5;O2 and MG 22:5;O, FC 0/72h of 2.07 and FC 0/168h of 1.83). No lipids were simultaneously affected for all three comparisons.

3.3.1. Intraparenchymal microdialysate fluid (MD)

Eleven internal standards were detected for positive ionization with a maximum mass-to-charge (m/z) error of 5.1 mDa or 6.9 ppm, while eleven internal standards were found for negative ionization with a maximum m/z error of 5.8 mDa or 8.8 ppm. The untargeted lipidomic analyses resulted in the identification of 8453 lipids, including 50 in tier 1, 49 in tier 2, 8409 in tier 3 and 35 internal standard peaks (different adducts for positive and negative ionization, Supp. Table 6). The MD samples displayed a high proportion of sterol lipids (1959 lipids, corresponding to 23.1% of all identified lipids), followed by total ceramides (1590, 18.7%), N-acyl amines (735, 8.6%) and diacylglycerols (638, 7.5%, Figure III-1). However, the most abundant lipid category was

glycerophospholipids, with 2458 identified compounds (29.0% of all identified species). When compared to serum and CSF, MD has a higher proportion of sterols, but smaller percentages of glycerolipids and glycerophospholipids (Figure III-1).

The identified species were filtered by RSD to match the limit of 5000 features for statistical analysis using MetaboAnalyst 4.0.[153] The 4135 lipids with RSD ≤30% were auto-scaled and normalized by quantile for statistical analysis. The PCA score plot displayed tightly clustered QC replicates (Figure III-6A), showcasing the reproducibility of the employed methods. The first collection time point, 4.75h post-injury, was fully separated from the others by PCA, indicating fast metabolism changes in the interstitial region around the injury that stabilized after a few hours. Unfortunately, the small number of samples (N = 2) and high complexity of the dataset once again resulted in an overfitted PLS-DA model (Figure III-6C), but the complete separation between the four time points indicate that further large-scale studies are justified for biomarker discovery and assessment of pathological pathways.

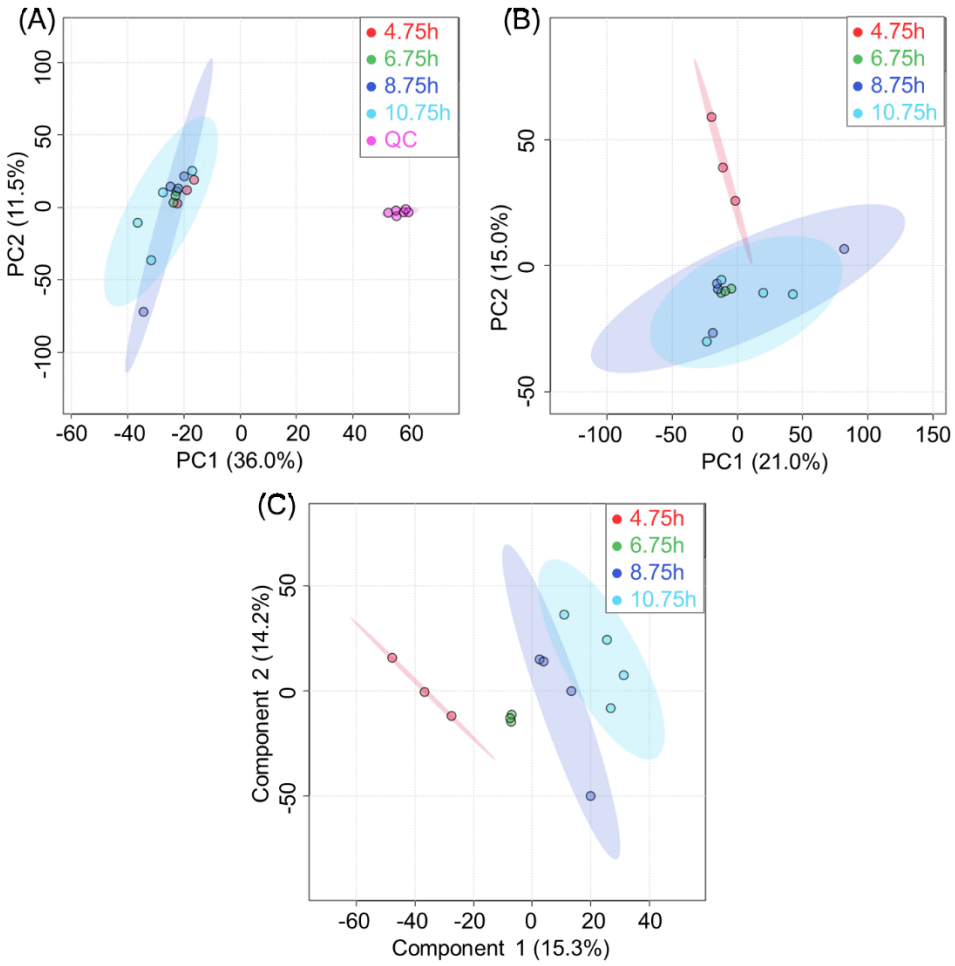

Figure III-6. Statistical analyses for intraparenchymal microdialysate fluid (MD) obtained from two pigs 4.75, 6.75, 8.75 and 10.75 hours after a controlled spinal cord injury. (A) PCA score plot with 6 QC replicates (a pool of serum from 100 healthy humans); (B) PCA score plot without QCs; (C) PLS-DA score plot with 3 components (R^2 of 0.9956, Q^2 of 0.5811 and p of for 0.13 for 1000 permutations).

The ANOVA for the 4 time points resulted in 159 significant lipids ($p < 0.05$), including 9 fatty acyls (7.9% of the fatty acyls used for statistics), 66 sphingolipids (5.2%) and 59 sterols (3.8%, Supp. Table 7). The samples collected 4.75 h post-injury were further compared with the remaining time points by Volcano plot analysis (Figure III-7, Supp. Table 7). The MD samples

showed a higher number of affected lipids when compared to serum and CSF. Whether that observation is a result of the closeness of the collected MD fluid with the injury site or the different time points (24 to 168h range for serum and CSF; 4.75 to 10.75h for MD) remains to be evaluated in future, large-scale studies. The comparison between MD samples collected at 4.75 h and 6.75 h post-injury resulted in the highest number of significantly altered lipids, with 145 species displaying decreased normalized intensities for the 6.75 h point, while 159 had increased intensities. The species with the most intense significant fold-changes were putatively identified in tier 3 as NAT 30:2 (or Car 28:6), with FC 4.75/6.75 of 0.02 and p of 3.0×10^{-8}), PE-Cer 30:4;O5 (or the isomers/isobars PE-Cer 28:1;O5, PA O-32:5, TG O-36:8 and PA O-29:0;O, with FC 4.75/6.75 of 13.3 and p of 6.7×10^{-4}) and Car 20:0;O (FC 4.75/6.75 of 0.1 and p of 5.5×10^{-3}). Acylcarnitine (Car) species are associated with energy metabolism, as they act in the transport of fatty acyl residues across the mitochondrial membrane for β-oxidation. The elevated levels of an acylcarnitine for the 6.75h pos-injury collection may corroborate the previous indication of alterations in the energy metabolism, although further studies are required to confirm the hypothesis.

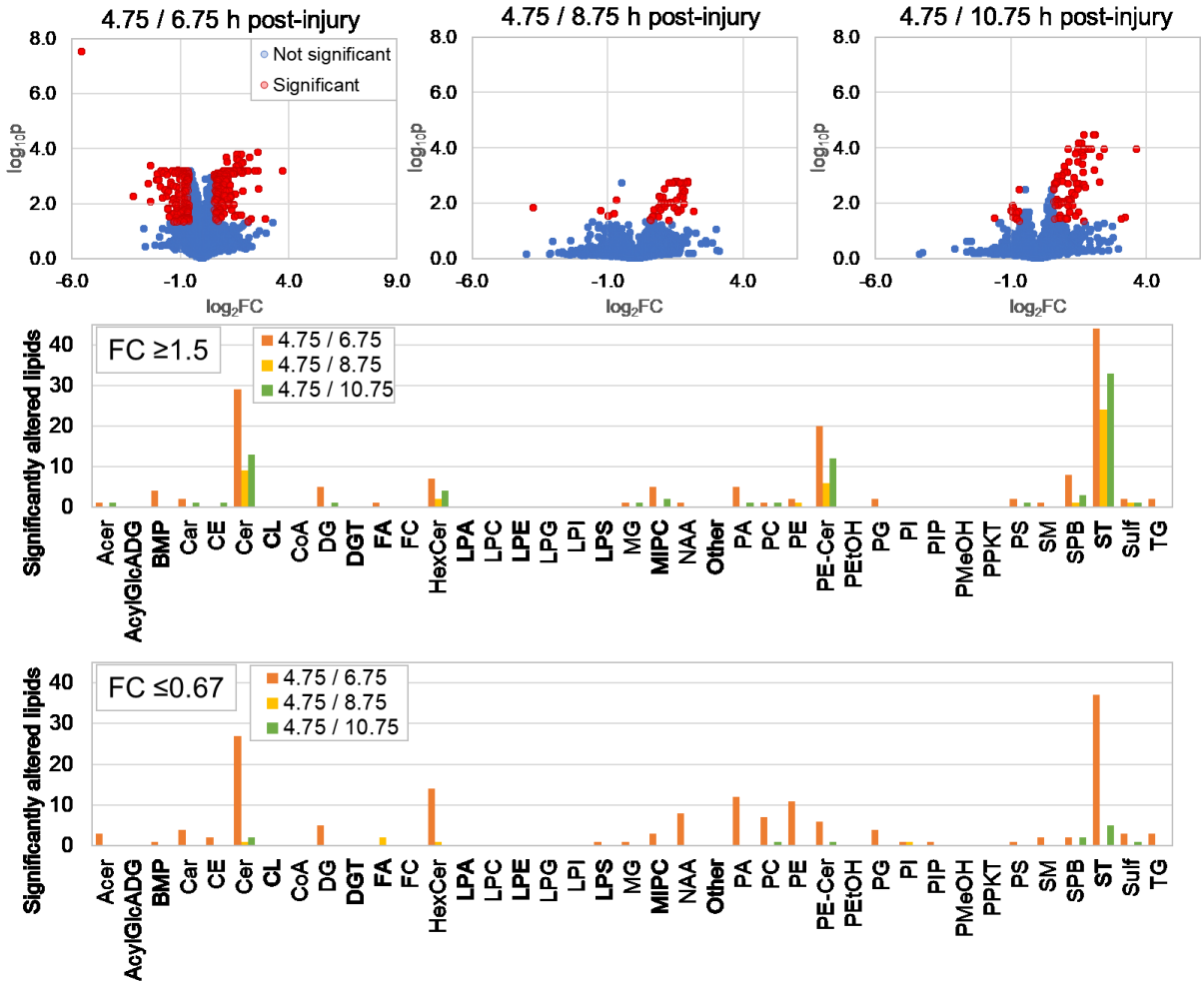

Figure III-7. Volcano plot analysis for intraparenchymal microdialysate collected around the controlled spinal cord injury region at 4.75, 6.75, 8.75- and 10.75 hours post-injury. Lipids were considered significantly altered for FC ≤0.67 or ≥1.5 and p <0.05.

Although the samples collected at 8.75 and 10.75 h post-injury had smaller numbers of significantly altered lipids (49 and 88, respectively), more than 86% of them had lower intensities when compared to the 4.75 h time point (Figure III-7, Supp. Table 7). Ceramides were highly affected, with approximately 90% of the significantly altered species displaying decreased normalized intensities for the later points. Furthermore, all the 24 significantly altered sterols for the 4.75 *versus* 8.75 h points were also decreased as time passed. It is worth mentioning that most

of the significantly altered lipids containing polyunsaturated fatty acyl (PUFA) residues were increased for 4.75h post-injury compared to both 8.75 and 10.75h (92.3% and 87.5% of the significantly altered PUFA lipids, respectively). Most of the oxidized lipids (91.8% of the significant lipids for 4.75/8.75 and 93.2% for 4.75/10.75) also displayed higher intensities for 4.75h compared to 8.75 and 10.75h post-injury. However, the same trend was not observed for the 6.75h time point, which had similar numbers of significant lipids with PUFA residues or oxidized species with increased and decreased intensities when compared to 4.75h (45.9% of PUFA lipids and 53.1% of oxidized lipids with FC 4.75/6.75 ≥1.5). Oxidation is tightly related to inflammation, a known secondary effect of spinal cord injuries. Furthermore, 35 lipids were altered for the three binary comparisons, all with higher intensities for 4.75 h, including 15 ceramides and 18 sterols. However, 232 lipids were altered only for the 4.75 *versus* 6.75h comparison, while 7 were altered only for 8.75h and 23 for 10.75h (Supp. Table 7). It seems that the first few hours after the injury are associated with major changes which are partially mitigated by the metabolism within 10.75h.

The most intense significant fold-changes for the comparison between 4.75 and 8.75 h post-injury were found for the putatively identified (tier 3) PI O-27:0 (or the isomers/isobars PA O-40:10, TG 41:9 and PA 34:1;O, FC 4.75/8.75 of 0.08 and p of 0.01) and PE 40:2 (FC 4.75/8.75 of 4.6 and p of 0.02). For the comparison between 4.75 and 10.75 h post-injury, the most intense fold-changes were found for PE-Cer 30:4;O5 (or the isomers/isobars PE-Cer 28:1;O5, PA O-32:5, TG O-36:8 and PA O-29:0;O, with FC 4.75/10.75 of 12.4 and p of 1.1×10^{-4} – also one of the most intense fold changes for 4.75 *versus* 6.75), PC 27:3;O (FC 4.75/10.75 of 9.4 and p of 0.03) and Car 18:0 (FC 4.75/10.75 of 8.6 and p of 0.04).

3.4 Discussion

This work described a pilot study for lipidomics of serum, cerebrospinal fluid (CSF) and intraparenchymal microdialysate (MD) of miniature pigs after a controlled spinal cord injury. Samples were collected from two pigs at different time points, *i.e.*, 0, 24, 72 and 168h post-injury for serum and CSF; and 4.75, 6.75, 8.75 and 10.75h post-injury for MD. Unfortunately, the small number of biological replicates (N = 2) restrains the conclusions that can be drawn from the statistical models. However, the main objective of this work was to evaluate the suitability of the methods for future applications to study spinal cord injuries. The nanoLC-MS approach developed in Chapter II allowed the identification of thousands of different lipid species using only 1.0 to 2.5 µL of each fluid. The evaluation of internal standards showcased the excellent mass accuracy (m/z error ≤8.8 ppm) obtained by a combination of high-resolution mass spectrometry and mass recalibration employing background features detected for all injections. Quality control (QC) samples composed by a pool of serum from 100 healthy humans were evaluated with each type of fluid. The QC replicates (extraction and injection) were tightly clustered in all PCA score plots (Figure III-2, Figure III-4 and Figure III-6), showcasing the reproducibility of the employed methods. Furthermore, more than 4500 lipids displayed relative standard deviations smaller than 30% for the QC replicates analyzed with each type of fluid.

The lipidome compositions of serum and CSF seem to be very similar. We observed comparable numbers of glycerolipids, glycerophospholipids, sphingolipids, fatty acyls and sterols in both fluids. The different time points were partially separated by PCA and fully separated by PLS-DA for all fluids. Samples collected at the time of injury only displayed minor differences when compared to 24, 72 and 168h post-injury. However, the serum samples collected at the moment of the injury were separated from samples collected after 24h by a simple PCA model,

whereas the samples for 72 and 168h post-injury were clustered together. The PCA analysis for serum indicates that the first 24h after a SCI may impose major metabolic alterations that are mitigated within the following days.

We observed a higher amount of significantly altered lipids for different time points in MD samples. The first collection point at 4.75h after the injury was fully separated from the later points by a simple PCA model, indicating that the striking effect of the injury upon the spinal cord tissue lipidome is partially mitigated by the metabolism within hours. Prior studies have also shown a dramatic drop in glucose and pyruvate levels in MD within minutes after a SCI, as well as an increase in lactate levels that could be related to a high energy demand that causes the activation of glycolysis for about 60 min.[134] We also observed minor alterations in acylcarnitines, which may corroborate the higher energy demand possibility, although further studies are required to confirm it. The observation agrees with the pathological development of SCI patients and emphasizes the need for immediate diagnosis and medical treatment right after a suspected injury.

Overall, MD samples resulted in the identification of lower numbers of lipids when compared to serum and CSF, although more significant changes were found. Interestingly, the MD samples had a higher proportion of sterols, accompanied by lower proportions of glycerolipids, glycerophospholipids and sphingolipids. The composition of MD obtained from the interstitial region around the spinal cord is likely related to the metabolism of the nerve tissue cells, although the collected samples are also affected by the employed probe and perfusate. The aqueous composition of the perfusate solution makes it unlikely that highly hydrophobic lipid droplets, triacylglycerols and cholesteryl esters will be profusely represented in MD samples, but the oxidized and modified sterols, ceramides and fatty acids may have a higher polarity that can promote better diffusion through the probe. Yet, myelin, a lipid-rich substance the surrounds nerve

tissue in white matter, is enriched in cholesterol and sterol lipids.[154] The acute phase that follows a severe SCI is characterized by death of nerve cells and oligodendrocytes, *i.e.* myelinating cells, as well as disruption of the blood-spinal cord barrier. The injury can destroy the white matter, mainly composed by myelinated axons.[132,155] The MD samples collected a few hours after the injury should reflect the inherent effects caused by the spinal cord contusion and rupture of the myelinated tissue with high content of sterols and ceramides. Following a SCI, phagocytic cells are activated to clear the debris from myelin damage.[132] Hence, we observed significantly affected levels of sterols and ceramides with increased intensities for the first time point that were reduced over the 10.75 hours of sample collection. About half of the significant sterols and ceramides had fold-changes higher than 1.5 for the comparison between 4.75 and 6.75h, but most of the significant sterols and ceramides showed elevated intensities for the 4.75 time point compared to 8.75 and 10.75h post-injury, indicating that the profusion of sterol release in the interstitial region caused by cell death is mitigated over time.

We will refrain from further discussing the biological importance of the significant lipid species as the small number of biological samples require more investigation to confirm the observations reported herein. Nevertheless, the nanoLC-MS method was highly suitable for the proposed application. A collaboration has been previously established between Dr. Brian Kwon (University of British Columbia, Vancouver, BC Canada), Dr. Leonard Foster (University of British Columbia, Vancouver, BC Canada), Dr. Liang Li (University of Alberta, Edmonton, AB, Canada) and the Rick Hansen Institute to study biomarkers of spinal cord injury through lipidomics, metabolomics, genomics and proteomics of serum and CSF samples obtained from 91 human patients (different time points and injury severity) and miniature pigs (animal model candidates). The application of different *omic* approaches to precious biological samples requires

the use of high-sensitivity methods to reduce sample volumes. This pilot study indicated that lipidomics of small volumes of serum, CSF and MD samples by nanoLC-MS has great potential to study the pathogenesis of spinal cord injury, while allowing higher sample aliquots to be store for further *omic* experiments. The large-scale lipidomics and metabolomics studies for human and pig samples are currently underway in Dr. Liang Li's lab at University of Alberta.

3.5 Conclusions

The pilot study described in this work aimed to verify the suitability of the nanoLC-MS method developed in Chapter II to study different biological fluids. The method allowed the reproducible detection and identification of thousands of lipids from 1.0 to 2.5 μL of serum, cerebrospinal fluid and intraparenchymal microdialysate, showcasing the potential of nanoLC-MS for untargeted lipidomics when only small volumes of diluted samples are available. Large-scale applications to study lipidome alterations following a spinal cord injury in humans and miniature pigs as candidate animal models are currently underway.

3.6 Acknowledgments

We thank all collaborators involved in the development of this pilot study, particularly Dr. Brian Kwon (Department of Orthopaedics) and Sarah Assadian (Director of Clinical Research at PROOF Centre of Excellence, University of British Columbia, Vancouver, BC, Canada). We also thank Dr. Ulrike Schweiger-Hufnagel and Dr. Aiko Barsch (Bruker Daltonics, Germany) for the support with MetaboScape data processing for nanoLC-MS lipidomics. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, Canadian

Institutes of Health Research, the Canada Research Chairs program, the Canadian Foundation for Innovations, Genome Canada and Alberta Innovates.

IV

Chapter IV: Lipidomics of Exosomes[6,7]

4.1 Introduction

Extracellular vesicles are nano-sized membrane structures secreted by cells that are composed of biologically active molecules such as proteins, metabolites, hormones, lipids and nucleic acids. Extracellular vesicles are divided into exosomes (30-100 nm particles arising from the multivesicular endosomal pathway), microvesicles (100-1000 nm particles resulting from membrane shedding of platelets, endothelial cells, red blood cells and others) and apoptotic bodies (100-5000 nm particles from cell apoptosis). Exosomes are single-membrane, small extracellular vesicles obtained from the fusion of multivesicular bodies with the plasma membrane of virtually all types of eukaryotic cells. When the endosome produced by the endoplasmic reticulum matures, a portion of its membrane invaginates and buds into its lumen, generating intraluminal vesicles. The mature endosome filled with intraluminal vesicles is called multivesicular endosome or multivesicular body, which contains molecules destined for degradation or secretion. [156] Its content can be degraded by fusion with lysosomes or released into the extracellular space by fusion with the plasma membrane in the form of membranous extracellular vesicles. Once secreted into the extracellular medium, the multivesicular bodies are named exosomes. [157–160]

Exosomes can communicate with proximal and distal target cells by releasing signaling molecules into the extracellular space, binding to specific sites in cell surfaces, fusing with

[6] Serum samples and exosome isolates were obtained by Dr. David M. Lubman, Dr. Jianhui Zhu and Jie Zhang (University of Michigan Medical Center, Ann Arbor, MI, USA).

[7] Supporting tables for this chapter are available with Dr. Liang Li.

membranes, and by endocytosis. [161] Hence, they are essential for inter-cell communication and can influence a variety of processes in the body, including immune response and pathogenesis. [162] Recent research has shown that exosomes may play a role in many pathophysiologies, being excellent candidates for biomarker discovery. Their diverse content and vital role in cell communication may result in exceptional, yet under-explored potential for biomarker research. However, the low concentrations, heterogeneity and isolation of exosomes constitute a challenge for untargeted *omic* techniques. Hence, the potential of exosomal lipids has not yet been confirmed and very little is known about its physiological and pathological roles. [160,162]

Exosomes are usually isolated from biological fluids or cell culture medium based on size, density or presence of marker molecules. There are different methods to isolate exosomes, *viz*. differential ultracentrifugation (UC), ultrafiltration, size-exclusion chromatography (SEC), immunoaffinity, precipitation and microfluidics. However, exosome fractions obtained by purification methods that depend on the presence of marker molecules, such as immunoaffinity, may result in the isolation of one subpopulation that may not be representative of the original sample. Although ultrafiltration and SEC may offer highly pure exosome isolates, the techniques may cause biases and fragmentation of larger vesicles under high pressure, affecting the analytical results obtained from the isolates. Alternatively, UC is one of the most common methods employed to study exosomes due to its simplicity and reliability. Samples are centrifuged at increasingly high speeds under refrigeration to separate particles according to their sizes and densities. However, co-isolation of lipid particles, lipoproteins and protein aggregates of similar sizes is a possibility that must be considered, particularly if cells have been ruptured during the isolation procedure. [159,161,162] However, multivesicular bodies are intermediate compartments of the endosome that are continuously changing. Consequently, there are no exclusive markers to characterize whether a

sample is purely composed of isolated exosome or other types of extracellular vesicles and particles. All the components of exosomes are derived from the cytosol or plasma membranes, hence not specific enough to serve as a marker of the type of extracellular vesicle present in the sample. However, a combination of markers may indicate the composition of the isolate obtained from a biological sample.[160]

Exosomes can be excreted by most types of cells into biological fluids and tissues. Their composition is highly variable according to the type of source cell. Lipids are critical in exosome biogenesis, as the excretion process is dependent upon membrane vesiculation. The lipid content of exosomes has been the focus of many studies in the past few years. Their membranes are typically enriched in glycosphingolipids, cholesterol, phosphatidylserines, phosphatidic acids, polyunsaturated fatty acids (PUFA), prostaglandins, sphingomyelins and ceramides. Further information on lipid membranes can be found in Chapter I (*1.2.3. Lipid membranes*, p. 8). Cholesteryl ester and triacylglycerol species are typically not found in membranous structures such as exosomes. However, they may be fount in high proportions in lipid droplets and lipoprotein particles due to their high hydrophobicity; hence, high concentrations of such species may indicate co-isolation of undesirable structures. [157,159,162,163] Nevertheless, exosomes are generated by invaginations of mature endosomes, which are in turn involved in the processing of low-density lipoprotein (LDL) particles, containing high amounts of cholesteryl esters and triacylglycerols surrounded by a layer of glycerophospholipids and apolipoprotein B-100. In fact, triacylglycerols and cholesteryl esters have been previously reported in exosome isolates obtained by UC of colorectal cancer cells.[164]

The biological role and composition of exosomal lipids remain unknown mostly due to a lack of comprehensive, high sensitivity techniques to study the low amounts of complex material

that can be obtained by isolation techniques. Most reports on exosome lipidic composition focus in a few lipid subclasses and rarely describe unique species, *i.e.* the fatty acyl components and other functional groups of lipids of interest within a subclass. Furthermore, the small size, low concentrations and high complexity of exosomes obtained from biological fluids require accurate methodologies with extremely high sensitivity for biomarker research, which has rarely been described.[162] The interest in exosomes as a potential source of biomarkers and vehicles for drug delivery has grown in recent years. The nano-sized structures are involved in immune function, tumor growth, metastasis, degenerative processes and regulation of cardiac function. [161] However, many questions remain unanswered, and the lack of reliable methods to isolate, purify and analyze exosomes still hamper their application in pathophysiology research and biomarker discovery.[159] The use of chemical isotope labeling to study exosome metabolomics by nanoLC-MS was previously reported by Luo *et al.*.[165] Now, we hypothesized that our previously developed high-sensitivity nanoLC-MS method for untargeted lipidomics (Chapter II) could be further adapted for exosome samples. Hence, we performed a pilot study with exosomes isolated from blood serum by ultra-centrifugation and size-exclusion chromatography. Our goal was to determine the composition of fractions obtained by both isolation techniques and further compare to the serum samples from which they were obtained. The samples employed for this work were also previously evaluated elsewhere for protein markers (CD9, CD63, CD81, albumin, apoliproteins, and total protein content), size distribution (transmission electron microscopy) and concentration of particles (NanoSight).[165,166]

4.2 Experimental

4.2.1 Sample preparation [8]

For the first phase of the pilot study, the evaluation of the isolation method was performed with lipid extracts obtained from serum exosome samples of healthy volunteers obtained by ultra-centrifugation (UC – 4 samples), size-exclusion chromatography (SEC – 1 sample) or a combination of both (UC&SEC – 1 sample). The extraction procedure previously optimized in Chapter II was adapted to the low concentrations of exosome samples, including sample volume, dilution of the extract, concentration of internal standards and injection volume (data not shown). For the optimized method, 30.0 µL of exosome isolate were mixed with 10.0 µL of the Splash Lipidomix Mass Spec Standard (Avanti Polar Lipids, Table II-1, p. 45) previously diluted 25× in methanol. The mixture was extracted with 256.7 µL of methanol and 533.3 µL of dichloromethane, with 20 s vortex between additions. A clean-up step was performed with 170.0 µL of water and 10 s vortex, followed by equilibration at room temperature for 10.0 min. The biphasic mixture was centrifuged for 10.0 min at 10,000 rpm and 4°C. The organic layer (453.3 µL) was evaporated to dryness on a SpeedVac for 30 min. The dried residue was immediately resuspended in 3.0 µL of mobile phase B (MPB), vortexed for 1.0 min and diluted with 27.0 µL of mobile phase A (MPA) before injection, resulting in no dilution of the exosome sample and 75× dilution of the internal standard mixture (Table II-1, p. 45).

Second, we compared the lipidome of serum samples and exosome fractions obtained by SEC. Exosome samples were prepared as previously described. For serum samples, the dilutions of the sample extract and the internal standard mixture (Table II-1) were further optimized. Serum

[8] Serum samples and exosome isolates were obtained by Dr. David M. Lubman, Dr. Jianhui Zhu and Jie Zhang (University of Michigan Medical Center, Ann Arbor, MI, USA).

samples cannot be analyzed with the same dilution factor as exosome isolates due to higher concentrations that would saturate the trap column of the nanoLC instrument and the MS detector, causing lipid aggregation, biases and potentially damaging the employed equipment. Hence, an aliquot of 9.0 µL of serum was mixed with 3.0 µL of the internal standard mixture (Table II-1, p. 45). The mixture was extracted with 77.0 µL of methanol and 160.0 µL of dichloromethane, with 20 s vortex between additions. A clean-up step was performed with 51.0 µL of water and 10 s vortex, followed by equilibration for 10 min at room temperature and centrifugation for 10 min at 10,000 rpm and 4°C. The organic phase (136.0 µL) was evaporated to dryness on a SpeedVac for 30 min and resuspended in 4.5 µL of MPB, followed by dilution with 40.5 µL of MPA. The resuspended extract was further diluted before injection (40× for serum and 120× for the internal standard mixture), *i.e.*, aliquots of 3.8 µL of the extract were vortexed with 2.6 µL of MPB and 23.6 µL of MPA.

All samples were extracted in triplicates in polypropylene microcentrifuge tubes (1.5 mL, Fisher Scientific). The resuspended extracts were stored in polypropylene inserts placed in amber autosampler vials capped with PTFE-lined septa at 4°C for a maximum of 24h before injection. The extraction and injection sequences were randomized to prevent bias. Pooled mixtures composed of aliquots of all exosome samples and all serum samples were prepared for quality control (QC). One pooled mixture of exosome, one pooled mixture of serum and one blank (water instead of sample and methanol instead of the internal standard mixture) were extracted with each group of 3 samples to control reproducibility and contamination. The pooled mixtures were also employed for the acquisition of MS/MS spectra for identification.

4.2.2 Analysis

Samples were analyzed by reversed-phase chromatography on a nanoLC instrument with C18 analytical and trap columns. The instrumentation and reagents were described in Chapter II (*2.2.1. Chemicals and reagents* and *2.2.2. Instrumentation*). The nanoLC conditions included: MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water; MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water; trapping at 7.0 µL/min for 0.25 min at 75% MPA; 50 min gradient (0 min – 5% MPB, 400 nL/min; 5 min – 30% MPB, 400 nL/min; 10 min – 50% MPB, 300 nL/min; 15 min – 70% MPB, 300 nL/min; 28 min – 98% MPB, 300 nL/min; 42 min – 0% MPB, 300 nL/min; 48 min – 0% MPB, 350 nL/min; 50 min – 0% MPB, 400 nL/min); 10 min re-equilibrium (50 min – 100% MPA, 400 nL/min; 59 min – 100% MPA, 400 nL/min; 59.5 min – 100% MPA, 50 nL/min); 45°C; and 5 µL injection. The previously developed nanoLC-MS method was further adapted to improve mass recalibration using the injection of a 10 mM sodium formate calibrant solution in 1:1 isopropanol/water immediately before each sample injection. The injection, trapping and analysis parameters were optimized to ensure the elution of a calibrant peak during the dead time of each sample chromatogram (5.0 µL injection with trapping at 3 µL/min for 1.0 min at 95% MPA; 1.6 min isocratic elution at 95% MPA and 420 nL/min). The mass spectrometer conditions were: electrospray ion source capillary voltage of 1375 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.15 bar, dry nitrogen gas flow rate of 2.5 L/min, ion source temperature of 190ºC and spectra acquisition rate of 1.44 Hz.

Samples were extracted in triplicates and each extraction was injected once in positive ionization and once in negative ionization. Injections were randomized, with each group of 6-9 sample extracts sandwiched between QC injections.

4.2.3 Data processing and statistics

The obtained chromatograms were re-calibrated using the sodium formate peak that eluted within the dead time of each injection to ensure high mass accuracy. Further lock mass calibration was applied to correct for minor mass drifts that may occur during the analytical run with erucamide for positive ionization, a plasticizer detected throughout all chromatograms (m/z 338.341741), and stearic acid for negative ionization (m/z 283.264254). Samples were aligned with an m/z tolerance of 5.0 mDa, retention time tolerance of 60 s, minimum intensity threshold of 6000 cts for positive ionization and 2500 cts for negative ionization and minimum peak length of 22 spectra. Features not found with the alignment parameters but detected in at least 5% of injections were searched again on the raw data for recursive extraction (no minimum intensity cut-off and minimum peak length of 11 spectra). The aligned features were filtered by detection in at least 80% of injections for one of the sample groups (exosome isolation method, serum or QC). The remaining missing values were substituted by the minimum intensity within each group (exosome isolation method) or type of sample (exosome or serum).

MS/MS identification was performed with the chromatograms acquired for the QC pooled samples with a different number of precursor ions (3, 5 or 7), variable collision energies (10 to 80 eV) and different MS/MS acquisition rates (2, 4 and 5 Hz). Precursor ions were automatically selected by the mass spectrometer software (Bruker oToF control) in auto-MS/MS mode, *i.e.*, the most intense ions are selected in precursor scans for fragmentation within each MS-MS/MS cycle of 1.2 s (0.69 s for the precursor survey scan at 1.44 Hz; and 0.51 s for CID of each selected precursor ion at 2, 4 or 5 Hz). Unfortunately, many lipids with lower intensity when compared to co-eluting molecules were not selected for fragmentation or provided fragments with a very low intensity that could not be used for confident identification. MS/MS spectra were matched to the aligned feature list with an m/z tolerance of 7 mDa and a retention time tolerance of 30 s. In the

future, potential biomarkers or molecules of interest could be further investigated with targeted analysis to improve their identification; however, this work was focused on untargeted, comprehensive lipidomics.

The lipid classes and subclasses of identified lipids followed the classification system proposed in 2005 and updated in 2009 by Fahy *et al.*, as well as the shorthand notation proposed by Liebisch *et al.* in 2013 and recently adopted by the database LipidMaps.[16,17,20] The identification procedure also followed the guidelines of the Lipidomics Standards Initiative (https://lipidomics-standards-initiative.org).[20,76] The MS/MS positive identification was performed with the MS-Dial LipidBlast[73,78], Human Metabolome Database (HMDB)[79,81,109,151] and MassBank of North America LC-MS/MS libraries (LipidBlast, HMDB, fatty acid esters of hydroxy fatty acids - FAHFA[82], IMS Oxidized Phospholipids[83] and all LC-MS/MS spectra) in combination with MetaboScape 4.0. MS/MS identifications were divided in two tiers, namely tier 1 for lipids with precursor m/z error ≤5.0 mDa, MS/MS score ≥500 and mSigma ≤150, and tier 2 for lipids with precursor m/z error ≤5.0 mDa, MS/MS score between 100 and 500 and mSigma ≤50. Unfortunately, many features could not be matched to MS/MS libraries due to the low intensity of precursor and fragment ions. Furthermore, many features were not chosen for fragmentation because of high-intensity co-eluting molecules. Unidentified features were searched in the LipidMaps database for exact m/z match (putative identification, tier 3). Lipids can have multiple isomers or isobars with similar structures, belonging to different lipid subclasses (Chapter I, *1.3.1.5. Identification and nomenclature*, p. 25). For this work, a six-tier filtering and scoring approach was employed to restrict the number of isomeric or isobaric compounds and select the best identification for each feature: (1) expected retention time range for each lipid subclass and fatty acyl/alkyl (FA) chain lengths (the total number of carbons); (2) expected adducts and

isoforms for each lipid subclass; (3) m/z error (ppm), (4) FA chain length, *i.e.*, even-chain fatty acids are more commonly found in nature and more likely to be detected in human biological fluids; (5) presence of functional groups other than the expected for each subclass; and (6) method sensitivity and expected intensity ratios between lipid subclasses. The procedure was described with more details in Chapter II (*2.2.7. Putative identification of lipids*, p. 50). The scores for each ranking tier were summed and the top choice for each feature (lowest score) was selected to determine lipid subclass, *i.e.,* the isomer or isobar with elution within the expected retention time range for the identified subclass; the most likely adduct and isoform; the smallest m/z error; an even number of carbons in FA chains; the expected structure for the subclass; and the lipid subclass more likely to be detected by the employed method and found in biological fluids. The top choice was used to determine the lipid subclass for normalization (Chapter I, *1.3.1.6. Ion suppression and normalization approaches*, p. 29), but other isomeric or isobaric possibilities that passed the retention time and adduct filters were also kept, ordered according to their scores.

The identified lipids were matched to one of the 14 deuterated internal standards according to subclass similarity and retention time range (Table II-1, p. 45). Normalized intensities were calculated by the peak intensity (*i.e.*, peak height) of each identified lipid divided by the intensity of the most similar deuterated internal standard to correct for ion suppression and other small differences that may occur during sample handling. Statistical analysis was performed in MetaboAnalyst 4.0 (www.metaboanalyst.ca). [153] The web platform restrains the number of uploaded features to less than 5000; hence, the normalized lipids were filtered by the relative standard deviation (RSD) before statistics. The dataset was further normalized by median (*i.e.,* the peak intensities were adjusted by a factor so that their median values for each sample were equal) and auto-scaled for statistical analysis, including Principal Component Analysis (PCA), Partial

Least Squares – Discriminant Analysis (PLS-DA), Random Forest, non-parametric Analysis of Variance (ANOVA) and Volcano plot analysis (fold change – FC *versus* p for t-test adjusted for false-discovery rate - p). Lipids were considered as significantly altered for $p < 0.05$ and $FC \leq 0.67$ or ≥ 1.5.

4.3 Results

The nanoLC-MS method allowed the MS/MS positive identification of 286 molecules and the putative identification of 8198 lipids (47.5% of the detected features for exosome and serum samples, Supp. Table 8). A high number of diacylglycerols (DG, 1135), phosphatidylcholines (PC, 849) and triacylglycerols (TG, 663) were identified (Figure IV-1). Abbreviations to lipid subclasses are described in Table II-2 (p. 48). Thirteen deuterated internal standards were detected with a maximum m/z error of 6.1 mDa or 7.6 ppm.

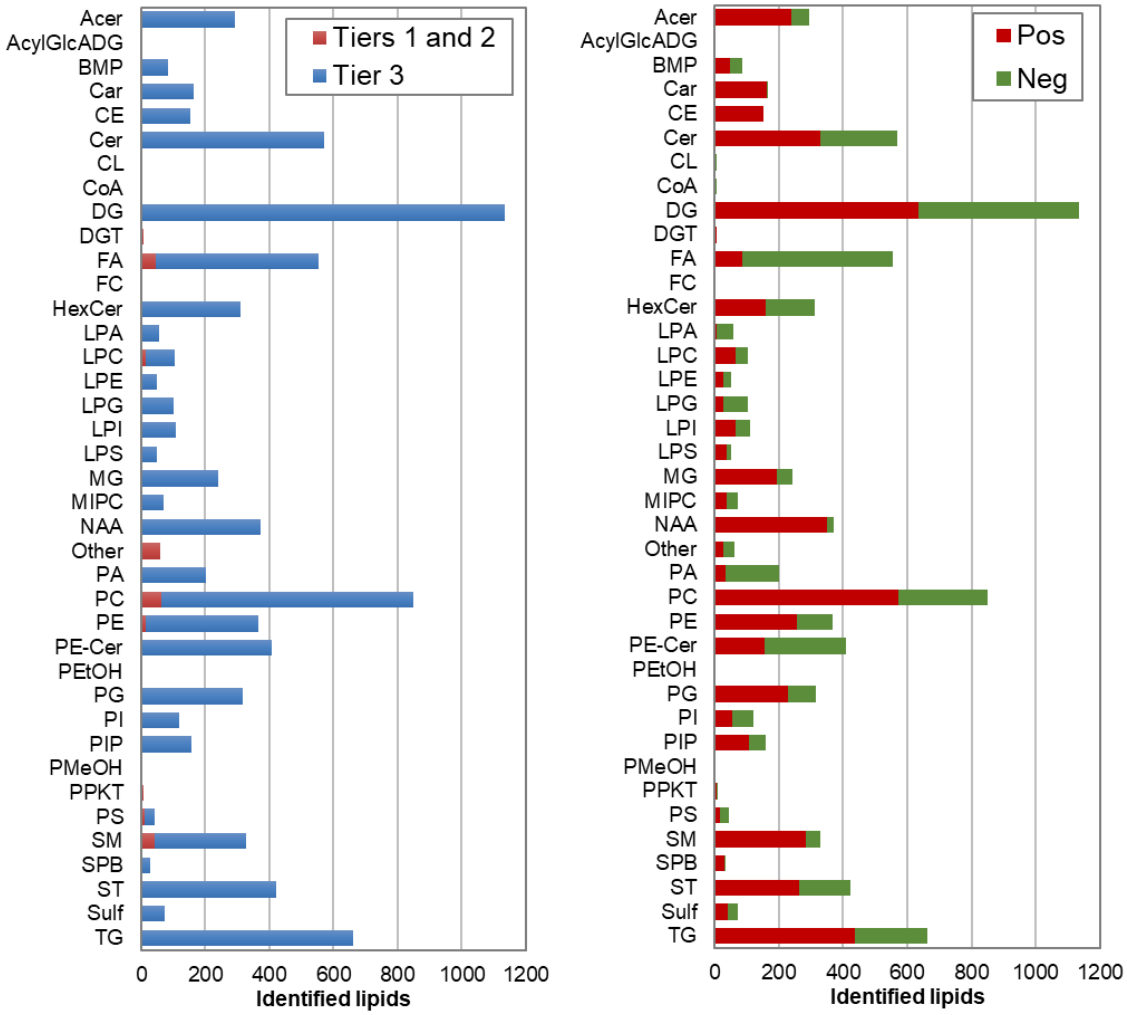

Figure IV-1. Subclass distribution of identified lipids for exosome samples, divided in identification level (tier 1: precursor m/z error ≤5.0 mDa, MS/MS score ≥500 and mSigma ≤150; tier 2: precursor m/z error ≤5.0 mDa, MS/MS score between 100 and 500, and mSigma ≤50; tier 3: putative identification by accurate m/z match with m/z error ≤5.0 mDa) and detection polarity (Pos: positive ionization, Neg: negative ionization). Abbreviations to lipid subclasses are defined in Table II-2.

4.3.1 Exosome isolation method

The lipidome of one exosome isolate obtained by size-exclusion chromatography (SEC), three obtained by ultracentrifugation (UC) and one obtained by ultracentrifugation followed by size-exclusion chromatography (SEC+UC) were first compared. The 11 QC extraction replicates (pooled exosome samples) were tightly clustered on the PCA score plot (Figure IV-2A), indicating good technical reproducibility for the employed methods. The extraction replicates for exosome samples were also clustered, as shown in the dendrogram (cluster analysis) in Figure IV-2B.

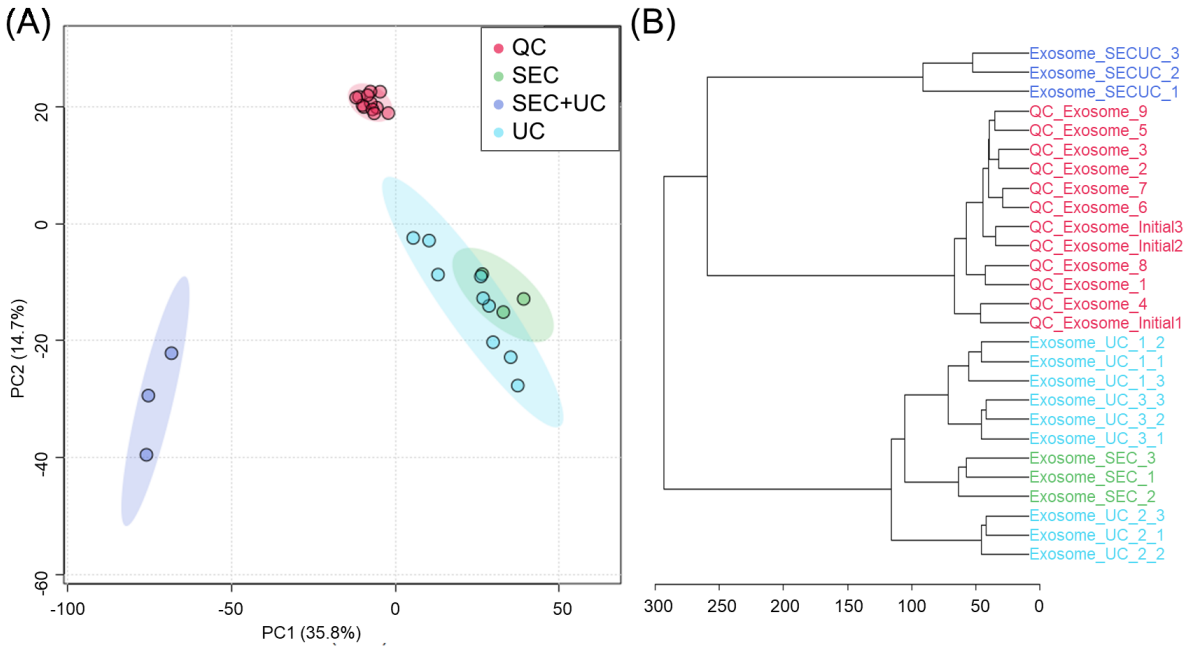

Figure IV-2. Statistical analysis for quality control of exosome isolated from serum samples of healthy humans by size-exclusion chromatography (SEC, green), ultracentrifugation (UC, light blue), a combination of both (SEC+UC, dark blue) and extraction replicates of a pool of all exosome samples used as quality control (QC, red). (A) PCA score plot showcasing clustered QC extraction and injection replicates; (B) dendrogram with extraction replicates for each sample and quality control injections clustered together.

Examples of chromatograms obtained for each type of sample are displayed in Figure IV-3. The results obtained with the previously developed high-sensitivity nanoLC-MS method were compared to chromatograms obtained under similar conditions by UHPLC-MS for undiluted exosome, 10× diluted blood serum and a blank extract (Figure IV-4). The comparison between Figure IV-3 and Figure IV-4 confirms the potential of nanoLC-MS to analyze low volumes of diluted samples.

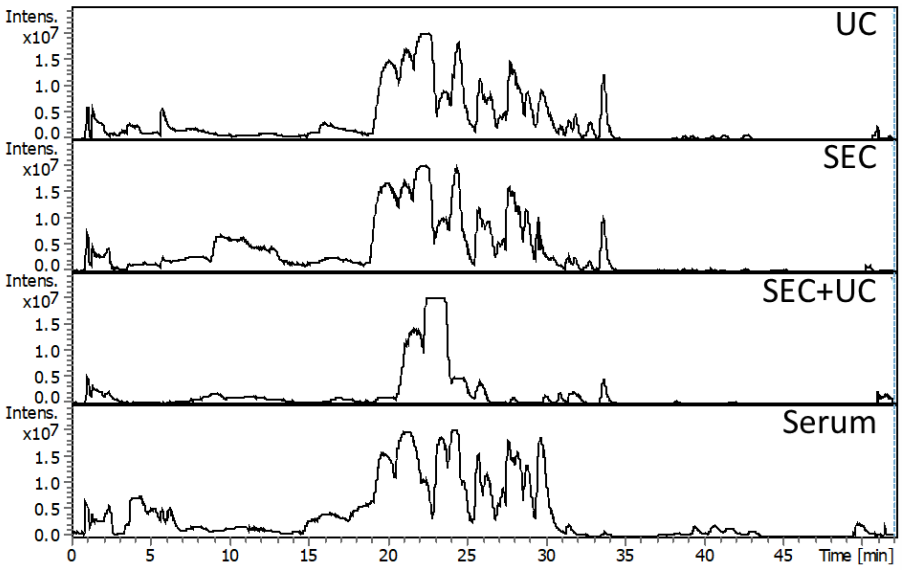

Figure IV-3. Examples of chromatograms obtained for blood serum (40× dilution) and exosome (no dilution) isolated from serum by UC, SEC and a combination of both (SEC+UC).

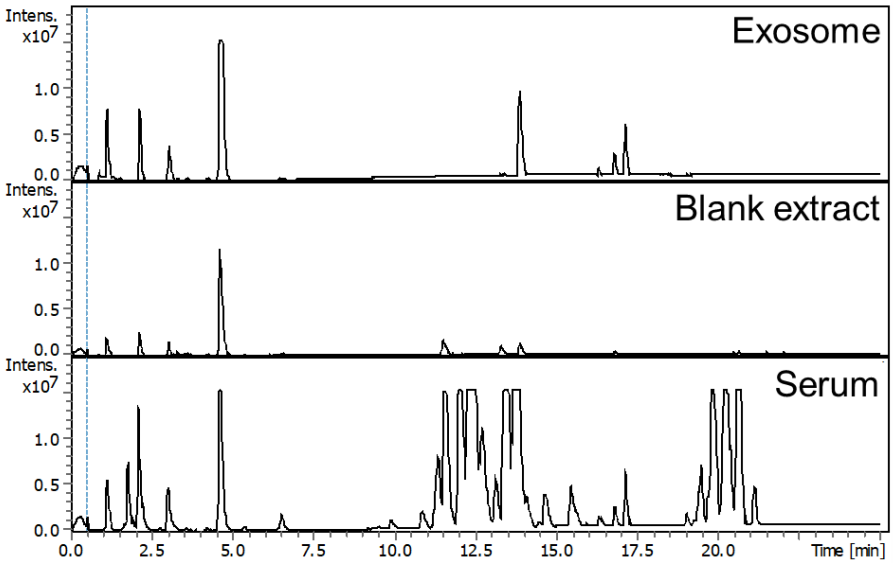

Figure IV-4. Chromatograms obtained for the analysis of undiluted exosome samples (top), a blank extract (extraction of water instead of samples) and blood serum diluted 10× using UHPLC-MS. Samples were extracted as described in the Experimental section. Analysis method: MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water; MPB - 10 mM NH_4COOH in 95:5 2-propanol/ water; 250 µL/min; 40 °C; 25 min gradient (0 min – 5% MPB; 1.8 min – 5% MPB; 8.5 min – 30% MPB; 18 min – 95% MPB; 25 min – 95% MPB) followed by 10 min of re-equilibrium (0 min – 95% MPB; 3 min – 95% MPB; 4 min – 5% MPB; 10 min – 95% MPB); and 4 µL injection.

The exosome isolate obtained by UC, SEC and SEC+UC were fully separated on PCA, indicating that the composition of the samples was affected by each technique (Figure IV-5A). Samples were also separated by PLS-DA (Figure IV-5B). Unfortunately, the PLS-DA model did not pass the permutation test ($p = 0.13$ for 1000 permutations) due to the overfitting caused by the high complexity of the dataset allied to an insufficient number of samples. Nevertheless, the results indicate that different isolation techniques will deeply affect the lipidome of exosome samples.

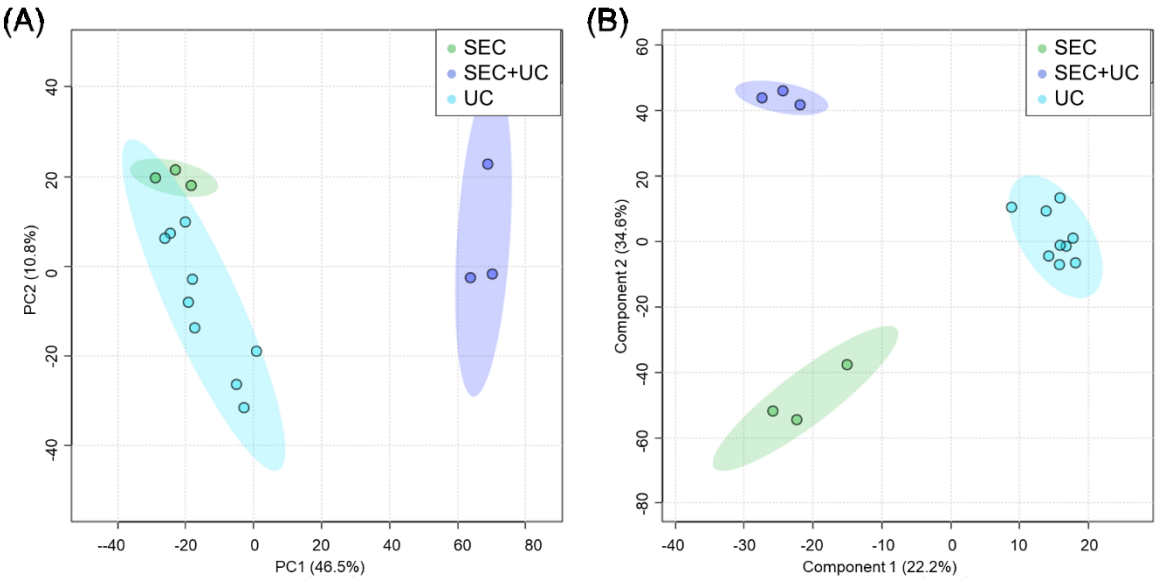

Figure IV-5. Statistical analysis for evaluation of exosome isolation methods from blood serum samples. (A) PCA score plot; (B) PLS-DA score plot with 2 components (R^2 = 0.9704, Q^2 = 0.7597, p = 0.13 for 1000 permutations).

The one-way ANOVA for the three studied isolation techniques resulted in 1750 significant lipids (Supp. Table 9, 62.8% of the lipids employed for statistics). Glycerolipids and glycerophospholipids had the highest number of significant lipids (19.3% and 51.2%, of the significant lipids, respectively).

Binary comparisons through Volcano plot analysis showed specific differences related to each method (Figure IV-6). The exosome samples obtained by UC were first compared with SEC, resulting in 251 significantly altered lipids, including 195 with higher normalized intensities for SEC (Figure IV-6A, Supp. Table 10). The most affected lipid subclasses for the comparison between UC and SEC were diacylglycerols (DG, 103 significant lipids with 99.0% significantly elevated for SEC); phosphatidylcholines (PC, 27 significant lipids with 70.4% elevated for UC); and ceramides (Cer, 24 significant lipids with 91.7% elevated for SEC). PCs and Cer are typically

enriched in exosome membranes, while triacylglycerols (TG) and cholesteryl esters (CE) are more commonly found in lipid droplets and lipoprotein particles that may be co-isolated. However, as the endosomes that form exosomes are related to LDL metabolism, it is not surprising to find CE and TG species in exosome isolates. For UC and SEC, 43.8% of the glycerolipids employed for statistics and 39.0% of the glycerophospholipids were significantly altered. Most of the glycerolipids were elevated for the SEC isolation, but only 5 TGs were significantly affected, all with higher normalized intensities for UC. CEs were also not deeply altered, with only 2 significant lipids (both elevated for UC). The glycerophospholipids were split, with 55.1% with higher normalized intensities for SEC, while the remaining 44.9% were elevated for UC. Sphingolipids were also enriched for SEC, with 94.1% of the 34 significantly altered lipids showing higher normalized intensities. Overall, the comparison between UC and SEC resulted in minor alterations for the lipids that are typically found in undesired particles.

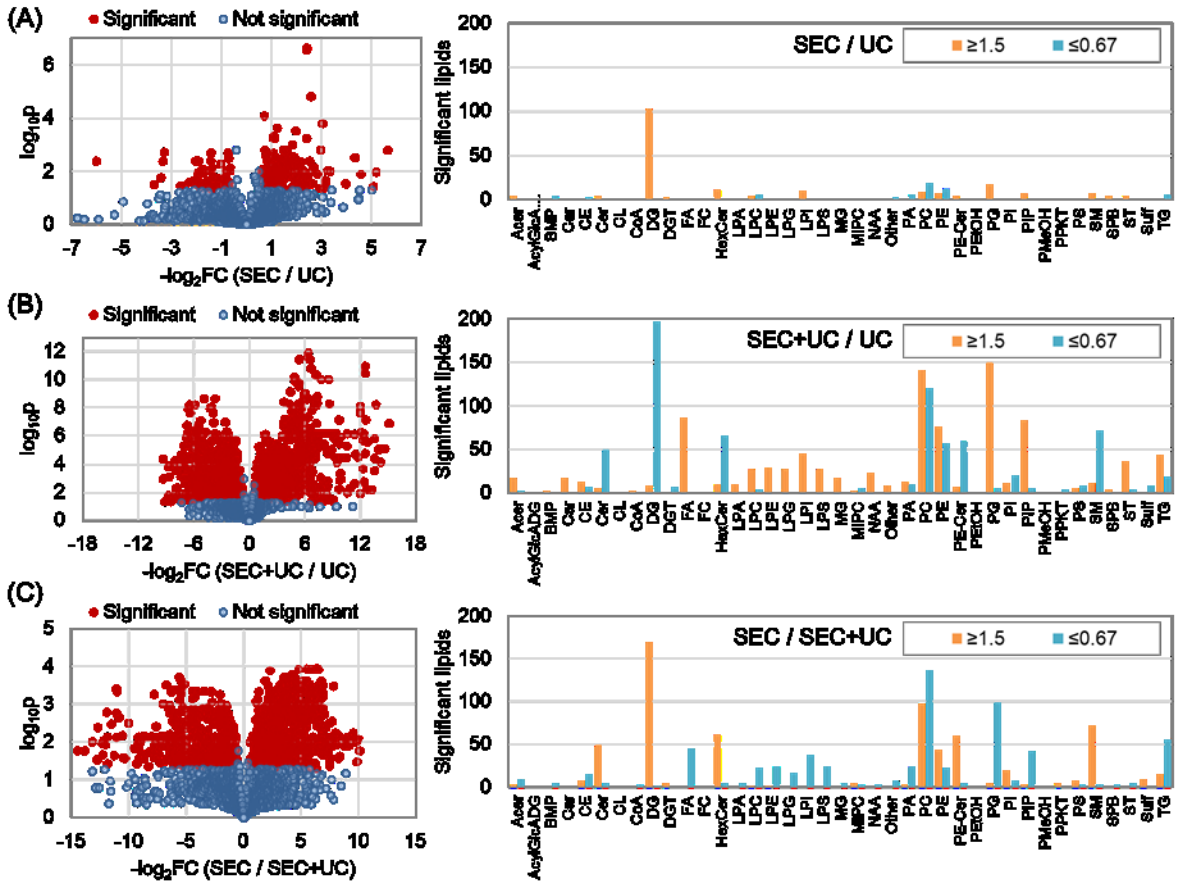

Figure IV-6. Volcano plot analysis for the binary comparisons between exosome samples obtained from blood serum by UC, SEC and a combination of both (SEC+UC). (A) SEC *versus* UC; (B) SEC+UC *versus* UC; (C) SEC *versus* SEC+UC. Abbreviations to lipid classes are defined in Table II-2 (p. 48).

Second, the exosome samples obtained by UC were compared to UC followed by SEC (UC+SEC). A total of 1661 lipids were significantly altered (59.6% of the lipids employed for statistics), with 56.5% displaying higher normalized intensities for the UC+SEC combination (Figure IV-6B, Supp. Table 11). The most affected lipid subclasses were also diacylglycerols (DG, 205 significant lipids with 197 elevated for UC); phosphatidylcholines (PC, 260 significant lipids with 140 elevated for SEC+UC); and ceramides (Cer, 219 significant lipids with 181 elevated for

UC). Glycerophospholipids were deeply affected for the UC/UC+SEC comparison, with 867 significantly altered lipids (73.6% with higher normalized intensities for UC+SEC). Most glycerolipids (76.5% of significantly altered glycerolipids) and sphingolipids (83.6%) were enriched for UC. All the 127 significantly altered fatty acyls were also found to be elevated for UC+SEC. Unfortunately, 70.5% of the 61 significantly altered triacyclglycerols and 66.7% of the 18 cholesteryl esters were elevated for UC+SEC, possibly indicating co-isolation of lipid droplets and particles.

Third, the exosome samples obtained by SEC were compared to UC followed by SEC (UC+SEC). A total of 1252 lipids were significantly altered (44.9% of the lipids employed for statistics), with 50.4% displaying higher normalized intensities for SEC (Figure IV-6C, Supp. Table 12). The most affected subclasses included once again diacylglycerols (DG, 170 significant lipids with 99.4% elevated for SEC), phosphatidylcholines (PC, 233 significant lipids with 58.4% elevated for UC+SEC); and ceramides (Cer, 176 significant lipids with 88.1% elevated for SEC). Similarly to the observations for UC *versus* UC+SEC, glycerophospholipids were deeply affected by the SEC and UC+SEC comparison, with 637 significant lipids (72.2% elevated for UC+SEC). All the 49 significantly altered fatty acyls were also increased for UC+SEC, while most sphingolipids (75.8%) and glycerolipids (90.4%) had higher intensities for SEC. Most significant triacylglycerols (TG, 79.7% of 69 lipids) and cholesteryl esters (CE, 68.2% of 22 lipids) were elevated for UC+SEC.

4.3.2 Comparison between serum and exosome lipidome

The lipidome of five exosome isolates obtained by SEC of blood serum samples was compared to the lipidome of the five original serum samples. The comparisons between exosome isolates and blood serum must be considered along with the different dilutions employed for

analysis (no dilution for exosome and 40× dilution for serum); however, both types of samples were normalized by the median, *i.e.,* all samples were scaled so that they have the same median normalized intensity to allow for direct comparisons. Even though exosome isolates are much more diluted than blood serum samples, the normalization procedures adopted herein are expected to compensate for the intensity differences so that the composition of the different types of samples are comparable by statistical models. Further information on data normalization and scaling for statistics of untargeted lipidomics is available in Chapter I (*1.3.1.7. Statistics applied to lipidomics*, p. 32)

The 8 QC extraction replicates were tightly clustered on the PCA score plot, indicating the suitability of the employed methods (Figure IV-7A). Extraction replicates were also clustered, as shown in the dendrogram (cluster analysis, Figure IV-7B). The exosome isolates were fully separated from serum samples by PCA and PLS-DA (Figure IV-8A and B). Unfortunately, the PLS-DA model did not pass the permutation test (p of 0.13 for 1000 permutations) due to the small number of samples. However, all samples were correctly classified by a Random Forest model, which is less prone to overfitting, with 7 predictors and 500 trees (out-of-bag, OOB error of zero).

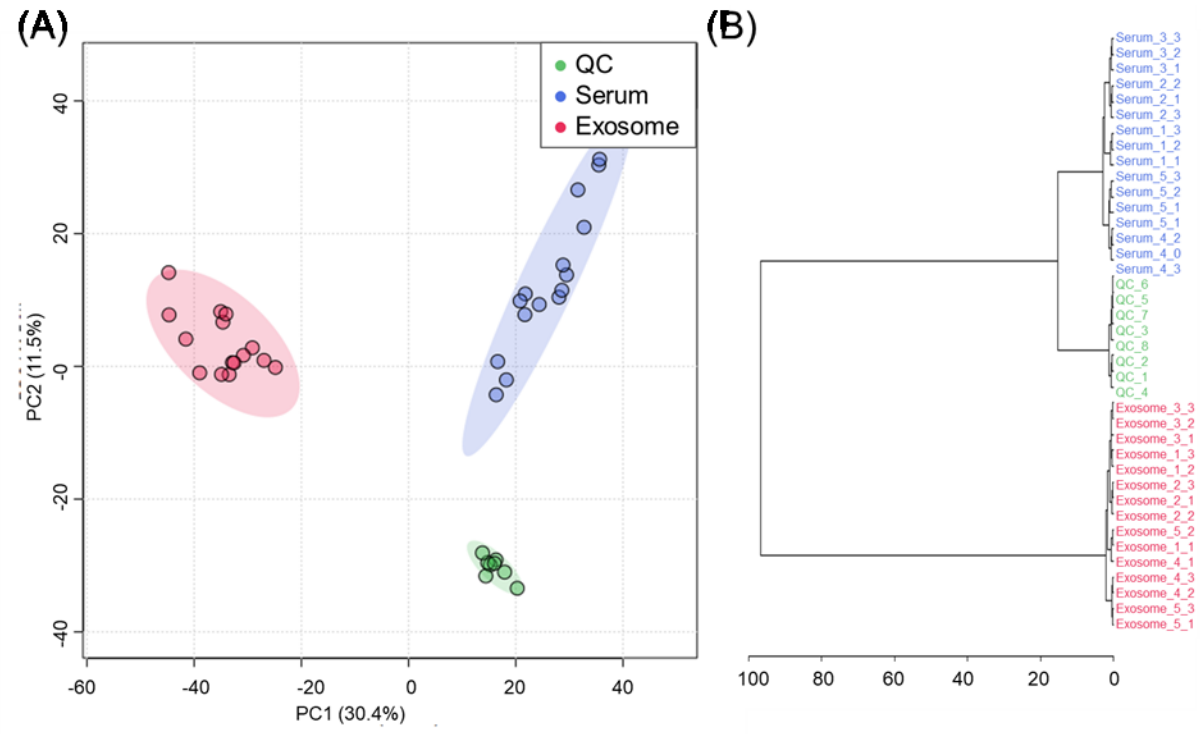

Figure IV-7. Statistical analysis for quality control of the comparison between the lipidome of exosome samples obtained from serum by UC and blood serum. (A) PCA score plot; (B) dendrogram (cluster analysis) with extraction triplicates for each of the five exosome and five serum samples.

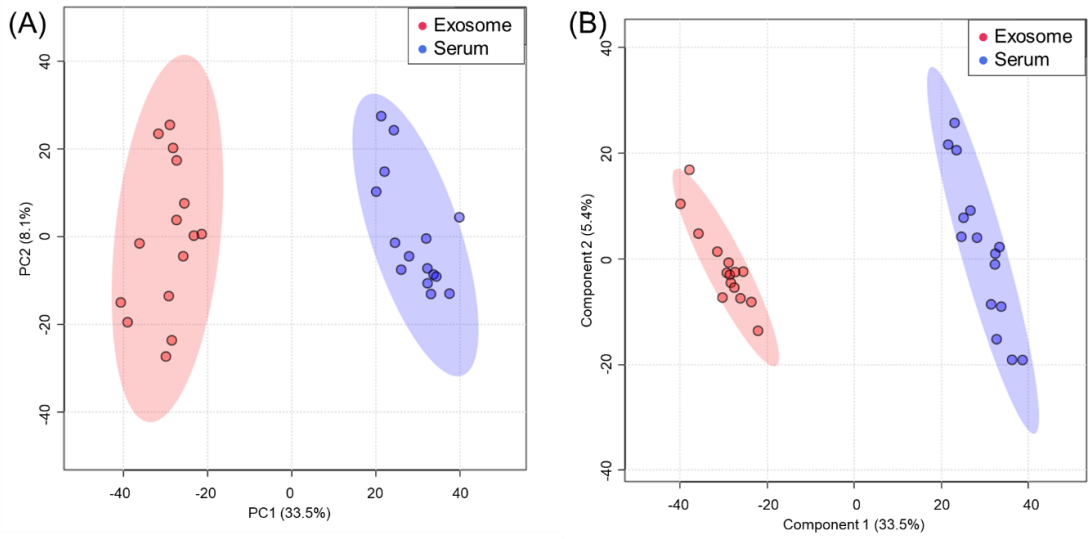

Figure IV-8. Statistical analysis for the comparison between exosome samples obtained from blood serum by UC and serum samples. (A) PCA score plot; (B) PLS-DA score plot with 2 components (R^2 of 0.9727, Q^2 of 0.9761 and p = 0.13 for 1000 permutations).

The Volcano plot analysis for serum and exosome samples resulted in 2102 significantly altered lipids (75.6% of the lipids employed for statistics, Figure IV-9, Supp. Table 13). As expected, we observed deep differences between the lipidomes. The most affected subclasses included ceramides (Cer, 352 significant lipids, 54.8% with smaller intensities for exosome); phosphatidylcholines (PC, 279 significant lipids, 59.5% with smaller intensities for exosome); and triacylglycerols (TG, 205 significantly altered lipids, 65.4% with higher intensities for exosome). We also observed that 92.8% of the significant sterols, 95.8% of the bis(monoacylglycero)phosphates (lysobisphosphatidic acids), 93.5% of fatty acids, and 90.3% of N-acyl amines displayed higher normalized intensities for the exosome samples.

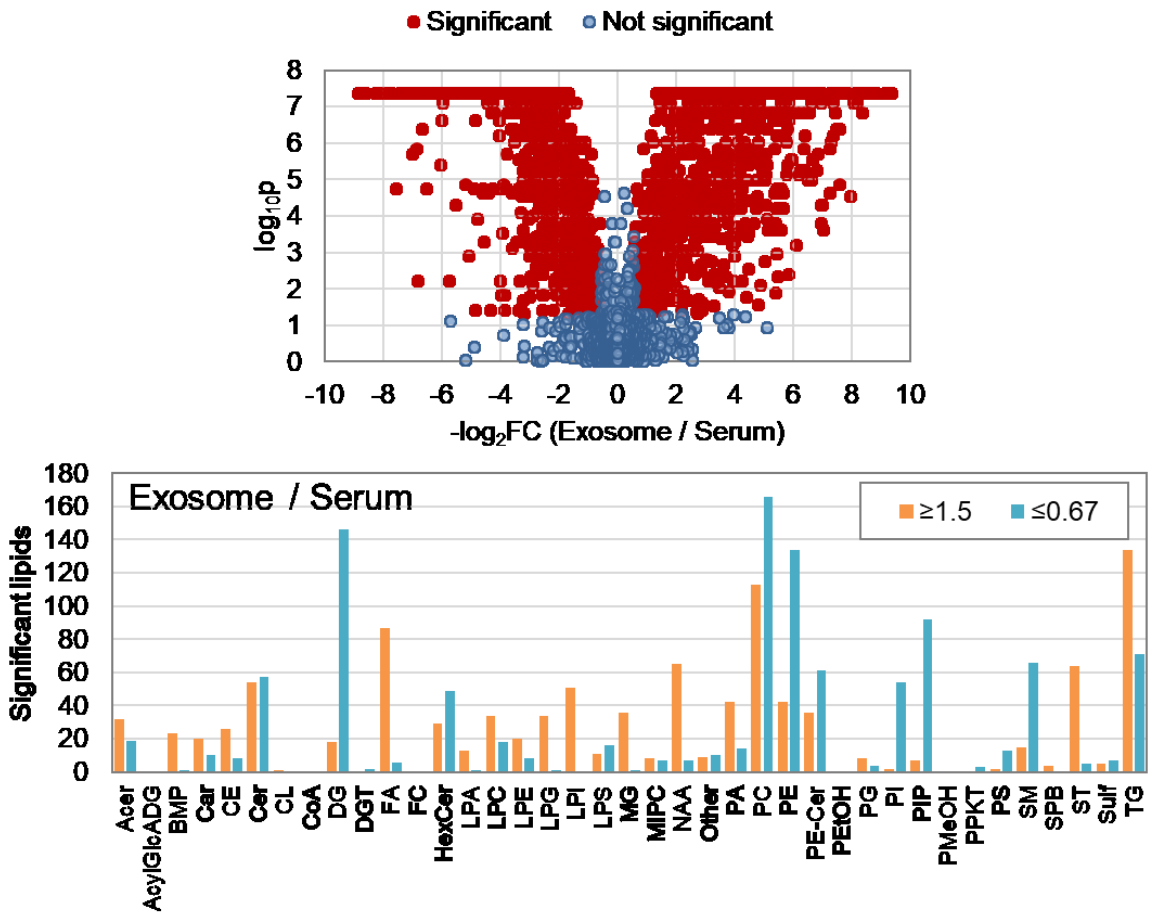

Figure IV-9. Volcano plot analysis for the comparison between exosome samples isolated from blood serum by UC and serum samples. Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

4.4 Discussion

We have performed untargeted, comprehensive lipidomics of exosome samples to explore the effect of different isolation techniques upon lipidic profile. First, we compared the lipidome of samples isolated by UC, SEC and a combination of both techniques (UC+SEC). UC and SEC resulted in similar lipidic profiles, but the combination of both caused a dramatic change. Second, we compared the lipidome of exosome samples isolated by UC from blood serum and the original

serum samples. The concentrations of both types of samples were normalized to the same median value and auto-scaled for accurate comparisons, as exosome samples are much more diluted than serum. The samples employed for this work were evaluated elsewhere for protein markers (CD9, CD63, CD81, albumin, apoliproteins, and total protein content), size distribution (transmission electron microscopy) and concentration of particles (NanoSight).[165,166]

The nanoLC-MS approach applied for this work includes novel aspects when compared to previous studies on the lipidome of exosome samples. The employed method offers extremely high sensitivities for the diluted exosomes, allowing the detection and identification of thousands of lipids belonging to a variety of subclasses. Comprehensive untargeted lipidomics of exosome using mass spectrometry has rarely been described due to the difficulty in analyzing the extremely diluted samples. High volumes of blood serum are required to provide enough exosome sample volume to match the limits of detection of traditional LC-MS *omics* methods. However, the nanoLC-MS method development described in Chapter II, with further adaptations for exosome, allowed the identification of thousands of lipids using only 30 μL of each sample. We found that undiluted exosome lipid extract provided peak intensities that were comparable to 40× diluted serum samples, a dilution factor that is not applicable for most UHPLC-MS methods.

The comparison between UC and SEC for isolation of exosomes from blood serum showed the effect of each technique upon the obtained samples. UC favored higher normalized intensities of 56 lipids, including all 5 significantly altered triacylglycerols and 5 cholesteryl esters. The hydrophobic lipid classes are not expected to be enhanced in extracellular vesicles but are more likely to be related to co-isolated lipid droplets and lipoprotein particles. However, a few triacylglycerol and cholesteryl ester species have been found in exosome fractions obtained from cancer cells by shotgun mass spectrometry analysis in the past, which corroborates our hypothesis

that previous assumptions about the composition of exosomes may not be entirely correct.[164] We also observed an enhancement for the intensities of all 3 significantly altered bis(monoacylglycero)phosphates (*i.e.*, lysobisphosphatidic acids, BMP), the 5 significantly altered phosphatidic acids, 70.4% of the phosphatidylcholines and 63.2% of the phosphatidylethanolamines. Glycerophospholipids are typically enriched in the membranes of exosomes; hence, these observations indicate better isolation of the desired extracellular vesicles by UC. However, the results show that both techniques are suitable to obtain reliable exosome isolates. On the other hand, the comparison of both isolation techniques with the combination of UC+SEC resulted in a high number of significantly altered lipids. The combination of SEC and UC should provide cleaner exosome samples, but the use of SEC columns demands expensive disposable materials. Overall, the highly hydrophobic triacylglycerols and cholesteryl esters were elevated for the combination of UC+SEC, while ceramides and glycerophospholipids, typically enhanced in exosome membranes, had higher normalized intensities for both separated techniques (no combination). Such observations indicate that the extra sample handling required by UC+SEC may cause contamination of the isolates. The obtained results show that the extra time and effort required by the combination of two isolation techniques is not advantageous.

We further compared the lipidome of exosome isolates with blood serum samples. About half of the 2101 significantly altered lipids (49.7%) were enhanced for the exosome samples. We found higher levels of most triacylglycerols (65.4% of the significantly altered species) and cholesteryl esters (76.5%). Although that is usually seen as an indication that the exosome isolates may also contain lipid droplets and particles, we can't dismiss the possibility of specific triacylglycerols and cholesteryl esters being packed inside the exosomes. The small vesicles originate from mature endosomes, which are also related to the metabolism of low-density

lipoproteins (LDL) composed by a hydrophobic core of triacylglycerols and cholesteryl esters. The hydrophobic lipid subclasses are not usually reported in exosomes, but the technique employed herein provides high sensitivity, which will result in the detection of different lipid species that are usually not easily ionized, such as the more hydrophobic lipids. This possibility is corroborated by our findings for the comparison between isolation by UC and SEC: only a very small number of triacylglycerols and cholesteryl esters were significantly altered for the techniques, indicating that both provide reliable results for exosome isolation. Further studies are required to confirm whether the significantly altered TGs and CEs are from the exosome vesicles or co-isolated lipid particles.

We also found a high proportion of fatty acyls (88.2% of the significantly altered fatty acyls), lysoglycerophospholipids (78.7%), phosphatidic acids (75%) and bis(monoacylglycero)phosphates (BMP, 95.8%) with increased intensities for exosome samples. Phosphatidic acids and BMPs can trigger membrane fusion between exosomes and target cells through hydrolysis, allowing the release of exosomal contents into the cell's cytosol.[167] The biological pathways of cholesteryl esters and BMPs are deeply interconnected. When low-density lipoproteins (LDL) are metabolized by endosomes and lysosomes, free cholesteryl esters and triacylglycerols are obtained. The LDL-derived cholesterol is then further processed and transported by endosomes, which can eventually mature into multivesicular bodies. BMPs are also abundant in the internal membranes of endosomes, being involved in their structure and functions. The production of exosomes starts with the biogenesis of multivesicular bodies in the endosome, followed by transport to the plasma membrane, fusion and further release of the intraluminal vesicles into the extracellular fluid. [162] Hence, the presence of BMPs in exosome, as well as cholesteryl esters and triacylglycerols, is probably a consequence of the composition of the

intraluminal vesicles that were eventually excreted as exosomes. The lipid subclass is also actively involved in cholesterol transport, distribution and oxidation.[168]

We recognize that this work is limited by a small number of samples. However, our goal was to complete a pilot study to confirm that the nanoLC-MS methodology that we previously optimized (Chapter II) could be employed to study the extremely diluted exosome samples with minor modifications. The results described herein confirm the potential of nanoLC-MS untargeted lipidomics to evaluate the composition of exosomes and other diluted samples.

4.5 Conclusions

Exosomes are a vital part of inter-cell communication and the trafficking of biologically active molecules. Although their impact upon the organism is highly recognized, their physiological and pathophysiological functions are not well known. The lipidome of exosome samples may play a role in homeostasis and pathogenesis of different diseases, but untargeted, comprehensive lipidomics has rarely been described due to the complexity and low concentrations of exosome isolates. We have previously developed a high sensitivity nanoLC-MS methodology for the lipidomics of small volumes of diluted biological samples (Chapter II) and confirmed its applicability to serum, cerebrospinal fluid and intraparenchymal microdialysate samples (Chapter III). Now, we completed a pilot study that shows that the method can be easily adapted to study exosome. The method was employed to study the differences between exosome samples isolated from blood serum by ultracentrifugation and size-exclusion chromatography, with both techniques generating similar results. We have also compared the lipidome of exosome with the parent blood serum, resulting in over 2000 significantly altered lipid species. In the future, we will improve the MS/MS conditions to obtain a higher proportion of positively identified lipids, which will provide

further insights into exosomal lipids. The method described herein can be further applied to study human diseases, *e.g*. cancer. The application of the methodology described in this work may help finally unravelling the role of exosomes in human physiology and pathogenesis.

4.6 Acknowledgments

We thank the collaboration of Dr. David M. Lubman and Dr. Jianhui Zhu (University of Michigan Medical Center, Ann Arbor, MI, USA) for the sample acquisition and discussions. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research, the Canada Research Chairs program, the Canadian Foundation for Innovations, Genome Canada and Alberta Innovates.

V

Chapter V: Comprehensive Lipidomic and Metabolomic Analysis of Lung Tissue for the Study of a Novel Vaccine Against Respiratory Syncytial Virus[9,10]

5.1 Introduction

Respiratory syncytial virus (RSV) is the main cause of acute lower respiratory infection in infants and young children. RSV incidence and mortality rates are difficult to assess due to the infrequent use of diagnostic tests, but the annual global burden is estimated at 33.1 million new cases in children younger than five years old and over 100,000 deaths every year.[169,170] The disease is transmitted by contact with nasal or oral secretions and can repeatedly infect the public, causing a particularly high risk for pediatric and elderly populations.[170,171] Adults with cardiac, pulmonary or immune-compromised conditions are also at high risk for severe infection.[169]

Treatment options for RSV infection are usually limited to supportive care, *e.g.,* supplemental oxygen and mechanical ventilation.[171] The humanized neutralizing monoclonal IgG1 antibody Palivizumab is used to prevent RSV in premature infants and high-risk populations; however, the high cost and monthly dosing required by the medication are prohibitive for widespread use. Although the virus has been known for several decades, no vaccines have yet been approved for human applications. A formalin-inactivated RSV vaccine was administered to infants and children during a clinical trial in 1966, resulting in abnormal immune responses that caused

[9] Versions of this chapter were published at ACS Infectious Diseases (Zardini Buzatto *et al.*, ACS Infectious Diseases 2020) and Journal of Proteome Research (Sarkar *et al*., Journal of Proteome Research 2019, 18(3), 1145-1161).[1,2]

[10] Supporting figures for this chapter are available in Appendix B. Supporting tables are available with Dr. Liang Li.

severe RSV infections and the death of two toddlers. The devastating results of this trial, as well as the increasing disbelief from the general population on vaccine safety and efficacy, warrant better characterization of vaccine-induced alterations in tissues and fluids before application in human subjects.[169]

RSV is a negative-sense RNA orthopneumovirus containing a lipid envelope with three transmembrane glycoproteins: fusion (F), attachment (G) and small hydrophobic protein.[172] The F protein facilitates the fusion of the virus with the host cell to promote viral penetration and is highly conserved amongst RSV viral strains, constituting a major candidate for novel vaccine formulations.[169,173] A vaccine consisting of the subunit F-protein is a safer alternative to the live-attenuated virus and is particularly interesting for elderly and maternal immunization. The RSV-neutralizing antibody has been previously shown to be transferred from mother to fetus through the placenta, immunizing the fetus before birth.[171,174,175] However, non-replicating vaccines based on subunit agents require adjuvants to augment their immunogenicity and stimulate long-lasting immune responses.[176] We have previously developed a novel intranasal RSV vaccine formulation composed of a truncated version of the F protein (ΔF) and a combination adjuvant (TriAdj) that promoted the production of cytokines, chemokines and interferons, as well as activation of dendritic cells, macrophages and neutrophils. The vaccine elicited complete protection from RSV challenge in mice.[1,173,175,177,178] The immune responses to the ΔF/TriAdj formulation, as well as the inflammatory responses of immunized and non-immunized mice following an intranasal challenge by RSV, have been recently evaluated elsewhere.[1]

In this study, we applied an integrated liquid chromatography – mass spectrometry lipidomics and metabolomics approach to study the metabolic changes induced by RSV in the lungs of ΔF/TriAdj-immunized and non-immunized mice. We hypothesized that the ΔF/TriAdj

formulation would mitigate the metabolic changes caused by RSV challenge in immunized mice, resulting in a metabolic and lipidic profile more similar to healthy controls. Lipids are not only a major energy source but also act as signaling molecules, hormones, modulators of gene expression and membrane components. The immune response depends on a cascade of signals that may include changes in cell membranes and phospholipid profiles, alterations in eicosanoid and polyunsaturated fatty acids (PUFA), and lipid oxidation and peroxidation. Exogenous and endogenous lipids and lipid-derived messengers play a critical role in the modulation of immune responses to pathogenic viruses and bacteria.[24,26,179] We aimed to further investigate the mechanism of action of the ΔF/TriAdj formulation by unraveling metabolic and lipidic alterations after intranasal immunization followed by RSV challenge, which may provide crucial information for the improvement of vaccine efficacy and safety, as well as for the future development of new interventions.

5.2 Experimental

5.2.1 *Immunization and RSV challenge*[11]

The ΔF/TriAdj vaccine formulation was prepared with a combination of three adjuvants, namely: polyinosinic:polycytidylic acid (poly(I:C)), innate defense regulator peptide 1002 (IDR1002), and poly[di(sodiumcarboxylatoethylphenoxy)]-phosphazene (PCEP). The preparation of the formulation was described elsewhere.[177] Briefly, an episomal vector was used to transfect HEK-293 cells with a truncated version of the RSV fusion protein (ΔF) without the transmembrane

[11] The vaccine preparation, immunization of animals, RSV challenge and collection of lungs were performed by Dr. Indranil Sarkar and Dr. Sylvia van Drunen Little-van den Hurk (Microbiology and Immunology, University of Saskatchewan, Saskatoon, Canada).

domain that was his_{10}-tagged at the carboxyl terminus. The truncated ΔF protein was then purified by affinity chromatography with TALON Superflow resin (Clontech, CA, USA) and formulated with poly(I:C) (Invivogen, CA, USA), IDR1002 (VQRWLIVWRIRK, Genscript, NJ, USA) and PCEP (Idaho National Laboratory, ID, USA) at a ratio of 1:2:1 poly(I:C) / IDR1002 / PCEP in PBS (pH 7.4, Life Technologies, ON, Canada).[177]

One cohort composed of three groups of five female, age-matched (six to eight-week-old) BALB/c mice (Charles River Laboratories, QC, Canada) were selected for this study (N = 5). Intranasal immunization of one group of five mice was performed with 20 µL of the vaccine formulation containing 1 µg of ΔF protein, 10 µg of poly(I:C), 20 µg of IDR1002 and 10 µg of PCEP (named group B). Three weeks after immunization, the immunized animals, along with five non-immunized mice (group C), were challenged with the RSV A2 strain (5×10^5 plaque-forming units in 50 µL, ATCC, VA, USA). Five untreated mice were used as controls (Group A). We selected only female mice to avoid cage effects, as males and females can not be intermingled in the same cage. As age and sex-matching are vital for reliable studies, we opted for only female mice to allow for comparisons with our previous related work. [1] The animal trial was carried out according to the guidelines established at the University of Saskatchewan following the Canadian Council on Animal Care.

The ten challenged animals and five non-immunized, age-matched healthy controls were sacrificed seven days post-RSV challenge, followed by the collection of both lung lobes. The lung tissue samples were washed with ice-cold physiological saline (0.85% NaCl), blot-dried with sterile gauze and snap-frozen in liquid nitrogen. All samples were kept at -80°C until analysis and thawed in a 4°C fridge immediately before preparation.

5.2.2 *Instrumentation*

Chromatographic separation was performed separately for lipidomics and metabolomics using a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific, Waltham, MA, USA) and a Waters Acquity BEH C18 column (5 cm × 2.1 mm with 1.7 µm particles; Waters Corporation, Milford, MA, USA). An ultra-high-resolution Maxis II QTOF mass spectrometer instrument (Bruker Daltonics, Billerica, MA, USA) equipped with an electrospray ionization (ESI) source was used for the detection of lipids and labeled metabolites.

5.2.3 *Tissue homogenization*

Both lung lobes from each animal were collected for this study. Each lung lobe was treated separately, *i.e*., one lobe was employed for lipidomics and dansyl chloride (DnsCl) labeling of amine and phenol-containing metabolites, while the other was used for p-(dimethylamino)phenacyl (DmPA) labeling of carboxylic acid-containing metabolites. All three procedures started with homogenization followed by liquid-liquid extraction (LLE) with organic solvents. The volumes of all reagents were normalized for the wet mass of tissue to ensure accurate relative comparisons.

Different sections of tissue samples may have a different distribution of metabolites and lipids; therefore, we opted not to divide the tissue samples for experimental replicates. Nevertheless, the introduction of an extra step to split and re-weigh tissue aliquots would require one extra freeze-thaw cycle and more time exposing the samples to room temperature, light and oxygen, which would affect the composition of the tissue samples. Hence, we opted for 5 biological replicates in each group (*i.e.*, five mice, N = 5), but we didn't perform full experimental replicates. Instead, we have performed labeling triplicates for metabolomics and LC-MS injection

duplicates for lipidomics. Sample extraction and analysis were randomized to minimize batch and confounding effects.

The first lung lobe (70 to 100 mg) from each animal was mixed with the Splash Lipidomix Mass Spec standard mixture (0.35 µL / mg of tissue, Table II-1, p. 45). Each lung lobe was then homogenized (Bio-Gen PRO200 Homogenizer, PRO Scientific Inc., Oxford, CT, USA) for three consecutive 30s intervals in methanol (4.0 µL / mg of tissue) and water (0.85 µL / mg of tissue), followed by 30 s rest periods in an ice bath. The homogenate was further processed by liquid-liquid extraction (LLE) with dichloromethane and water (4.0 and 2.0 µL / mg of tissue, respectively). After three 30 s vortex cycles with 30 s intervals in an ice bath, the mixture was kept in ice for equilibration for 15 min and centrifuged for 10 min at 4°C (10,000 rpm). The organic layer was employed for lipidomics, while the aqueous layer was used for DnsCl labeling of amine and phenol-containing metabolites. The remaining lung lobe from each animal was employed for the chemical isotope labeling of carboxylic acid groups with DmPA.

The sample preparation steps were performed in 2 mL polypropylene microcentrifuge tubes (Rose Scientific, Edmonton, AB, Canada). We also prepared one blank extract for each group of five samples for controlling possible chemical contaminations from plastic tubes by substituting the tissue samples for water (1.0 µL / mg of tissue) and the internal standard solution for pure methanol. The blank extracts had low intensities and only displayed a small number of reproducible chromatographic peaks (Appendix B - Figure B- 1). Lipid internal standard mix and amino acid standard extracts were also prepared with water instead of tissue samples for quality control of the extraction, labeling and analysis procedure.

5.2.4 *Lipidomics*

After homogenization and LLE of the first lung lobe with dichloromethane, methanol and water, the upper aqueous layer was separated for metabolomics via dansyl chloride labeling of amine and phenol groups, while the organic layer was evaporated to dryness on a SpeedVac for 30 min for untargeted lipidomics. The residue was immediately resuspended in a mixture of 6:4 mobile phase A (MPA) / mobile phase B (MPB) (0.20 µL / mg of tissue), vortexed for 30 s and diluted with 0.50 µL / mg of tissue of 9:1 MPA / MPB. Each resuspended sample was divided into two vials for randomized injection duplicates. A pooled mixture composed by aliquots of all sample extracts was prepared for quality control (QC) injections and CID-MS/MS for lipid identification. As there are no other sample preparation steps after homogenization besides drying and resuspension, we opted for injection duplicates from separate vials for lipidomics. Samples were kept at 4°C in polypropylene inserts (Canadian Life Science, Peterborough, ON, Canada) in amber injection vials (Agilent Technologies, Santa Clara, CA, USA) capped with polytetrafluoroethylene (PTFE)-lined septa (Waters Corporation, Milford, MA, USA), and injected within a maximum of 2 days after the extraction.

Samples were analyzed by reversed-phase ultra-high-performance liquid chromatography coupled with electrospray ionization and quadrupole time-of-flight mass spectrometry (RP-UHPLC-ESI-QTOF-MS) under positive and negative ionization. LC-MS lipid analysis was performed with the following conditions: MPA – 10 mM ammonium formate in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM ammonium formate in 95:5 isopropanol/ water (v/v); 250 µL/min; column temperature of 40°C; injection volume of 1.0 µL for positive ionization and 4.0 µL for negative ionization; 22 min gradient (0 min – 5% MPB; 1.8 min – 5% MPB; 8.5 min – 30% MPB; 18 min – 95% MPB; 22 min – 95% MPB) followed by 10 min of re-

equilibrium (0 min – 95% MPB; 3 min – 95% MPB; 4 min – 5% MPB; 10 min – 95% MPB); and ESI-QToF detection (capillary voltage of 4500 V; end plate offset of 500 V, nebulizer gas pressure of 1.0 bar, dry gas flowrate of 8.0 L/min; dry temperature of 230°C; spectra acquisition rate of 1 Hz, and m/z range of 150 to 1500 Da). QC pooled samples composed of equal aliquots of all sample extracts were injected six times before the sample sequence to stabilize chromatographic retention and MS signal. The 15 sample extracts (N = 5 for the three groups) were randomly divided into two sets, sandwiched between QC injections. Each set was injected twice, alternating between positive and negative ionization.

MS/MS spectra were acquired for identification with the QC pooled mixture employing variable collision energies (10 to 80 eV) and injection volumes (1.0 to 5.0 µL). MS/MS fragmentation was performed on auto MS/MS mode, *i.e.,* the mass spectrometer software (oToF Controls, Bruker Daltonics) automatically selected the most intense ions during precursor scans for MS/MS fragmentation. Hence, we have also acquired MS/MS chromatograms with variable mass ranges, *i.e*., the mass range for selection of precursor ions was reduced to 100 Da intervals to allow for fragmentation of ions that would not be chosen otherwise due to lower intensity when compared to other co-eluting molecules (*e.g.* only ions with precursor m/z between 150 and 250 were selected for the first chromatogram; 250 to 350 for the second chromatogram, etc.). Unfortunately, many lipids with lower intensity when compared to co-eluting molecules were still not selected for fragmentation or provided fragments with very low intensities that could not be used for confident identification. In the future, potential biomarkers or molecules of interest could be further investigated with targeted analysis to improve their identification; however, this work was focused on untargeted, comprehensive lipidomics for relative quantification.

The obtained chromatograms were re-calibrated based on the direct infusion of 1.0 mmol/L sodium formate calibrant solution in 1:1 water/ isopropanol for the first minute of each injection. The data was aligned on MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA) with minimum intensity cut-off of 3000 cts; minimum peak length of 6 spectra; m/z tolerance of 5.0 mDa; retention time tolerance of 10 s; maximum cluster overlap of 0.25; and maximum isotope pattern error of 0.1. Missing values were substituted by recursive extraction, *i.e*., features detected in at least 10% of injections were searched again on the raw data with a minimum peak length of 1 spectrum and without a minimum intensity cut-off. If not found, the remaining missing values were substituted by the minimum group intensity for features detected in at least 50% of within-group injections, or global minimum intensity for features detected in less than 50% of within-group injections (A, B, C or QC).

The lipid classes and subclasses of identified lipids followed the classification system proposed in 2005 and updated on 2009 by Fahy *et al.* under the leadership of the International Lipid Classification and Nomenclature Committee (ILCNC).[16,17] Lipid nomenclature followed the LipidMaps database; however, this work was completed before the recent update that took place between 2019 and 2020 (for further details, please refer to Chapter I - *1.3.1.5. Identification and nomenclature*, p. 25). Hence, lipid subclasses and abbreviations are slightly different from Chapters II to IV, as described in Table V-1.

The MS/MS positive identification was performed with the MS-Dial LipidBlast, Human Metabolome Database (HMDB) and MassBank of North America (MoNA) LC-MS/MS libraries in combination with MetaboScape 4.0.[73,78,79,81–83,109,151] Unfortunately, many features could not be matched to MS/MS libraries due to the low intensity of precursor and fragment ions. Furthermore, many features were not chosen for fragmentation because of high-intensity co-eluting molecules.

The manual investigation of MS/MS spectra was performed for some of the lipids selected during statistical analysis; however, not all lipid classes display characteristic fragments that can be used to manually confirm their identification. Nevertheless, the manual check of thousands of lipids would be unreasonable. Hence, only lipids considered important for the discussion were verified.

Unidentified features were searched in the LipidMaps database for accurate m/z match (putative identification). One of the biggest challenges in lipid identification by mass spectrometry is the differentiation of multiple isomers or isobars with similar structures. Each detected feature can usually be m/z-matched to multiple isomers or isobars that might belong to different lipid subclasses. For this work, a five-tier filtering and scoring approach was employed to restrict the number of isomeric and isobaric compounds and select the best identification to determine lipid subclasses for normalization, as described in Chapter II (*2.2.7. Putative identification of lipids*, p. 50). The scores for each ranking tier were summed and the top choice for each feature was selected by the smallest total score, *i.e.,* the isomer or isobar with elution within the expected retention time range for the identified subclass; the most likely adduct and isoform; the smallest m/z error; an even number of carbons in FA chains; and no extra functional groups except for the expected for the subclass. The top choice was used to determine the lipid subclass for normalization, but other isomeric or isobaric possibilities that passed the retention time and adduct filters were also kept, ordered according to their scores. The position of double bonds and the stereospecific configuration of glycerol derivatives were not determined.

Table V-1. List of lipid subclass abbreviations employed for this study. Lipid nomenclature followed the LipidMaps database (2018).

Subclass	Category	Abbreviation
Bis(monoacylglycero)phosphate or hemibismonoacylglycerophosphate (lysobisphosphatidic acid)	Glycerophospholipids	BMP
Carnitine or acylcarnitine	Fatty acyls	Car
Ceramide	Sphingolipids	Cer
Ceramide non-hydroxy fatty acid-sphingosine	Sphingolipids	Cer-NS
Cholesterol and derivatives	Sterol Lipids	Chol
Cholesteryl ester	Sterol Lipids	CholE
Cardiolipin	Glycerophospholipids	CL
Acyl Coenzyme A	Fatty acyls	CoA
Diacylglycerol	Glycerolipids	DG
Diacylglyceryltrimethylhomoserine	Glycerolipids	DGDMS
Fatty acid or fatty acid ester	Fatty acyls	FA
Diacylglycosylglycerol, monogalactosyldiacylgylcerols or digalactosyldiacylgylcerol	Glycerolipids	GDG
Sulfoquinovosylmonoacylglycerols	Glycerolipids	GMG
Hexosylceramide (gluco- and galactosylceramide)	Sphingolipids	HexCer
Lactosylceramide	Sphingolipids	LacCer
Lysophosphatidic acid	Glycerophospholipids	LPA
Lysophosphatidylcholine	Glycerophospholipids	LPC
Lysophosphatidylethanolamine	Glycerophospholipids	LPE
Lysophosphatidylglycerol	Glycerophospholipids	LPG
Lysophosphatidylinositol	Glycerophospholipids	LPI
Lysophosphatidylserine	Glycerophospholipids	LPS
Lysoglycosphingolipid	Glycerophospholipids	LSL
Monoacylglycerol	Glycerolipids	MG
N-acyl amine, ethanolamine or taurine	Fatty acyls	NAA, NAE, NAT
Phosphatidic acid	Glycerophospholipids	PA
Phosphatidylcholine	Glycerophospholipids	PC
Phosphatidylethanolamine	Glycerophospholipids	PE
Phosphatidylethanolamine-ceramide	Sphingolipids	PE-Cer
Phosphatidylglycerol	Glycerophospholipids	PG
Phosphatidylinositol or phosphatidylinositol-monophosphate	Glycerophospholipids	PI
Phosphatidylinositol-ceramide, Mannosyl-phosphatidylinositol-ceramide	Sphingolipids	PI-Cer
Phosphatidylserine	Glycerophospholipids	PS
Sphingomyelin	Sphingolipids	SM

Subclass	Category	Abbreviation
Sphingosine 1-phosphate	Sphingolipids	Sph
Sphingosine or sphingoid base	Sphingolipids	Sph
Sulfoquinovosyldiacylglycerols	Glycerolipids	SQDG or SulfDG
Sulfatide (3-O-sulfogalactosylceramide)	Sphingolipids	Sulf
Triacylglycerol	Glycerolipids	TG
Wax ester	Fatty acyls	WE

We employed a mixture of 14 deuterated lipid standards for internal normalization of the identified lipids (Table II-1, p. 45). The normalization procedure chosen for lipidomics in this work followed the Lipidomics Standards Initiative guidelines (https://lipidomics-standards-initiative.org). The LC-MS separation gradient was optimized to ensure that lipids belonging to the same class eluted in a relatively narrow retention time window, while different classes are separated to reduce ion suppression and lipid aggregation as much as possible (Appendix B - Figure B- 1). The class of each identified lipid was determined and matched to one of the internal standards for normalization according to subclass, structural similarity and retention time range. Hence, lipids that belong to the same class were normalized by a deuterated internal standard that shares structural similarities and elutes in a comparable matrix. The intensity of each identified lipid was divided by the intensity of the matched internal standard to obtain normalized intensity ratios. This widely used approach for normalization of lipidomics data allows the correction of ion suppression and small variations that may occur during sample preparation.[4,12,152,180] A detailed discussion on normalization strategies for lipidomics is available in Chapter I (*1.3.1.6. Ion suppression and normalization approaches*, p. 29). Normalized intensities, *i.e.*, the intensity of each lipid (peak height) divided by the intensity of the matched internal standard, were auto-scaled and filtered by a relative standard deviation (RSD) smaller than 30% for QC injections before

statistical analysis. Features that were not identified as lipids by either accurate mass or MS/MS were not employed for statistical analysis.

5.2.5 *Metabolomics: amine and phenol-containing metabolites*

After LLE of the first lung lobe, the aqueous phase was employed for DnsCl labeling of amine and phenol groups.[1,181] Synthesis of ^{13}C-dansyl chloride (DnsCl) was performed as previously published.[182–184] The chemical isotope labeling procedure employed for this work has been previously described in several publications.[1,181,184–186] The accuracy and precision of the techniques for relative quantification of metabolites were also previously evaluated.[181,184,185,187]

Briefly, aliquots of 25.0 µL of the aqueous layer obtained after LLE of the first lung lobe were mixed with 12.5 µL of acetonitrile, 12.5 µL of 250 mM $NaHCO_3$/Na_2CO_3 aqueous buffer (pH 9.4) and 25.0 µL of 18 mg/mL DnsCl in acetonitrile. After 45 min incubation at 40°C, samples were cooled down to 4°C and vortexed with 5.0 µL of 250 mM NaOH in water, followed by 10 min incubation at 40°C to quench the labeling reaction. Finally, the pH of the labeled samples was adjusted with 25.0 µL of 425 mM formic acid in 1:1 acetonitrile/ water. Three aliquots of each sample were labeled with ^{12}C-DnsCl, whereas a pooled mixture composed of 25% of the total volume of the aqueous layer for each sample was labeled with ^{13}C-DnsCl. The total dansyl-labeled metabolite concentration was determined by UHPLC-UV (Waters Acquity UPLC with a photodiode array detector, Waters Corporation, Milford, MA, USA) with an Agilent Kinetex C18 column (50 X 2.1 mm, 1.7 µm, Agilent Technologies, Santa Clara, CA, USA), using a 6 min gradient elution (0 min – 0% MPB; 1 min – 95% MPB; 2.5 min – 95% MPB; 3.0 min – 0% MPB; 6 min – 0% MPB), with MPA: 0.1% formic acid in 95:5 water/ acetonitrile, MPB: 0.1% formic acid in 95:5 acetonitrile/ water; 4.0 µL injection; 450 µL/min; 30°C; and UV absorbance detection

at 338 nm.[181] A calibration curve was previously obtained with a mixture of 17 amino acid standards (Amino Acid Standard, Sigma Aldrich, St. Louis, MO, USA), labeled with ^{12}C-DnsCl using the same procedure. The absorbance at the chosen wavelength is mostly related to the dansyl group and the gradient was developed to fully co-elute all labeled metabolites, allowing the accurate estimation of total metabolite concentrations by the calibration curve. The individually ^{12}C-DnsCl labeled samples were mixed with the ^{13}C-DnsCl labeled pool in a 1:1 ratio of total metabolite concentration for normalization. The performance of the normalization procedure using total metabolite concentrations was previously assessed elsewhere.[181,185,188] Labeled samples were kept at 4°C in polypropylene inserts placed in amber vials capped with PTFE-lined septa for a maximum of 5 days before the injection.

LC-MS analysis for DnsCl-based metabolomics was performed employing: MPA - 0.1% formic acid in 95:5 water/ acetonitrile; MPB - 0.1% formic acid in 95:5 acetonitrile/ water; gradient elution (0 min – 80% MPA, 3.5 min – 65% MPA, 18 min – 35% MPA, 24 min – 1% MPA); 180 µL/min; 30 °C; 10 min re-equilibrium (0 min – 1% MPA, 2 min – 1% MPA, 2.5 min – 80% MPA, 10 min – 80% MPA); and ESI-QToF detection (capillary voltage of 4500 V; endplate offset of 500 V, nebulizer gas pressure of 1.0 bar, dry gas flow rate of 8.0 L/min; dry temperature of 230°C; spectra acquisition rate of 1 Hz, m/z range of 220 to 1000 Da). The obtained chromatograms were re-calibrated based on the direct infusion of 1.0 mmol/L sodium formate calibrant solution in 1:1 isopropanol/ water (v/v) for the first minute of each injection using Compass DataAnalysis 4.4 (Bruker Daltonics, Billerica, MA, USA). Each chromatogram was exported as a comma-separated value (CSV) file before alignment using our in-house developed script IsoMS Shiny 0.3.1 (www.mycompoundid.org).[178,186] Peak pairs, *i.e.*, the ^{12}C-labeled metabolites from individual samples and the co-eluting ^{13}C-labeled metabolites from the pool, only differ in two carbon atoms

(two ^{12}C or two ^{13}C atoms in the dansyl group); hence, they co-eluted perfectly. The intensity of each ^{12}C-labeled metabolite from the individual samples was divided by the intensity of the corresponding ^{13}C-labeled metabolite from the pool to obtain normalized intensity ratios. Since all samples were normalized to the same metabolite concentration with the pooled mixture, the ^{13}C-labeled pool acted as an "internal standard" mixture for each detected metabolite, ensuring normalization of ion suppression and other small differences that may occur during sample analysis.

Definitive identification of metabolites was performed with the MyCompoundID retention time library (www.mycompoundid.org), based on m/z tolerance of 10 ppm and retention tolerance of 30 s.[186,189,190] The non-identified $^{12}C/^{13}C$ peak pairs were searched on the MyCompoundID MS library for putative identification by mass-match (neutral mass tolerance of 5.0 mDa). The MyCompoundID library is composed by human endogenous metabolites and their predicted metabolic products after one or two common metabolic reactions. For this work, we considered identifications of endogenous metabolites without any reactions (endogenous metabolite standards in the MyCompoundID library identified without any modifications) and after one predicted metabolic reaction, *i.e.*, the product of a metabolic reaction involving one of the endogenous metabolites in the MyCompoundID library, *e.g.* hydrolysis or oxidation. [191,192]

5.2.6 *Metabolomics: carboxylic acid-containing metabolites*

The first lung lobe of each animal was used for lipidomics and dansyl chloride labeling of amine and phenol groups. The second lung lobe (80 to 150 mg) was employed for p-dimethylaminophenacyl (DmPA) bromide labeling of carboxylic acid groups. Syntheses of ^{12}C- and ^{13}C-DmPA were performed as previously published.[182–184] Samples were homogenized in a

mixture of ice-cold water (0.80 µL / mg of tissue), HCl 6.0 mol/L (0.20 µL / mg of tissue) and saturated NaCl (0.20 µL / mg of tissue) for three 30 s intervals, followed by 30 s intervals in an ice bath. The homogenate was immediately extracted with ethyl acetate (6.0 µL / mg of tissue), vortexed for 30 s and centrifuged for 10 min at 4°C (10,000 rpm). The organic layer, *i.e.,* 85.0% of the initial ethyl acetate volume, was employed for the labeling of carboxylic acid groups with DmPA bromide, following a previously published method with minor modifications for tissue samples.[183] Briefly, the organic layer obtained after homogenization and LLE with ethyl acetate was vortexed with 180 mg/mL triethylamine in acetonitrile (0.80 µL/mg of tissue) and evaporated to dryness on a SpeedVac. The residue was resuspended in 20 mg/mL triethylamine in acetonitrile (1.20 µL/mg of tissue). Three aliquots of each sample were individually and randomly labeled with ^{12}C-DmPA, while a pooled mixture was prepared with 45.0 µL of each sample and labeled with ^{13}C-DmPA. Each sample aliquot was vortexed with 15.0 µL of 20.0 mg/mL ^{12}C- or ^{13}C-DmPA bromide in acetonitrile and incubated at 85°C for 55 min. After cooling down to room temperature, the labeled samples were diluted with 30.0 µL of 1:1 MPA/MPB (v/v). Each individually labeled sample was mixed with the ^{13}C-DmPA-labeled pool in a 1:1 ratio (v/v) before LC-MS analysis. Labeled samples were kept at 4°C in polypropylene inserts placed in amber vials capped with PTFE-lined septa for a maximum of 5 days before injections.

The labeled samples were analyzed by RPLC-ESI-QTOF-MS (positive ionization). LC-MS analyses were performed with MPA: 0.1% formic acid in 95% water/ acetonitrile; MPB: 0.1% formic acid in 95:5 acetonitrile/ water; 6.0 µL injection; gradient elution (0 min – 70% MPA; 4 min – 70% MPA; 18 min – 10% MPA; 30 min – 1% MPA; 40 min – 1% MPA); 10 min re-equilibrium (0 min – 1% MPA; 2 min – 1% MPA, 2.5 min – 70% MPA, 10 min – 70% MPA); 180 µL/min; 30°C; and ESI-QToF detection (capillary voltage of 4500 V; endplate offset of 500

V, nebulizer gas pressure of 1.0 bar, dry gas flow rate of 8.0 L/min; dry temperature of 230°C; spectra acquisition rate of 1 Hz, and m/z range of 150 to 1000 Da). The obtained chromatograms were re-calibrated based on the direct infusion of 1.0 mmol/L sodium formate calibrant solution in 1:1 isopropanol/ water (v/v) for the first minute of each injection using Compass DataAnalysis 4.4 (Bruker Daltonics, Billerica, MA, USA). Each chromatogram was exported as a comma-separated value (CSV) file before alignment using our in-house developed script IsoMS Shiny 0.3.1 (www.mycompoundid.org).[186] Peak pairs were selected as a ^{12}C- DmPA labeled metabolite from the individual sample and the corresponding ^{13}C-labeled metabolite from the pooled mixture. Since the ^{12}C- and ^{13}C-labeled metabolites only differ in two isotope-labeled carbon atoms, they co-elute perfectly. The normalization procedure was similar to the description for dansyl chloride labeling, *i.e.*, normalized intensities were calculated by dividing the intensity of each ^{12}C-labeled peak from the individual sample by the corresponding ^{13}C-labeled peak from the pool. The ^{13}C-labeled metabolites were used as internal standards for each sample, ensuring normalization of ion suppression effect and other small differences that may occur during sample injection. Definitive identification of metabolites was performed with our recently extended MyCompoundID retention time library (www.mycompoundid.org), as described for metabolomics of amine and phenol-containing metabolites.

5.2.7 *Statistics*

Statistical analysis and pathway analysis were performed using MetaboAnalyst 4.0 (https://www.metaboanalyst.ca).[153] Statistical tests were based on five biological replicates for each group (N=5). Metabolomics experiments included labeling triplicates for each labeling group, while lipidomics statistical models had LC-MS injection duplicates for each ionization polarity.

The PLS-DA models were evaluated by cross-validation (R^2 and Q^2) and permutation tests (p for 1000 permutations). Student's t-tests, Volcano plots and ANOVA were performed with non-parametric tests with unequal variances and p value adjusted for false-discovery rate (p). Metabolites and lipids were considered significantly altered for $p < 0.05$ and fold-change (FC) ≤ 1.5 or ≥ 0.67. Pathway search was based on the Human Metabolome Database (HMDB, http://www.hmdb.ca) ID descriptors with the Kyoto Encyclopedia of Genes and Genomes database (KEGG, https://www.genome.jp/kegg/pathway.html) and Small Molecule Pathway Database (SMPDB, http://smpdb.ca).[79,81,109,151,153,193–195]

5.3 Results and Discussion

5.3.1. Lipidomics

LC-MS lipidomic analysis of healthy control mice (group A), ΔF/TriAdj-immunized, RSV- challenged animals (group B) and non-immunized, RSV-challenged mice (group C) (Figure V-1, N = 5) resulted in 8377 detected features (92.9% with relative standard deviation - RSD <30% for QC injections), divided into 7883 features for positive ionization, 494 for negative ionization and 140 features commonly detected for both polarities (m/z tolerance of 10.0 ppm and retention time tolerance of 10 s). We putatively identified 354 lipids by MS/MS and 3723 lipids by mass-match (m/z error ≤5.0 mDa, Supp. Table 14). Eighteen features were also positively identified as internal standards (Table II-1), *i.e.,* twelve deuterated internal standards were detected for positive ionization with m/z error ≤2.5 ppm or 1.8 mDa, while six standards with m/z error ≤3.1 ppm or 2.5 mDa were obtained for negative ionization. Representative chromatograms obtained for a quality control injection (QC, a pool of extracts from all samples), blank extracts (water instead of sample and methanol instead of internal standard mix) and each studied group are shown in

Appendix B - Figure B- 1, while the subclass distribution of identified lipids is provided in Figure V-2.

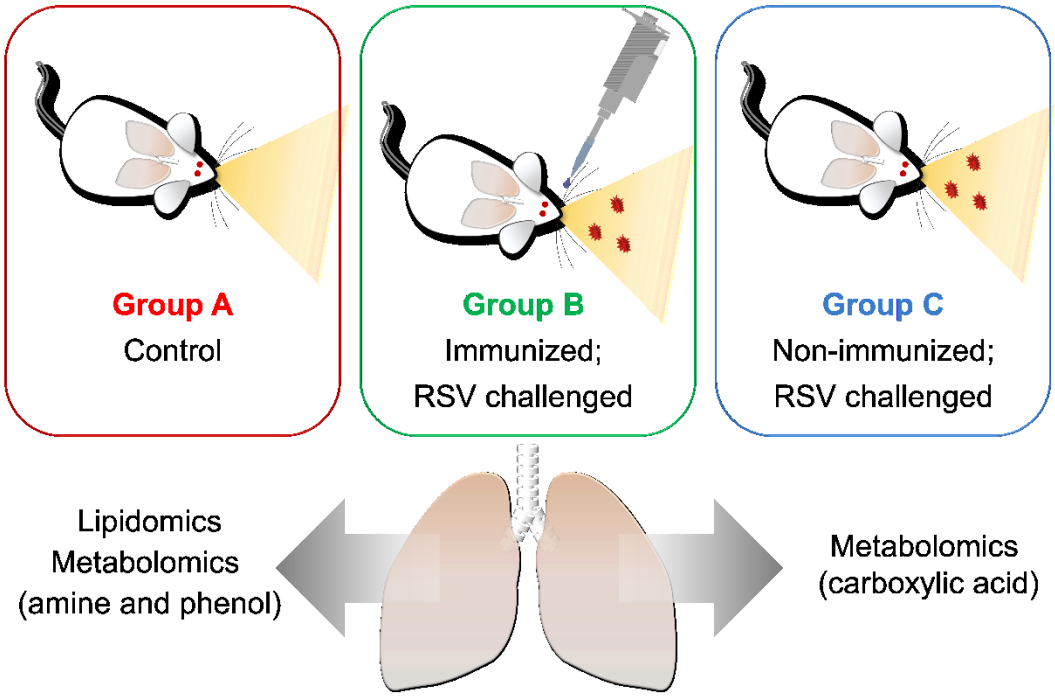

Figure V-1. Experimental design employed for lipidomics and metabolomics of rat lung tissue. Group A: five healthy control mice (not immunized or RSV-challenged, red); Group B: five immunized, RSV-challenged animals (green); Group C: five non-immunized, RSV-challenged mice (blue).

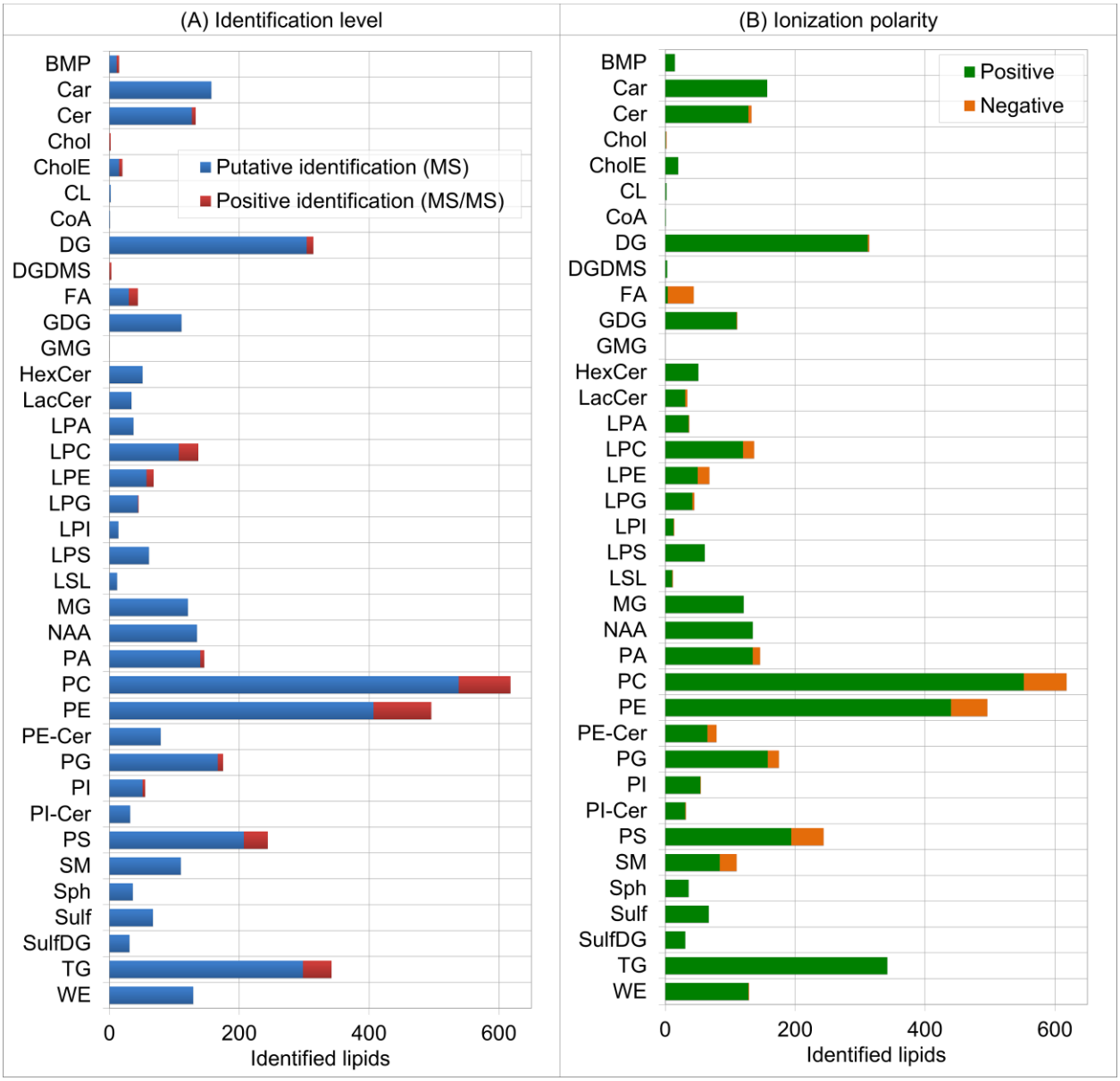

Figure V-2. Subclass distribution for putatively (m/z match, MS) and positively (MS/MS) identified lipids. (A) Number of identified lipids divided into identification level: positive (MS/MS, red) or putative (mass-match, blue). (B) Number of identified lipids divided into positive (green) and negative ionization (orange). Abbreviations to lipid subclasses are provided in Table V-1.

By applying non-parametric Analysis of Variance (ANOVA) to the three groups, we found 2316 significantly altered lipids with p value adjusted for false-discovery rate (p) <0.05 (Appendix B - Figure B- 2, Supp. Table 15). The most affected lipid subclasses included phosphatidylcholines (PC), phosphatidylethanolamines (PE), diacylglycerols (DG), phosphatidylserines (PS), phosphatidylglycerols (PG) and triacylglycerols (TG, Figure V-3A). A comprehensive list of abbreviations employed for lipid identifications and subclasses is provided in Table V-1. Boxplots for 10 of the most significantly altered lipids are displayed in Figure B- 3.

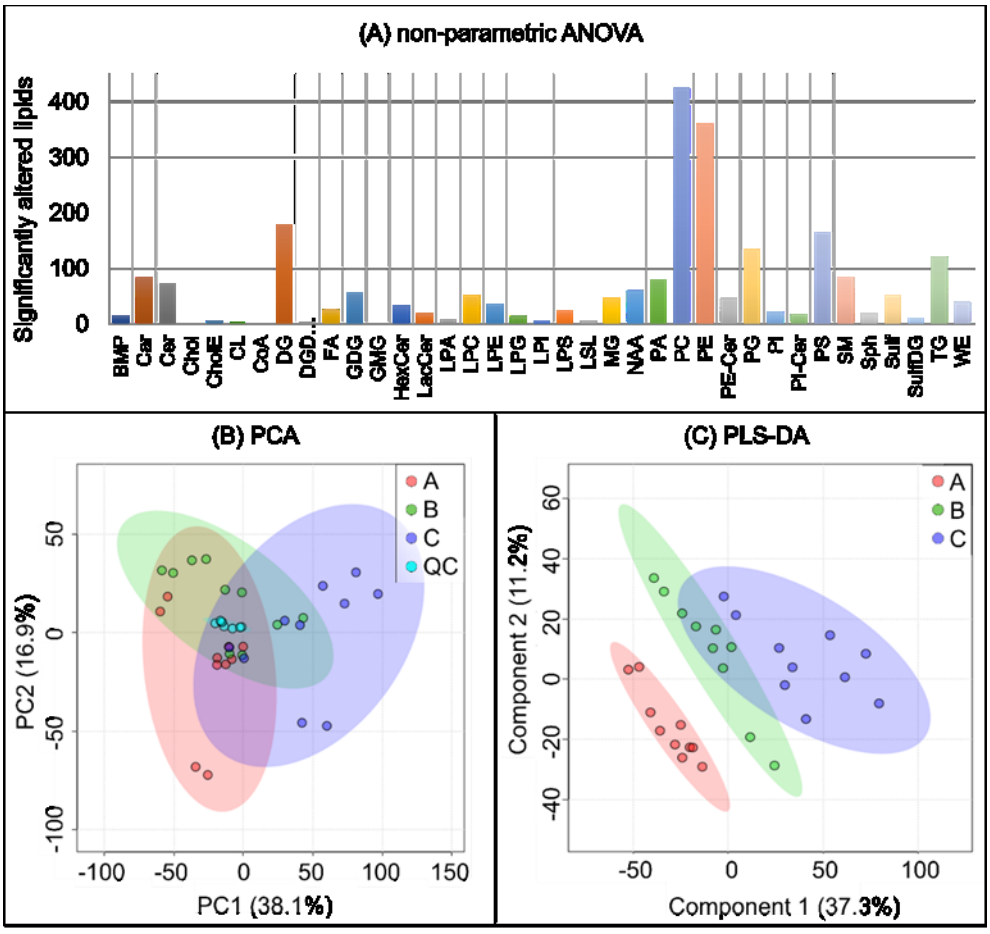

Figure V-3. Statistical analysis for lipidomics of lung tissue for healthy controls (group A, red); immunized, RSV-challenged animals (group B, green); and non-immunized, RSV-challenged

mice (group C, blue). (A) Subclass distribution of significantly altered lipids with $p < 0.05$ by non-parametric ANOVA (Kruskal-Wallis test). (B) PCA score plot with quality control (QC, cyan) injections. (C) PLS-DA score plot built with five components, resulting in R^2 of 0.9980, Q^2 of 0.9663 and p=0.002 for 1000 permutations. Abbreviations to lipid subclasses are provided in Table V-1.

The PCA score plot in Figure V-3B displays all QC injections in a tightly clustered group, indicating good technical reproducibility of the employed methods. The PCA score plot without QCs is shown in Figure B- 4 with circled duplicate injections for each sample to emphasize their proximity. The healthy controls (A, red) and non-immunized, RSV-challenged (C, blue) mice are separated on the first principal component (PC1), whereas the ΔF/TriAdj-immunized group (B, green) is overlayered with the healthy controls (A, red), with a slight separation only in the second principal component. We didn't expect a full separation between the three groups in the PCA score plot due to the high complexity of the dataset, but the observed trend demonstrates the efficacy of the formulation, as well as the suitability of the employed methods. The groups were fully separated by PLS-DA (Figure V-3C), indicating that the ΔF/TriAdj formulation and RSV challenge profoundly affect the lipid composition of lung tissue. All samples were correctly classified by a Random Forest model with 20 predictors and 1000 trees, with out-of-bag (OOB) error of zero.

The healthy control mice were compared to the non-immunized, RSV-challenged mice (A/C) using a Volcano plot (fold change, FC, *versus* p for non-parametric Student's t-test corrected for false-discovery rate, p). The Volcano plot for healthy *versus* non-immunized animals, along with PCA and PLS-DA score plots, are shown in Appendix B - Figure B- 5. Compounds were

considered significantly altered for FC ≥1.5 or ≤0.67 and p <0.05 A total of 1423 lipids were significantly altered (Figure V-4A; Supp. Table 15), demonstrating the dramatic change in the lung lipidic composition when non-immunized animals were compared to healthy controls. RSV challenge without prior immunization resulted in increased normalized intensities (*i.e.*, the intensity of each identified lipid divided by the intensity of an internal standard belonging to the same or the most similar lipid subclass) for 94.1% of the significantly altered lipids in lung tissue. We have previously reported a higher influx of immune cells in the lungs of non-immunized mice compared to immunized or healthy control animals, which corroborates our current observations. [1] The phosphatidyl glycerolipid and acyl glycerol metabolisms were profoundly affected by the virus.

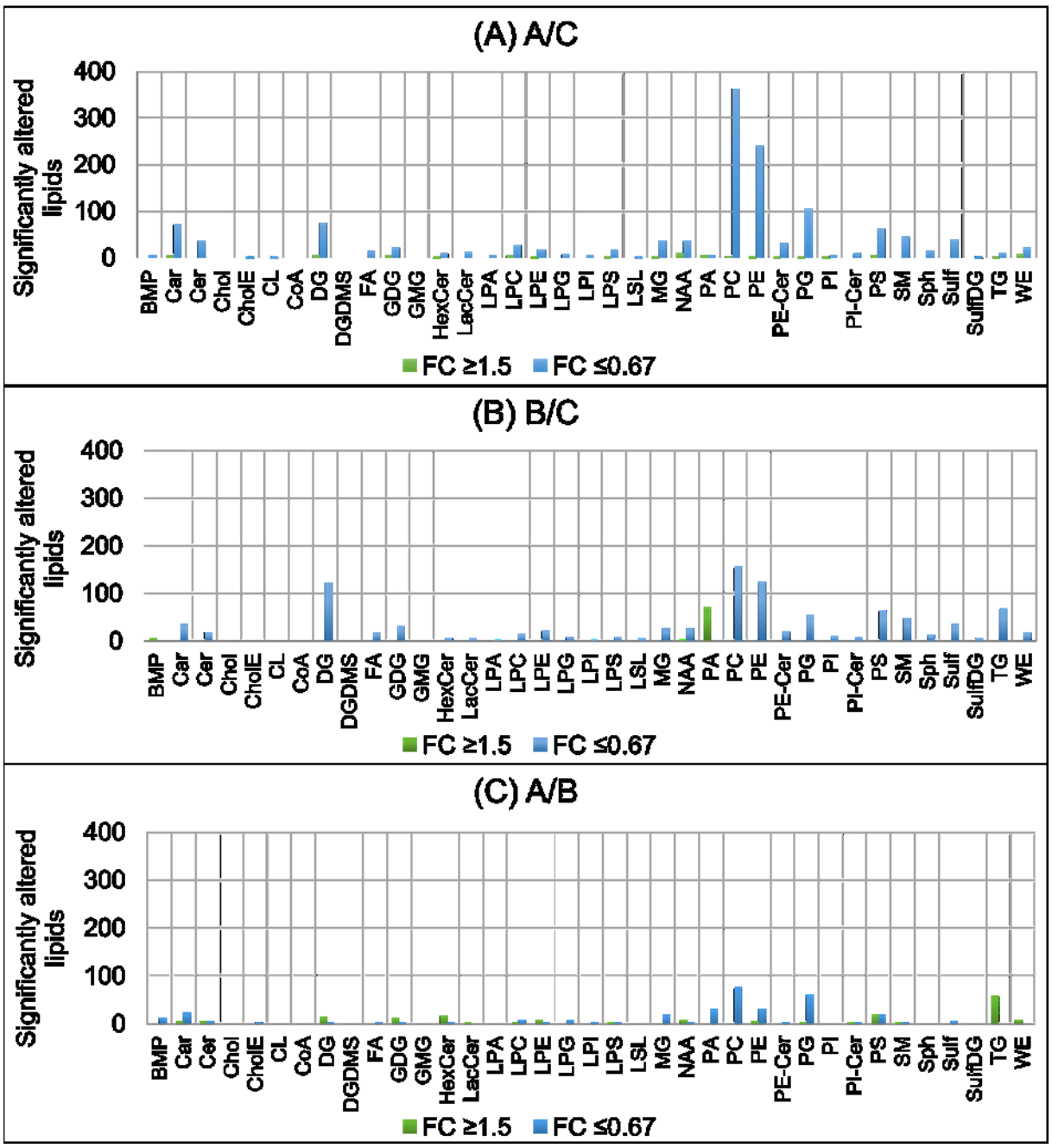

Figure V-4. Classes of significantly altered lipids ($p < 0.05$ and FC ≥ 1.5 or ≤ 0.67) for the binary comparisons between (A) healthy control *versus* non-immunized, RSV-challenged mice (A/C); (B) immunized *versus* non-immunized, RSV-challenged mice (B/C); and (C) healthy control *versus* immunized, RSV-challenged mice (A/B). Abbreviations of lipid classes are described in Table V-1 (p. 154).

The non-immunized, RSV-challenged mice were also compared to immunized animals (B/C) by a Volcano plot, resulting in 1053 significantly altered lipids (Figure V-4B; Supp. Table 15). The Volcano, PCA and PLS-DA plots are shown in Figure B- 6. When compared to the immunized mice (group B), 92.5% of the significantly altered lipids displayed higher normalized intensities for the non-immunized mice (FC B/C ≤0.67*,* 974 lipids). Overall, we have observed that RSV challenge without prior immunization led to significantly higher levels of lipids when compared to immunized mice or healthy controls. Immunization with the ΔF/TriAdj formulation seems to mitigate the striking lipidic changes caused by RSV challenge. A total of 647 lipids were affected for the non-immunized animals compared to both the healthy controls or the immunized group (A/C and B/C), including 637 with significantly smaller normalized intensities for both (Figure V-5A and C). The Venn diagram on Figure V-5A emphasizes the effect of immunization: the comparison between non-immunized animals and healthy controls (A/C) had 574 lipids uniquely altered, *i.e.,* related to RSV challenge, but modulated by the ΔF/TriAdj formulation.

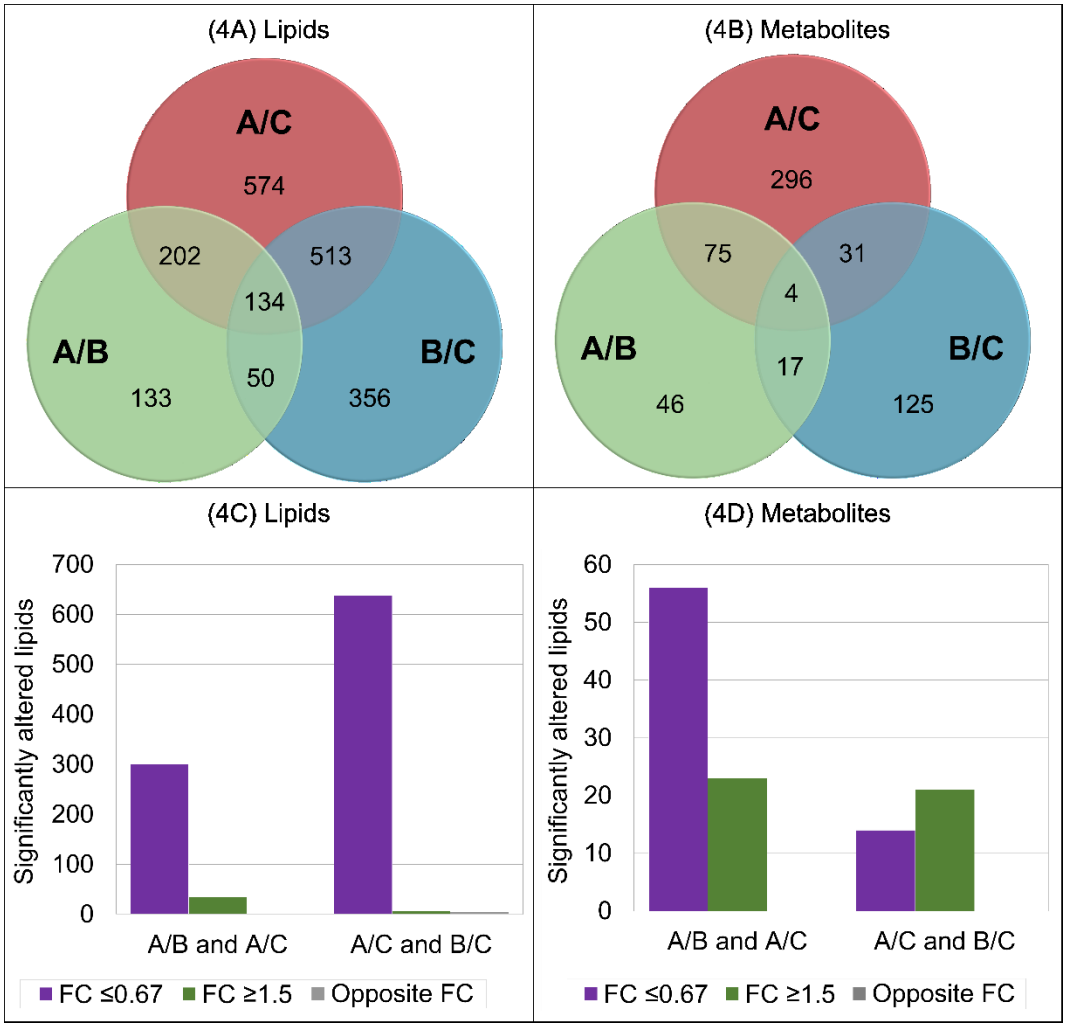

Figure V-5. Venn diagrams and bar plots for binary comparisons of healthy control *versus* non-immunized, RSV-challenged mice (A/C); immunized *versus* non-immunized, RSV-challenged mice (B/C); and healthy control *versus* immunized, RSV-challenged mice (A/B) for (A, C) significantly altered lipids (FC ≥1.5 or ≤0.67 and $p < 0.05$); and (B, D) significantly altered amine, phenol and carboxylic acid-containing metabolites. The Venn diagrams (A and B) include all significantly altered compounds, whereas the bar plots (C and D) separates them according to fold change (FC) ≤0.67 (purple) or FC ≥1.5 (green). No molecules displayed opposite FC for the A/B and A/C comparisons, *i.e.*, all the significantly altered lipids and metabolites had either FC ≤0.67 or FC ≥1.5 for both the immunized and non-immunized groups.

The healthy control mice were further compared to the immunized, RSV-challenged animals (A/B), and 519 lipids were significantly altered (Figure V-4C; Supp. Table 15). The Volcano, PCA and PLS-DA plots are shown in Figure B- 7. The number of altered lipids was much smaller than what was found for the comparison of healthy controls with non-immunized, RSV-challenged mice, confirming that the ΔF/TriAdj formulation modulates the effect of RSV challenge (Figure V-5A). Amongst the 519 significantly altered lipids, 65.5% displayed higher normalized intensities for the immunized group (Figure V-4C). However, 64.7% of the lipids altered for the immunized animals (group B) were also significantly affected for the non-immunized group (C), suggesting a small effect of the ΔF/TriAdj formulation upon these lipids (Figure V-5C). Still, the immunized animals had the least amount of significant changes when compared to healthy controls (A/B, Figure V-5A), emphasizing the modulation of the ΔF/TriAdj formulation.

A total of 183 lipids displayed significantly higher normalized intensities for the immunized group compared to healthy controls (A/B), but not for the non-immunized mice (A/C); hence related only to immunization. The positively identified bis(monoacylglycero)phosphate (BMP) 18:1_18:1, a common minor constituent of animal tissues, is an example of the significantly increased lipids for the immunized group but not for the non-immunized mice. BMPs stimulate the activity of lysosomal lipid-degrading enzymes, including acid sphingomyelinase, acid ceramidase and acid lipases. Accumulation of BMP has been previously linked to lipid storage disorders and dysregulation of cholesterol homeostasis. [168] Such alterations possibly introduced by immunization must be taken into consideration for future vaccine developments.

5.3.2. Metabolomics

The chemical isotope labeling of amine and phenol groups by dansyl chloride (DnsCl) resulted in the detection of 2600 peak pairs. The PCA score plot obtained for the amine and phenol-containing metabolites can be observed in Figure V-6. The MyCompoundID retention time library resulted in the identification of 120 unique metabolites by mass and retention time, while 2021 metabolites were identified by mass-match (m/z error ≤5.0 mDa, Supp. Table 16).[190–192]

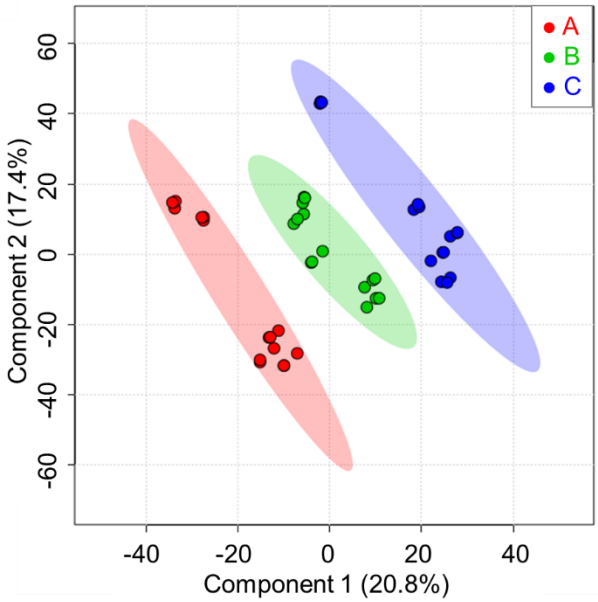

Figure V-6. PCA score plot obtained for dansyl chloride labeling of amine and phenol-containing metabolites from lung tissue obtained from for healthy controls (group A, red); immunized, RSV-challenged animals (group B, green); and non-immunized, RSV-challenged mice (group C, blue).

The chemical isotope labeling of carboxylic acid groups with p-(dimethylamino)phenacyl (DmPA) bromide resulted in the detection of 3781 peak pairs, amongst which 47 were uniquely identified and 3061 were identified by mass-match (m/z error ≤5.0 mDa, Supp. Table 16). The

PCA score plot in Appendix B - Figure B- 8 displays all QC injections in a tightly clustered group, demonstrating good reproducibility for the employed method.

The results for DnsCl labeling were combined with the DmPA data to generate one unique, comprehensive metabolite list containing 6381 peak pairs. The PCA score plot obtained after combining the datasets is shown in Figure V-7A, with QC injections and labeling triplicates tightly clustered together (Appendix B - Figure B- 9). The complexity of the dataset containing 6381 peak pairs prevented a complete separation between the three studied groups by the simple PCA model; however, the healthy controls and immunized animals are fully overlayered, whereas the non-immunized group shows a slight separation trend. Nevertheless, a PLS-DA model displayed excellent separation (Figure V-7B), while a Random Forest model with 7 predictors and 500 trees resulted in outstanding classification performance, with an OOB error of zero. The non-parametric ANOVA analysis had 928 significant peak pairs for the three studied groups ($p < 0.05$) (Appendix B - Figure B- 10; Supp. Table 17).

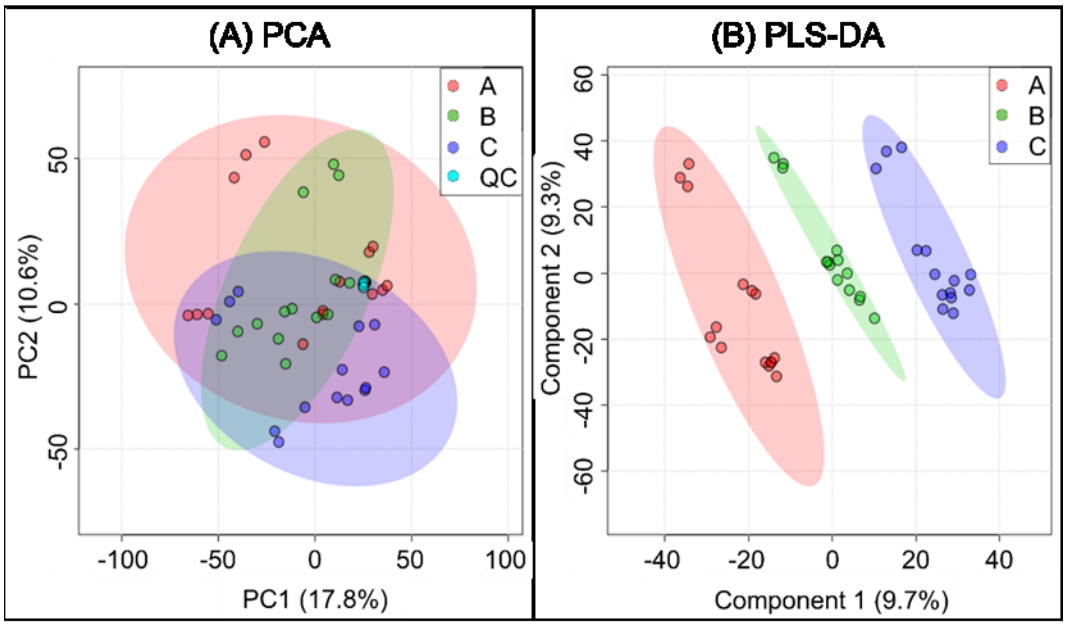

Figure V-7. PCA (A) and PLS-DA (B) score plots obtained for the combination of the DnsCl with the DmPA results for lung tissue samples from healthy control (A, red); immunized, RSV-challenged (B, green); and non-immunized, RSV-challenged (C, blue) mice. The PLS-DA model with 4 components resulted in R^2 of 0.9945, Q^2 of 0.9512 and p=0.005 (1000 permutations).

Univariate binary comparisons were also evaluated through Volcano plots. First, the healthy control mice were compared to the non-immunized, RSV-challenged mice (A/C). A total of 406 peak pairs were significantly altered (Figure V-8A; Supp. Table 17). The PCA and PLS-DA score plots are presented in Appendix B - Figure B- 11. Interestingly, we found that lipids and metabolites were not affected to the same extent by RSV challenge. While 94.1% of the significantly altered lipids for the non-immunized animals had higher normalized intensities when compared to healthy controls (Figure V-4A), only 48.8% of peak pairs for significantly altered metabolites displayed the same behavior (FC A/C ≤0.67).

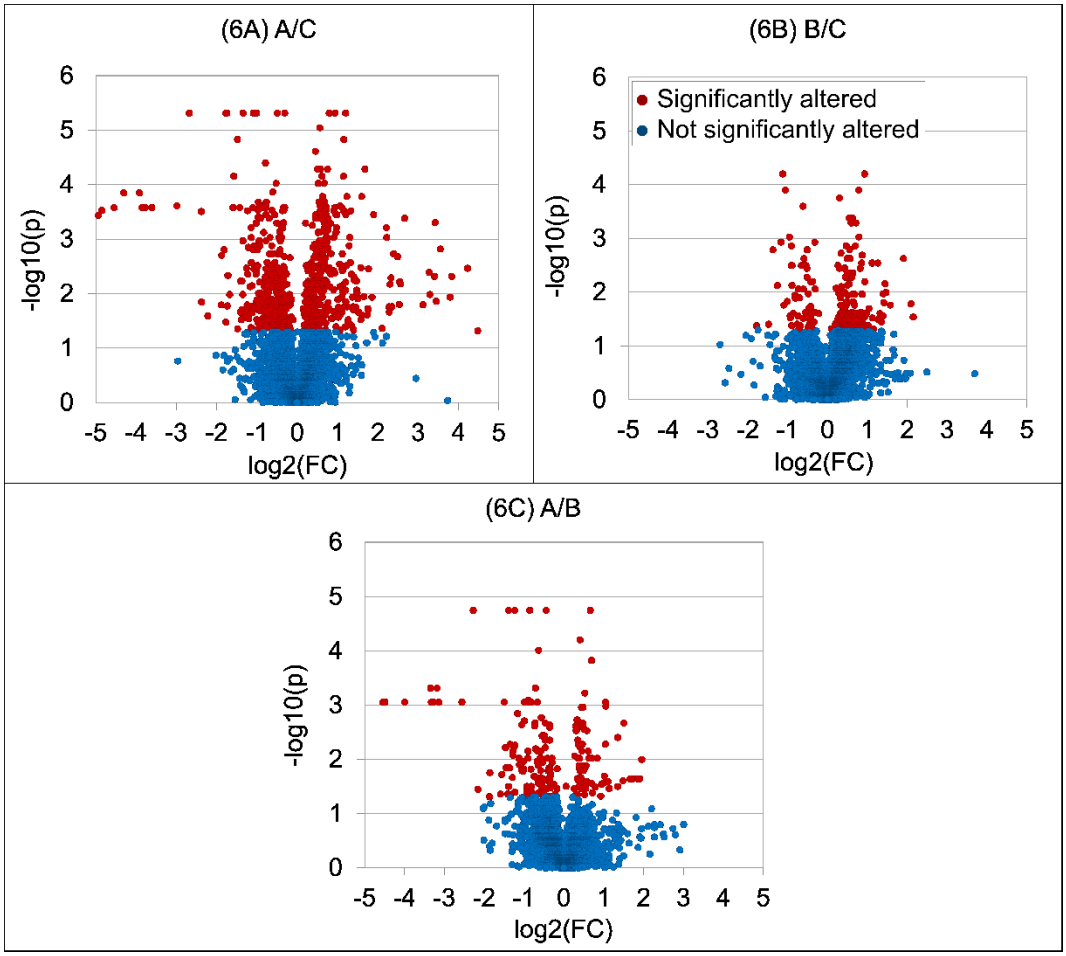

Figure V-8. Volcano plots for metabolites containing amine, phenol and carboxylic acid groups for the binary comparisons between (A) healthy control *versus* non-immunized, RSV-challenged mice (A/C); (B) immunized *versus* non-immunized, RSV-challenged mice (B/C); and (C) healthy control *versus* immunized, RSV-challenged mice (A/B). Significant peak pairs are indicated in red (FC ≤0.67 or ≥1.5 and p <0.05), while non-significant peak pairs are in blue.

The non-immunized, RSV-challenged mice were also compared to immunized, RSV-challenged animals (B/C) by Volcano plot, resulting in 177 significantly altered peak pairs (Figure V-8B; Supp. Table 17). The PCA and PLS-DA score plots are presented in Figure B- 12. While most significantly altered lipids had higher normalized intensities for the non-immunized group

(FC B/C ≤0.67, Figure V-4B), 81.9% of the significant peak pairs for metabolomics displayed higher intensities for the immunized animals (FC B/C ≥1.5). A total of 35 peak pairs were similarly affected for the non-immunized animals compared to both the healthy controls and immunized mice (A/C and B/C, Figure V-5B and D).

The healthy controls were further compared to the immunized, RSV-challenged mice (A/B), and 142 peak pairs were significantly altered (Figure V-8C, Figure B- 13, Supp. Table 17). Most of them had higher normalized intensities for the immunized group. 23 peak pairs resulted in FC ≥1.5 for both the immunized and non-immunized mice compared to healthy controls (A/B and A/C), while 56 had FC ≤0.67 (Figure V-5B and D); hence, these 79 metabolites were not fully modulated by immunization. No peak pairs displayed opposite fold changes for the A/B and A/C comparisons.

Analogous to lipidomics, the number of altered peak pairs for the immunized animals compared to healthy controls was smaller than what was found for the non-immunized group, once again emphasizing the reduced metabolic alterations of RSV challenge after immunization (Appendix B - Figure B- 14). However, 63 peak pairs were significantly altered for the immunized animals but not for the non-immunized group, indicating immunization-related alterations that cannot be explained by RSV challenge. For example, the positively identified 3-hydroximethylglutaric acid (FC A/B of 0.41), an intermediate in the leucine degradation pathway, was significantly increased for the immunized animals, but not for non-immunized mice compared to healthy controls. This observation must be carefully noticed for further vaccine developments, as high levels of organic acids in tissue can lead to metabolic acidosis.

5.3.3. Reduction of metabolic changes caused by RSV after vaccination

The untargeted lipidomics approach resulted in more significantly altered molecules than the chemical isotope labeling of metabolites, *i.e.,* 60.2% of the lipids employed for statistics were significant in the non-parametric ANOVA test, whereas only 18.9% of the metabolites showed significant changes. This observation may be correlated with the higher influx of immune cells in the lungs of non-immunized animals when compared to immunized mice and healthy controls, as described by Sarkar *et al*.[1] The control *versus* non-immunized, RSV-challenged mice (A/C) comparison led to the highest number of significantly altered metabolites and lipids, while immunized, RSV-challenged mice (A/B) had the lowest number, showcasing the ΔF/TriAdj modulation (Appendix B - Figure B- 14A). The observed changes for the comparisons of immunized and non-immunized mice with healthy controls are caused by a combination of immune response and viral replication. We have previously shown that there is increased and sustained inflammation in the lungs of the non-immunized, challenged group and controlled inflammation in the immunized group.[1] Our previous work showed that, at day 7 post-RSV-challenge (when lungs were collected for this study), the numbers of inflammatory cells were higher in the non-immunized group than in the immunized mice. Pulmonary homeostasis seemed to be restored for the immunized group seven days after RSV challenge. [1] For this work, we have observed increased levels of lipids for the non-immunized group compared to healthy controls (A/C), but the same was not valid for the immunized group (A/B), corroborating our previous findings.

336 lipids and 79 metabolites were altered for both the immunized and non-immunized mice compared to healthy controls (A/B and A/C). 35 lipids and 23 metabolites showed lower normalized intensities after RSV challenge (FC ≥1.5, $p < 0.05$ for A/B and A/C), while 301 lipids

and 56 metabolites had higher normalized intensities, regardless of vaccination (Figure B- 14). Such compounds were not regulated by the ΔF/TriAdj formulation, including the positively identified kynurenine, itaconic acid, mesaconic acid and several dipeptides, bis(monoacylglycero)phosphates (BMP), phosphatidylcholines (PC), phosphatidylglycerols (PG) and phosphatidylserines (PS). Itaconic and mesaconic acid are small, branched fatty acids involved in lipid metabolism, which can be directly related to the high number of significantly altered lipids found in this study. Furthermore, dysregulation of the kynurenine pathway may be triggered by inflammatory cytokines as an important part of the immune response. The result is a systematic upregulation when the immune system is activated, leading to the depletion of tryptophan. Although kynurenine displayed higher intensity ratios for both the immunized and non-immunized animals, other metabolites involved in its pathway were regulated by immunization, *e.g.* tryptophan and 5-hydroxyindoleacetic acid. In fact, 1087 lipids and 327 metabolites that were significantly altered for the non-immunized group were regulated in the immunized mice, including positively identified fatty acids, ceramides, acylglycerols, PCs, PEs, PSs, tryptophan, 5-hydroxyindoleacetic acid, serotonin, cholic acid and hypoxanthine. The comparison between immunized and non-immunized, RSV-challenged animals (B/C) resulted in 974 lipids and 32 metabolites with significantly higher normalized intensities for the non-immunized group (FC B/C $\leq$0.67, $p < 0.05$), while 79 lipids and 145 metabolites showed significantly lower normalized intensities (FC B/C $\geq$1.5, $p < 0.05$) (Figure B- 14).

Several compounds related to immune response and lung function were found to be significantly altered for the non-immunized, RSV-challenged mice (A/C), but not for the immunized animals compared to healthy controls (A/B, Appendix B - Figure B- 14). Dipalmitoyl-phosphatidylcholine, also known as PC(16:0/16:0) or DPPC, was positively identified by MS/MS

(Figure V-9A). DPPC had significantly higher normalized intensities for the non-immunized group compared to control animals (FC(A/C)=0.65, p=0.001), but not for the immunized group (FC(A/B)=0.90, p=0.7, Figure V-9A). It was also significant for the comparison between immunized and non-immunized, RSV-challenged mice, emphasizing the ΔF/TriAdj modulation (FC(B/C)=0.72, p=0.02). DPPC is the major constituent of the pulmonary surfactant layer, the first point of interaction between the virus and the host innate immune system.[15,196,197] Even though the mechanisms of the immune response in pulmonary surfactant are relatively well-characterized, secretion and regulation of the lipid fraction are still not well known.[198] Surfactant degradation has been previously linked to higher susceptibility to lung inflammation and infection, while its synthesis and distribution is supposedly impaired during RSV infection due to damage to alveolar type II cells.[197–199] Interestingly, DPPC normalized intensities were increased in lung tissue for the non-immunized, RSV-challenged group when compared to both the immunized animals and healthy controls. Palmitic acid, a substrate for the synthesis of DPPC, was also positively identified with significantly higher normalized intensities for the non-immunized group (FC B/C of 0.58, p=0.01), along with PC(30:0) and PC(32:1), minor components of the pulmonary surfactant layer (Figure V-9A). Even though PC(30:0) and PC(32:1) were not identified by MS/MS match with the employed databases, the putative identifications were manually confirmed by the characteristic neutral loss of 183.066 Da for phosphatidylcholines, as well as neutral losses corresponding to FA(14:0) and FA(16:0) for PC(30:0); and FA(16:0) and FA(16:1) for PC(32:1). Hence, we have found indications that the immunization by the ΔF/TriAdj formulation may help to maintain the integrity of the pulmonary surfactant upon RSV challenge.

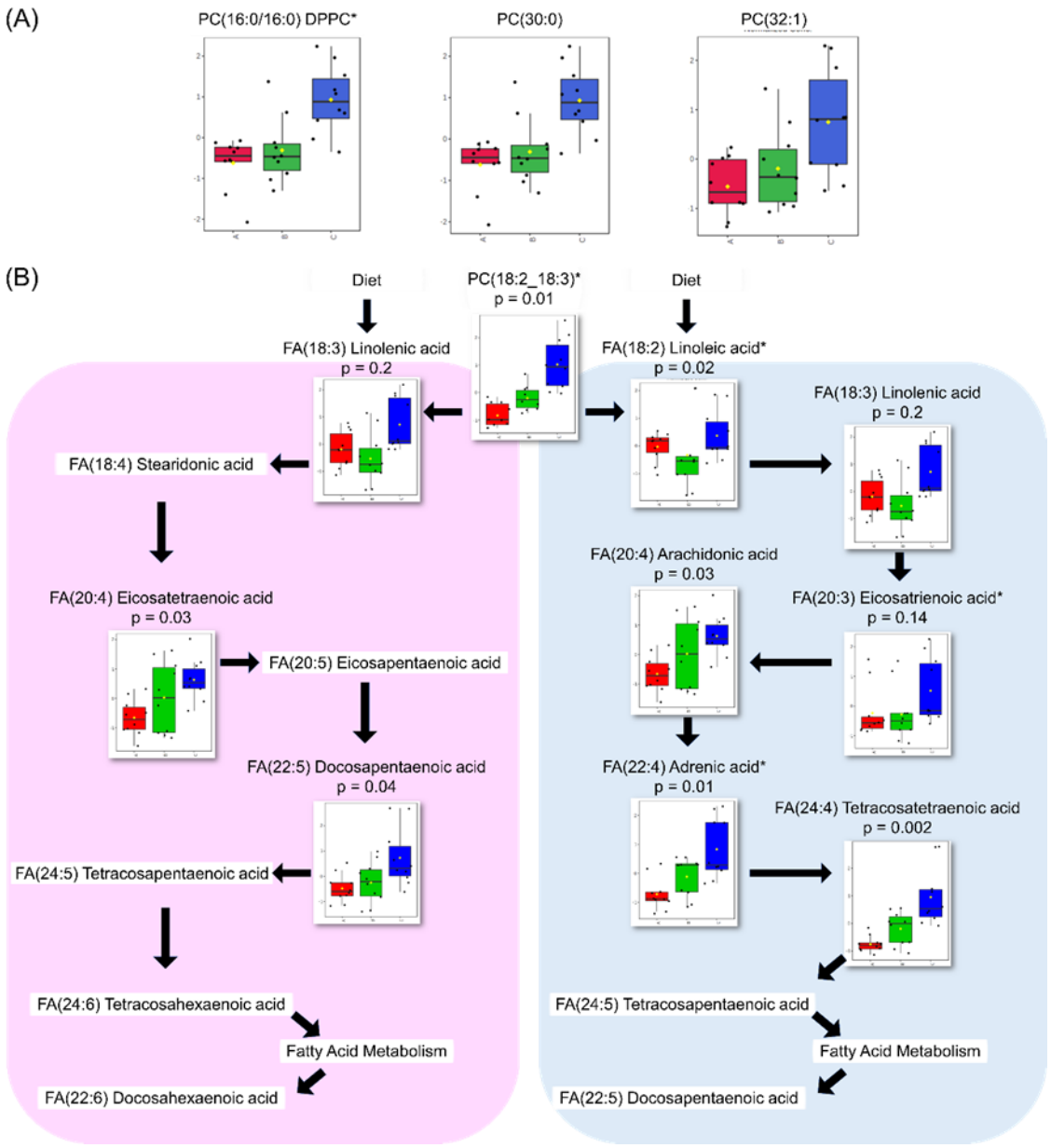

Figure V-9. (A) Boxplots for lipids involved in the composition of the lung surfactant layer. (B) Linoleic acid and alpha-linolenic acid metabolism. Group A (control) is displayed in red; B (immunized, RSV-challenged) in green, and C (non-immunized, RSV- challenged) in blue. The displayed p values were obtained by non-parametric ANOVA for normalized, auto-scaled intensities. *Positive identification (MS/MS).

Pathway analysis was performed with the positively identified lipids (MS/MS match) and metabolites (mass and retention time match). Unfortunately, pathways that include specific lipid molecules are not yet well characterized. Furthermore, matching multiple complex lipids and metabolites to pathways requires the use of common database identifiers, but the currently available databases are limited and a high number of positively identified compounds could not be used. The 186 positively identified lipids (MS/MS) and 163 metabolites registered on the Human Metabolome Database (HMDB, Supp. Table 18) were employed for pathway analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Small Molecule Pathway Database (SMPDB).[79–81,109,153,193–195,200] We employed MetaboAnalyst 4.0 to compare (1) healthy controls and non-immunized, RSV-challenged animals (A/C, Figure V-10A); (2) immunized and non-immunized, RSV-challenged animals (B/C, Figure V-10B); and (3) healthy controls and immunized, RSV-challenged animals (A/B, Figure V-10C).[153] A total of 79 pathways had at least one matched metabolite or lipid (Table V-2). Forty-four pathways displayed $p < 0.05$ for enrichment analysis of at least one of the three studied comparisons, *viz.,* significant coordinated changes in normalized intensities of matched metabolites and lipids.[200] However, only 7 showed pathway impact, the sum of importance of all matched metabolites and lipids within the pathway, ≥0.5, including amino acid-related biosynthesis and metabolism (alanine, aspartate, glutamate, glutamine, phenylalanine, tyrosine, tryptophan, valine, leucine and isoleucine), linoleic acid metabolism and phosphatidylethanolamine biosynthesis (Table V-2).

Table V-2. Metabolic pathways matched for the 349 positively identified lipids (MS/MS) and metabolites (retention time library) with HMDB IDs for binary comparisons of the studied groups using MetaboAnalyst 4.0. Group A: healthy controls; group B: immunized, RSV-challenged animals; group C: non-immunized, RSV-challenged animals.

Table V-2

Pathway name	Source	Impact	Match status (A/B)	p (A/B)	Match status (A/C)	p (A/C)	Match status (B/C)	p (B/C)
Alanine, aspartate and glutamate metabolism	KEGG	0.716	7/24	9.66×10^{-3}	7/24	3.82×10^{-3}	7/24	6.01×10^{-1}
Alpha-linolenic acid and linoleic acid metabolism	SMPDB	0.179	2/17	2.03×10^{-1}	2/17	5.48×10^{-3}	2/17	6.30×10^{-2}
Alpha-linolenic acid metabolism	KEGG	0.000	1/9	3.56×10^{-2}	1/9	6.60×10^{-3}	1/9	2.44×10^{-1}
Aminoacyl-tRNA biosynthesis	KEGG	0.129	16/69	3.31×10^{-2}	16/69	5.68×10^{-3}	16/69	2.68×10^{-1}
Ammonia recycling	SMPDB	0.057	4/25	2.43×10^{-3}	4/25	2.56×10^{-4}	4/25	9.02×10^{-1}
Arachidonic acid metabolism	KEGG	0.000	2/65	2.31×10^{-2}	1/36	6.60×10^{-3}	1/36	2.44×10^{-1}
Arginine and proline metabolism	KEGG	0.304	7/44	1.89×10^{-2}	7/44	3.06×10^{-3}	7/44	1.85×10^{-1}
Aspartate metabolism	SMPDB	0.800	3/34	5.57×10^{-4}	3/34	2.94×10^{-4}	3/34	9.30×10^{-2}
Beta-alanine metabolism	KEGG	0.000	2/26	5.82×10^{-3}	1/17	3.82×10^{-3}	1/17	9.11×10^{-1}
Betaine metabolism	SMPDB	0.000	1/18	4.07×10^{-1}	1/18	2.05×10^{-1}	1/18	4.91×10^{-2}
Bile acid biosynthesis	SMPDB	0.019	5/60	4.53×10^{-2}	5/60	3.76×10^{-2}	5/60	7.95×10^{-2}
Biosynthesis of unsaturated fatty acids	KEGG	0.000	10/42	3.72×10^{-1}	10/42	1.82×10^{-3}	10/42	1.05×10^{-2}
Biotin metabolism	KEGG	0.000	1/5	4.13×10^{-1}	1/5	5.61×10^{-1}	1/5	3.46×10^{-1}
Butanoate metabolism	KEGG	0.029	3/22	2.78×10^{-1}	3/22	2.38×10^{-1}	3/22	7.92×10^{-1}
Butyrate metabolism	SMPDB	0.000	1/16	4.28×10^{-1}	1/16	8.81×10^{-1}	1/16	4.50×10^{-1}
Carnitine synthesis	SMPDB	0.000	2/16	4.85×10^{-1}	2/16	8.51×10^{-1}	2/16	4.21×10^{-1}
Catecholamine biosynthesis	SMPDB	0.000	1/14	1.20×10^{-1}	1/14	6.14×10^{-1}	1/14	3.82×10^{-1}
Citrate cycle (TCA cycle)	KEGG	0.064	2/20	2.61×10^{-1}	2/20	3.87×10^{-1}	2/20	6.34×10^{-1}
Citric acid cycle	SMPDB	0.102	2/25	2.92×10^{-1}	2/25	4.01×10^{-1}	2/25	6.30×10^{-1}

Table V-2

Pathway name	Source	Impact	Match status (A/B)	p (A/B)	Match status (A/C)	p (A/C)	Match status (B/C)	p (B/C)
Cyanoamino acid metabolism	KEGG	0.000	1/6	1.41×10^{-2}	1/6	2.70×10^{-4}	1/6	9.66×10^{-1}
Cysteine and methionine metabolism	KEGG	0.138	3/27	1.89×10^{-2}	3/27	7.67×10^{-4}	3/27	4.34×10^{-1}
Cysteine metabolism	SMPDB	0.000	1/24	4.28×10^{-1}	1/24	5.93×10^{-1}	1/24	9.02×10^{-1}
D-glutamine and d-glutamate metabolism	KEGG	1.000	2/5	9.80×10^{-2}	2/5	4.01×10^{-1}	2/5	6.34×10^{-1}
Drug metabolism - cytochrome p450	KEGG	0.000	1/56	7.67×10^{-1}	1/56	9.42×10^{-1}	1/56	9.20×10^{-1}
Fatty acid biosynthesis	SMPDB	0.075	3/33	4.07×10^{-1}	3/33	1.87×10^{-1}	3/33	1.32×10^{-1}
Fatty acid elongation in mitochondria	SMPDB	0.000	1/27	1.04×10^{-1}	1/33	7.83×10^{-2}	1/27	3.90×10^{-2}
Fatty acid metabolism	SMPDB	0.000	1/39	1.04×10^{-1}	1/40	7.83×10^{-2}	1/39	3.90×10^{-2}
Glutamate metabolism	SMPDB	0.006	4/45	5.82×10^{-3}	4/45	3.43×10^{-3}	4/45	8.52×10^{-1}
Glutathione metabolism	KEGG	0.104	4/26	2.64×10^{-1}	4/26	2.76×10^{-1}	4/26	5.37×10^{-2}
Glycerolipid metabolism	KEGG	0.126	2/18	1.41×10^{-2}	2/18	2.38×10^{-1}	2/18	2.92×10^{-1}
Glycerophospholipid metabolism	KEGG	0.498	7/30	1.09×10^{-2}	7/30	1.39×10^{-3}	7/30	1.56×10^{-1}
Glycine and serine metabolism	SMPDB	0.072	4/50	2.85×10^{-2}	4/50	9.68×10^{-3}	4/50	4.21×10^{-1}
Glycine, serine and threonine metabolism	KEGG	0.299	3/31	4.65×10^{-2}	3/31	1.19×10^{-2}	3/31	2.68×10^{-1}
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	KEGG	0.044	1/14	1.89×10^{-1}	1/14	1.42×10^{-4}	1/14	1.02×10^{-2}
Glyoxylate and dicarboxylate metabolism	KEGG	0.129	1/18	1.70×10^{-1}	1/18	2.13×10^{-1}	1/18	9.82×10^{-1}
Histidine metabolism	KEGG	0.242	4/15	3.56×10^{-2}	4/15	5.90×10^{-3}	4/15	2.36×10^{-1}
Homocysteine degradation	SMPDB	0.000	1/7	1.13×10^{-2}	1/7	2.56×10^{-4}	1/7	9.47×10^{-1}
Ketone body metabolism	SMPDB	0.000	1/12	4.28×10^{-1}	1/12	8.81×10^{-1}	1/12	4.50×10^{-1}
Limonene and pinene degradation	KEGG	0.000	1/8	7.22×10^{-1}	1/8	7.97×10^{-1}	1/8	9.48×10^{-1}
Linoleic acid metabolism	KEGG	1.000	2/6	9.80×10^{-2}	2/6	4.12×10^{-3}	2/6	2.20×10^{-1}
Lysine biosynthesis	KEGG	0.000	2/4	4.13×10^{-1}	2/4	4.25×10^{-2}	2/4	1.31×10^{-1}
Lysine degradation	SMPDB	0.237	2/20	4.28×10^{-1}	2/20	3.47×10^{-2}	2/20	7.32×10^{-2}

Table V-2

Pathway name	Source	Impact	Match status (A/B)	p (A/B)	Match status (A/C)	p (A/C)	Match status (B/C)	p (B/C)
Malate-aspartate shuttle	SMPDB	0.143	1/7	6.08×10^{-3}	1/7	3.43×10^{-3}	1/7	9.02×10^{-1}
Methane metabolism	KEGG	0.400	1/9	1.41×10^{-2}	1/9	2.70×10^{-4}	1/9	9.66×10^{-1}
Methionine metabolism	SMPDB	0.039	7/39	4.53×10^{-2}	6/39	2.36×10^{-3}	6/39	4.91×10^{-2}
Mitochondrial beta-oxidation of long chain saturated fatty acids	SMPDB	0.000	1/24	5.42×10^{-1}	1/24	7.04×10^{-1}	1/24	9.02×10^{-1}
Mitochondrial electron transport chain	SMPDB	0.000	1/15	4.28×10^{-1}	1/15	8.81×10^{-1}	1/15	4.50×10^{-1}
Nitrogen metabolism	KEGG	0.000	3/9	3.56×10^{-2}	3/9	8.06×10^{-2}	3/9	7.96×10^{-1}
Oxidation of branched-chain fatty acids	SMPDB	0.000	1/21	4.28×10^{-1}	1/21	8.81×10^{-1}	1/21	4.50×10^{-1}
Pantothenate and CoA biosynthesis	KEGG	0.000	1/15	9.80×10^{-2}	1/15	7.74×10^{-1}	1/15	2.36×10^{-1}
Phenylalanine and tyrosine metabolism	SMPDB	0.116	3/25	1.84×10^{-2}	3/25	6.54×10^{-3}	3/25	6.30×10^{-1}
Phenylalanine metabolism	KEGG	0.407	3/11	3.56×10^{-2}	3/11	2.69×10^{-2}	3/11	6.34×10^{-1}
Phenylalanine, tyrosine and tryptophan biosynthesis	KEGG	1.000	2/4	3.56×10^{-2}	2/4	4.25×10^{-2}	2/4	4.34×10^{-1}
Phosphatidylcholine biosynthesis	SMPDB	0.220	2/18	1.84×10^{-2}	2/18	2.36×10^{-3}	2/18	6.29×10^{-1}
Phosphatidylethanolamine biosynthesis	SMPDB	0.500	3/13	2.84×10^{-3}	3/13	5.86×10-5	3/13	8.02×10^{-1}
Phospholipid biosynthesis	SMPDB	0.100	3/25	8.17×10^{-1}	3/25	1.37×10^{-2}	3/25	5.69×10^{-2}
Phytanic acid peroxisomal oxidation	SMPDB	0.000	1/19	4.28×10^{-1}	1/19	8.81×10^{-1}	1/19	4.50×10^{-1}
Plasmalogen synthesis	SMPDB	0.000	1/16	5.42×10^{-1}	1/16	7.04×10^{-1}	1/16	9.02×10^{-1}
Porphyrin and chlorophyll metabolism	KEGG	0.000	1/27	3.87×10^{-1}	1/27	5.61×10^{-1}	1/27	9.41×10^{-1}
Porphyrin metabolism	SMPDB	0.000	1/36	1.11×10^{-1}	1/36	8.51×10^{-1}	1/36	2.31×10^{-1}
Primary bile acid biosynthesis	KEGG	0.030	3/46	1.52×10^{-1}	3/46	4.39×10^{-2}	3/46	4.34×10^{-1}
Propanoate metabolism	KEGG	0.000	1/20	4.19×10^{-1}	1/20	9.18×10^{-1}	1/20	4.34×10^{-1}
Purine metabolism	KEGG	0.077	6/68	3.56×10^{-2}	6/68	5.87×10^{-2}	6/68	2.87×10^{-1}
Pyrimidine metabolism	KEGG	0.023	1/54	9.45×10^{-1}	1/41	2.68×10^{-1}	1/41	4.34×10^{-1}
Pyruvate metabolism	KEGG	0.106	1/23	9.22×10^{-2}	1/23	9.45×10^{-1}	1/23	2.36×10^{-1}
Selenoamino acid metabolism	SMPDB	0.000	2/28	1.91×10^{-2}	2/28	1.14×10^{-4}	2/28	9.30×10^{-2}

Table V-2

Pathway name	Source	Impact	Match status (A/B)	p (A/B)	Match status (A/C)	p (A/C)	Match status (B/C)	p (B/C)
Spermidine and spermine biosynthesis	SMPDB	0.086	2/14	1.36×10^{-1}	2/14	2.16×10^{-2}	2/14	6.23×10^{-2}
Sphingolipid metabolism	KEGG	0.296	3/21	1.49×10^{-3}	3/21	2.70×10^{-4}	3/21	3.46×10^{-1}
Steroid biosynthesis	KEGG	0.000	1/35	9.04×10^{-2}	1/35	3.85×10^{-1}	1/35	4.07×10^{-1}
Steroid hormone biosynthesis	KEGG	0.000	1/72	7.22×10^{-1}	1/72	5.13×10^{-1}	1/72	2.44×10^{-1}
Taurine and hypotaurine metabolism	KEGG	0.429	1/8	1.89×10^{-1}	1/8	8.42×10^{-1}	1/8	8.47×10^{-1}
Threonine and 2-oxobutanoate degradation	SMPDB	0.000	1/13	4.07×10^{-1}	1/13	3.71×10^{-1}	1/13	1.05×10^{-1}
Tryptophan metabolism	KEGG	0.371	4/40	1.49×10^{-3}	4/40	4.60×10-6	4/40	1.83×10^{-3}
Tyrosine metabolism	KEGG	0.140	1/44	1.09×10^{-1}	1/44	5.68×10^{-1}	1/44	3.49×10^{-1}
Ubiquinone and other terpenoid-quinone biosynthesis	KEGG	0.000	1/3	1.09×10^{-1}	1/3	5.68×10^{-1}	1/3	3.49×10^{-1}
Urea cycle	SMPDB	0.266	2/23	1.44×10^{-2}	2/23	3.43×10^{-3}	2/23	6.23×10^{-2}
Valine, leucine and isoleucine biosynthesis	KEGG	0.667	2/11	5.02×10^{-2}	2/11	1.47×10^{-1}	2/11	2.77×10^{-1}
Valine, leucine and isoleucine degradation	KEGG	0.000	3/38	6.17×10^{-2}	3/38	1.47×10^{-1}	3/51	6.23×10^{-2}
Vitamin b6 metabolism	KEGG	0.490	1/9	6.44×10^{-1}	1/9	1.62×10^{-1}	1/9	4.34×10^{-1}

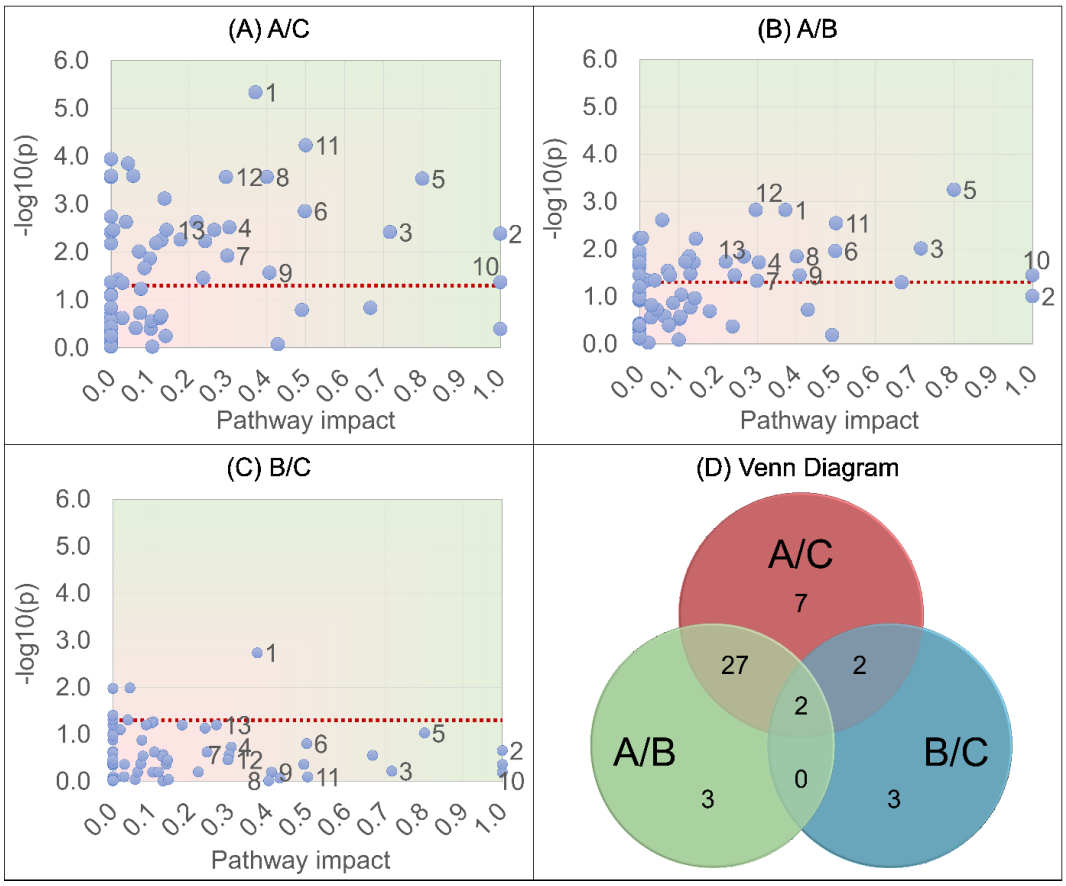

Figure V-10. Pathway analysis for the positively identified lipids and metabolites. Significantly affected pathways ($p < 0.05$ and pathway impact >0.25 for at least one comparison) are labeled as: 1. Tryptophan metabolism; 2. Linoleic acid metabolism; 3. Alanine, aspartate and glutamate metabolism; 4. Arginine and proline metabolism; 5. Aspartate metabolism; 6. Glycerophospholipid metabolism; 7. Glycine, serine and threonine metabolism; 8. Methane metabolism; 9. Phenylalanine metabolism; 10. Phenylalanine, tyrosine and tryptophan biosynthesis; 11. Phosphatidylethanolamine biosynthesis; 12. Sphingolipid metabolism; 13. Urea cycle. (A) A/C comparison (healthy controls *versus* non-immunized, RSV-challenged animals); (B) A/B comparison (healthy controls *versus* immunized, RSV-challenged animals); (C) B/C comparison (immunized *versus* non-immunized, RSV-challenged animals); (D) Venn diagram for the number of significantly affected pathways ($p < 0.05$), regardless of impact value.

The comparison between healthy controls and non-immunized, RSV-challenged mice (A/C, Figure V-10A) resulted in 38 significantly affected pathways ($p < 0.05$), including five with impact ≥0.5: alanine, aspartate and glutamate metabolism; linoleic acid metabolism; phenylalanine, tyrosine and tryptophan biosynthesis; aspartate metabolism; and phosphatidylethanolamine biosynthesis. Similarly, the comparison between healthy controls and immunized, RSV-challenged mice (A/B, Figure V-10B) resulted in 32 significantly affected pathways, including four with impact ≥0.5: alanine, aspartate and glutamate metabolism; aspartate metabolism; phenylalanine, tyrosine and tryptophan biosynthesis; and phosphatidylethanolamine biosynthesis. The comparison between immunized and non-immunized, RSV-challenged mice (B/C, Figure V-10C) significantly affected 7 pathways, which were modulated by ΔF/TriAdj immunization (Table V-2, Table V-3): tryptophan metabolism, methionine metabolism, glycosylphosphatidylinositol(GPI)-anchor biosynthesis, betaine metabolism, biosynthesis of unsaturated fatty acid, fatty acid elongation in mitochondria and fatty acid metabolism. The tryptophan metabolism, affected for all three comparisons and further discussed in our related work, resulted in an impact of 0.37 and $p=0.002$ for the comparison between immunized and non-immunized mice, while the remaining pathways had smaller impacts.[1]

Table V-3. Pathways related to ΔF/TriAdj immunization and RSV-challenge that were significantly affected (enrichment p <0.05) for at least one of the comparisons (A/B, A/C or B/C), including the identified metabolites and lipids. The remaining matched pathways are presented in Table V-2.

Pathway	p			Identified metabolites/lipids
	A/B	A/C	B/C	
Glycerophospholipid metabolism	0.01	0.001	0.03	phosphatidylethanolamine[a], 3-sn-phosphatidylcholine[a], 1-acyl-sn-glycero-3-phosphocholine, ethanolamine phosphate[a], ethanolamine[a], phosphatidic acid, 3-O-sn-phosphatidyl-L-serine
Methionine metabolism	0.04	0.002	0.049	L-serine, L-homoserine, sarcosine, putrescine, 5'-methyladenosine, adenosine
Biosynthesis of unsaturated fatty acids	0.4	0.002	0.01	palmitic acid[a], stearic acid[a], oleic acid[a], linoleic acid[a], gamma-linoleic acid, eicosenoic acid[a], eicosadienoic acid[a], eicosatrienoic acid, arachidonic acid*[a], erucic acid[a], docosadienoic acid*[a], docosatetraenoic acid[a], clupadonic acid*[a], docosahexaenoic acid*[a], nervonic acid*[a]
Linoleic acid metabolism	0.1	0.004	0.2	PC(18:2_18:3)[a], linoleic acid[a], linolenic acid, eicosatetraenoic acid*[a], eicosatrienoic acid, arachidonic acid*[a], docosapentaenoic acid*[a], adrenic acid[a], tetracosatetraenoic acid*[a]
Lysine biosynthesis and degradation	0.4	0.04	0.1 (syn.); 0.07 (degr.)	lysine, (S)-2,3,4-tetrahydropiperidine-2-carboxylate*, allysine*, aminoadipic acid[a], oxoglutaric acid*, oxoadipic acid*[a], glutamic acid
Phospholipid biosynthesis	0.8	0.01	0.06	LPC(16:0)[a], LPE(16:0)[a], ethanolamine[a], PA(32:0)*, PC(P-32:0)*[a], PE(32:0)*, PG(P-32:0)*[a]
Primary bile acid biosynthesis	0.2	0.04	0.3	Cholic acid[a], taurine, chenodeoxycholic acid
Spermidine and spermine biosynthesis	0.1	0.02	0.01	putrescine[a], 5'-methyladenosine[a]
Alanine, aspartate and glutamate metabolism	0.01	0.004	0.6	aspartic acid[a], asparagine[a], oxaloacetic acid, alanine[a], pyruvic acid*[a], (S)-1-pyrroline-5-carboxylate*[a], glutamic acid*, 2-keto-glutamic acid*[a], oxoglutaric acid*, gamma-aminobutyric acid, succinic acid semi-aldehyde*[a], glutamine, succinic acid

Pathway	p			Identified metabolites/lipids
	A/B	A/C	B/C	
Phosphatidylethanolamine biosynthesis	0.003	0.0001	0.8	ethanolamine[a], serine[a], O- ethanolamine phosphate[a], 3-O-sn-phosphatidyl-L-serine, 1,2-diacylglycerol[a], phosphatidylethanolamine[a], 3-sn-phosphatidylcholine[a]
Sphingolipid metabolism	0.001	0.0003	0.3	L-serine[a], ethanolamine phosphate[a], ceramide[a], sphinganine 1-phosphate*[a], ceramide 1-phosphate*, lactosyl ceramide*[a], galactosyl ceramide*[a], glucosyl ceramide*[a],
Purine	0.04	0.06	0.3	hypoxanthine[a], glutamine, adenosine monophosphate[a], 2'deoxyguanosine-5'-monophosphate[a], guanosine[a]
Glycerolipid metabolism	0.01	0.2	0.3	Triacylglycerol[a], phosphatidic acid, 1,2-diacylglycerol[a], monoacylglycerol, fatty acid[a], lysophosphatidic acid*, monogalactosyl-diacylglycerol*, digalactosyl-diacylglycerol*[a]
Nitrogen metabolism	0.04	0.06	0.3	Glutamine, glutamic acid

*putatively identified based on mass matches. [a]molecules with $p < 0.05$ for non-parametric ANOVA.

Seven pathways were significantly altered only for the non-immunized, RSV-challenged group compared to healthy controls (A/C), but not for the immunized group (A/B); hence, they were modulated by the ΔF/TriAdj formulation: alpha-linoleic acid and linoleic acid metabolisms; lysine biosynthesis and degradation; and phospholipid, bile acid, spermidine and spermine biosynthesis. However, we found three pathways that were affected only for the immunized, RSV-challenged mice compared to healthy controls (A/B, glycerolipid, nitrogen and purine metabolisms), which were considered as a result of immunization with the proposed ΔF/TriAdj formulation. Finally, 29 pathways were altered for both the A/C and A/B comparisons (Figure V-10D, $p < 0.05$), thereby not modulated by immunization.

5.3.4. Linoleic acid and alpha-linolenic acid pathways

Linoleic acid is an essential PUFA that acts as a precursor for longer omega-6 fatty acids, used in the biosynthesis of prostaglandins via arachidonic acid. The alternative form alpha-linolenic acid is the precursor of omega-3 fatty acids via stearidonic acid. Both compounds affect the composition of membranes, which may influence cellular function, including inflammation and immune response. A scheme depicting the linoleic acid and alpha-linolenic acid metabolism, along with boxplots for the positively and putatively identified molecules, is displayed in Figure V-9B. The linoleic acid metabolism was exclusively affected for the non-immunized, RSV-challenged mice *versus* healthy controls (A/C, Table V-3). No significant changes were observed for the immunized animals (A/B); thus, the metabolic reactions seem to be regulated by the ΔF/TriAdj formulation. Overall, metabolites involved in the alpha-linolenic acid and linoleic acid metabolisms had higher normalized intensities for the non-immunized, RSV-challenged animals, with similar values for the healthy and immunized mice. Linoleic acid was significantly increased for the non-immunized mice compared to immunized animals (FC(B/C)=0.61, p=0.02) and healthy controls (FC(A/C)=0.72, p=0.0495); however, it was not significant for the immunized group *versus* healthy controls (FC(A/B)=1.17, p=0.50), suggesting ΔF/TriAdj modulation of RSV challenge. Other lipids involved in the alpha-linolenic acid and linoleic acid metabolisms were putatively or positively identified, *i.e.*, tetracosatetraenoic acid (p=0.002), adrenic acid, (p=0.01), arachidonic acid (p=0.03), eicosatetraenoic acid (p=0.03), eicosatrienoic acid (p=0.14) and alpha- or gamma-linolenic acid (p=0.17). Most of them also displayed higher normalized intensities for the non-immunized animals, with regulation by the vaccine formulation in the immunized group (Figure V-9B). The PUFA docosapentaenoic acid (DPA) and docosahexaenoic acid (DHA) were

also significantly increased for the non-immunized animals compared to healthy controls (FC(A/C) of 0.59 and 0.86, respectively, with $p < 0.05$), but not for the immunized mice.

Overall, we observed modulation of the linoleic and alpha-linolenic acid pathway by the ΔF/TriAdj formulation. The relationship between unsaturated fatty acids, such as linoleic acid and alpha-linolenic acid, and the activation and proliferation of T-cells have a delicate balance. PUFA are immune modulators that can induce inflammatory responses and control cell homeostasis. There is an indication that low dietary intake of fatty acids is essential for the proliferation of T-cells and the production of cytokines for an adequate immune response to pathogens, but excessive levels of PUFA may induce apoptosis of immune cells. [22] We have now found evidence that RSV challenge significantly affects the linoleic and alpha-linolenic acid pathways, including the PUFA DHA and DPA, with higher levels of most identified species in the lung tissue of mice challenged by RSV without prior immunization. The regulation observed for immunized animals may be a result of memory cells triggering a faster, stronger immune response, with increased uptake of fatty acids and stimulation of their oxidative metabolism immediately after viral challenge. Non-immunized animals should display a less intense, long-lasting immune response that is still active when the lungs were collected, 7 days after RSV challenge. We have previously reported a high influx of immune cells in the lungs of both immunized and non-immunized animals one day after RSV challenge; however, the levels of immune cells decreased to numbers more similar to healthy mice 7 days after challenge.[1] Hence, the composition of the lung tissue studied herein, collected 7 days after RSV challenge, reflects the increased levels of PUFA precursors in non-immunized mice due to the slower, continuing immune response to the virus. The higher influx of immune cells whose membranes contain higher concentrations of PUFA precursors results in the upregulation of the linoleic and alpha-linolenic acid pathways and stimulation of the production

of PUFA as precursors of eicosanoids and prostaglandins, which act as important signaling molecules. We can infer that the mechanism of action of the proposed vaccine formulation involves the regulation of fatty acids and PUFA biosynthetic pathways, resulting in levels that are more similar to healthy controls for immunized mice compared to non-immunized animals.

5.3.5. Biosynthesis of unsaturated fatty acids

The biosynthesis of unsaturated fatty acids was significantly altered for both the healthy controls *versus* non-immunized, RSV-challenged group (A/C) and the immunized *versus* non-immunized mice (B/C); however, it was not significant for the healthy controls *versus* immunized mice (A/B, Table V-3). The boxplots for the 15 positively and putatively identified fatty acids involved in the biosynthesis pathway are provided in Figure V-V-11. The biosynthesis and β-oxidation of lipids is further discussed in Chapter I (*1.2.4.2. Energy production* and *1.2.4.3. Lipogenesis and metabolism*, p. 11). The breakdown of triacylglycerols into fatty acids can be stimulated upon viral infection to increase the host energy supply for viral replication and immune response.[201] Fatty acids are major energy sources for cells involved in the immune system and are required for cellular membranes. Also, PUFA may act as signaling molecules or their precursors during an immune response, as previously discussed.[24,27,202] Overall, RSV challenge without prior immunization (C) seems to increase normalized intensities of fatty acids in lung tissue, whereas the effect is minimized for ΔF/TriAdj-immunized animals (B). Similarly to our findings for the linoleic and alpha-linolenic acid pathways, most of the identified lipids involved in the biosynthesis of unsaturated fatty acids displayed significantly higher normalized intensities for the non-immunized, RSV-challenged group (C), with FC (A/C) <0.67 and $p < 0.05$ for six fatty acids, while the remaining nine had fold changes (A/C) between 0.68 and 0.86. The immunized animals

had normalized intensities that were more similar to the healthy control group, suggesting modulation of the pathway, with only one significantly altered lipid, FA(22:2) (FC(A/B)=0.54, p=0.01). Furthermore, the comparison between immunized and non-immunized, RSV-challenged mice resulted in 8 significantly altered fatty acids (FC(B/C) <0.67, p <0.05), and all the remaining lipids displayed fold changes between 0.67 and 0.87.

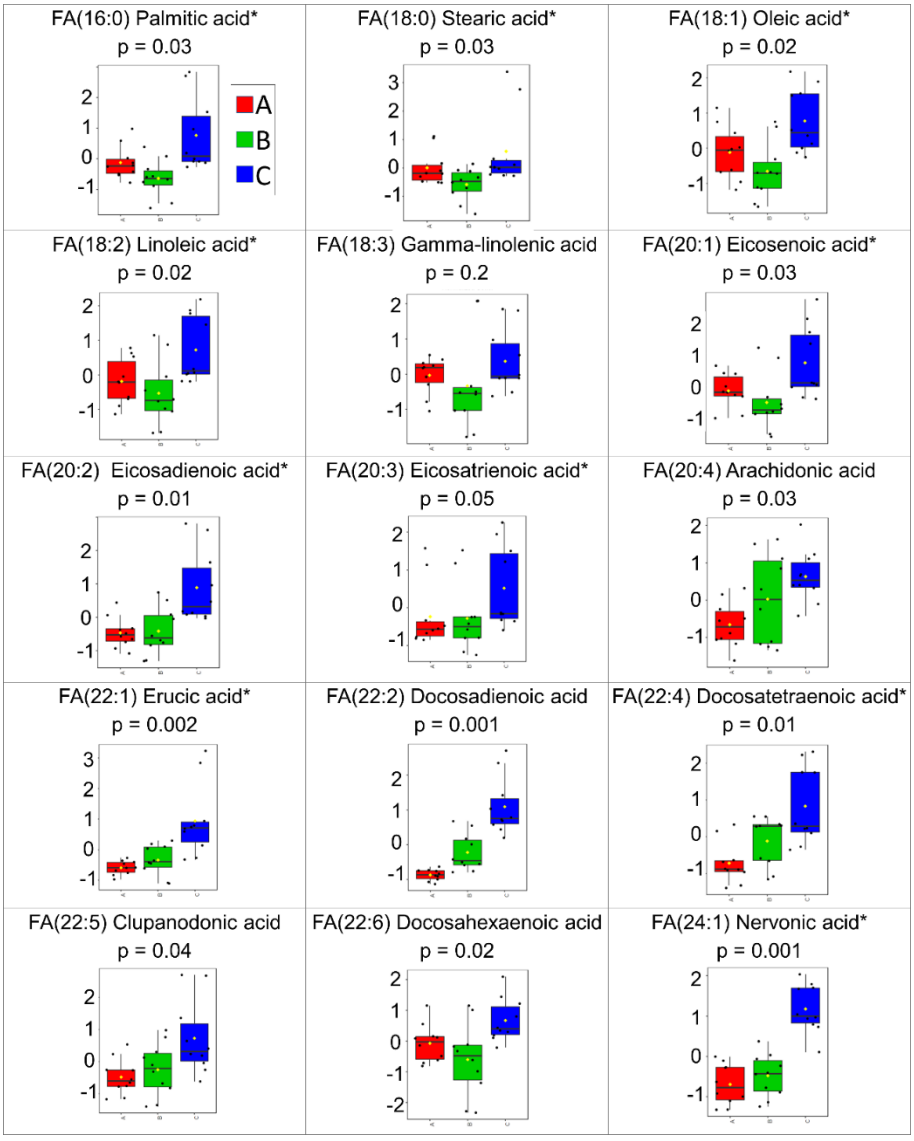

Figure V-V-11. Boxplots for lipids involved in the biosynthesis of unsaturated fatty acids. Group A (control) is displayed in red; B (immunized, RSV-challenged) in green, and C (non-immunized,

RSV-challenged) in blue. The p values were obtained by non-parametric ANOVA for normalized, auto-scaled intensities. *Positive identification (MS/MS).

5.3.6. Lysine biosynthesis and degradation

The lysine biosynthesis and degradation pathways were also modulated by the ΔF/TriAdj formulation. Both pathways were significantly affected for the non-immunized group when compared to healthy controls (A/C), but not for the immunized animals (A/B, Table V-3, Figure V-12). Lysine is an essential amino acid that can induce an inflammatory and immune response. Although lysine was not significantly altered by immunization or RSV challenge, its oxidized derivative allysine had higher normalized intensities for the non-immunized, RSV-challenged mice compared to the control group (FC(A/C)=0.70, p=0.03), but not for the immunized group (FC(A/B)=0.76, p=0.22). The higher levels of the oxidized form may indicate higher oxidative stress due to the immune response. Allysine is involved in the stabilization of elastin and collagen, with increased levels being possibly related to pulmonary fibrosis. The observed regulation by the vaccine formulation suggests mitigation of oxidative processes upon immunization, with a possible protective effect against lung damage. Furthermore, aminoadipic acid (p=0.0002) and oxoadipic acid (p=0.048) were significantly affected by the three studied groups. Both acids are part of the kynurenine/alpha-aminoadipate aminotransferase mitochondrial reaction, *i.e.*, conversion of aminoadipic and oxoglutaric acid into oxoadipic and glutamic acid. Aminoadipic acid, an intermediate in the breakdown of lysine, displayed lower normalized intensities for the non-immunized group when compared to the healthy controls (FC(A/C)=1.35, p=0.001) and immunized animals (FC(B/C)=1.24, p=0.0001). However, it was not significant for the immunized animals (FC(A/B)=1.09, p=0.7); hence, modulated by the vaccine. Oxoadipic acid, a product in

the aminotransferase reaction, had the opposite trend, with significantly higher normalized intensities for the non-immunized group (FC(A/C)=0.59, p=0.02), but also not significantly altered for the immunized mice. The aminotransferase enzyme is also involved in the tryptophan metabolism, significantly affected for all three comparisons (A/B, A/C and B/C). The lower normalized intensities found for aminoadipic and oxoglutaric acid, combined with higher intensities for oxoadipic acid, suggest higher activity of the aminotransferase enzyme upon viral challenge, with subsequent regulation by the vaccine formulation.

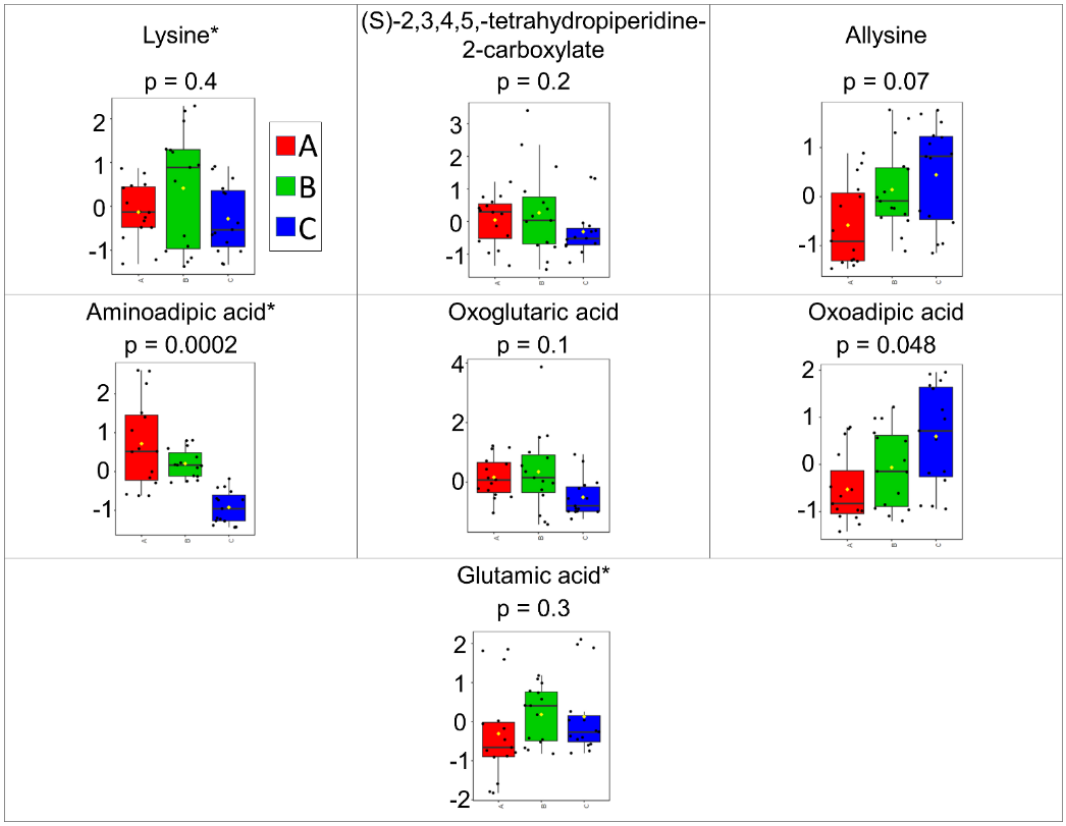

Figure V-12 .Boxplots for metabolites involved in the biosynthesis and degradation of lysine. Group A (control) is displayed in red; B (immunized, RSV-challenged) in green, and C (non-immunized, RSV-challenged) in blue. The p values were obtained by non-parametric ANOVA for normalized, auto-scaled intensities. *Positive identification (retention time and accurate mass).

5.3.7. Alanine, aspartate and glutamate metabolisms

The alanine, aspartate and glutamate metabolisms were not significantly affected for the immunized *versus* non-immunized, RSV-challenged mice (B/C), but resulted in enrichment $p < 0.05$ for both groups compared to healthy controls (A/C and A/B, Table V-3, Figure V-13). The metabolism for the three amino acids seems to be affected by RSV challenge, but not modulated by ΔF/TriAdj immunization. Aspartic acid and asparagine, involved in the aspartate pathway, displayed higher normalized intensities for the RSV-challenged groups regardless of immunization. Both metabolites had significantly increased levels for immunized and non-immunized animals compared to healthy controls (A/B and A/C $p < 0.05$, FC between 0.71 and 0.79), but they were not significantly altered in a direct comparison (B/C $p = 0.6$ and FC of 1.0 for both). For the alanine metabolism, pyruvic acid was significantly altered for the immunized animals compared to non-immunized mice (FC(B/C)=1.99, $p = 0.02$), while alanine was not significant for any of the three comparisons. The glutamate metabolism displayed no significant alterations for the immunized *versus* non-immunized mice (B/C), but 2-ketoglutaramic acid was significantly less intense for both RSV-challenged groups when compared to healthy controls (FC ≥1.5, $p < 0.05$), whereas succinic acid semi-aldehyde was significant only for the non-immunized group (FC (A/C)=1.29, $p = 0.004$).

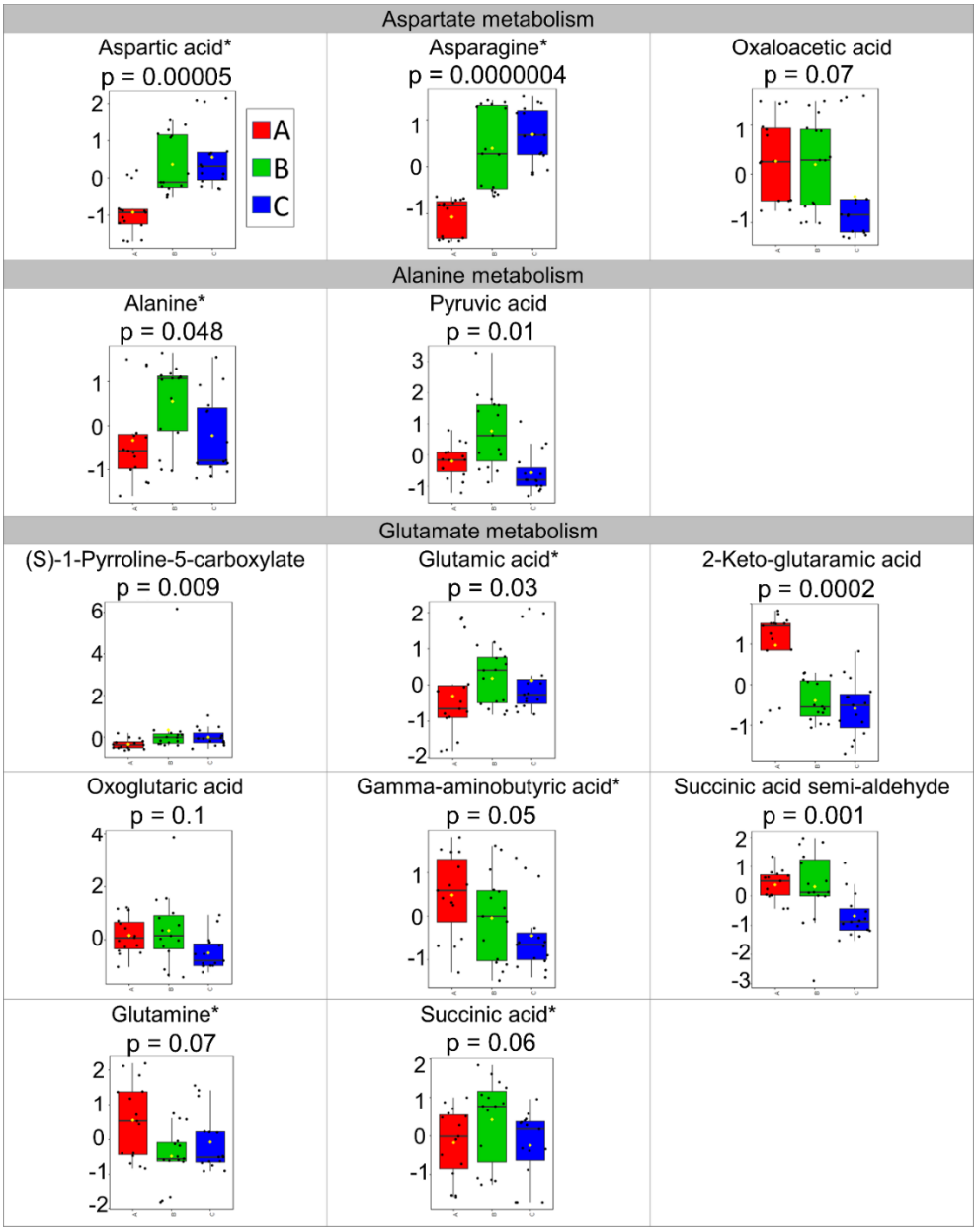

Figure V-13. Boxplots for metabolites involved in the alanine, aspartate and glutamate metabolism. Group A (control) is displayed in red; B (immunized, RSV-challenged) in green, and C (non-immunized, RSV- challenged) in blue. The p values were obtained by non-parametric ANOVA for normalized, auto-scaled intensities. *Positive identification (retention time and accurate mass).

We observed increased levels of pyruvic acid upon immunization, but similar normalized intensities for healthy and non-immunized animals. This alpha-keto acid is the output of glycolysis, acting as an energy source in the citric acid cycle during aerobic respiration by conversion into acetyl-CoA. Acetyl-CoA, in turn, is involved in lipid biosynthetic pathways (Chapter I – *1.2.4.3. Lipogenesis and metabolism*, p. 12), which displayed extensive alteration due to immunization and RSV challenge. Dysregulation in pyruvate metabolism may be related to many diseases, including cancer and neurodegenerative conditions; however, the higher levels observed herein for immunized animals may be a consequence of immune response, which requires proliferation of immune cells and activation of signaling processes. [203]

5.3.8. Glycerolipid, nitrogen and purine metabolisms

The glycerolipid, nitrogen and purine metabolisms were significantly affected exclusively for the immunized, RSV-challenged mice compared to healthy controls (A/B, Table V-3, Figure V-14 and Figure V-15). The significant coordinated alterations upon immunization for the three pathways may be related to the vaccine formulation. The nitrogen and purine metabolisms are deeply interconnected, with both being significantly affected (Figure V-14). The purine metabolism regulates nucleotide biosynthesis and degradation, involving different types of nitrogen-containing heterocyclic molecules, *e.g.,* adenine, guanine, hypoxanthine, xanthine, theobromine, caffeine and uric acid. Dysregulation in purine metabolism is related to autoimmune conditions and neurological dysfunction.[204] Even though we found $p < 0.05$ for enrichment analysis for the comparison between immunized mice and healthy controls (A/B), the univariate analysis by Volcano plot did not result in any significantly altered metabolites for the purine metabolism, *i.e.,* metabolic changes between the matched molecules were coordinated, but individual changes

were below the previously established significance threshold. Nevertheless, the comparison for healthy controls and non-immunized mice resulted in p=0.06 for pathway enrichment analysis, just slightly above the arbitrary limit of 0.05. Hypoxanthine and 2'-deoxyguanosine-5'-monophosphate were significantly altered for the non-immunized group *versus* healthy controls (A/C) with opposite behaviors: hypoxanthine was more intense for the RSV-challenged group (FC(A/C)=0.64), while 2'-deoxyguanosine-5'-monophosphate had higher normalized intensities for the healthy group (FC(A/C)=1.57). For the immunized group *versus* healthy controls, adenosine monophosphate and 2'-deoxyguanosine-5'-monophosphate displayed FC(A/B)=1.5 with p=0.07, while the remaining metabolites had higher p values.

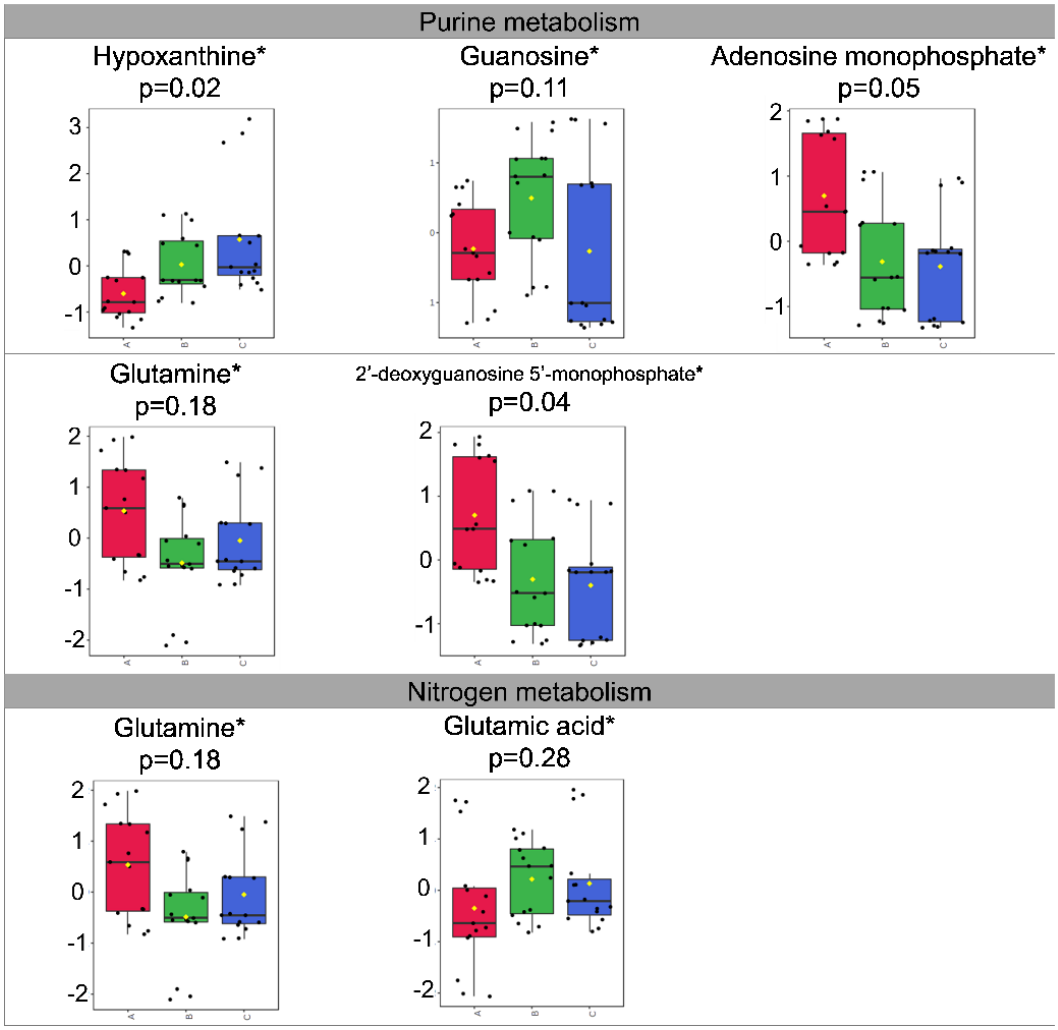

Figure V-14. Boxplots for metabolites involved in the purine and nitrogen metabolisms. Group A (control) is displayed in red; B (immunized, RSV-challenged) in green, and C (non-immunized, RSV- challenged) in blue. The p values were obtained by non-parametric ANOVA for normalized, auto-scaled intensities. *Positive identification (retention time and accurate mass).

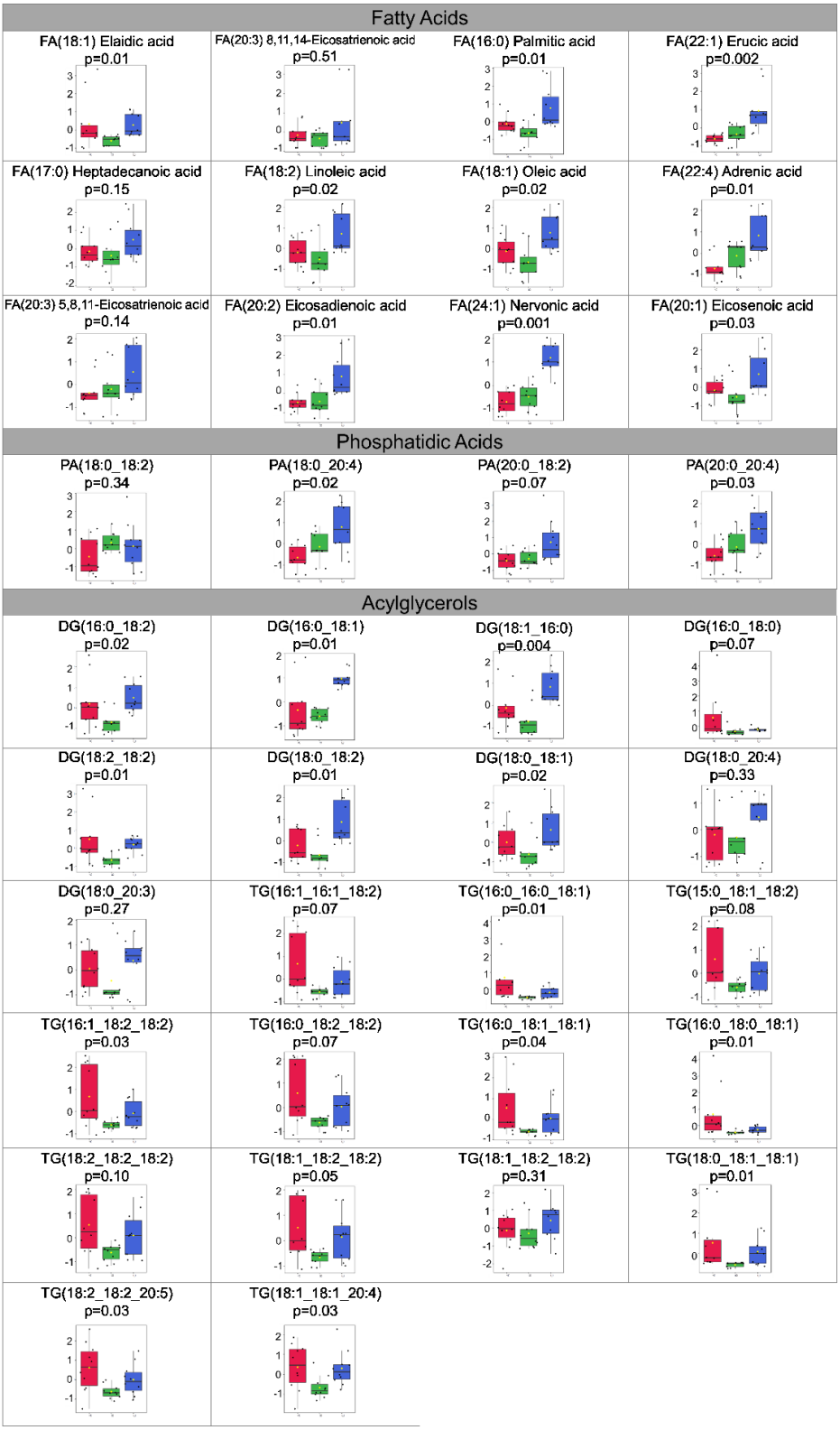

Figure V-15. Boxplots for positively identified lipids involved in the glycerolipid metabolism. Group A (control) is displayed in red; B (immunized, RSV-challenged) in green, and C (non-

immunized, RSV- challenged) in blue. The p values were obtained by non-parametric ANOVA for normalized, auto-scaled intensities.

The glycerolipid metabolism involves the biosynthesis of acylglycerols, beginning with the combination of glycerol-3-phosphate with acyl groups from acyl-CoA molecules. Further details are provided in Chapter I (*1.2.4.3. Lipogenesis and metabolism*, p. 12) The resulting phosphatidic acids may have variable fatty acyl lengths and can be dephosphorylated to produce diacylglycerols. The addition of one extra fatty acyl moiety via acyl-CoA or acyltransferases results in triacylglycerols, which are stored in the organism as a source of energy. Processes that require extra energy may trigger the breakdown of triacylglycerols into glycerol and fatty acids. The β-oxidation of fatty acids has acetyl-CoA or propionyl-CoA as final products (Chapter I, *1.2.4.2. Energy production*). We have previously discussed that pyruvic acid or pyruvate, also a source of acetyl-CoA through glycolysis, had higher normalized intensities for the immunized animals, but not for healthy or non-immunized mice (Figure V-13). Now, we also observed a dramatic change in the glycerolipid composition of lung tissue upon immunization and RSV challenge (Figure V-15). The immunized group had significantly lower levels of triacylglycerols when compared to healthy controls (FC(A/B) $\geq$1.5, $p<0.05$). Even though the levels were also reduced for the non-immunized animals (FC(A/C) $\geq$1.5), the difference was less pronounced and none of them was significant ($p>0.05$). The boxplots for triacylglycerols in Figure V-15 emphasize a trend of lower levels for the immunized group compared to both the non-immunized and healthy controls, *i.e.*, the animals were either synthesizing fewer triacylglycerol molecules or consuming them for energy-demanding processes after immunization and RSV-challenge. However, the positively identified diacylglycerols, phosphatidic acid and fatty acids were not significantly altered for the

immunized animals (A/B); in fact, we observed similar levels of both lipid subclasses for the control and immunized groups, but higher levels for the non-immunized animals. The lower levels of fatty acids and triacylglycerols found for immunized animals, combined to higher levels of pyruvic acid, indicate a higher rate of β-oxidation of lipids. Immune cell differentiation is highly dependent on lipid metabolism, with fatty acids residues being vital for their activation, function and survival. It has been previously demonstrated that memory T cells employ lipid oxidative metabolism, *i.e.*, cleavage of triacylglycerols, as a main source of energy. Hence, the observed results suggest higher uptake of energy by the action of memory cells in the immunized animals. [25,205,206]

5.4 Conclusions

Our previous work has shown the immunological and protective effects of the ΔF/TriAdj formulation against RSV. [1] Now, we characterized its effects upon metabolites and lipids, as well as the affected metabolic pathways in lung tissue. The combination of two chemical isotope labeling techniques for metabolomics with untargeted lipidomics allowed a more complete evaluation of metabolic changes induced by the vaccine formulation and RSV challenge, greatly expanding our previous knowledge. This work provided information to elucidate the mechanism of action for the vaccine formulation, as well as the alterations caused in lung lipidic and metabolic composition upon RSV challenge. Linoleic acid, fatty acids and amino acid metabolisms resulted in significantly coordinated changes between matched molecules for healthy controls, immunized and non-immunized RSV-challenged mice. We also observed significantly higher intensities for lipids involved in the lung surfactant layer for the non-immunized, RSV-challenged animals compared to healthy controls, whereas immunization with the proposed formulation seems to

stabilize its composition. The healthy *versus* immunized, RSV-challenged comparison showed the least amount of significant changes, with only three affected pathways: glycerolipid, nitrogen and purine metabolism. The characterization of metabolic changes due to the ΔF/TriAdj formulation shows modulation of RSV infection, as well as minimal alterations when compared to healthy animals. In the future, we intend to evaluate the vaccine formulation effects for an immunized, non-challenged group, as well as evaluation of sex-related differences that may appear for male *versus* female mice.

The pathology due to RSV infection is not essentially caused by the viral replication, but by the immune responses occurring in the lungs. In this study, the lung metabolites and lipids that were found to be altered in response to RSV infection involved glycerophospholipids, purine, pyrimidine and amino acids, which can be correlated to the production of pro-inflammatory cytokines, including IL-1β, IL-6, IL-10, TNF-α, and IFN-γ. The modulation of metabolic and lipidic alterations operated by the ΔF/TriAdj formulation in immunized animals when compared to unvaccinated, RSV-challenged mice supports the reduction of the immunopathology resulting from RSV infection. Some of the key lipids that were significantly altered include linoleic acid and alpha-linolenic acid, as well as the PUFA docosapentaenoic acid (DPA) and docosahexaenoic acid (DHA). These lipids have potential roles in pathophysiological events but were regulated by the triple-adjuvant vaccine formulation. Therefore, they are potential candidates as biomarkers for vaccine mechanism and efficacy. Future development in vaccine / adjuvant formulation for RSV may focus on the significantly altered lipids and metabolites found in this study. Also, novel adjuvant formulations targeting these lipids can be incorporated into subunit vaccines to improve vaccine efficacy.

We recognize that this work is limited by a small number of animals (N = 5) and the inherent differences found for mice as animal models of human conditions. Nevertheless, we achieved a broad characterization of the metabolic effects of RSV challenge and ΔF/TriAdj immunization that, when considered in conjunction with our previous immunological work, depicts a clear image of the immune response to the virus and provides crucial information for future developments and therapies.

5.5 Acknowledgments

We thank the animal care team at VIDO-InterVac (University of Saskatchewan, Saskatoon, SK, Canada) for the care and handling of the mice, as well as Dr. Sylvia van Drunen Littel-van den Hurk and Dr. Indranil Sarkar for the productive collaboration. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research, the Canada Research Chairs program, the Canadian Foundation for Innovations, Genome Canada and Alberta Innovates.

VI

Chapter VI: Lipidomics of Parkinson's Disease and Dementia[12,13]

6.1 Introduction

Parkinson's disease (PD) is a chronic disorder characterized by progressive motor deficits, *e.g.*, rest tremor, rigidity and bradykinesia.[33,207,208] The disease onset is associated with the formation of Lewy bodies, *i.e.*, abnormal aggregation of α-synuclein proteins, in the *substantia nigra pars compacta* region of the midbrain, leading to progressive loss of dopaminergic neurons and microgliosis. [33] As the pathology spreads to the cerebral cortex, cognitive decline and dementia may ensue in up to 80% of patients.[207,209] About 6.1 million people were living with Parkinson's disease worldwide in 2016 and the risk of mortality for affected individuals is approximately 1.5 times higher when compared to a healthy population. [210,211] However, there are no objective biological tests for definitive diagnosis of PD and even less is known about differential transition to dementia. Clinical diagnosis and pathological examination of the brain after death remains the gold standard, but symptoms become evident only when the dopaminergic neuronal loss reaches 50 to 80% of *pars compacta* neurons. Consequently, the condition is often undiagnosed in the early stages, when neuroprotective treatments would likely be most beneficial.[209,212] Recent systematic reviews note the accuracy of PD clinical diagnosis between 79.6 and 83.9%. Only 53% of early PD cases were correctly identified and the diagnosis accuracy at the first visit varied

[12] The experimental design and sample collection were performed in collaboration with Dr. Richard Camicioli (Neuroscience and Mental Health Institute and Department of Medicine - Neurology) and Dr. Roger A. Dixon (Neuroscience and Mental Health Institute and Department of Psychology, University of Alberta, Edmonton, Canada).

[13] Supporting figures for this chapter are available in Appendix D. Supporting tables are available with Dr. Liang Li.

between 26 and 88%. [213,214] Early and more accurate diagnosis of PD and prediction of differential transition to dementia through a biomarker panel could allow the use of effective neuroprotective treatments to prevent dopaminergic neuron death and the spread of the α-synuclein pathology.

The information currently available for PD onset and transition to dementia is fragmented. Hence, additional approaches are required to further advance our understanding of the pathology. A combination of different *omics* technologies to study Parkinson's disease and its progression to dementia may provide crucial information for biomarker discovery and pathogenesis research. Lipidomics is the large-scale and unbiased study of lipids in a given biological system, as well as their biochemical interaction with metabolites, proteins and other lipids. [11,215] Lipid concentrations are the products of complex and changing interactions; thus, unlike proteins and genes, they are not usually directly associated with a specific biological pathway. Recent findings showed that lipid metabolism may play a significant role in neurodegenerative disorders.[32,216] Multiple mutations and single-nucleotide polymorphisms in genes involved in lipid metabolism have been linked to PD, *e.g.*, glucosylceramidase beta (*GBA*)[217–219]; sterol regulatory element-binding transcription factor 1 (SREBF1)[220], diacylglycerol kinase (DGKQ)[221], N-acyl sphingosine amidohydrolase 1 (ASAH1)[222], phospholipase A2 (PLA2G6)[37] and sphingomyelin phosphodiesterase 1 (SMPD1).[222] Also, the lipid-binding α-synuclein protein may interact with phospholipids and fatty acids. [37] However, even though targeted analyses of lipids have been reported for brain tissue [223–225], plasma [226,227] and serum [228]; untargeted comprehensive lipidomics analyses for biomarker discovery is yet to be performed.

The metabolic profiling of serum samples from PD patients and healthy controls was previously evaluated by chemical isotope labeling.[229] Now, we performed a two-phase study of untargeted lipidomics for the same sample set. First, we hypothesized that there would be

alterations in the lipidome of PD patients compared to healthy controls. We compared lipidomic profiles from 43 healthy controls and 43 patients diagnosed with Parkinson's disease with no baseline symptoms of dementia. Second, we hypothesized the lipidome of baseline samples would discriminate PD patients who progressed to dementia (PDD) after a 3-year interval from those who remained cognitively stable (PDND). No sample collection was performed for the 3-year *post-hoc* assessment; therefore, we expected that later dementia could be predicted using the baseline lipidome up to three years before noticeable symptoms.

6.2 Experimental

6.2.1 *Participants*[14]

We recruited 43 patients aged between 65 and 82 years old with a clinical diagnosis of PD, no features of atypical Parkinsonism and no unstable health conditions compromising survival from the University of Alberta Movement Disorder Clinics, community neurologists and the Parkinson's Society of Alberta (Edmonton, AB, Canada) (Table VI-1). Parkinson's disease was diagnosed by experienced neurologists based on the presence of two out of three key symptoms (rest tremor, bradykinesia and rigidity), as described elsewhere. [230] The 43 healthy controls aged between 65 and 84 years old were recruited via seniors' centers, advertisements, clinics, and patient contacts. Participants were matched on age and sex. The University of Alberta Health Ethics Review Board approved this study and participants provided informed consent according to the Declaration of Helsinki.

[14] The recruitment and assessment of patients and healthy controls were performed in collaboration with Dr. Richard Camicioli and Dr. Roger A. Dixon (University of Alberta, Edmonton, AB, Canada).

Table VI-1. Summary of baseline demographic information for the samples used for lipidomics of Parkinson's disease and dementia.

	Healthy control	Parkinson's disease (PD)	p	PD with no dementia (PDND)	PD with incipient dementia (PDD)	p
N	43	43	N/A	27	16	N/A
Age (years)	71.5 (±5.0)	70.7 (±4.1)	0.44	69.6 (±3.5)	72.6 (±4.5)	0.028
Gender (F/M)	19/24	19/24	1.0	12/15	7/9	0.97
Levodopa equivalents (mg)	N/A	644.0 (±360.1)	N/A	611.8 (±392.9)	703.8 (±293.4)	0.41
Levodopa duration (years)	N/A	4.7 (±4.2)	N/A	3.5 (±3.6)	6.7 (±4.3)	0.020
PD duration (years)	N/A	8.4 (±4.7)	N/A	7.6 (±4.3)	9.6 (±5.1)	0.21
UPDRS - part III	2.2 (±3.2)	16.1 (±7.9)	3.81×10^{-15}	16.7 (±8.1)	15.2 (±7.8)	0.56
B12 (ng/mL)	388.5 (±198.7)	293.25 (±112.8)	0.0080	295.7 (±92.2)	289.1 (±144.5)	0.87

Notes: values are displayed as average for each group, accompanied by one standard deviation within brackets; $p < 0.05$ was deemed significant, and calculated for two-tailed, unpaired Student's t-tests; Levodopa equivalents – contribution of each PD drug taken by the patient that has a similar effect of Levodopa; UPDRS - part III: motor deficit score for the Unified Parkinson's Disease Rating Scale.

Participants performed standardized assessments at 0, 18 and 36 months, including motor deficit and cognitive function (Table VI-2 and Table VI-3). However, blood samples were collected only at the first visit. No patients or controls displayed cognitive deficits at the time of sample collection. Volunteers were reassessed after 18 and 36 months for cognitive and functional decline by the study neurologist (Dr. Richard Camicioli, University of Alberta, Canada) and a research assistant through interviews with patients and caregivers, as well as cognitive and

functional assessments. Each visit (0, 18 and 36 months) included assessment of the standardized Mini-Mental Status Examination (MMSE), Frontal Assessment Battery (FAB), Dementia Rating Scale (DRS-II), Clinical Dementia Rating (CDR), Geriatric Depression Scale (GDS), Blessed-Orientation-Concentration Test, Unified Parkinson's Disease Rating Scale (UPDRS), Hoehn and Yahr Rating Scale (HY), and Cumulative Illness Rating Scale (CIRS). [229–231] Demographic information (age, sex, education), vital signs, levodopa-equivalent dosages and other medications were also recorded (Table VI-1, Table VI-2 and Table VI-3). Patients that did not meet criteria for PD or that developed other conditions in the follow-up visits were excluded before the analysis phase, as well as any participants with baseline dementia (CDR <1.0; MMSE <23), stroke, cerebellar signs, early autonomic dysfunction, supranuclear gaze Palsy, conditions that might affect cognition (recent myocardial infarction, angina, congestive heart failure), severe depression, psychosis, cognitive sequelae, alcoholism, drug abuse, or atypical Parkinsonism. No volunteers were taking cognitive enhancing drugs during the study. [229–232]

Table VI-2. Demographic information acquired at 0 (baseline) and 36 months after sample collection for the study participants. PD: Parkinson's disease; PDD: Parkinson's disease with dementia; PDND: Parkinson's disease with no dementia diagnosis; NA: not available; mo.: months.

Table VI-2

ID	PD	Dementia	Dementia score (36 mo.)	Sex	Age (baseline)	Education (years)	Apolipoprotein E4 Count	B12 (ng/mL, baseline)	Folate (ng/mL, baseline)	Homocysteine (µmol/L, baseline)
2	PD	PDND	1	M	65.7	10	0	327	883	12.2
3	PD	PDND	1	M	73.7	12	0	333	672	25.1
4	PD	PDND	2	M	65.6	20	0	249	789	11.8
6	PD	PDD	3	M	70.7	13	0	214	505	19.0
7	PD	PDND	2	M	65.1	14	0	289	936	13.3
8	PD	PDD	4	F	66.2	15	0	350	1237	11.4
9	PD	PDD	3	M	75.9	12	1	551	912	11.7
10	Control	Control	1	F	67	16	0	1050	990	11.1
12	PD	PDND	1	F	68.9	25	0	466	1180	5.0
13	PD	PDND	1	F	73	16	0	281	971	11.1
14	Control	Control	1	F	65	20	NA	309	725	9.3
15	PD	PDD	4	M	72.3	12	0	137	706	18.8
17	PD	PDND	1	M	72.2	18	0	309	620	15.3
18	Control	Control	1	M	66.2	18	1	424	852	9.5
19	PD	PDD	4	M	75.2	9	0	289	1025	9.4
20	PD	PDND	1	M	67.5	17	0	389	945	9.4
21	PD	PDND	2	F	69.2	14	1	592	859	19.0
22	PD	PDND	1	M	70.2	16	0	331	810	11.2
23	PD	PDND	1	M	72.1	16	0	384	717	12.3
24	PD	PDD	3	F	67.8	12	0	172	902	12.8
25	PD	PDD	4	M	69.5	15	2	663	1012	10.9

Table VI-2

ID	PD	Dementia	Dementia score (36 mo.)	Sex	Age (baseline)	Education (years)	Apolipoprotein E4 Count	B12 (ng/mL, baseline)	Folate (ng/mL, baseline)	Homocysteine (µmol/L, baseline)
26	PD	PDND	1	M	69.6	13	0	190	493	13.0
27	PD	PDD	4	F	75.1	16	0	171	800	12.8
28	PD	PDD	3	M	69.7	16	0	397	762	19.2
29	PD	PDND	1	F	68.7	12	1	251	792	15.8
30	PD	PDND	1	M	66	16	0	235	927	10.1
32	PD	PDD	3	M	70.8	12	0	253	723	10.1
33	Control	Control	1	M	67.6	21	0	494	945	9.2
34	Control	Control	1	F	73.4	16	1	240	959	8.9
35	Control	Control	1	M	68	14	1	376	872	10.6
36	PD	PDD	4	F	82.5	12	0	257	1349	14.2
37	Control	Control	1	F	72.1	13	0	176	723	15.0
38	Control	Control	1	M	72.8	13	0	235	799	11.7
40	Control	Control	1	M	68.4	18	0	598	844	7.2
42	PD	PDND	1	M	65.2	17	0	208	643	16.0
43	PD	PDD	4	F	79.5	16	1	293	1082	13.1
45	PD	PDND	1	M	77.1	10	0	173	705	16.8
46	Control	Control	1	F	73.1	14	0	250	824	12.2
47	PD	PDND	1	F	66.2	12	0	245	658	10.7
48	Control	Control	1	M	79.1	16	0	435	754	8.7
49	Control	Control	1	F	76.2	14	0	186	942	14.0
50	Control	Control	1	M	70.4	16	0	281	613	14.0
51	Control	Control	1	F	78.5	12	0	682	1388	12.8
52	PD	PDND	2	F	66.2	11	0	315	877	9.4
53	Control	Control	1	M	71.6	9	0	536	862	8.6
54	Control	Control	1	M	65.3	8	1	268	666	13.6
55	Control	Control	1	M	77.6	11	1	554	760	9.1
56	Control	Control	1	F	72.1	9	1	472	931	6.1

Table VI-2

ID	PD	Dementia	Dementia score (36 mo.)	Sex	Age (baseline)	Education (years)	Apolipoprotein E4 Count	B12 (ng/mL, baseline)	Folate (ng/mL, baseline)	Homocysteine (µmol/L, baseline)
57	Control	Control	1	M	69.1	16	0	322	621	6.0
59	Control	Control	1	F	66.6	23	0	601	940	10.5
60	Control	Control	1	F	68.7	15	0	497	1076	8.0
63	PD	PDND	1	F	77.3	17	0	283	823	10.9
64	PD	PDD	4	M	67.9	16	0	300	750	17.4
65	Control	Control	1	F	67.8	13	1	295	727	9.6
66	Control	Control	2	F	83.3	17	0	1023	684	14.8
67	Control	Control	1	F	68.3	19	0	399	983	5.6
68	Control	Control	1	F	74.7	15	1	393	982	8.4
69	Control	Control	1	F	68.5	10	0	467	715	9.6
70	Control	Control	1	F	69	11	0	132	541	12.9
71	PD	PDD	3	M	69.7	12	1	150	835	18.2
72	PD	PDD	4	F	73.8	14	0	224	817	15.7
73	Control	Control	1	F	71.3	18	0	469	1083	10.0
74	Control	Control	1	M	66	19	0	404	931	12.5
75	Control	Control	1	F	71.5	21	0	338	893	9.6
76	Control	Control	2	M	81.5	18	0	256	816	6.6
77	PD	PDND	1	M	71.6	12	0	194	539	18.5
78	Control	Control	1	M	69.6	16	0	312	859	11.5
79	Control	Control	1	M	66.3	12	0	429	580	11.6
81	PD	PDND	2	F	71.4	11	0	328	715	6.1
82	PD	PDND	1	F	69.7	15	0	214	792	15.5
83	Control	Control	1	F	81.1	14	0	514	1177	8.9
84	Control	Control	1	M	72.9	16	0	374	1323	11.9
86	Control	Control	1	M	70.5	19	0	168	1328	10.2
87	PD	PDND	1	M	69.8	18	0	371	735	12.2
89	PD	PDD	4	F	75.5	14	0	205	1240	16.6

Table VI-2

ID	PD	Dementia	Dementia score (36 mo.)	Sex	Age (baseline)	Education (years)	Apolipoprotein E4 Count	B12 (ng/mL, baseline)	Folate (ng/mL, baseline)	Homocysteine (µmol/L, baseline)
91	PD	PDND	1	F	68.6	16	0	239	903	10.5
92	PD	PDND	2	F	75.1	12	0	196	885	15.3
93	Control	Control	1	M	71.8	15	0	151	685	14.1
94	Control	Control	2	M	83.5	15	1	190	1752	15.8
95	Control	Control	1	M	65.1	14	1	401	1120	8.0
96	PD	PDND	2	F	66.6	14	0	325	1265	11.1
98	PD	PDND	1	M	66.4	14	0	267	439	15.6
99	Control	Control	1	M	73.5	11	0	227	529	12.8
100	Control	Control	1	M	68.5	16	0	168	1007	9.6
103	Control	Control	1	M	70.9	16	0	369	764	8.8
104	Control	Control	1	M	68.6	12	0	242	715	8.7

Table VI-3. Demographic information acquired at 0 (baseline), 18 and 36 months (mo.) after sample collection for the study participants (cont.). NA: not available; MMSE: Mini-Mental Status Examination; HY: Hoehn and Yahr Rating Scale; UPDRS: Unified Parkinson's Disease Rating Scale; mo.: months

Table VI-3

ID	Dementia	MMSE			HY			UPDRS part 3 (motor)			UPDRS total score			PD duration (years)	Levodopa duration (years)	Levodopa equivalent dosage (mg)		
		0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	0 mo.	0 mo.	18 mo.	36 mo.
002	PDND	30	28	28	2.0	2.0	1.5	12	12	11	15	14	12	5.6	0.11	225.0	900.0	450.0
003	PDND	29	30	29	3.0	2.0	2.0	31	33	21	42	41	33	10.8	3.80	675.0	937.5	875.0
004	PDND	29	26	27	2.0	1.0	2.0	17	5	9	36	14	16	4.0	0.00	100.5		
006	PDD	30	29	29	2.0	2.5	2.0	22	31	30	38	44	43	5.7	3.71	175.0	175.0	250.5
007	PDND	30	28	26	1.5	1.0	1.0	16	13	13	23	26	19	3.2	1.65	600.0	651.0	1095.0
008	PDD	29	29	28	2.0	1.5	2.0	12	11	8	22	13	19	5.7	2.15	617.0	1668.0	483.6
009	PDD	25	24	29	2.0	2.0	2.0	8	6	11	12	20	20	9.4	8.92	675.0	575.0	0.0
010	Control	27	30	30	0.0	0.0	0.0	0	0	0	0	0	0			0.0	0.0	0.0
012	PDND	29	27	27	2.5	3.0	4.0	16	15	26	27	25	43	5.8	1.24	750.0	1150.0	1875.0
013	PDND	30	29	28	3.0	2.0	3.0	11	9	20	14	12	23	3.4	0.26	550.0	700.0	700.0
014	Control	29	28	29	0.0	0.0	0.0	7	7	6	10	9	12	0.0	0.00	0.0	0.0	0.0
015	PDD	27	24	27	2.0	2.0	2.0	20	13	14	35	29	26	7.9	6.89	625.5	700.5	850.5
017	PDND	29	29	30	2.0	2.0	2.0	17	20	21	26	31	32	2.5	0.11	300.0	300.0	300.0
018	Control	30	29	29	0.0	0.0	0.0	3	3	3	9	9	9	0.0	0.00	0.0	0.0	0.0
019	PDD	28	29	30	2.0	2.0	2.0	15	12	27	32	25	47	14.0	11.96	500.0	437.5	550.0
020	PDND	29	28	28	2.0	2.0	2.0	22	21	21	31	25	29	4.6	2.40	601.0	601.0	601.0
021	PDND	30	27	29	3.0	2.5	2.5	14	11	15	21	19	23	11.0	10.02	800.0	600.0	600.0
022	PDND	29	27	28	2.0	2.0	2.0	29	27	19	39	35	26	5.6	0.44	225.0	750.0	1125.0
023	PDND	30	28	27	2.0	2.0	2.0	17	18	14	24	27	25	3.2	1.78	450.0	750.0	1000.0
024	PDD	29	29	26	2.0	2.0	2.0	3	2	4	6	5	10	13.1	2.87	825.1	1100.1	800.0

Table VI-3

ID	Dementia	MMSE			HY			UPDRS part 3 (motor)			UPDRS total score			PD duration (years)	Levodopa duration (years)	Levodopa equivalent dosage (mg)		
		0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	0 mo.	0 mo.	18 mo.	36 mo.
025	PDD	26	20	16	2.0	3.0	4.0	17	28	45	31	49	79	5.1	4.06	450.0	375.0	400.0
026	PDND	30	30	29	2.0	2.0	2.0	16	10	11	29	27	27	7.1	2.14	500.0	600.0	800.0
027	PDD	NA	NA	NA	3.0	NA	NA	25	NA	NA	NA	NA	NA	17.2	11.88	NA	NA	NA
028	PDD	27	26	26	2.0	2.5	2.5	7	17	16	20	28	28	12.6	9.31	801.0	913.5	1026.0
029	PDND	28	30	29	2.0	1.0	2.0	10	16	20	23	26	33	9.3	2.27	450.0	450.0	450.0
030	PDND	29	28	29	2.0	2.0	2.0	12	14	14	22	25	24	3.3	0.00	301.5	526.5	526.5
032	PDD	29	29	26	2.0	3.0	2.5	8	15	33	12	21	45	15.4	12.35	1367.0	1875.0	824.1
033	Control	28	28	30	0.0	0.0	0.0	1	4	0	2	5	4	0.0	0.00	0.0	0.0	0.0
034	Control	27	29	28	0.0	0.0	0.0	2	3	3	2	5	4	0.0	0.00	0.0	0.0	0.0
035	Control	28	29	27	0.0	0.0	0.0	0	0	2	0	0	2	0.0	0.00	0.0	0.0	0.0
036	PDD	27	28	21	2.5	2.0	5.0	11	14	31	23	28	69	14.4	5.38	975.0	1112.8	0.0
037	Control	30	29	29	0.0	0.0	0.0	1	0	1	1	1	2	0.0	0.00	0.0	0.0	0.0
038	Control	28	28	29	0.0	0.0	0.0	0	1	0	0	2	2	0.0	0.00	0.0	0.0	0.0
040	Control	29	29	30	0.0	0.0	0.0	1	0	4	1	0	9	0.0	0.00	0.0	0.0	0.0
042	PDND	27	28	26	3.0	3.0	3.0	22	19	23	24	28	29	18.4	13.42	2100.0	1376.3	1326.0
043	PDD	26	27	26	2.5	2.5	2.5	24	30	29	42	50	43	3.5	2.46	400.0	400.0	750.0
045	PDND	25	26	23	2.0	2.0	2.5	14	34	42	24	46	48	14.5	8.50	1125.0	1562.5	0.0
046	Control	29	30	28	0.0	0.0	0.0	0	1	2	1	2	4	0.0	0.00	0.0	0.0	0.0
047	PDND	28	30	28	1.0	1.0	2.0	6	8	5	6	10	7	5.2	0.48	600.0	750.0	1000.0
048	Control	28	28	25	0.0	0.0	0.0	2	0	0	2	1	2	0.0	0.00	0.0	0.0	0.0
049	Control	29	28	27	0.0	0.0	0.0	0	0	1	0	0	3	0.0	0.00	0.0	0.0	0.0
050	Control	30	28	29	0.0	0.0	0.0	0	0	2	0	1	8	0.0	0.00	0.0	0.0	0.0
051	Control	NA	NA	NA	0.0	NA	NA	12	NA	NA	NA	NA	NA	0.0	0.00	NA	NA	NA
052	PDND	29	26	28	2.0	2.0	2.0	11	0	13	20	8	23	6.5	3.50	450.0	450.0	450.0
053	Control	29	29	28	0.0	0.0	0.0	3	5	2	3	5	2	0.0	0.00	0.0	0.0	0.0
054	Control	27	25	26	0.0	0.0	0.0	2	0	2	3	1	4	0.0	0.00	0.0	0.0	0.0

Table VI-3

ID	Dementia	MMSE			HY			UPDRS part 3 (motor)			UPDRS total score			PD duration (years)	Levodopa duration (years)	Levodopa equivalent dosage (mg)		
		0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	0 mo.	0 mo.	18 mo.	36 mo.
055	Control	25	26	29	0.0	0.0	0.0	3	4	4	7	5	7	0.0	0.00	0.0	0.0	0.0
056	Control	29	26	30	0.0	0.0	0.0	0	1	0	0	1	1	0.0	0.00	0.0	0.0	0.0
057	Control	28	28	29	0.0	0.0	0.0	0	0	1	0	0	1	0.0	0.00	0.0	0.0	0.0
059	Control	30	28	28	0.0	0.0	0.0	1	2	0	2	4	2	0.0	0.00	0.0	0.0	0.0
060	Control	30	28	29	0.0	0.0	0.0	4	3	6	4	7	8	0.0	0.00	0.0	0.0	0.0
063	PDND	26	29	27	2.0	2.5	2.0	33	27	36	47	37	53	14.5	2.52	100.0	200.0	100.0
064	PDD	NA	NA	NA	2.0	NA	NA	12	NA	NA	NA	NA	NA	0.7	0.65	NA	NA	NA
065	Control	30	29	29	0.0	0.0	0.0	0	0	0	1	0	0	0.0	0.00	0.0	0.0	0.0
066	Control	29	29	26	0.0	0.0	0.0	15	9	15	15	12	26	0.0	0.00	0.0	0.0	0.0
067	Control	30	28	28	0.0	0.0	0.0	0	0	0	1	1	0	0.0	0.00	0.0	0.0	0.0
068	Control	28	29	24	0.0	0.0	0.0	6	8	8	6	10	8	0.0	0.00	0.0	0.0	0.0
069	Control	29	27	29	0.0	0.0	0.0	2	0	3	3	0	4	0.0	0.00	0.0	0.0	0.0
070	Control	28	27	28	0.0	0.0	0.0	0	0	1	0	1	2	0.0	0.00	0.0	0.0	0.0
071	PDD	28	25	23	2.5	2.5	2.5	11	19	23	20	26	30	12.8	12.79	817.0	750.0	750.0
072	PDD	23	22	13	3.0	2.5	4.0	16	11	36	20	14	61	3.0	1.51	650.0	1312.5	225.0
073	Control	NA	NA	NA	0.0	NA	NA	3	NA	NA	NA	NA	NA	0.0	0.00	NA	NA	NA
074	Control	30	28	29	0.0	0.0	0.0	1	0	2	3	1	3	0.0	0.00	0.0	0.0	0.0
075	Control	29	29	29	0.0	0.0	0.0	0	0	1	1	0	1	0.0	0.00	0.0	0.0	0.0
076	Control	29	27	28	0.0	0.0	0.0	3	2	5	4	3	8	0.0	0.00	0.0	0.0	0.0
077	PDND	NA	NA	NA	3.0	NA	NA	39	NA	NA	NA	NA	NA	4.9	2.44	NA	NA	NA
078	Control	27	25	27	0.0	0.0	0.0	2	3	3	2	4	3	0.0	0.00	0.0	0.0	0.0
079	Control	29	27	26	0.0	0.0	0.0	0	2	0	3	5	3	0.0	0.00	0.0	0.0	0.0
081	PDND	28	27	27	2.0	2.0	2.5	19	25	27	30	41	43	5.9	5.94	601.5	601.5	701.5
082	PDND	29	28	27	1.5	1.0	1.5	12	7	9	29	21	25	9.8	3.74	750.0	600.0	501.0
083	Control	25	24	26	0.0	0.0	0.0	6	3	4	10	3	7	0.0	0.00	0.0	0.0	0.0
084	Control	NA	NA	NA	0.0	NA	NA	0	NA	NA	NA	NA	NA	0.0	0.00	NA	NA	NA

Table VI-3

ID	Dementia	MMSE			HY			UPDRS part 3 (motor)			UPDRS total score			PD duration (years)	Levodopa duration (years)	Levodopa equivalent dosage (mg)		
		0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	18 mo.	36 mo.	0 mo.	0 mo.	0 mo.	18 mo.	36 mo.
086	Control	30	29	28	0.0	0.0	0.0	1	1	1	2	2	2	0.0	0.00	0.0	0.0	0.0
087	PDND	30	30	30	1.0	1.0	2.0	11	9	19	16	13	25	14.1	11.47	876.5	1001.5	1001.5
089	PDD	29	27	28	3.0	3.0	4.0	32	27	40	40	37	61	13.1	10.09	975.0	1800.0	1800.0
091	PDND	29	27	28	1.0	1.0	2.0	8	10	6	17	19	13	5.0	2.96	450.0	550.0	750.0
092	PDND	28	28	28	2.0	1.0	2.0	12	6	10	16	10	14	8.2	5.17	700.0	501.0	701.0
093	Control	30	26	28	0.0	0.0	0.0	3	0	0	4	3	3	0.0	0.00	0.0	0.0	0.0
094	Control	24	24	25	0.0	0.0	0.0	3	3	5	6	6	9	0.0	0.00	0.0	0.0	0.0
095	Control	30	30	27	0.0	0.0	0.0	0	0	0	1	2	1	0.0	0.00	0.0	0.0	0.0
096	PDND	30	28	30	1.0	2.0	2.0	8	10	9	10	16	13	13.9	2.14	751.5	751.5	751.5
098	PDND	30	28	0	2.0	2.0	0.0	15	16	0	17	19	0	6.1	6.12	875.0	875.0	0.0
099	Control	29	28	26	0.0	0.0	0.0	0	2	1	0	3	2	0.0	0.00	0.0	0.0	0.0
100	Control	30	30	28	0.0	0.0	0.0	0	0	0	1	1	1	0.0	0.00	0.0	0.0	0.0
103	Control	29	30	30	0.0	0.0	0.0	5	4	7	8	8	11	0.0	0.00	0.0	0.0	0.0
104	Control	29	29	30	0.0	0.0	0.0	1	1	0	2	1	0	0.0	0.00	0.0	0.0	0.0

Parkinson's disease dementia (PDD) was defined at the 36-month visit through a cognitive decline in at least two domains (a 3-point change on the MMSE when compared to baseline, 6-point change on the DRS-II or impairment in two or more domains on the CDR) during the 3-year study, along with functional impairment due to cognitive decline. Dementia diagnosis was confirmed by the study neurologist (Dr. Richard Camicioli, University of Alberta) through interviews, cognitive evaluations and functional assessments. Patients that declined to below age- and education-based cut-offs were included in the dementia subgroup, but volunteers without functional impairment or clear cognitive decline were grouped with the non-impaired subjects.

6.2.2 *Instrumentation*

Samples were analyzed by a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific, Waltham, MA, USA) employing a Waters Acquity BEH C18 column (5 cm × 2.1 mm with 1.7 μm particles; Waters Corporation, Milford, MA, USA) and Maxis II QTOF mass spectrometer instrument (Bruker Daltonics, Billerica, MA, USA) equipped with an electrospray ionization (ESI) source.

6.2.3 *Sample preparation and analysis*

Blood serum samples were prepared according to a modified Folch liquid-liquid extraction protocol with a combination of dichloromethane, methanol and water. [62,63] The employed chemicals and reagents were described in Chapter II (*2.2.1. Chemicals and reagents*, p. 42). An internal standard mixture composed of 14 deuterated lipids was added to samples before extraction for normalization of ion suppression and small differences that may occur during sample handling

(Table II-1, p. 45). Different sample volumes and final dilutions were employed for detection under positive and negative electrospray ionization due to the inherent characteristics of each polarity. For positive ionization, aliquots of 6.0 µL of human blood serum were vortexed with 2.0 µL of the internal standard mixture, 54.0 µL of methanol, 111.0 µL of dichloromethane and 28.0 µL of water. After a 10 min equilibration at room temperature, samples were centrifuged for 10 min at 10,000 rpm and 4°C. The bottom organic layer (95% of the dichloromethane volume, 105.4 µL) was evaporated to dryness using a SpeedVac for 30 min. The residue was immediately resuspended in 12.0 µL of 1:1 mobile phase A (MPA)/mobile phase B (MPB) and diluted with 108.0 µL of MPA (20-fold dilution). For negative ionization, 30.0 µL aliquots of blood serum and 10.0 µL of the internal standard mixture were extracted with 46.0 µL of methanol, 111.0 µL of dichloromethane and 4 µL of water, followed by resuspension in 3 µL of 1:1 MPA/MPB and dilution with 27.0 µL of MPA (no dilution). Samples were kept under 4°C for a maximum of 2 days before the injection. Experimental duplicates were performed, and samples were randomized for preparation and analysis. Quality control (QC) samples were prepared with a pool of human serum obtained from 100 healthy individuals.

The extractions and injections were performed in 250 µL polypropylene inserts (Canadian Life Science, Peterborough, ON, Canada) placed inside 2 mL amber vials (Agilent Technologies, Santa Clara, CA, USA) with PTFE/silicone septa caps (Waters Corporation, Milford, MA, USA). We recognize that the use of plastic ware with organic solvents is not ideal, although much more practical and affordable than glassware. For this study, all samples were prepared using the same type and batch of inserts, septa, vial caps and pipet tips, thus minimizing the effects of plasticizers and other contaminants upon the statistical analysis. Blank extractions of water instead of the sample were performed with each sample batch to ensure quality control. A detailed discussion on

contamination sources and plastic/glassware for lipidomics is available in Chapter II (*2.3.3. Evaluation of contamination sources*, p. 63)

Reversed-phase chromatography was performed by UHPLC-ESI-QToF-MS. Analysis conditions were: mobile phase A (MPA) – 10 mM ammonium formate in 50:40:10 acetonitrile/ methanol/ water (v/v/v); mobile phase B (MPB) – 10 mM ammonium formate in 95:5 isopropanol/ water (v/v); 250 µL/min; column temperature of 40°C; injection volume of 5.0 µL for positive ionization and 9.0 µL for negative ionization; and 22 min gradient (0 min – 5% MPB; 1.8 min – 5% MPB; 8.5 min – 30% MPB; 18 min – 95% MPB; 22 min – 95% MPB) followed by 10 min of equilibration (0 min – 95% MPB; 3 min – 95% MPB; 4 min – 5% MPB; 10 min – 95% MPB). A ultra-high-resolution Maxis II QTOF mass spectrometer instrument equipped with an ESI source was used for detection under positive and negative ionization (capillary voltage of 4500 V; endplate offset of 500 V, nebulizer gas pressure of 1.0 bar, dry gas flow rate of 8.0 L/min; dry temperature of 230°C; spectra acquisition rate of 1 Hz, m/z range of 150 to 1500 Da). A 1.5 min segment for mass re-calibration was inserted at the beginning of each chromatogram during which 1.0 mmol/L sodium formate calibrant solution in 1:1 isopropanol/ water (v/v) was infused into the ion source using a peristaltic pump. Quality control (QC) samples, *i.e.,* the extract of a pool of serum samples obtained from 100 healthy individuals, were injected six times before the sample sequence to stabilize retention time and MS signal, and then re-injected after every ten sample injections to account for technical variation, totaling 18 QC replicates.

6.2.4 *Data processing*

The detected features were aligned using MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA) with minimum intensity cut-off of 5000 cts for positive and 2000 cts for negative

ionization; minimum peak length of 6 spectra; retention time tolerance of 15 s; m/z tolerance of 5.0 mDa; missing value substitution by recursive extraction for features detected in at least 10% of injections with a minimum peak length of 6 spectra; filtering by features detected in at least 80% of injections for each group (control, Parkinson's disease, or QC); and adduct correlation cut-off of 0.8. Features detected under positive and negative ionization were merged into a unique list with m/z tolerance of 5.0 mDa and retention time tolerance of 15 s.

6.2.5 *Identification*

A comprehensive classification system for lipids was proposed in 2005 and updated in 2009 by Fahy *et al*. under the leadership of the International Lipid Classification and Nomenclature Committee and has been adopted for the present work.[16,17] Abbreviations to lipid classes and nomenclature followed the LipidMaps database (https://www.lipidmaps.org; database nomenclature from 2018) and the MS-DIAL LipidBlast spectral library (Table V-1).[16,17,20,73,78]

MS/MS spectra were acquired for identification using a pool of all samples with collision energies between 20 and 60 eV. Positive MS/MS identification was performed with the MS-Dial LipidBlast MS/MS library, Bruker Human Metabolome Database (HMDB) Metabolite Library 2.0 and MassBank of North America (MoNA) LC-MS/MS libraries, in combination with MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA).[73,78–83,109] Lipids with MS/MS score higher than 50, combined with precursor mass error smaller than 5.0 mDa and mSigma (isotopic pattern match factor) smaller than 100, were considered as positively identified (Supp. Table 19).

Features not identified by MS/MS were inputted in the Lipid Maps database for putative mass match (tolerance of 5.0 mDa).[75] Lipids can have a high number of isomers and isobars, particularly for mass-based identification. A mass search for lipids can define (1) lipid class, (2)

the sum of components, *i.e.,* total number of carbons in fatty acyl/alkyl chains and double bonds, and (3) the bond type. The individual compositions of specific fatty acyl/alkyl chains require MS/MS match, and cannot be determined by a mass search. [20] Furthermore, isomeric/isobaric overlap for different lipid classes and adducts may occur. Each detected feature can be m/z matched to a high number of isomeric and isobaric lipids that may belong to different classes. Hence, we employed a 5-tier filtering and scoring approach to select the best identification possibilities, as described in Chapter II (*2.2.7. Putative identification of lipids*, p. 50). The scores for each ranking tier were summed, and the top choice for each feature was selected by the smallest m/z error and total score. The top choice was used to determine lipid class for normalization, but other isomeric or isobaric possibilities that passed the exclusion filters for retention time range and adducts were kept, ordered by score values (Supp. Table 20).

6.2.6 *Normalization and statistics*

Lipids can suffer strong ion suppression in ESI due to their characteristic high hydrophobicity and amphiphilic behavior. We adapted a widely used approach for normalization of lipidomics data to correct ion suppression, as well as small variations that may occur during sample preparation, by using a set of 14 deuterated internal standards belonging to different lipid classes (Table II-1, p. 45). The putatively identified lipids were matched to one of the 14 internal standards according to lipid class similarity and expected retention time range for each class. Normalized intensity ratios, *i.e.*, the intensity of each lipid divided by the intensity of the matched internal standard, were calculated for normalization. Normalized intensity ratios were auto-scaled, normalized for the median value for all samples and filtered by a relative standard deviation (RSD) smaller than 30% for QC injections (2633 lipids) for statistical analysis on MetaboAnalyst 4.0

(www.metaboanalyst.ca), *viz.* PCA, PLS-DA, orthogonal partial least square discriminant analysis (OPLS-DA) and Random Forest. [153] Lipids not detected for QCs, but found in healthy control or PD samples, were also included. Volcano plots, *i.e.*, fold change (FC) *versus* p adjusted for false discovery rate (p), were used to select significantly altered lipids with FC ≥1.2 or ≤0.84 and p <0.05 (two-tailed, non-parametric test for unequal variances). Biomarker panels were selected by classification performance using Random Forest and evaluated through Receiver Operating Characteristic (ROC) curves.

6.2.7 *Confounding factors*

Confounding factors are characteristics that may cause biological differences within the studied population, but not expected to be specifically related to the disease status. They may lead to severe false discovery and bias if not statistically evaluated. We examined the effect of age, sex and B12 serum levels as potential confounders for the separation between control individuals and PD patients. For PD patients without dementia (PDND) and with incipient dementia (PDD), we evaluated age, sex, time since PD diagnosis, medication dosages (Levodopa equivalents), treatment period and motor deficit score (UPDRS part 3). Healthy control individuals (n=43) were age and sex-matched to PD patients (n=43) (Table VI-1). Similarly, the PDND (n=27) and PDD subgroups (n=16) were matched on sex, but the PDND subgroup was significantly younger than PDD (p of 0.03, Table VI-1). B12 serum levels were significantly higher for healthy controls when compared to PD patients (p of 0.008), but not significantly altered for the PDND/PDD comparison (p of 0.9, Table VI-1). PDND and PDD patients were not significantly different regarding disease duration (time since PD diagnosis, p of 0.2) or motor deficit score (UPDRS – part III, p of 0.6, Table VI-1). Furthermore, all PD patients were treated with varying doses of Levodopa or

equivalent drugs before and throughout the study period, according to medical instructions. Medication dosages were therapeutically adjusted during the 3-year evaluation due to disease progression, but treatment courses were not altered on account of participation in this study. The Levodopa equivalent dosages at baseline sample collection were not significantly different for PDND and PDD patients (p of 0.4), even though the PDD patients received medication for longer periods (p of 0.02, Table VI-1).

The seven potentially confounding factors were evaluated by separating groups (healthy control and PD; or PDND and PDD) into two subgroups at the median value for age, sex, levodopa equivalent dosage, the period of treatment, time since PD diagnosis, B12 serum levels and motor deficit score (UPDRS - part III, Table VI-1). We employed statistical analysis through Volcano plots, two-way ANOVA, PCA, PLS-DA and ROC curves to gauge the effect of each factor upon the PD diagnosis or dementia classification. The evaluation of confounding factors followed a four-step procedure:

1. Evaluation of the overall effect of each potentially confounding factor on the lipidome, regardless of disease status, by Volcano plots, PCA and PLS-DA.
2. Evaluation of the potential interference or interaction of each confounding factor and the control/PD or PDND/PDD classification by two-way ANOVA. Each sample was given two labels: (1) disease status (control/PD or PDND/PDD), and (2) above or below the median value for the evaluated confounder. Two-way ANOVA interaction models were calculated. For the null hypothesis, we considered that there was an additive relationship between disease status (control/PD, or PDND/PDD) and the confounder, *e.g.*, null hypothesis = the mean intensity ratios for control individuals are equal for males and females. For the

alternative hypothesis, we considered that the effect of changes due to disease status depended on the confounding factor. The null hypothesis was rejected for interaction $p < 0.05$. If there was no significant interaction ($p > 0.05$), the effects of changes in disease status were not significantly different at each level of the evaluated confounder and the 1-way ANOVA model was adequate.

3. Evaluation of the performance of the proposed biomarker panels to classify samples as below or above the median for each factor. ROC curves were constructed using the two biomarker panels (control/PD or PDND/PDD) but applied to classify the samples as "below the median" or "above the median". Uni- and multivariate AUC values were compared to the ones found for the control/PD or PDND/PDD comparisons. When the performance for median-split classification surpassed the performance for disease classification (higher uni- and multivariate AUC values), the biomarker candidate was further investigated.
4. Evaluation of the performance of the proposed biomarker panels to classify the median-split samples as control/PD or PDND/PDD. Each median-split subgroup of samples was employed for the classification of samples as control/PD or PDND/PDD by Random Forest, using the proposed biomarker panels. Confounding was indicated by multivariate AUC values outside the 95% confidence interval (CI) for the ROC curve with the complete sample cohort (control/PD or PDND/PDD).

We recognize that the current analyses of confounding factors were thorough but not exhaustive. Yet, confounders cannot be easily controlled, especially if the goal is to achieve a statistically representative number of human patients. Restricting the patient cohort to prevent

minor effects would substantially decrease the sample set available for the study. The only option to completely control confounders is by using cellular or animal models; however, as often found in clinical research, models do not always reflect the same biological and biochemical alterations as studies with human patients.

6.3 Results

6.3.1. Lipidomic profile generated by UHPLC-QToF-MS

The healthy control, Parkinson's disease and QC samples resulted in the detection of 6032 features for positive ionization (82.1% with RSD ≤30% for QC injections) and 1001 for negative ionization (84.6% with RSD ≤30% for QC injections). Examples of chromatograms obtained for healthy individuals, PD patients, QC injections and the internal standard mixture are displayed in Figure VI-1. Positive and negative ions were merged to a list of 7043 features, with 318 positive identifications by MS/MS (Supp. Table 19). Furthermore, 2768 features were putatively mass-matched to lipids with a mass tolerance of 5.0 mDa (Supp. Table 20). All 14 deuterated internal standards (Table II-1, p. 45) were detected and identified with an m/z error ≤3.7 ppm (1.3 mDa, Figure VI-1G and H). The PCA plot in Figure VI-2 displays all QC injections in a tightly clustered group, as well as duplicate experiments for each sample, indicating good technical reproducibility.

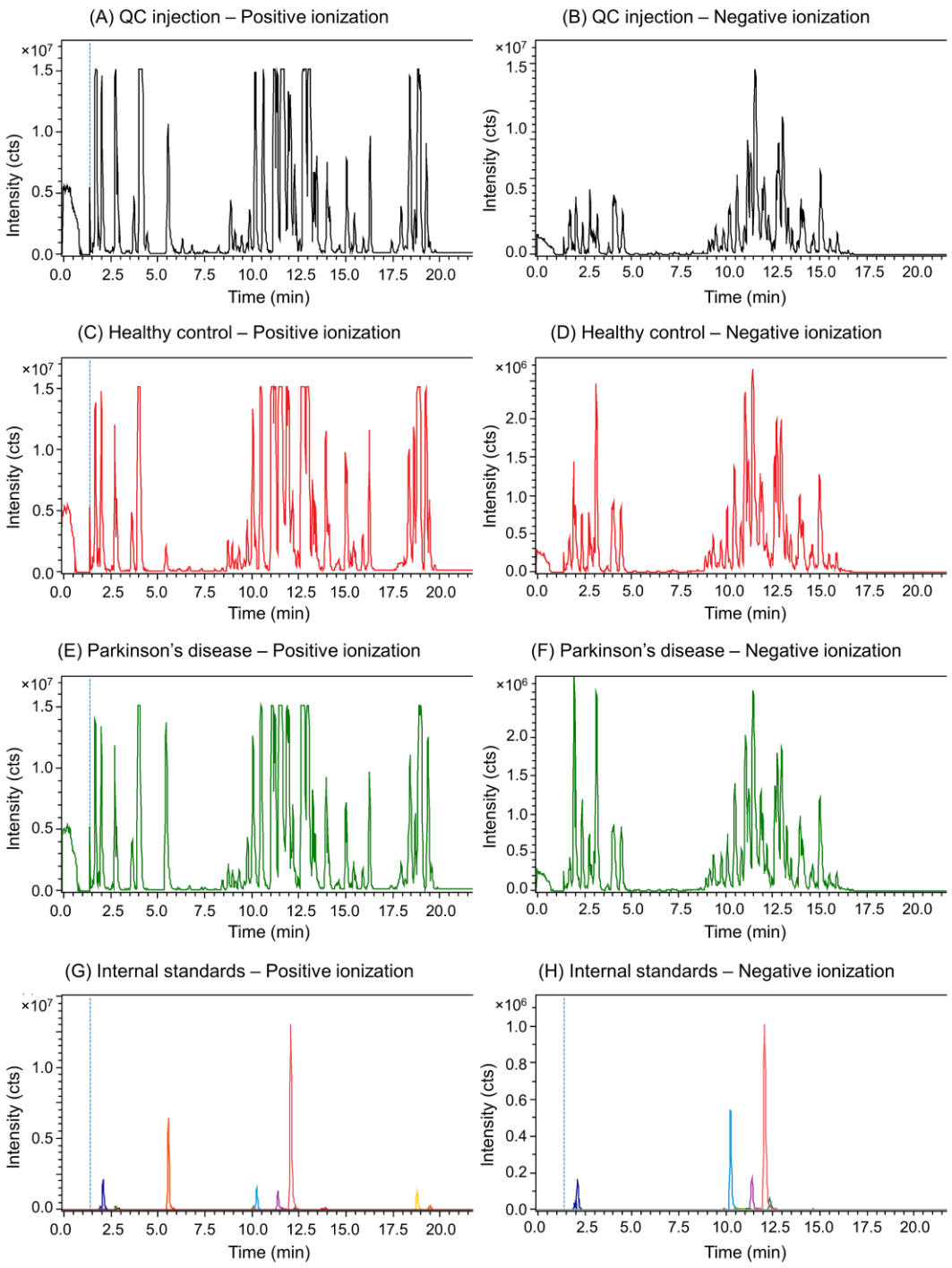

Figure VI-1. Examples of base peak chromatograms (BPC) obtained under positive (A, C, E and G) and negative ionization (B, D, F and H) for lipidomics or Parkinson's disease and dementia. (A, B) QC injections; (C, D) blood serum extract of a healthy control subject; (E, F) serum extract

of a Parkinson's disease patient; and (G, H) extracted ion chromatograms for the internal standard mixture, composed by 14 deuterated lipids belonging to different classes (Table II-1).

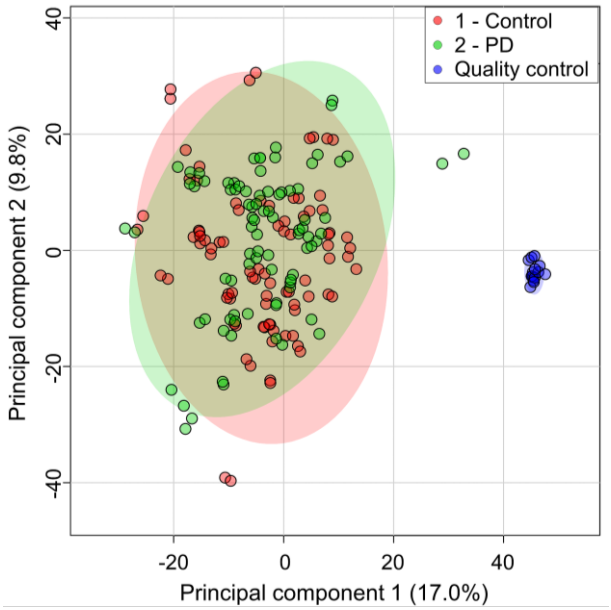

Figure VI-2. PCA score plot for the separation between healthy control subjects (red), PD patients (green) and quality control (QC) replicates (blue). The tight clustering of QCs and duplicate experiments for each sample indicates a high reproducibility for the chosen methods.

6.3.2. Parkinson's disease patients versus healthy controls

We first hypothesized that samples from PD patients displayed a different lipid composition than healthy controls. The Volcano plot analysis for the lipidome of PD patients compared to healthy controls showed a significant decrease in intensity ratios for PD for 34 lipids (FC for healthy control/PD $\geq$1.2 and p $\leq$0.05) and increased values for 95 lipids (FC for control/PD $\leq$0.84) (Figure VI-3A, Supp. Table 21). Boxplots for some of the significantly altered lipids are displayed in Appendix C - Figure C-1. The significantly altered lipids included 16 sphingomyelins

(SM), 7 ceramides (6 ceramide-phosphoethanolamines and one hexosyl ceramide) and 5 triacylglycerols (TG) with lower normalized intensities for the PD patients, whereas 38 phosphatidylserines (PS), 15 phosphatidic acids (PA) and 14 fatty acids (FA) displayed higher normalized intensities when the disease was diagnosed (Figure VI-3B).

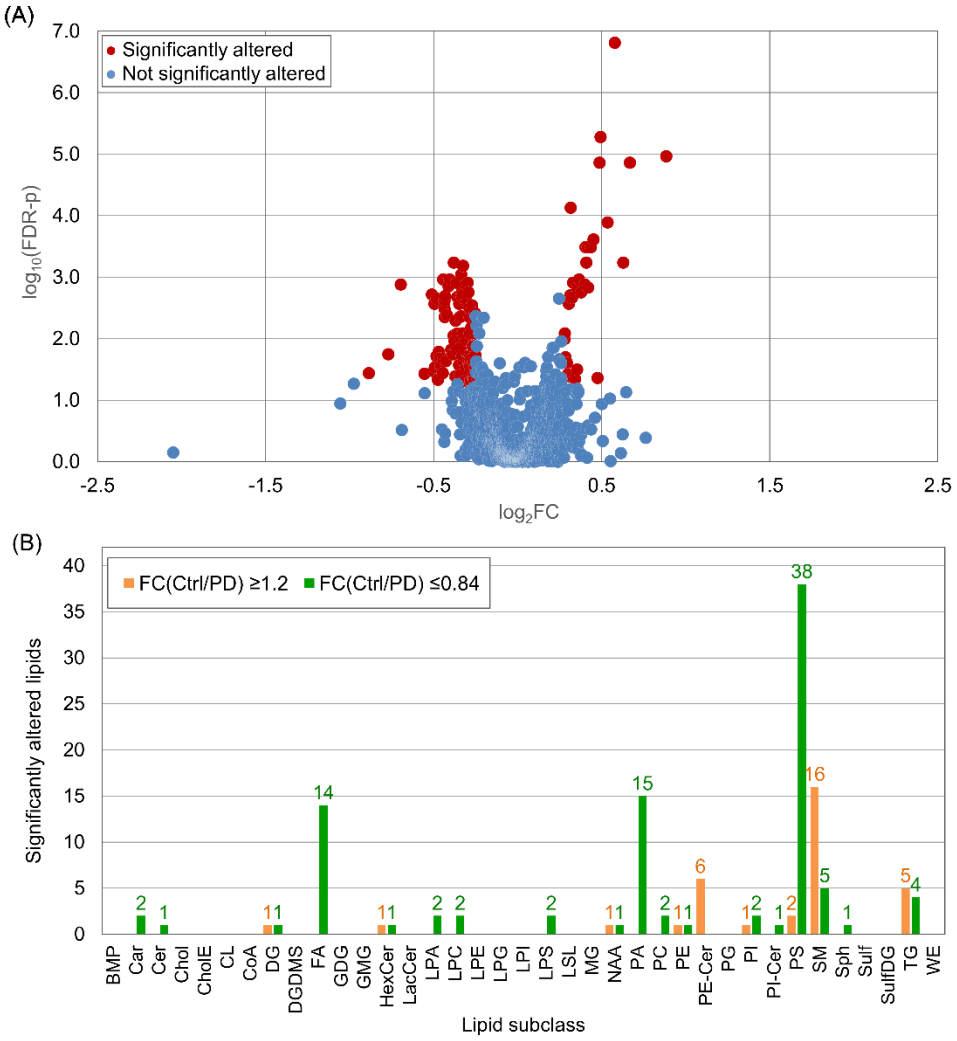

Figure VI-3. Volcano plot analysis for lipidomics of Parkinson's disease. (A) Volcano plot with significantly altered lipids displayed in red (FC ≥1.2 or <0.84, p ≤0.05), and non-significant lipids in blue. (B) Classes of significantly altered lipids for the control/PD Volcano plot analysis. Abbreviations to lipid subclasses are defined in Table V-1 (p. 154).

A PCA model did not result in a good separation between PD patients and healthy controls due to the complexity of the dataset (Figure C-2), but duplicate experiments were tightly clustered. The PLS-DA analysis with 8 components resulted in good separation (Figure VI-4A), as well as the OPLS-DA analysis (Figure VI-4B). Moreover, a Random Forest model resulted in an out-of-bag error (OOB) of 0.02 with 1000 trees and 12 predictors (classification error or 3.5% for controls and 0% for PD).

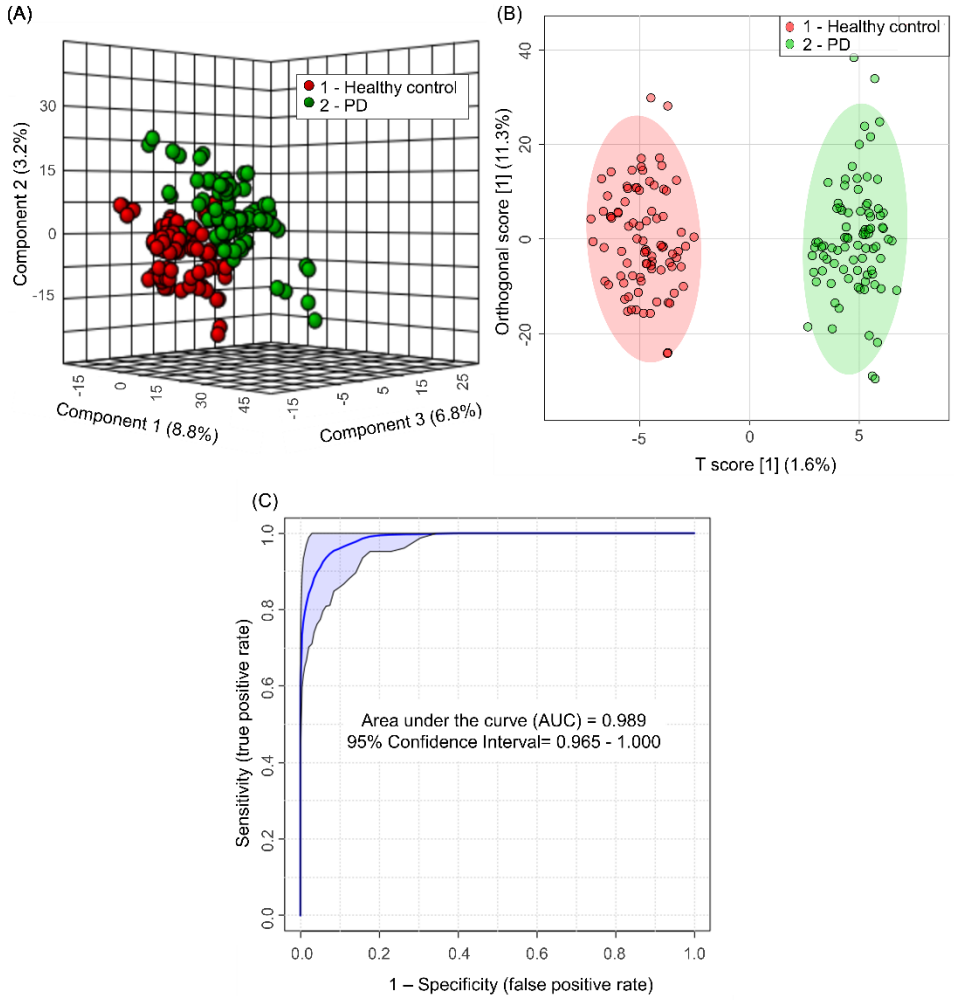

Figure VI-4. Statistical analysis for lipidomics of Parkinson's disease. (A) PLS-DA score plot with 8 components, R^2 of 0.9573, Q^2 of 0.7783 and $p < 0.001$ for 1000 permutations. (B) OPLS-DA

score plot with R^2Y of 0.957, Q^2 of 0.823 and $p < 0.001$ for 100 permutations. (C) ROC curve obtained for the control/PD separation with the proposed 7-lipid biomarker panel, as presented in Table VI-4.

One-quarter of the control and PD samples (11 controls and 11 PD) was randomly retained as a validation set, while the remaining three-quarters (32 controls and 32 PD) were employed as a training set for the selection of potential biomarkers. Lipids were ranked regarding their ability to correctly classify the training samples from PD patients or healthy controls by Random Forest. The area under the Receiver Operating Characteristic (ROC) curve (AUC) were used to gauge the classification performance. Multivariate AUC values increased for biomarker panels composed by two to seven lipids but, as more compounds were added, AUC values became steady. Hence, we constructed a biomarker panel with the 7 most important lipids for the control/PD classification by Random Forest using the training set (Table VI-4). The proposed 7-lipid biomarker panel resulted in AUC of 0.989 (95% confidence interval - CI of 0.965 – 1.000), sensitivity of 95.3%, specificity of 93.8% and average accuracy of 93.8% for 100 cross-validations (Figure VI-4C). The 7-lipid biomarker panel for the control/PD classification (Table VI-4) was further employed to predict the disease status for the validation set. We achieved sensitivity of 90.9%, specificity of 90.9% and overall accuracy of 90.9%.

Table VI-4. Proposed biomarker panel for the classification of samples as (1) control or PD; or (2) PDND or PDD. Abbreviations to lipid subclasses are defined in Table V-1 (p. 154).

Classification	Polarity m/z / retention time (min)	Univariate AUC	Identification
(1) Control/PD	N802.60161/ 13.26	0.719	HexCer(t37:1) | PS(P-38:0) | PS(O-38:1) | PC(O-35:2) | PC(P-35:1)
	P320.25676/ 1.74	0.739	NAE(18:4) | NAE(16:1) | Sph(d18:2)
	N764.54187/ 11.78	0.753	PS(O-34:0(OH)) | PC(31:0) | PE(34:0) | PE-NMe2(32:0) | PC(O-31:1(OH))
	P775.66759/ 15.06	0.787	PE-Cer(d42:0) | SM(d39:0) | CerP(d44:1)
	P809.65147/ 13.23	0.685	SM(d40:1) | PE-Cer(d43:1) | MGDG(39:1)
	P803.69882/ 15.92	0.770	PE-Cer(d44:0) | SM(d41:0) | CerP(d46:1)
	P845.65181/ 12.96	0.626	MGDG(42:4)
(2) PDND/PDD	N814.59749/ 13.47	0.854	PC(P-36:2) | PC(O-36:3) | PE(40:2(OH)) | PS(P-39:1) | PC(38:2(OH))
	P320.25676/ 1.74	0.760	NAE(18:4) | NAE(16:1) | Sph(d18:2)
	P740.55611/ 8.34	0.734	LPC(34:4) | PC(O-34:4) | PC(P-34:3) | LPC(32:1) | PC(O-32:1)
	P786.59935/ 12.92	0.832	PC(18:1/18:1)
	P723.56329/ 11.36	0.769	PE-NMe(32:0) | PC(30:0) | PC(O-30:1(OH)) | PC(P-30:0(OH)) | PE(33:0)

Notes: For putative identifications, all isomeric and isobaric possibilities with m/z error ≤5.0 mDa that passed the retention time and adduct filters were kept, but a maximum of five are shown; AUC: area under the ROC curve; P: positive ionization; N: negative ionization.

6.3.3. Parkinson's disease dementia

In the second phase, baseline lipidome discrimination of PDND and PDD was evaluated. All samples were collected three years before dementia diagnosis, when patients displayed no

signals of dementia; hence, the second phase consisted in an evaluation of prediction of transition to PDD up to three years before clinical symptoms. A volcano plot resulted in 66 lipids with decreased normalized intensity ratios for dementia patients (FC for PDND/PDD $\geq$1.2 and p $\leq$0.05, Figure VI-5A), including 25 triacylglycerols (TG) and 20 diacylglycerols (DG). In contrast, 48 lipids had increased intensity ratios for dementia patients (FC for PDND/PDD $\leq$0.84 and P $\leq$0.05), including 18 phosphatidylcholines (PC), 9 fatty acids (FA) and 4 phosphatidylethanolamines (PE) (Figure VI-5B, Supp. Table 22). Boxplots for some of the significantly altered lipids are displayed in Appendix C - Figure C-3.

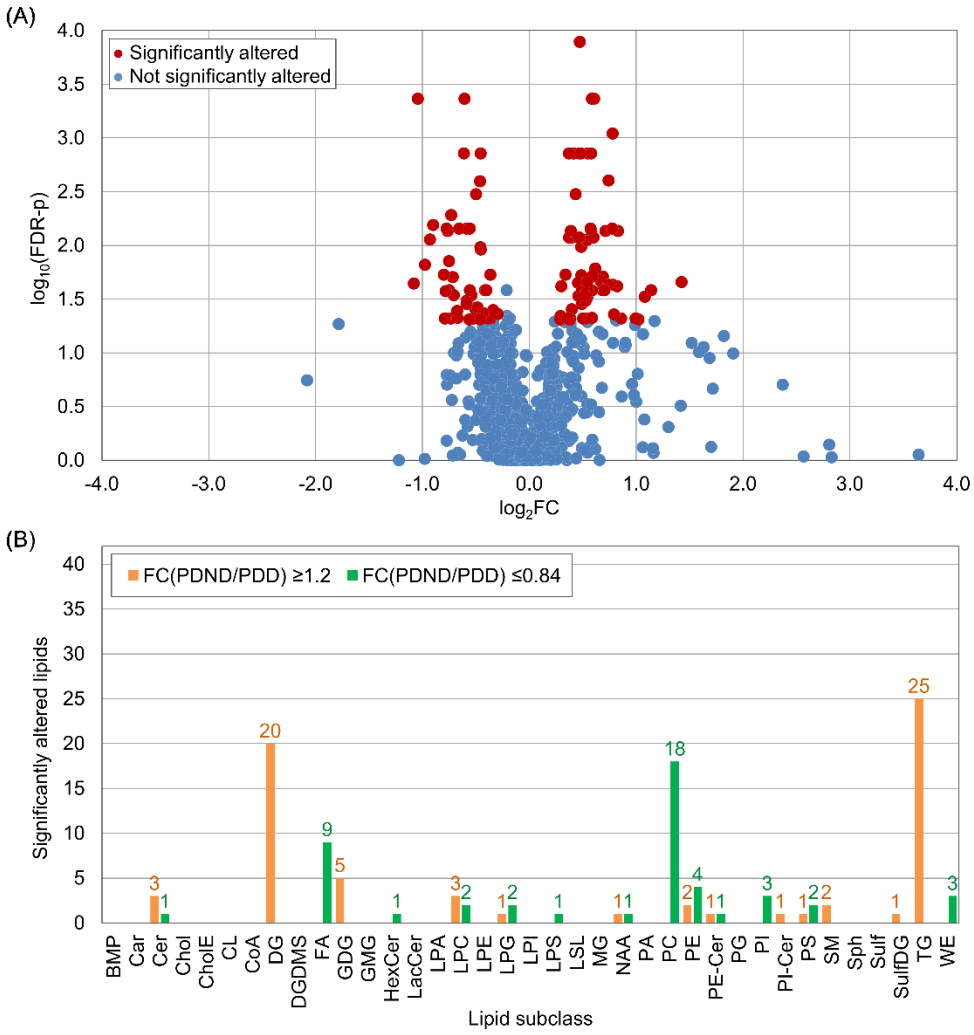

Figure VI-5. Volcano plot analysis for lipidomics of PD patients that didn't develop dementia in the 3-year interval after sample collection compared to PD patients clinically diagnosed with dementia in the *post-hoc* assessments. (A) Volcano plot with significantly altered lipids displayed in red (FC ≥1.2 or <0.84, P ≤0.05) and non-significant lipids in blue. (B) Classes of significantly altered lipids for the PDND/PDD Volcano plot analysis (FC ≥1.2 or <0.84, P ≤0.05). Abbreviations to lipid subclasses are defined in Table V-1 (p. 154).

A PCA model did not result in a good separation between PDND and PDD patients due to the complexity of the dataset (Figure C-4), but experimental duplicates were tightly clustered. The

PLS-DA model with 6 components and the OPLS-DA score plot showed a clear separation between the two subgroups (Figure VI-6A and B). The Random Forest model with 12 predictors and 5000 trees resulted in an OOB error of 0.06.

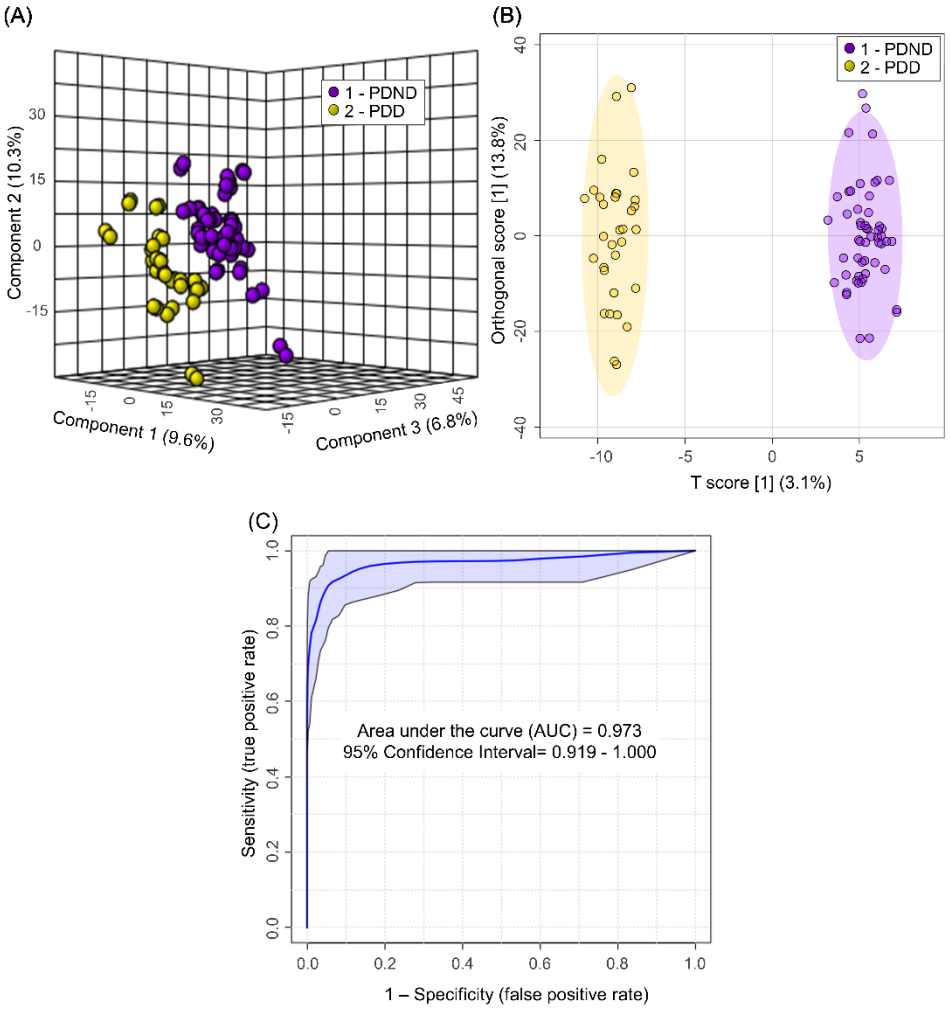

Figure VI-6. Statistical analysis for lipidomics of PD patients that didn't develop dementia in the 3-year interval after sample collection (PDND, purple) compared to PD patients clinically diagnosed with dementia in the post-hoc assessments (PDD, yellow). (A) PLS-DA score plot with 6 components, R^2 of 0.9700, Q^2 of 0.7640 and p of 0.04 for 1000 permutations. (B) OPLS-DA score plot with R^2Y of 0.986, Q^2 of 0.884 and p <0.001 for 1000 permutations. (C) ROC curve

obtained for the PDND/PDD separation with the proposed 5-lipid biomarker panel, as presented in Table VI-4.

As in the previous phase, we compared PDND with PDD by selecting the most important lipids for classification by Random Forest as biomarker candidates. The sample set was randomly divided into a training set, composed by 20 PDND and 12 PDD samples, and a validation set, composed by 7 PDND and 4 PDD samples. We constructed a biomarker panel with the 5 most important lipids for the PDND/PDD classification by Random Forest using the training set, as shown in Table VI-4. The ROC analysis resulted in multivariate AUC of 0.973 (95% CI of 0.919 – 1.000), sensitivity of 95.8%, specificity of 90.0% and overall accuracy of 91.9% (Figure VI-6C). The 5-lipid biomarker panel (Table VI-4) was further employed to classify the PD samples randomly retained for a validation set as PDND or PDD. All samples were correctly classified, resulting in excellent sensitivity of 100%, specificity of 100% and overall accuracy of 100%.

6.3.4. Confounding factors

We evaluated the effect of seven potentially confounding factors on the lipidome, *viz*., age, sex, time since PD diagnosis, medication dosages (levodopa equivalents), the period of treatment, motor deficit score (UPDRS - part III) and B12 serum levels (Table VI-1). Using a median split, we divided the dataset (control and PD; or PDND and PDD) into two subgroups for each potential confounder (*i.e.*, above/below the median). The detailed description and results for the evaluation of confounding factors are provided on the following sub-sections.

In summary, although we detected significant fold changes for selected lipids for each potential confounder (Figure VI-7), the classification of samples as healthy control or PD by the

proposed 7-lipid biomarker panel was only minimally affected. As previously mentioned, the classification of the complete sample set as control or PD resulted in AUC of 0.989 with 95% confidence interval between 0.965 and 1.000 (Figure VI-4C). For age, we examined the multivariate AUCs obtained when the control and PD individuals were separated as above and below the age median before disease classification. Both groups were within the 95% confidence interval for the complete control/PD dataset, *i.e.,* the classification of individuals above the age median as PD or control resulted in AUC of 0.972, while people below the median resulted in AUC of 0.987. For sex, stratified classification as control or PD resulted in slightly reduced AUC values for females (0.985) and males (0.942). We also evaluated the performance of the 7-lipid biomarker panel for individuals with B12 levels below (AUC of 0.949) and above the median (0.993), which only showed minor differences. Overall, AUC values were only slightly below the 95% confidence interval for sex (males) and B12 levels (below the median value). The 2-way ANOVA analysis for each potential confounder and PD status showed no significant interference for the 7 lipids in the biomarker panel, *i.e.*, interaction p was above the arbitrary significance level of 0.05.

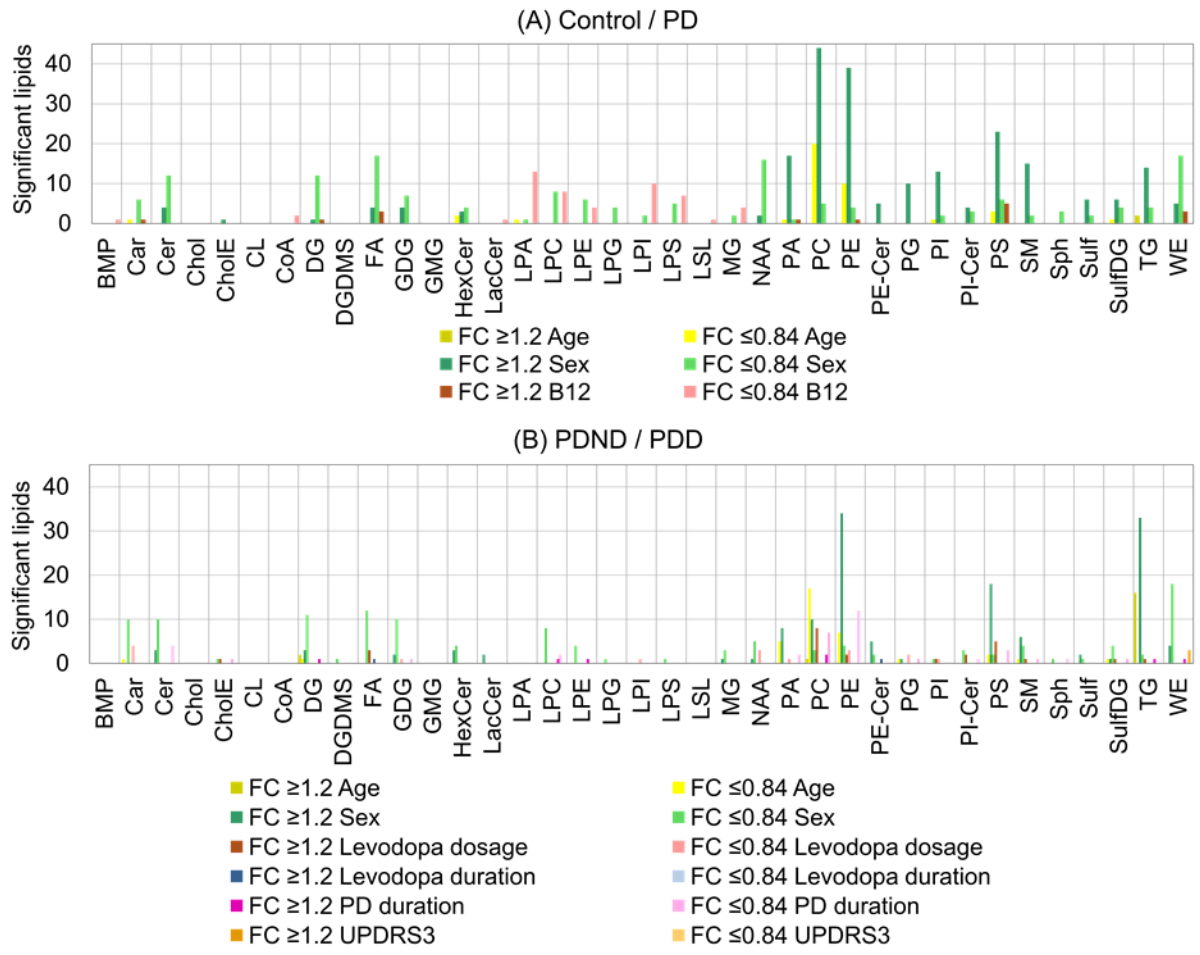

Figure VI-7. Classes of significantly altered lipids (FC ≥1.2 or <0.84 and p <0.05) for the evaluated potential confounding factors by Volcano plot. (A) control/PD comparison; (B) PDND/PDD comparison. Abbreviations to lipid subclasses are defined in Table V-1 (p. 154).

For the second study phase, the classification of samples as PDND or PDD using the 5-lipid biomarker panel (Table VI-4, p. 232) was also not strongly affected by age, sex, medication dosage, period of treatment, time since PD diagnosis or UPDRS-III score. The classification of all samples as PDND or PDD resulted in AUC of 0.973 with 95% confidence interval of 0.919 – 1.000, as previously described (Figure VI-6C). The AUC values for the median-split subgroups were within the 95% confidence interval of the complete PDND/PDD subgroups for most

confounders. The period of treatment with Levodopa or equivalent medication was the only exception, with AUC of 0.894 for patients below the median value (8.0% smaller than the complete PDND/PDD cohort) and 1.000 for patients above the median (2.9% higher), indicating a mild influence. However, the five lipids in the PDND/PDD biomarker panel resulted in $p > 0.05$ for the interference between all potential confounders and disease status, *i.e.*, not significant by 2-way ANOVA.

6.3.4.1 Age

The age of healthy control subjects was not significantly different from the PD group (p of 0.44, Table VI-1), but the PDND and PDD subgroups were not age-matched (p of 0.03). It is expected that the aging processes will influence the lipidome of blood serum samples; however, the changes caused by PD as the patient ages are not entirely understood, particularly as dementia develops. Therefore, the effect of age on the lipidome was evaluated for the control and PD subjects.

6.3.4.2 Age effect on the lipidome of control and PD subjects

The 43 healthy control and 43 PD samples were separated into two subgroups: (1) below the median (20 control and 23 PD samples) and (2) above the median of 69.97 years old (23 control and 20 PD samples). For the control group, univariate analysis through a volcano plot resulted in two triacylglycerols with significantly decreased normalized intensity ratios for the older subgroup (FC for below/above median ≥ 1.2 and $p \leq 0.05$), while 40 lipids showed increased values (FC < 0.84), including 20 phosphatidylcholines and 10 phosphatidylethanolamines (Supp. Table 23). Two of the significant lipids with FC (below/above the median) < 0.84 were also significantly

altered for the control/PD comparison: one phosphatidylserine and one hexosyl ceramide. Furthermore, five significant lipids for age were also significantly altered for the PDND/PDD comparison, with all of them displaying higher intensity ratios for individuals above the age median and diagnosed with PDD (FC below/above the age median and PDND/PDD <0.84). The lipids selected for the control/PD and PDND/PDD biomarker panels were not significantly affected by age. The multivariate statistical analysis by PCA did not result in a discernible separation between individuals above and below the age median, but a PLS-DA model was developed with 8 components, resulting in R^2 of 0.9428, Q^2 of 0.6837 and p <0.001 for 1000 permutations (Appendix C - Figure C-5).

The PDND and PDD subgroups were not age-matched (p of 0.03), increasing the importance of significant effects caused by age differences. For the PD group, the Volcano plot for lipids detected below the median (16 PDND and 6 PDD) and above the median of 69.74 years old (11 PDND and 10 PDD) resulted in 19 lipids with decreased intensities for the older subgroup, while 36 had increased values (Supp. Table 24). The lipids with lower intensity ratios for the older subgroup include 16 triacylglycerols and 2 diacylglycerols; whereas for higher intensity ratios, 17 phosphatidylcholines, 7 phosphatidylethanolamines and 5 phosphatidic acids, amongst others. Two of the significantly altered lipids were also selected for the PDND/PDD comparison, both with FC (below/above median) <0.84 and FC (PDND/PDD) <0.84, *i.e.*, they were affected by age regardless of disease status. However, the lipids selected for the biomarker panels for control/PD and PDND/PDD were not significantly altered. The PCA score plot displayed no discernible differentiation for the median-split PD subgroups, but PLS-DA with 6 components resulted in R^2 of 0.9769, Q^2 of 0.8359 and p of 0.03 for 1000 permutations (Figure C-6).

6.3.4.3 Age effect on PD and dementia classification

The age-matched control and PD samples were split into two subgroups at the median value of 69.97 years old. The two-way ANOVA analysis for age (above/below median) and disease status (control/PD) resulted in 15 lipids with significant interactions between the two factors (interaction $p < 0.05$, Supp. Table 25), but none of them were part of the 7-lipid control/PD biomarker panel (Table VI-4). Even though the PDND and PDD subgroups were not age-matched, two-way ANOVA analysis for dementia status (PDND/PDD) and age (below/above median) revealed that none of the identified lipids had significant interaction between the two factors (Supp. Table 26).

The 7-lipid biomarker panel for the control/PD comparison (Table VI-4, p. 232) was employed to generate a ROC curve for the classification of PD samples as below or above the median of 69.97 years old. The obtained multivariate AUC was 0.701, with an accuracy of 65.3% (Table VI-5, Figure VI-8A). The values were much smaller than the ones for the sample classification as control or PD (AUC of 0.989 and accuracy of 93.8%, Figure VI-4C), indicating little effect. Similarly, the five-lipid biomarker panel for the PDND/PDD comparison (Table VI-4, p. 232) was employed to generate a second ROC curve for the classification as below or above the age median for the PD group (69.74 years old). The obtained multivariate AUC was 0.712 with an accuracy of 67.4% (Table VI-6, Figure VI-9A), once again smaller than the values found for the original PDND/PDD classification (AUC of 0.972 and accuracy of 91.9%, Figure VI-6C). The low performance of both biomarker panels to classify samples as below/above the age median, when compared to their excellent performance for classification as control/PD or PDND/PDD, are indications that age is not an important confounding factor for the selected lipids.

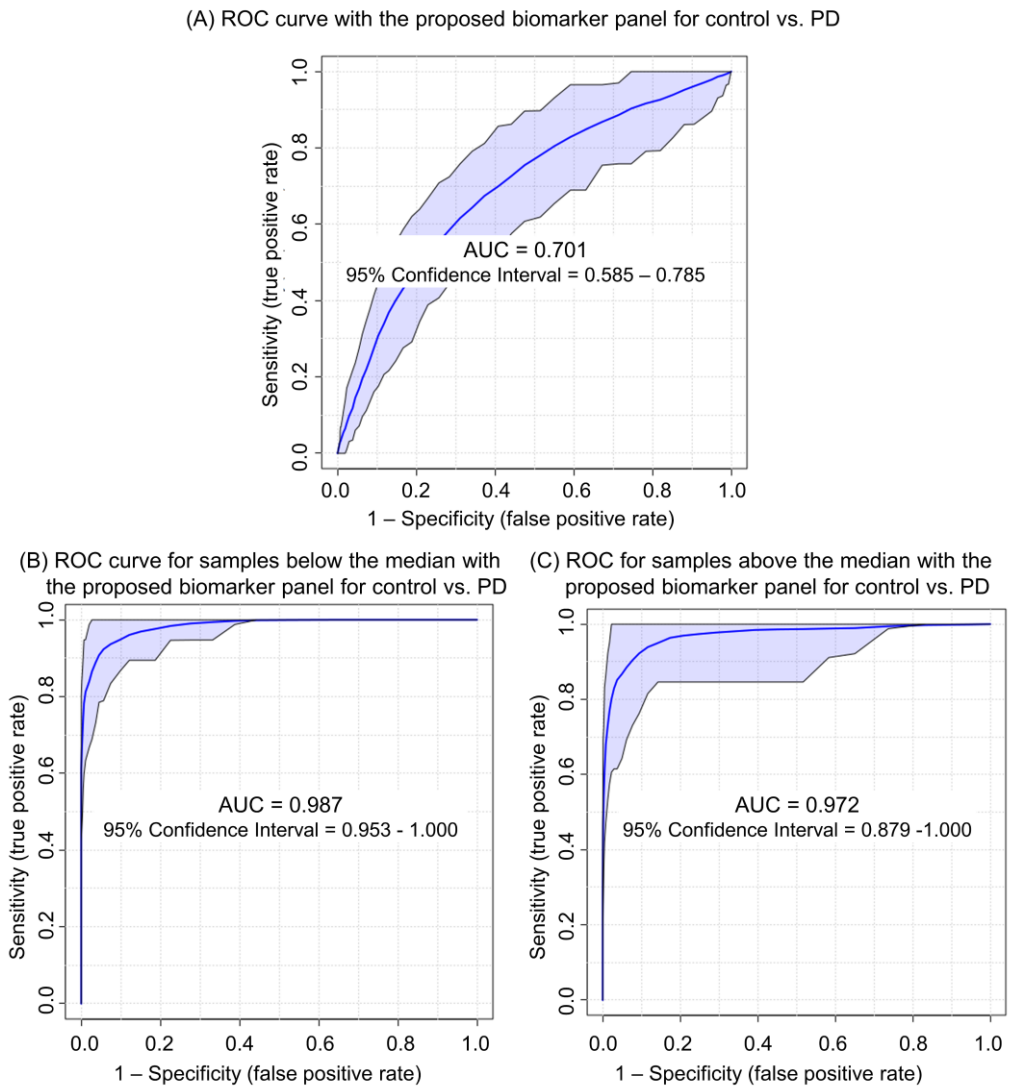

Figure VI-8. ROC curves using the 7-lipid biomarker panel for the control/PD comparison (Table VI-4, p. 232) to evaluate age as a confounding factor. The biomarker panel was used to classify: (A) all control and PD samples as above or below the age median of 69.97 years old, with sensitivity of 62.8%, specificity of 72.1% and overall accuracy of 65.3% by 100 cross-validations; (B) only samples from individuals aged below the median of 69.97 years old as control or PD, with sensitivity of 91.3%, specificity of 100% and accuracy of 93.0%; (C) only samples for individuals aged above the median of 69.97 years old as control or PD, with sensitivity of 95.0%, specificity of 93.5% and accuracy of 92.0%.

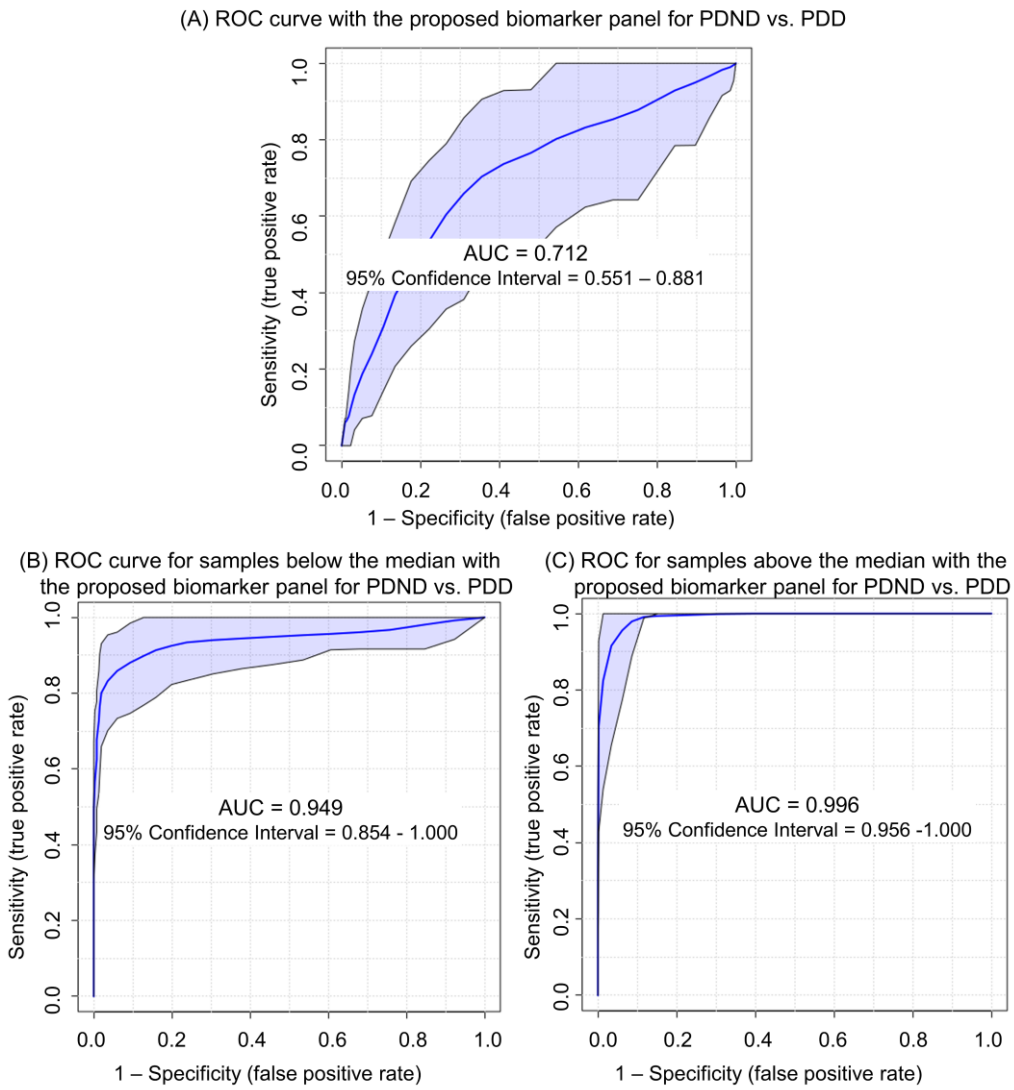

Figure VI-9. ROC curves for PD patients (PDND and PDD) using the 5-lipid biomarker panel (Table VI-4, p. 232) to evaluate age as a confounding factor. The biomarker panel was used to classify: (A) all PD samples as above or below the age median of 69.74 years old, with sensitivity of 63.6%, specificity of 76.2% and overall accuracy of 67.4% by 100 cross-validations; (B) only samples from PD patients aged below the median of 69.74 years old as PDND or PDD, with sensitivity of 100.0%, specificity of 87.5% and accuracy of 87.9%; (D) only samples from PD patients aged above the median of 69.74 years old as PDND or PDD, with sensitivity of 95.0%, specificity of 100.0% and accuracy of 96.4%.

Table VI-5. Uni- and multivariate AUCs for the proposed biomarker panel for diagnosis of PD (Table VI-4, p. 232), when applied for the classification of samples as below or above the median for age, sex (male or female) and B12 serum level. For putative identifications, all isomeric and isobaric possibilities with the same m/z error that passed the retention time and adduct filters were kept, but a maximum of five are shown. The ionization mode is described as "P" for positive or "N" for Negative. Lipid class abbreviations are described in Table V-1 (p. 154).

Polarity m/z / retention time	**Identification**	**Class**	**Univariate AUC**			
			Control/ PD	Age	Sex	B12
N764.54187/ 11.78	HexCer(t37:1) | PS(P-38:0) | PS(O-38:1) | PC(O-35:2) | PC(P-35:1)	HexCer	**0.753**	0.573	0.520	0.548
N802.60161/ 13.26	NAE(18:4) | NAE(16:1) | Sph(d18:2)	NAA	**0.719**	0.685	0.555	0.547
P320.25676/ 1.74	PS(O-34:0(OH)) | PC(31:0) | PE(34:0) | PE-NMe2(32:0) | PC(O-31:1(OH))	PS	**0.739**	0.527	0.558	0.548
P775.66759/ 15.06	PE-Cer(d42:0) | SM(d39:0) | CerP(d44:1)	PE-Cer	**0.787**	0.567	0.527	0.516
P803.69882/ 15.92	SM(d40:1) | PE-Cer(d43:1) | MGDG(39:1)	SM	**0.771**	0.548	0.508	0.565
P809.65147/ 13.23	PE-Cer(d44:0) | SM(d41:0) | CerP(d46:1)	PE-Cer	**0.685**	0.520	0.532	0.606
P845.65181/ 12.96	MGDG(42:4)	GDG	**0.626**	0.510	0.572	0.594
Multivariate AUC (7-lipid biomarker panel)			**0.989**	0.701	0.836	0.749
95% confidence interval for multivariate AUC			**0.965-1.000**	0.585-0.785	0.702-0.923	0.611-0.879

Table VI-6. Uni- and multivariate AUCs for the proposed biomarker panel for prediction of PDD (Table VI-4, p. 232), when applied for the classification of PD samples as below or above the median for age, sex (male or female), time since PD diagnosis, Levodopa equivalent dosage, period of treatment, and UPDRS-part 3 (motor assessment). The ionization mode is described as "P" for positive or "N" for Negative. Lipid class abbreviations are described in Table V-1 (p. 154).

Polarity m/z / retention time	**Identification**	**Class**	**Univariate AUC**						
			PDND/ PDD	Age	Sex	Levodopa dosage	Levodopa duration	PD duration	UPDRS
N814.59749/ 13.47	PC(P-36:2) | PC(O-36:3) | PE(40:2(OH)) | PS(P-39:1) | PC(38:2(OH))	PC	**0.854**	0.531	0.527	0.579	0.621	0.584	0.502
P320.25676/ 1.74	NAE(18:4) | NAE(16:1) | Sph(d18:2)	NAA	**0.760**	0.538	0.556	0.631	0.601	0.586	0.543
P723.56329/ 11.36	PE-NMe(32:0) | PC(30:0) | PC(O-30:1(OH)) | PC(P-30:0(OH)) | PE(33:0)	PE-Cer	**0.769**	0.532	0.555	0.514	0.554	0.593	0.543
P740.55611/ 8.34	LPC(34:4) | PC(O-34:4) | PC(P-34:3) | LPC(32:1) | PC(O-32:1)	LPC	**0.734**	0.652	0.531	0.576	0.530	0.540	0.639
P786.59935/ 12.92	PC(18:1/18:1)	PC	**0.832**	0.523	0.581	0.627	0.569	0.700	0.526
Multivariate AUC (5-lipid biomarker panel)			**0.961**	0.712	0.677	0.602	0.711	0.750	0.702
95% confidence interval for multivariate AUC			**0.919-1.000**	0.551-0.881	0.510-0.826	0.458-0.723	0.547-0.882	0.566-0.902	0.537-0.826

The median-split subgroups were further employed to generated ROC curves for the classification of disease status using each distinct subgroup, *i.e.*, only samples below or only samples above the age median. For the control/PD classification, the samples from individuals aged above the median resulted in AUC of 0.972 (95% confidence interval between 0.879 and 1.000) and accuracy of 92.0% (Figure VI-8C, Table VI-7). For individuals below the median, the AUC was 0.987 (95% confidence interval between 0.953 and 1.000) and accuracy of 93.0% (Figure VI-8B, Table VI-8). The performances for classification of the subgroups composed only by individuals below or above the age median as control or PD were similar to each other, and not better than the results found for the complete control/PD sample set (AUC of 0.989 with 95% confidence interval between 0.965 and 1.000, and accuracy of 93.8% - Figure VI-4C). Likewise, the distinct median-split subsets for PDND and PDD resulted in AUC of 0.996 (95% confidence interval between 0.956 and 1.000) with an accuracy of 96.4% for above the median (Figure VI-9C); and AUC of 0.949 (95% confidence interval between 0.854 and 1.000) with an accuracy of 87.9% for below the median (Figure VI-9B). The performance for the subgroup above the age median was slightly better than for people below the median, but both AUC values were within the 95% confidence interval for the complete PDND/PDD data set (AUC of 0.972 with 95% confidence interval between 0.919 and 1.000, Figure VI-6), indicating that the separation of patients according to age is not required for the selected biomarker panels.

Table VI-7. Uni- and multivariate AUCs for the proposed biomarker panel for the classification of samples as healthy control or PD, when applied for median-split subgroups for age, sex and B12 serum levels. The 7-lipid biomarker panel was used to classify only samples from (a) individuals below the age median; (b) individuals above the age median; (c) males; (d) females; (e) individuals below the median for B12 serum level; and (f) individuals above the median for B12 serum level as control or PD. The ionization mode is described as "P" for positive or "N" for Negative. Lipid class abbreviations are described in Table V-1.

Polarity m/z / retention time	Identification	Class	Univariate AUC						
			Control/ PD	Age		Sex		B12 serum level	
				Below median	Above median	Males	Females	Below median	Above median
N764.54187/ 11.78	HexCer(t37:1) | PS(P-38:0) | PS(O-38:1) | PC(O-35:2) | PC(P-35:1)	HexCer	**0.753**	0.777	0.658	0.673	0.777	0.744	0.707
N802.60161/ 13.26	NAE(18:4) | NAE(16:1) | Sph(d18:2)	NAA	**0.719**	0.764	0.734	0.705	0.797	0.728	0.775
P320.25676/ 1.74	PS(O-34:0(OH)) | PC(31:0) | PE(34:0) | PE-NMe2(32:0) | PC(O-31:1(OH))	PS	**0.739**	0.773	0.659	0.737	0.721	0.584	0.871
P775.66759/ 15.06	PE-Cer(d42:0) | SM(d39:0) | CerP(d44:1)	PE-Cer	**0.787**	0.730	0.873	0.688	0.891	0.793	0.796
P803.69882/ 15.92	SM(d40:1) | PE-Cer(d43:1) | MGDG(39:1)	SM	**0.771**	0.671	0.847	0.655	0.862	0.750	0.829
P809.65147/ 13.23	PE-Cer(d44:0) | SM(d41:0) | CerP(d46:1)	PE-Cer	**0.685**	0.643	0.722	0.716	0.676	0.726	0.620
P845.65181/ 12.96	MGDG(42:4)	GDG	**0.626**	0.783	0.554	0.596	0.607	0.586	0.712
Multivariate AUC (7-lipid biomarker panel)			**0.989**	0.987	0.972	0.942	0.985	0.949	0.993
95% confidence interval for multivariate AUC			**0.965-1.000**	0.953-1.000	0.879-1.000	0.859-0.992	0.935-1.000	0.863-0.999	0.965-1.000

Table VI-8. Uni- and multivariate AUCs for the proposed biomarker panel for the classification of samples as PDND or PDD, when applied for the classification of median-split PD samples for age, sex (male or female), time since PD diagnosis, Levodopa equivalent dosage, period of treatment, and UPDRS-part 3 (motor assessment). The 5-lipid biomarker panel was used to classify only samples from (a) individuals below the median or (b) individuals above the median as PDND or PDD. The ionization mode is described as "P" for positive or "N" for Negative. Lipid class abbreviations are described in Table V-1 (p. 154).

Polarity m/z / retention time	**Univariate AUC**												
	PDND /PDD	**Age**		**Sex**		**Levodopa dosage**		**Levodopa duration**		**PD duration**		**UPDRS3**	
		Below median	**Above median**	**Males**	**Females**	**Below median**	**Above median**	**Below median**	**Above median**	**Below median**	**Above median**	**Below median**	**Above median**
N814.59749/ 13.47	**0.854**	0.724	0.909	0.724	0.857	0.664	0.950	0.750	0.843	0.672	0.893	0.730	0.808
P320.25676/ 1.74	**0.760**	0.737	0.830	0.769	0.810	0.813	0.808	0.756	0.768	0.776	0.807	0.851	0.808
P723.56329/ 11.36	**0.769**	0.646	0.750	0.741	0.640	0.652	0.718	0.618	0.784	0.693	0.689	0.690	0.709
P740.55611/ 8.34	**0.734**	0.836	0.564	0.741	0.702	0.504	0.988	0.521	0.920	0.560	0.873	0.887	0.536
P786.59935/ 12.92	**0.832**	0.802	0.845	0.730	0.935	0.652	0.950	0.750	0.852	0.635	0.889	0.847	0.745
Multivariate AUC	**0.961**	0.949	0.996	1.000	0.960	0.965	1.000	0.894	1.000	0.995	0.977	0.982	0.978
95% CI	**0.919-1.000**	0.854-1.000	0.956-1.000	0.995-1.000	0.813-1.000	0.759-1.000	1.000-1.000	0.551-1.000	1.000-1.000	0.936-1.000	0.833-1.000	0.906-1.000	0.894-1.000

6.3.4.4 Sex

The control/PD groups and PDND/PDD subgroups were sex-matched ($p > 0.05$, Table VI-1). Sex is an important moderator in PD development, and it has been suggested that the disease may affect males and females differently. More males are diagnosed with PD than females, and there are a few clinical characteristics that are observed more often in men, *e.g.*, rigidity and rapid eye movement disorder. The possible neuroprotective effect of the female hormone estrogen has been cited as a cause for the gender-related symptoms, which may also alter the metabolome and lipidome of patients.[233] Hence, we investigated the effect of sex on the lipidome, as well as the possibility of interference for control/PD and PDND/PDD classifications.

6.3.4.5 Sex effect on the lipidome of control and PD subjects

For the control/PD sample set, univariate analysis through a volcano plot resulted in 220 lipids with significantly decreased intensity ratio for the males when compared to the females (FC for females/males ≥ 1.2 and $p \leq 0.05$), and 155 with increased values (FC for female/male < 0.84) (Supp. Table 23, Figure VI-7). Out of the 375 significantly altered lipids, 49 were identified as phosphatidylcholines, 43 phosphatidylethanolamines, 29 phosphatidylserines, 35 ceramides and 22 wax esters. The glycerophospholipid metabolism was strongly gender-dependent, with mostly higher intensity ratios for female subjects. We also observed a strong influence of sex on the sphingomyelin and ceramide metabolisms. Nineteen lipids were significantly changed for sex and the control/PD comparison, including two with FC for control/PD ≥ 1.2 and 17 with FC < 0.84 ($P \leq 0.05$). The seven lipids selected for the control/PD biomarker panel (Table VI-4, p. 232) were not significantly altered by sex. A partial separation for males and females was observed for the

multivariate statistical analysis by PLS-DA with 8 components, resulting in R^2 of 0.9597, Q^2 of 0.8402 and $p < 0.01$ (100 permutations, Figure C-7).

For the PD group, similar results were observed. 137 lipids had significantly smaller intensities for the male subgroup (FC for females/males in PD group ≥ 1.2 and $p \leq 0.05$), while 126 had a significant increase (FC < 0.84 and $p \leq 0.05$, Supp. Table 24). The glycerophospholipid metabolism was once again affected, with 38 phosphatidylethanolamines, 20 phosphatidylserines and 13 phosphatidylcholines significantly altered; however, we also found significant alterations for 35 triacylglycerols, 14 diacylglycerols, 32 ceramides and 22 wax esters. Six lipids were significantly altered for both the sex and PDND/PDD comparisons, with one displaying FC (female/male) and FC (PDND/PDD) ≥ 1.2, while the rest had FC (female/male) and FC (PDND/PDD) < 0.84. None of the five lipids selected for the PDND/PDD biomarker panel were significantly affected by sex. A good PLS-DA separation for males and females with PD was achieved with 6 components, resulting in R^2 of 0.9746, Q^2 of 0.8472 and $p < 0.01$ for 100 permutations (Figure C-8).

6.3.4.6 Sex effect on the PD and PDD diagnosis

Two-way ANOVA for disease status (control/PD) and sex (male/female) resulted in 61 lipids with significant interference (interaction $p < 0.05$, Supp. Table 25), meaning that both the disease status and sex influence their intensity ratios. However, the 7 lipids selected for the control/PD biomarker panel (Table VI-4, p. 232) did not show significant interference for disease status and sex ($p > 0.05$). Similarly, 50 lipids displayed significant interference for sex and PDND/PDD classification, but none of the five selected for the PDND/PDD biomarker panel (Supp. Table 26).

ROC curves using the 7-lipid biomarker panel for control/PD (Table VI-4, p. 232) to classify samples as females or males resulted in multivariate AUC of 0.836 (95% confidence interval between 0.702 and 0.923), with an accuracy of 76.6% and univariate AUCs between 0.508 and 0.572 (Figure VI-10A, Table VI-5). Likewise, we used the 5-lipid PDND/PDD biomarker panel (Table VI-4, p. 232) to classify samples as females or males, resulting in a multivariate AUC of 0.677 (95% confidence interval between 0.510 and 0.826, Figure VI-11A, Table VI-6) and accuracy of 63.8%. Univariate AUCs for the 5 lipids were between 0.527 and 0.581 (Table VI-6). The low multi- and univariate AUC values, combined with lower accuracies and no significant interference, indicate that sex is not an important confounding factor for the selected biomarker panels.

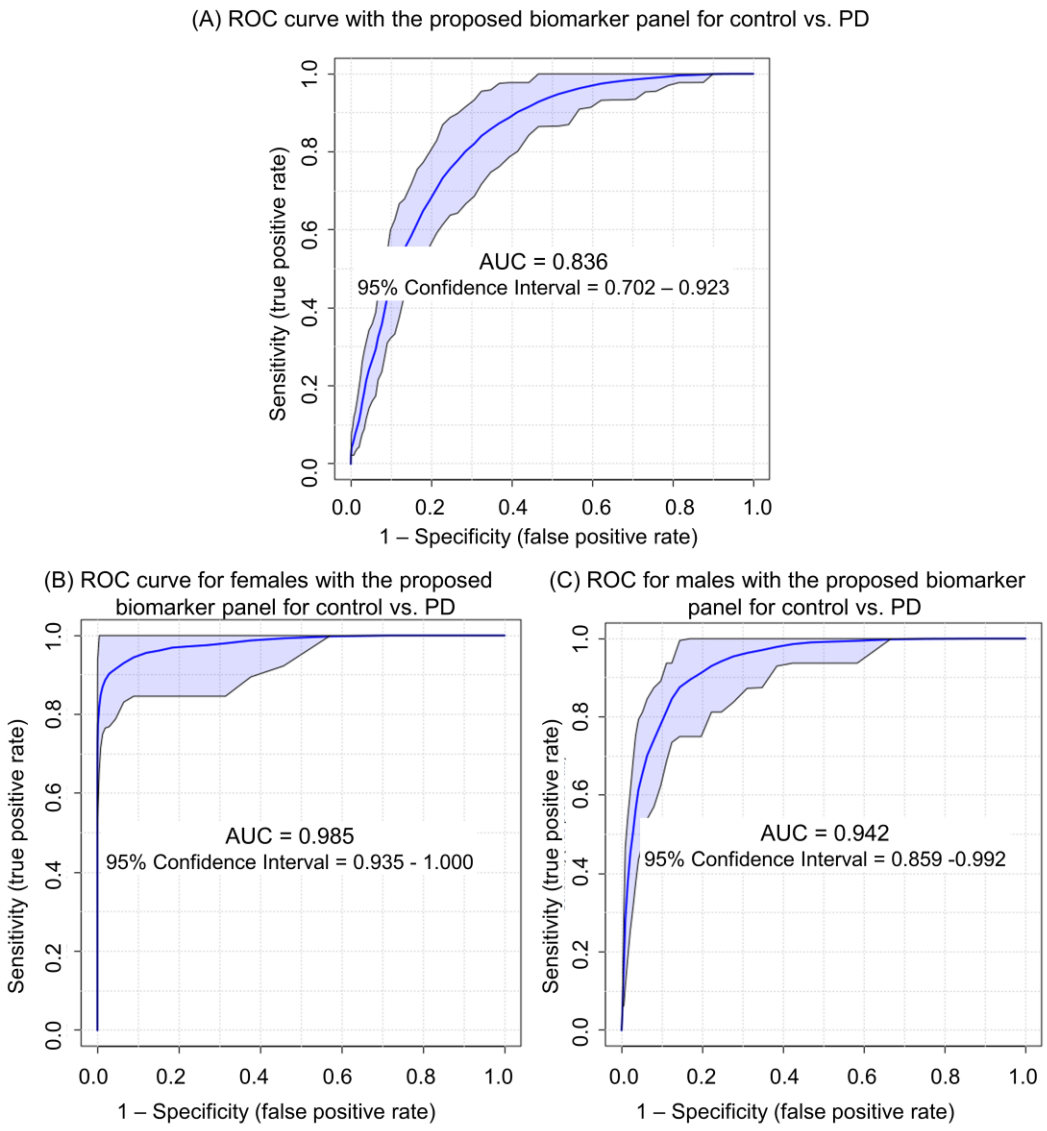

Figure VI-10. ROC curves using the 7-lipid biomarker panel for the control/PD comparison (Table VI-4, p. 232) to evaluate sex as a confounding factor. The biomarker panel was used to classify: (A) all samples (control and PD) as male or female, with sensitivity of 84.4%, specificity of 75.0% and overall accuracy of 76.6%; (B) only samples from female individuals as control or PD, with sensitivity of 89.5%, specificity of 97.4% and accuracy of 93.8%; (C) only samples from male individuals as control or PD, with sensitivity of 89.6%, specificity of 87.5% and accuracy of 86.6%.

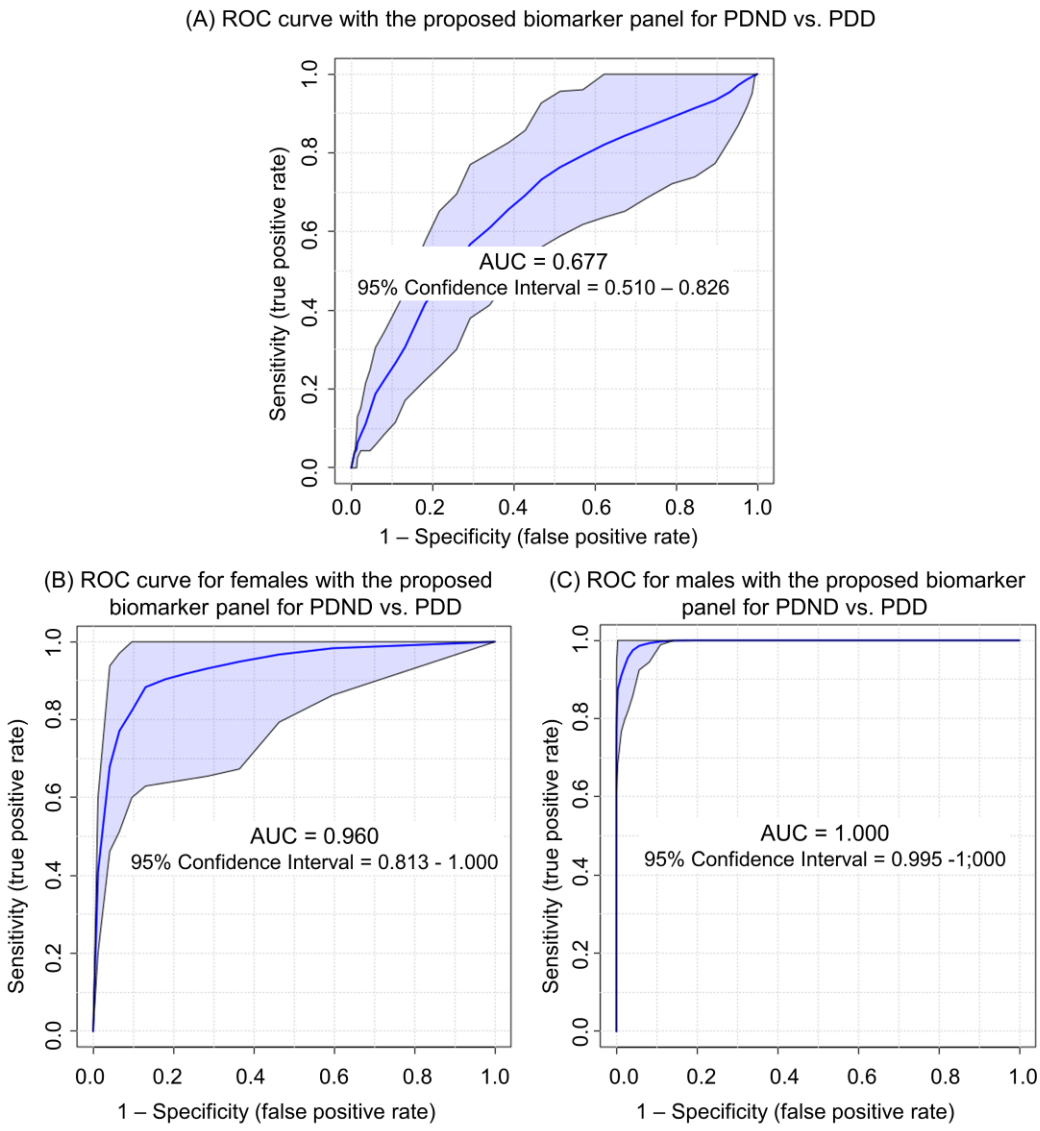

Figure VI-11. ROC curves for PD patients, including PDND and PDD, using the 5-lipid biomarker panel (Table VI-4, p. 232) to evaluate sex as a confounding factor. The biomarker panel was used to classify: (A) all PD samples as male or female, with sensitivity of 68.4%, specificity of 64.6% and overall accuracy of 63.8%; (B) only samples from female PD patients as PDND or PDD, with sensitivity of 100%, specificity of 96.7% and overall accuracy of 94.8%; (C) only samples from male PD patients as PDND or PDD, with sensitivity of 85.7%, specificity of 91.7% and overall accuracy of 88.6%.

The distinct male and female subgroups were examined for classification as healthy control or PD using the 7-lipid biomarker panel (Table VI-4, p. 232). For the female subgroup, the ROC analysis resulted in a multivariate AUC of 0.985 (95% confidence interval between 0.953 and 1.000) and accuracy of 93.8% (Figure VI-10B, Table VI-7). For the males, we observed multivariate AUC of 0.942 (95% confidence interval between 0.859 and 0.992) and accuracy of 86.6% (Figure VI-10C, Table VI-7). The performance for females was very similar to the results found when all samples were employed, with multivariate AUC within the 95% confidence interval for the complete dataset (AUC of 0.989 with 95% CI between 0.965 and 1.000), but the performance for males was not as good.

The male and female subsets were also distinctly classified as PDND or PDD by Random Forest using the 5-lipid biomarker panel (Table VI-4). The ROC curve for the PDND/PDD classification using only male subjects resulted in a multivariate AUC of 1.000 (95% CI between 0.995 and 1.000) and accuracy of 94.8% (Figure VI-11B, Table VI-8). For female patients, the multivariate AUC was 0.960 (95% CI between 0.813 and 1.000) and accuracy of 88.6% (Figure VI-11C, Table VI-8). The ROC classification of dementia status divided by sex resulted in slightly better performance for males when compared to the results obtained regardless of sex (multivariate AUC of 0.972 with 95% confidence interval between 0.919 and 1.000, and accuracy of 91.9%). However, the multivariate AUC values for separated subgroups for males and females were within the 95% confidence interval for the complete dataset. Hence, the separation between males and females is not required for the accurate diagnosis of PDD using the 5-lipid biomarker panel. However, we highly recommend sex-matching for future research.

6.3.4.7 Time since PD diagnosis

The PDND and PDD subgroups were statistically matched regarding disease duration, *i.e.*, time since PD diagnosis (p of 0.2). Nevertheless, PD progression and dementia development are closely related; thus, we investigated the lipidome differences between patients diagnosed (1) below the median of 6.5 years before baseline sample collection (16 PDND and 6 PDD) and (2) above the median value (11 PDND and 10 PDD).

6.3.4.8 Time since PD diagnosis: effect on the lipidome of PD patients

Univariate analysis through a volcano plot resulted in 37 lipids with significantly increased intensity ratios for the subgroup diagnosed for more than 6.5 years at sample collection (FC for below/above median <0.84 and p ≤0.05), but only seven showed decreased values (FC for below/above median ≥1.2, Figure VI-7, Supp. Table 24). Out of the 44 increased lipids, we identified 9 phosphatidylcholines, 12 phosphatidylethanolamines, 3 phosphatidylserines and 5 ceramides. Such compounds might be related to PD progression; however, further investigation is needed to confirm this hypothesis. Interestingly, eight of the significantly increased lipids for the subgroup diagnosed for longer periods (FC below/above median <0.84) also had significantly smaller intensity ratios for the PDD subgroup (FC PDND/PDD ≥1.2). None of 7 lipids selected for the control/PD biomarker panel (Table VI-4, p. 232) were significantly altered for time since PD diagnosis. However, one of the 5 lipids selected for the PDND/PDD biomarker panel displayed significant FC (below/above median) ≥1.2, identified by MS/MS as PC(18:1/18:1) (P786.59935/12.92, p of 0.049). Although the p value was just barely below the arbitrary significance limit of p<0.05, the lipid was further evaluated by 2-factor ANOVA but no significant interaction between dementia status and time since PD diagnosis was observed (p of 0.40), *i.e.*, no

interference between the below/above the median status and dementia classification. There was a slight separation for the multivariate statistical analysis by PCA, but a PLS-DA model with 6 components resulted in R^2 of 0.9752, Q^2 of 0.8554 and p of 0.03 for 1000 permutations (Figure C-9).

6.3.4.9 Time since PD diagnosis: effect on the PDD diagnosis

The two-way ANOVA interaction model for dementia status (PDND/PDD) and time since PD diagnosis (below and above the median of 6.5 years) resulted in 91 lipids with significant interference ($p < 0.05$, Supp. Table 26). The five lipids selected for the PDND/PDD biomarker panel were not significantly affected by the interaction between dementia status and time since PD diagnosis ($p > 0.05$, Table VI-4).

The ROC curve using the 5-lipid PDND/PDD biomarker panel (Table VI-4) to classify samples as above or below the median value of 6.5 years since PD diagnosis showed multivariate AUC of 0.750 and accuracy of 68.1%. The univariate AUCs for the 5 lipids in the biomarker panel were smaller than the values originally found for the PDND/PDD classification (Figure VI-12A, Table VI-6). The low values indicate that the median-split classification for time since PD diagnosis employing the 5-lipid PDND/PDD biomarker panel was rather random.

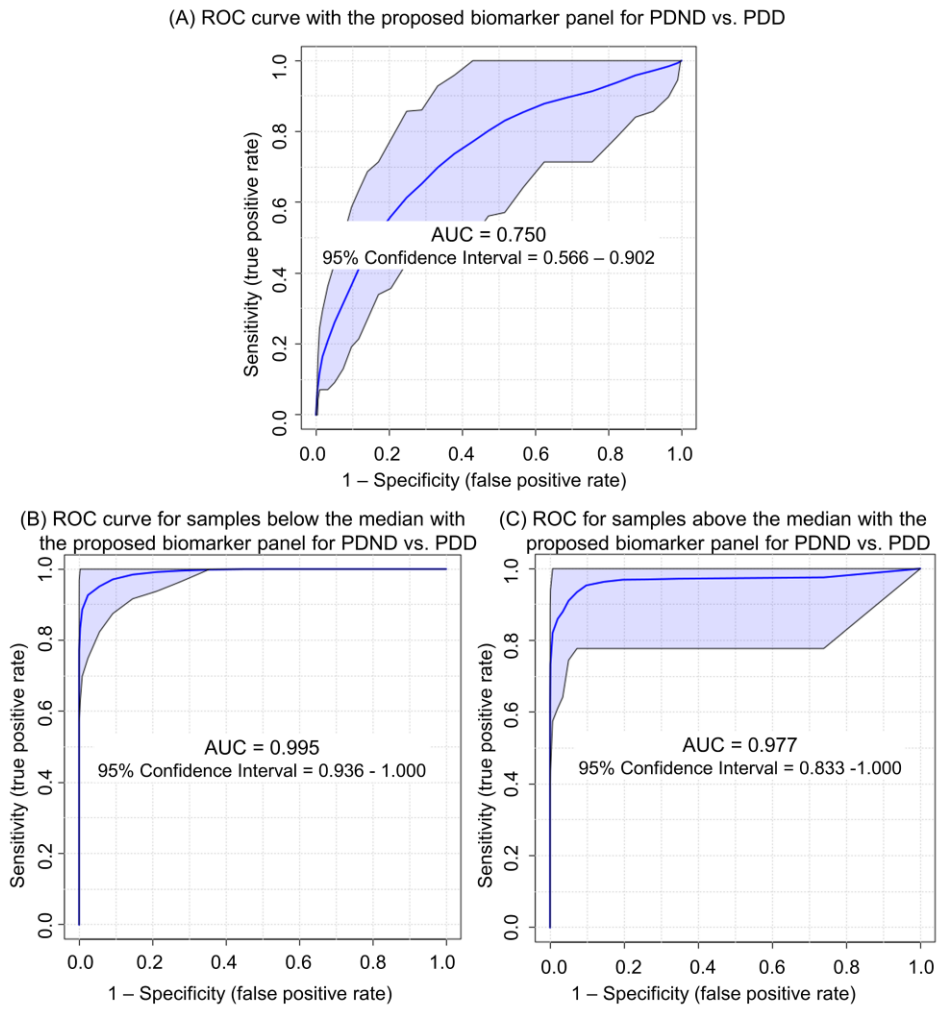

Figure VI-12. ROC curves for PD patients, including PDND and PDD, using the 5-lipid biomarker panel (Table VI-4, p. 232) to evaluate time since PD diagnosis as a confounding factor. The biomarker panel was used to classify: (A) all PD samples as above or below the median of 6.5 years since PD diagnosis at baseline sample collection, with sensitivity of 75.0%, specificity of 71.4% and overall accuracy of 68.1%; (B) only samples from PDND and PDD patients diagnosed with PD for less than the median of 6.5 years before sample collection, with sensitivity of 100%, specificity of 100% and accuracy of 92.8%; (C) only samples from PDND and PDD patients diagnosed with PD for more than 6.5 years before sample collection, with sensitivity of 100%, specificity of 96.7% and accuracy of 99.3% by cross-validation.

The ROC analysis for PDND/PDD classification for each subgroup resulted in similar multivariate AUC values that were within the 95% confidence interval of the complete dataset (0.919 to 1.000, Figure VI-6C). For patients diagnosed with PD for more than 6.5 years at sample collection, we found multivariate AUC of 0.977 and accuracy of 99.3% (Figure VI-12C, Table VI-8). For patients diagnosed for less than the median of 6.5 years, the multivariate AUC was 0.995 with an accuracy of 92.8% (Figure VI-12B, Table VI-8). The results were slightly better for the subgroup above the median. If dementia emerges in PD, it does so over extended periods. Regarding the current results, a reasonable expectation is that longer follow-up periods (beyond the three years in the present study) would produce more distinct time-based profiles.

6.3.4.10 Medication (Levodopa equivalent dosage and period of treatment)

Levodopa, also known as L-DOPA, is a dopamine precursor that can cross the blood-brain barrier, allowing for higher dopamine concentrations and, consequently, mitigation of PD symptoms. The compound (or similar versions) is a component of most PD treatments. All patients in this study were treated with Levodopa or equivalent drugs. PDND and PDD patients were not given significantly different dosages (p of 0.41, Table VI-1), but PDD patients received treatment for longer periods (p of 0.02). Therefore, statistical models were developed to evaluate the effect of medication dosage and treatment period for the PD group. First, the PD patients were divided into two subgroups delimited by the median dosage of 609.25 mg (16 PDND and 4 PDD patients below; 10 PDND and 10 PDD patients above the median). Second, PD patients were divided into two subgroups delimited by the median period of 2.96 years receiving treatment at the time of sample collection (17 PDND and 5 PDD patients below; 10 PDND and 11 PDD patients above the median).

6.3.4.11 Medication (Levodopa equivalent dosage and period of treatment): effect on the lipidome of PD patients

Univariate analysis through a volcano plot resulted in 16 lipids with significantly increased intensity ratios for the subgroup taking dosages above the median (FC below/above the median ≤0.84 and p ≤0.05), and 25 lipids with significantly decreased ratios, including 8 phosphatidylcholines, 5 phosphatidylethanolamines, 5 phosphatidylserines, 3 fatty acids and 3 N-acyl amines (Supp. Table 24). Four lipids were also significant for the PDND/PDD comparison, but none were selected for the 5-lipid PDND/PDD biomarker panel (Table VI-4, p. 232). There was no visual separation for the multivariate statistical analysis by PCA, and a PLS-DA model with 5 components resulted in R^2 of 0.9517, Q^2 of 0.7392 and p of 0.003 (1000 permutations) (Figure C-10).

The period receiving levodopa-based treatment did not result in major differences. Only two lipids were significantly altered for the below/above the median comparison through a volcano plot (Supp. Table 24), both with FC (below/above the median) ≥1.2, but neither was selected for the 5-lipid PDND/PDD biomarker panel (Table VI-4, p. 232). One lipid was also significantly altered for the PDND/PDD comparison (FC PDND/PDD <0.84), but there was no overlap between the significantly altered lipids for medication dosage and period of treatment. There was no visual separation for the multivariate statistical analysis by PCA for treatment period, and a PLS-DA model with 6 components resulted in R^2 of 0.9734, Q^2 of 0.7986 and p of 0.02 (1000 permutations) (Figure C-11).

6.3.4.12 Medication (Levodopa equivalent dosage and period of treatment): effect on the PDD diagnosis

The two-way ANOVA for Levodopa dosage and dementia status resulted in only one lipid with significant interference (interaction $p < 0.05$, Supp. Table 26). However, 17 showed interference for the period of treatment and dementia status ($p < 0.05$, Supp. Table 26), *i.e.*, the intensity ratios for PDND or PDD patients is affected by the period of treatment for the selected lipids. None of them were selected for the 5-lipid PDND/PDD biomarker panel (Table VI-4, p. 232).

The ROC curve using the 5-lipid PDND/PDD biomarker panel (Table VI-4) for classification of PD samples as below or above the dosage median resulted in a multivariate AUC of 0.602 and accuracy of 59.1% (p of 0.005 for 1000 permutations, Figure VI-13A), indicating random classification. For the period of treatment, we found a multivariate AUC of 0.711 and an accuracy of 66.9% (Figure VI-14A, Table VI-6).

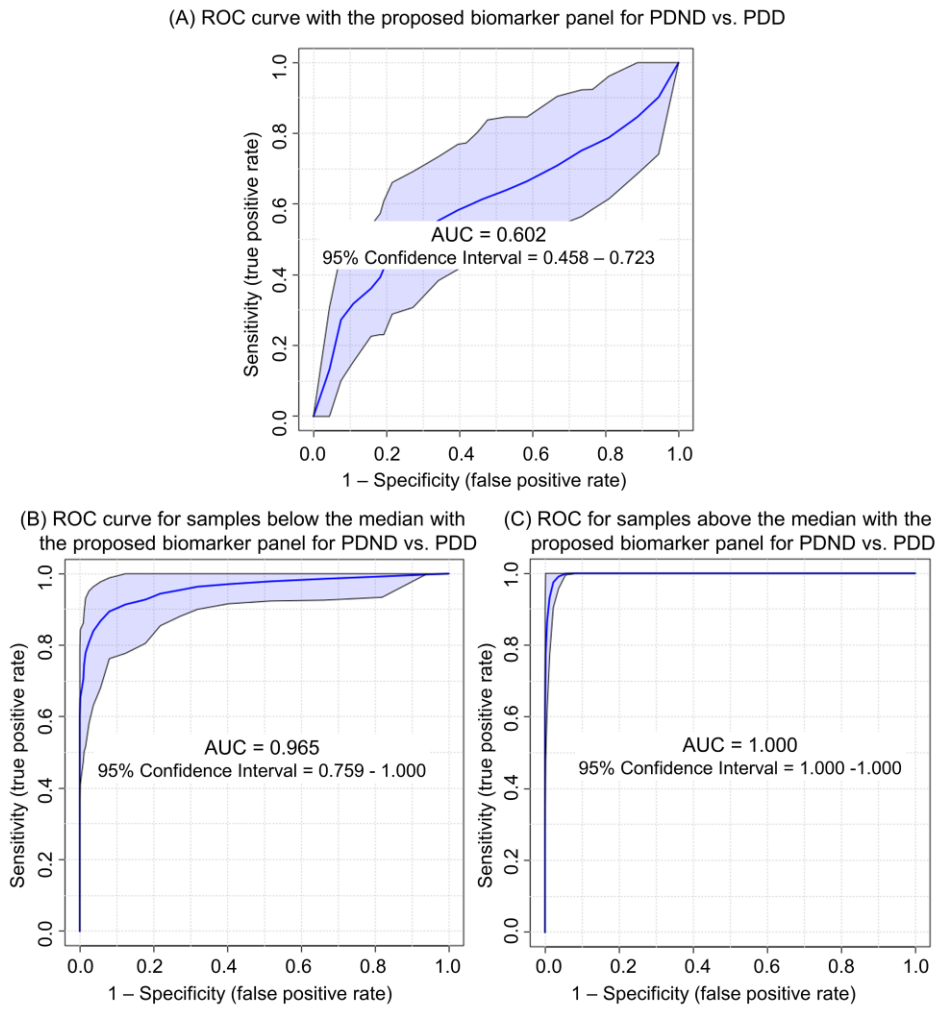

Figure VI-13. ROC curves for PD patients, including PDND and PDD, using the 5-lipid biomarker panel (Table VI-4, p. 232) to evaluate Levodopa equivalent dosage as a confounding factor. The biomarker panel was used to classify: (A) all PD samples as above or below the median of 609.3 mg of Levodopa or equivalent medication, with sensitivity of 67.5%, specificity of 60.0%, and overall accuracy of 59.1%; (B) samples from PD patients taking levodopa equivalent dosages smaller than the median of 609.3 mg as PDND or PDD, with sensitivity of 100%, specificity of 87.5% and accuracy of 82.1%; (C) samples from PD patients taking Levodopa equivalent dosages higher than the median of 609.3 mg at the moment of sample collection as PDND or PDD, with sensitivity of 100%, specificity of 100.0% and accuracy of 98.9%.

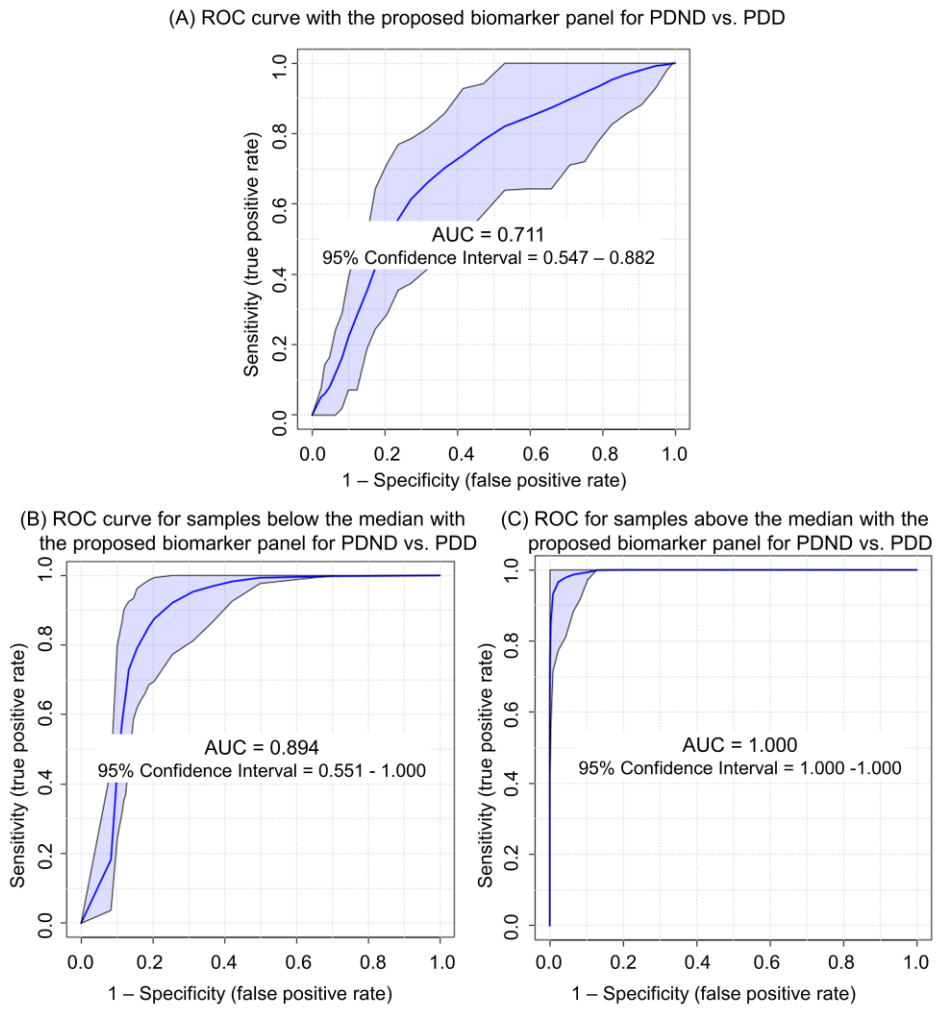

Figure VI-14. ROC curves for PD patients, including PDND and PDD, using the 5-lipid biomarker panel (Table VI-4, p. 232) to evaluate to evaluate period of treatment with Levodopa or equivalent medications as a confounding factor. The biomarker panel was used to classify: (A) all PD samples as above or below the median of 2.96 years of treatment with Levodopa or equivalent medication at the moment of sample collection, with sensitivity of 63.6%, specificity of 71.4%, and overall accuracy of 66.9%; (B) samples from PD patients taking levodopa or equivalent medication for less than 2.96 years at sample collection as PDND or PDD, with sensitivity of 80%, specificity of 88.2% and accuracy of 76.6%; (C) samples from PD patients taking Levodopa or equivalent

medication for more than 2.96 years at the moment of sample collection as PDND or PDD, with sensitivity of 100%, specificity of 100.0% and accuracy of 97.1%.

The ROC curves for classification of each distinct dosage subgroup as PDND or PDD resulted in multivariate AUC of 0.965 and accuracy of 82.1% for the subgroup below the median (Figure VI-13B, Table VI-8); and AUC of 1.000 and accuracy of 98.9% for above the median (Figure VI-13C, Table VI-8). For the treatment period subgroups, the PDND/PDD classification with the proposed 5-lipid biomarker panel resulted in multivariate AUC of 0.894 and accuracy of 76.6% for below the median (Figure VI-14B, Table VI-8); and multivariate AUC of 1.000 and accuracy of 97.1% for above the median (Figure VI-14C, Table VI-8). The results for the Levodopa dosage and treatment period subgroups above the median were slightly better than the values found for all the diseased subjects (AUC of 0.972 and accuracy of 91.9%, Figure 2E), but AUC values were within the 95% confidence interval for the complete dataset of 0.919 to 1.000 (Figure VI-6C, p. 235), except for the treatment period below the median.

We found evidence that the lipidome and dementia diagnosis is mildly affected by medication dosages and period of treatment. However, it is not viable or ethical to perform studies for advanced PD with untreated patients. Even though PD patients usually do not require treatment in the early stages, PDD is a condition that develops as the disease progresses. PD symptoms can be severe, and it would be unethical to keep impaired human patients suffering without treatment when it is available. The only option to completely exclude confounding factors is through studies with animal models or *in vitro*, but the results are often not reproducible in human patients. We previously published a metabolomics study using the same sample set that excluded any compounds involved in the dopamine pathway to minimize the Levodopa interference, but lipid

pathways and interactions with other molecules are not yet well known, which hinders the possibility of exclusion for this work. Nevertheless, diagnosis tests for PDD will probably be applied to patients that are receiving treatment for PD; hence, the elimination of this confounder for biomarker discovery would be unreasonable. [229]

6.3.4.13 UPDRS part III (motor examination) score

The Unified Parkinson's Disease Rating Scale (UPDRS) is a series of questions developed for patients and caregivers and used to classify the impairment level of PD patients. In this study, only Part III – Motor Examination was considered, covering 14 items scored from 0 to 4. PDND and PDD patients were not significantly different regarding the UPDRS-III score (p of 0.56, Table VI-1); however, the impairment level may be related to disease development. Therefore, the PD group was once again divided into two subgroups, delimited by the median UPDRS-III score of 15.0 (14 PDND and 9 PDD patients below the median; and 13 PDND and 7 PDD patients above the median).

6.3.4.14 UPDRS - part III (motor examination) score: effect on the lipidome of PD patients

Univariate analysis through a volcano plot resulted in three wax esters with significantly decreased intensity ratios for the subgroup with higher scores, *i.e.,* more severe motor impairment (FC below/above the median ≥1.2 and $p \leq 0.05$, Supp. Table 24). None of the studied lipids were also significantly affected for the PDND/PDD comparison nor selected for the 5-lipid biomarker panel (Table VI-4, p. 232). There was no visual separation for the multivariate statistical analysis

by PCA, but a PLS-DA model showed a partial separation with 6 components, resulting in R^2 of 0.9714, Q^2 of 0.7838 and p of 0.03 for 1000 permutations (Figure C-12).

6.3.4.15 UPDRS - part III (motor examination) score: effect on PDD diagnosis

The two-way ANOVA analysis resulted in five lipids with significant interference between the UPDRS-III score and dementia status (Supp. Table 26). The 5 lipids in the PDND/PDD biomarker panel displayed p >0.05 for the interaction, *i.e.,* the null hypothesis of equal means for the biomarker candidates for PD patients with UPDRS-III above or below the median cannot be discarded (Table VI-4, p. 232).

A ROC curve using the 5-lipid PDND/PDD biomarker panel (Table VI-4) for the classification of samples as below or above the UPDRS-III median resulted in a multivariate AUC of 0.702 and accuracy of 64.0% (Table VI-6, Figure VI-15A). Univariate AUC values were smaller than the ones found for the PDND/PDD classification (0.502 to 0.639); therefore, the severity of motor impairment did not interfere with the proposed 5-lipid biomarker panel.

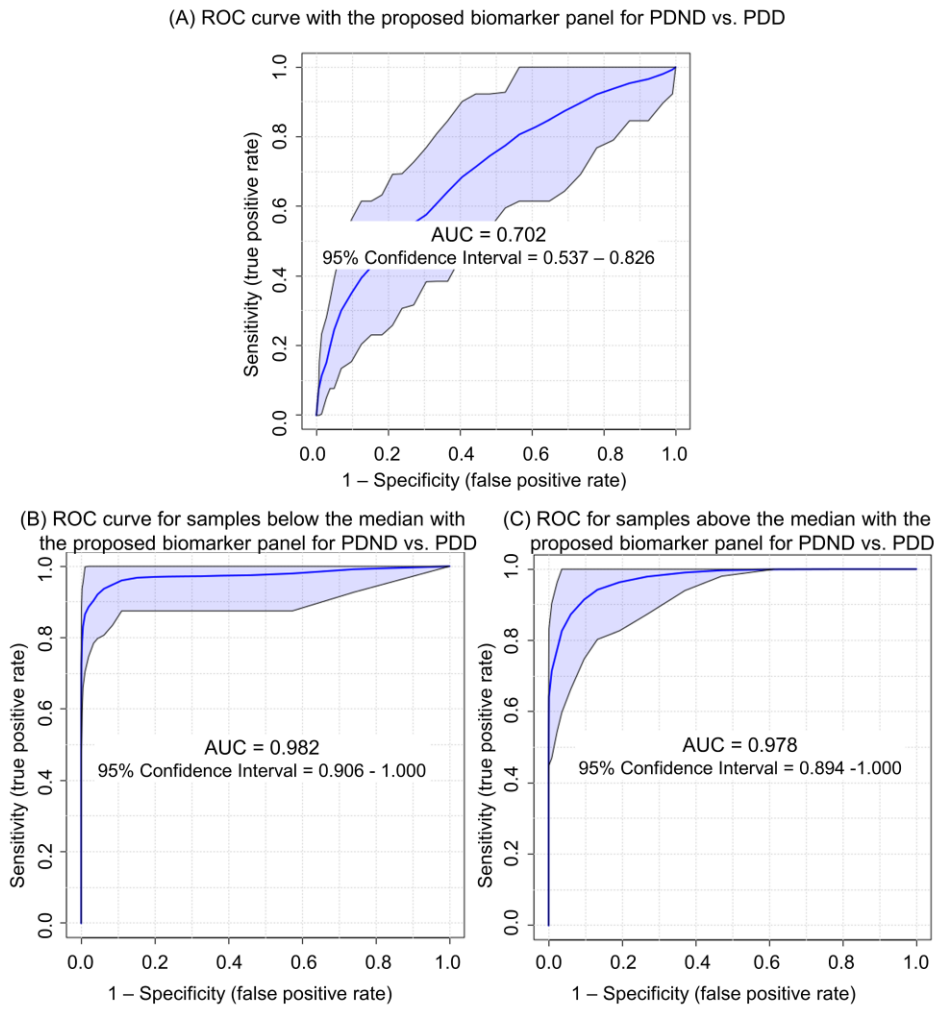

Figure VI-15. ROC curves for PD samples, including PDND and PDD, using the 5-lipid biomarker panel (Table VI-4, p. 232) to evaluate UPDRS – part III (motor examination) as a confounding factor. The biomarker panel was used to classify: (A) all PD samples as above or below the UPDRS - part III (motor examination) median score of 15, with sensitivity of 65.2%, specificity of 60.0% and overall accuracy of 64.0%; (B) only PD samples from patients with UPDRS – part III scores below the median of 15 as PDND or PDD, with sensitivity of 100.0%, specificity of 92.9% and accuracy of 93.3%; (C) only PD samples from patients with UPDRS – part III scores above the median of 15 as PDND or PDD, with sensitivity of 100.0%, specificity of 88.5% and accuracy of 87.3%.

The ROC curve constructed for the classification of the subgroup below the median as PDND or PDD led to an AUC of 0.982 and accuracy of 93.3% (Figure VI-15B, Table VI-8); whereas the subgroup above the median resulted in an AUC of 0.978 and accuracy of 87.3% (Figure VI-15C, Table VI-8). Both subgroups had a good classification performance, but AUC values were within the 95% confidence interval for the complete dataset (AUC of 0.972 with 95% confidence interval between 0.919 and 1.000, Figure VI-6C). Interestingly, the subgroup with UPDRS-III scores above the median of 15.0, *i.e.*, more severe motor impairment, resulted in slightly lower accuracy for PDD diagnosis than the patients with lower scores.

6.3.4.16 B12 levels

Vitamin B12 deficiency is common in PD patients, and there is some evidence linking low levels with more severe motor and cognitive decline. For this study, the healthy controls had significantly higher levels of B12 when compared to PD patients (p of 0.008), but PDND and PDD patients were not significantly different (p of 0.87, Table VI-1). Therefore, the healthy controls and PD individuals were separated at the median value of 304.5 ng/mL (16 controls and 27 PD patients below the median; and 27 controls and 16 PD patients above the median).

6.3.4.17 B12 levels: effect on the lipidome of PD patients

Univariate analysis through a volcano plot resulted in 15 lipids with significantly decreased intensity ratios for the subgroup with higher B12 levels (FC below/above the median ≥1.2 and p ≤0.05), including 5 phosphatidylserines, 3 fatty acids and 3 wax esters (Supp. Table 23). Conversely, 51 lipids had significantly increased ratios (FC below/above the median <0.84), including 13 phosphatidic acids, 10 phosphatidylglycerols, 8 phosphatidylcholines, 7

phosphatidylinositols and 4 phosphatidylserines. The 7 lipids in the control/PD biomarker panel were not significantly affected by B12 levels (Table VI-4, p. 232). There was no visual separation for the multivariate statistical analysis by PCA, but a PLS-DA model showed a partial separation with 8 components, resulting in R^2 of 0.9534, Q^2 of 0.6422 and p of 0.001 for 1000 permutations (Figure C-13).

6.3.4.18 B12 levels: effect on PD diagnosis

The two-way ANOVA analysis resulted in six lipids with significant interference between the B12 levels and disease status (Supp. Table 25). The 7 lipids in the control/PD biomarker panel and the 5 lipids in the PDND/PDD biomarker panel (Table VI-4, p. 232) displayed interaction p >0.05, *i.e.*, no significant interference.

A ROC curve using the 7-lipid control/PD biomarker panel (Table VI-4) for the classification between below and above the B12 level median resulted in a multivariate AUC of 0.749 and an accuracy of 68.2% (Table VI-5, Figure VI-16A). The ROC curve constructed for the classification of the subgroup below the median as control or PD led to multivariate AUC of 0.949 and accuracy of 86.9% (Figure VI-16B, Table VI-7); whereas the subgroup above the median resulted in multivariate AUC of 0.993 and accuracy of 95.7% (Figure VI-16C, Table VI-7). The median-split subgroup with higher B12 levels produced a slightly better classification, although within the 95% confidence interval for the complete dataset (AUC of 0.989 with 95% confidence interval between 0.965 and 1.000). The subgroup with lower levels had the opposite effect, with AUC slightly below the 95% confidence interval.

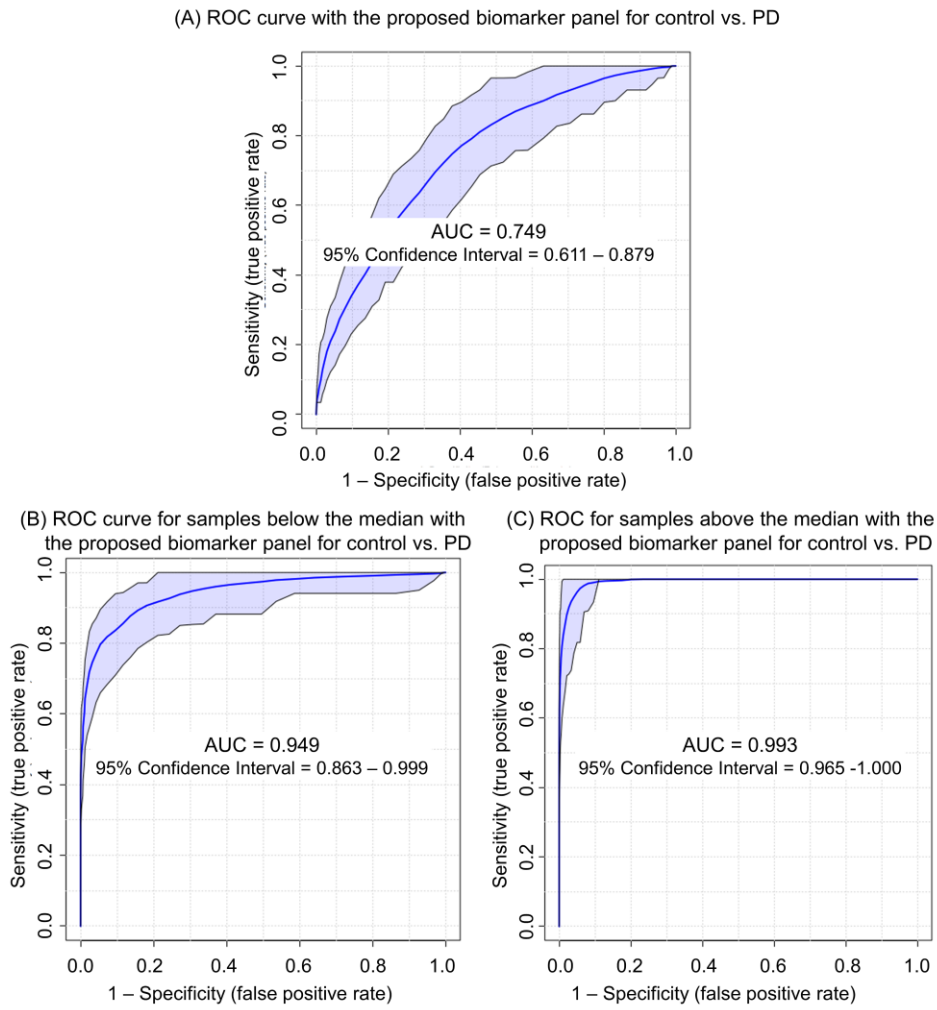

Figure VI-16. ROC curves for control and PD samples using the 7-lipid biomarker panel (Table VI-4, p. 232) to evaluate B12 serum levels as a confounding factor. The biomarker panel was used to classify: (A) all samples as above or below the median of 304.5 ng/mL for B12 serum level, with sensitivity of 72.1%, specificity of 62.8% and overall accuracy of 68.2%; (B) only samples from individuals with B12 serum levels below the median of 304.5 ng/mL as control or PD, with sensitivity of 88.9%, specificity of 90.0% and accuracy of 86.9%; (C) only samples from individuals with B12 serum levels above the median of 304.5 ng/mL as control or PD, with sensitivity of 97.0%, specificity of 100.0% and accuracy of 95.7%.

6.4 Discussion

PD is a complex neurodegenerative disorder for which new biomarker approaches are being harnessed to improve early detection and prediction of dementia. We used untargeted lipidomics of blood serum applied in two phases. For the first phase, healthy control individuals were separated from PD patients. In the second phase, the PD group was divided into patients that either progressed to dementia (PDD) or did not show a significant cognitive decline (PDND) after 3 years. The LC-MS method, combined with the developed data processing routine, provided good lipid coverage with reproducible results. We constructed two distinct biomarker panels for (1) diagnosis of PD with 7 lipids; and (2) prediction of dementia using 5 lipids. Validation with independent sample cohorts resulted in excellent sensitivities, specificities and accuracies. The selected biomarker panels are promising and may improve the clinical diagnosis of PD and prediction of dementia.

In the first phase, 129 lipids were found to be significantly altered in the serum of PD patients when compared to healthy controls, including sphingomyelins, ceramides and sphingosine. The 7-lipid control/PD biomarker panel also emphasized the effect of PD upon the sphingolipid metabolism: out of seven selected lipids, four are sphingolipids. Our findings indicate that the regulation of sphingomyelin and ceramide metabolisms is affected by PD and might be useful diagnostic and therapeutic targets. It is often unclear whether significant changes in the levels of specific lipid species are a result of a pathological or a compensatory mechanism. Mutations in the glucosylceramidase beta (*GBA*) gene, which codes for the lysosomal enzyme glucocerebrosidase (GCase), are commonly associated with PD. *GBA* mutations have been previously related to autophagy-lysosome pathway dysfunction, which is one of the main degradation mechanisms for α-synuclein. The enzyme GCase catalyzes the break-down of

glucosylceramides into glucose and ceramide. Hence, the observed altered levels of ceramides may be linked to deficiency or decreased activity of the enzyme, possibly affected by the mutations in the *GBA* gene.[234] The accumulation of α-synuclein into Lewy bodies, a hallmark of Parkinson's disease development, is therefore correlated with altered levels of ceramides in blood serum. [234] It is still not clear if a GCase deficiency causes the accumulation of α-synuclein or an excess of the protein decreases GCase activity. However, dysregulation in ceramide metabolism has been previously suggested to contribute to the accumulation of α-synuclein protein, enhancing its toxicity. Abnormal concentrations of ceramides have also been linked to increased neuronal death, a characteristic marker of PD.[235–237] Furthermore, dopaminergic neurons regulate the activity of the DNase enzyme known as neutral sphingomyelinase, responsible for breaking sphingomyelin into phosphocholine and ceramide, which may be affected by neuronal death. [238] Our findings confirm that the regulation of sphingolipids and ceramide metabolisms is affected during Parkinson's disease and might be useful diagnosis and therapeutic targets. [237]

The glycerophospholipid metabolism was also affected by the control/PD comparison. The role of phospholipids for PD onset is still not clear, but recent studies suggest that they may play a part in α-synuclein folding and aggregation. The protein has an amphipathic α-helix domain at the N-terminus and a hydrophobic central domain, suggesting lipid-binding activity. The N-terminal domain adopts the helical structure upon interaction with lipid membranes, particularly the negatively charged species such as PS and PI. The interaction with PS, a subclass highly affected in the control/PD comparison (38 significantly elevated PS species for PD patients), has been previously shown to cause increased oligomerization of α-synuclein.[239] Hence, interaction with glycerophospholipids can control protein folding and modulate structural changes that could provoke or avoid aggregation. Also, α- and β-synucleins can inhibit phospholipases, enzymes that

catalyze the hydrolysis of phospholipids into lysophospholipids, phosphatidic acids, diacylglycerols and fatty acids. It is worth noting that 14 fatty acids had higher intensity ratios for the PD group, while none resulted in lower values. We also observed higher levels of 15 phosphatidic acids when PD was diagnosed, which is consistent with the expected alterations in phospholipase activity. Recessive mutations in the phospholipase PLA2G6 gene (PARK14), related to the cleavage of the phosphate group from phosphatidyl lipids, have also been previously related to early-onset and sporadic PD. [223,240,241]

In the second phase, the baseline discrimination of PDND and PDD (with incipient dementia) was examined. This approach led to insights into the transition of PD to dementia. Specifically, 114 lipids were significantly altered for the PDND and PDD subgroups. The glycerolipid, glycerophospholipid and ceramide levels were markedly affected in patients who developed dementia. Elevated levels of diacylglycerols, essential precursors of glycerophospholipids, have been previously reported in brain tissue for PD patients, as well as individuals with Lewy body dementia. [242] We now found evidence that reduced levels of both diacylglycerols and triacylglycerols in serum may be correlated with incipient dementia in PD patients, indicating alterations in energy metabolism. The impact of PD progression to dementia upon the glycerolipid metabolism contrasts with the first phase of this study: only 2 diacylglycerols and 9 triacylglycerols were significantly affected for the control/PD comparison, whereas 20 diacylglycerols and 25 triacylglycerols had significant changes for PDND/PDD. We also observed that glycerophospholipid metabolism was highly involved in PD progression to dementia, particularly for membrane lipids, which may play an important role in α-synuclein aggregation. The 5-lipid PDND/PDD biomarker panel reflects the previous observations, with two phosphatidylcholines, one lysophosphatidylcholine and one phosphatidylethanolamine.

An important aspect of this study is the evaluation of possible confounders, rarely presented in biomarker discovery studies. We evaluated the effects of age, sex, medication dosage and period, PD duration, motor deficit severity and B12 levels on the lipidome of healthy controls and PD patients, as well as the possible interference in the proposed biomarker panels. Fortunately, the biomarker panels were not majorly affected. However, we detected significant fold changes for lipids that were not part of the panels due to each confounding factor, emphasizing that an evaluation of confounders should be an essential step for biomarker discovery. Yet, such factors cannot be easily controlled, especially if the goal is to achieve a representative number of patients. Restrictions for only minor effects would substantially decrease the sample size of the cohorts available for future studies. Although we recommend matching the compared groups regarding age, sex, medication and disease progression, the benefits of excluding significant portions of the patient cohort to ensure reliable conclusion must be evaluated against the reduction in the number of available subjects.

Although this study presents very strong evidence for two novel biomarker panels discriminating healthy subjects from Parkinson's disease patients and predicting differential transition to dementia, a few limitations must be noted. First, stereoisomers and positions of unsaturation were not addressed, as these usually require targeted, specific methodologies. Instead, we chose to emphasize untargeted, comprehensive lipidomics. In the near future, we will confirm the identifications of the selected biomarkers by targeted MS/MS analysis. Second, biological and demographic differences for the patient cohort are potential confounders. We evaluated the effects of seven confounders, but the proposed biomarker panels were only mildly affected. Third, PD patients did not have autopsy confirmation of diagnosis as they were alive at the time this study was completed. PD and PDD diagnosis were based on clinical evaluation confirmed by the study

neurologist (Dr. Richard Camicioli, University of Alberta). Fourth, the specificity of the proposed biomarker panels has not yet been tested against similar conditions or on a second, large-scale patient cohort. This study aimed for an initial biomarker discovery evaluation, which obviously requires a follow-up to address possible issues and validate the proposed biomarkers. Nevertheless, the presented results demonstrate the potential of lipidomic analysis for biomarker discovery and pathogenesis study. Lipids are involved in a variety of different processes in our organism, granting the massive importance of evaluating their alterations to study physiological and pathological processes. The information described herein emphasizes that untargeted lipidomics may provide essential information for the selection of biomarkers and future therapeutic targets.

6.5 Conclusions

This work highlights that lipids play a more significant role in Parkinson's disease and dementia than previously recognized. We selected two biomarker panels that resulted in excellent AUC and diagnosis accuracy for PD and progression to dementia, surpassing the currently available clinical diagnosis. The two distinct biomarker panels allowed (1) the diagnosis of PD and (2) prediction of transition to dementia up to three years before clinical diagnosis. Validation of the proposed biomarkers showed an excellent performance. Furthermore, the untargeted lipidomics approach resulted in further clues to unravel the mechanisms of the pathologies. Our future work will extend the current results for blood serum with lipidomics of other fluids and tissues, as well as comparison to similar diseases, to further study the pathogenesis of PD and PD dementia.

6.6 Acknowledgments

We thank the contribution of Jaspaul Tatlay, Dorothea Mung, Barinder Bajwa (Department of Chemistry, University of Alberta) and Krista Nelles (Study Coordinator at the Faculty of Medicine & Dentistry of the University of Alberta), as well as the excellent collaboration with Dr. Richard Camicioli (Neuroscience and Mental Health Institute and Department of Medicine - Neurology) and Dr. Roger A. Dixon (Neuroscience and Mental Health Institute and Department of Psychology, University of Alberta, Edmonton, Canada). This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research (CIHR), the Canada Research Chairs (CRC) program, Genome Canada, Alberta Innovates and the Canadian Consortium on Neurodegeneration in Aging (CCNA).

VII

Chapter VII: Tissue Lipidomic Alterations Induced by Long-Term Exposure to High Dosages of Dexamethasone[15,16]

7.1 Introduction

Dexamethasone is an anti-inflammatory and immunosuppressant glucocorticoid synthetic drug, a class of corticosteroid medications composed by molecules analogous to steroid hormones produced by the adrenal cortex. Corticosteroids, including glucocorticoids and mineralocorticoids, regulate a myriad of essential biochemical and physiological processes in various tissues in the organism, with intense effects upon inflammation and immune response. The discovery of adrenal corticoid hormones in the 1950s was awarded a Nobel Prize in Physiology or Medicine due to the exceptional results achieved when patients with rheumatoid arthritis were treated with "Compound E", a synthetic version of the steroid hormone cortisol isolated from animal adrenal glands, at the Mayo Clinic (Rochester, MN, USA).[243] To this day, research into the immune regulation and adverse effects of synthetic glucocorticoids remains an active topic of research. [244]

Glucocorticoids play a role in the organism's response to stress and are part of the feedback mechanism of the immune response. Cortisol is a key example of natural corticosteroid action, which induces adaptive responses to stressful situations. Endogenous glucocorticoids are synthesized from cholesterol through steroidogenesis in the adrenal cortex upon activation of the

[15] Supporting tables are available with Dr. Liang Li.

[16] The sample collection and experimental design for this work was performed in collaboration with Dr. Anas M. Abdel Rahman and Dr. Majed Dasouki (King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia). The sample preparation was performed by the author (AZB) and Xiaohang Wang, who later worked on a related application for metabolomics of the aqueous fraction obtained after homogenization and liquid-liquid extraction of tissue samples.[3]

hypothalamic-pituitary-adrenal (HPA) axis, which coordinates endocrine responses to external stimuli, *e.g.*, psychological stress or trauma. The steroid molecules bind to glucocorticoid receptors (GR) in virtually all types of cells in the organism. GR modulate the immune response through alterations in gene expression, modulation of transcription factors and regulation of T cell activation, suppressing the production of inflammatory mediators such as eicosanoids. Low dosages of glucocorticoids have been linked to increased immune response, while high concentrations suppress the immune system. [243] Glucocorticoids also display non-genomic effects, including alteration of cation transport through membranes and cytoplasmic signaling complexes. Their synthetic versions, including dexamethasone, prednisone and prednisolone, are used to treat conditions related to an exaggerated immune response, *e.g.,* allergies, asthma, rheumatoid arthritis, anaphylactic shock and sepsis. They can also be used for cancer treatment due to an interference in the mechanism of abnormal cell proliferation in malignant tumors. Despite their beneficial effects, glucocorticoids increase the turnover of stored energy and mitochondrial oxidation, leading to dyslipidemia, obesity, insulin resistance and oxidative stress.[245] Systemic, long-term exposure to glucocorticoids also includes suppression of the HPA axis, causing an adrenal crisis.[246] The long-term production of high concentrations of cortisol or treatment with high dosages of synthetic glucocorticoids for long periods is the cause of Cushing's syndrome, a condition characterized by a fatty hump between the shoulders ("buffalo hump"), rounded ("moon") face, weight gain, slow healing of wounds, high blood pressure, heart disease and bone loss.[244]

Dexamethasone is a synthetic corticosteroid derived from cortisol used for the treatment of arthritis, allergic reactions, asthma, inflammation, septic shock, cerebral edema, anaphylactic reactions, eczema, psoriasis and selected types of cancer, amongst others. It acts as an agonist of GR in the cytoplasm of cells and has minor mineralocorticoid activity. The protective and adverse

effects of dexamethasone are mediated through activation of GR in the liver, adipose tissue, heart and skeletal muscle, as well as minor effects upon the mineralocorticoid receptors in the kidneys and heart.[245] GRs are expressed in most types of tissues; hence, the systemic metabolic effects of dexamethasone may affect different organs. Recently, dexamethasone has been suggested as a potential treatment for COVID-19, with a reported decrease in mortality rates between 8 and 26% for patients with hypoxemia.[247] Although the results were not peer-reviewed and more information is still required (Aug/2020), the preliminary results were received by governments with optimism and led to immediate alterations in the COVID-19 treatment protocols followed by the National Health Service (NHS) of the United Kingdom (Aug/2020).[248]

Even though dexamethasone is on the World Health Organization's List of Essential Medicines and is one of the most commonly prescribed drugs in the USA, it is accompanied by an extensive list of potential adverse effects.[249] The long-term, high-dosage use of dexamethasone has been related to bone loss, cataracts, glaucoma, myopathy, candidiasis, weight gain, euphoria, irritability, hyperactivity, psychosis, nausea, hypertension, persistent headache, growth suppression in children, and others.[244,246,250] Furthermore, sudden withdraw after long-term treatment may lead to hypotension, fever, adrenal insufficiency, cardiovascular collapse and death.[250] The effects of long-term glucocorticoid treatment upon skeletal muscle (weakness, atrophy, fatigue), heart (cardiac failure, ischemic heart disease, atrial fibrillation), brain (mood changes, euphoria, anxiety, psychosis, headaches), kidney (nephrocalcinosis, nephrolithiasis, increased uric acid) and liver (hepatic steatosis or "fatty liver," deposition of glycogen) are relatively well known.[3,244,251] However, the biological processes that cause these undesirable symptoms are yet to be explored.

Recently, the alterations caused by long-term, high-dosage dexamethasone treatment upon different tissues was investigated by proteomics and metabolomics in a rat model.[3,252] However, as blood serum dyslipidemia is a hallmark of excessive exposure to glucocorticoids, we hypothesized that untargeted, comprehensive lipidomics could provide further insight into the mechanisms of dexamethasone adverse effects in different types of tissue.[244] Little is known about glucocorticoid-induced lipidic alterations, besides reported elevated levels of serum total triglycerides and cholesterol.[253] There is some evidence that dexamethasone may promote lipid accumulation in muscle tissue, affect fatty acid β-oxidation (Chapter I – *1.2.4.2. Energy production*, p. 11) and induce oxidative stress.[253,254] To investigate the relationship between dexamethasone and tissue lipids, we performed comprehensive, untargeted lipidomic analysis of skeletal muscle, liver, kidney, heart and brain tissue obtained from rats that were injected with high dosages of dexamethasone or saline solution (controls) for 14 weeks to promote controlled adverse effects of prolonged treatment. The lipidic composition of the selected types of tissue may provide further information into the biological mechanisms that cause the undesirable effects of dexamethasone, as well as insights into glucocorticoid-related pathways.

7.2 Experimental

7.2.1 Animal model[17]

The protocol employed for this study was approved by the animal ethics committee of the King Faisal Specialist Hospital and Research Center (KFSHRC, approval number 2150016). Male

[17] Animal handling and sample collection were performed in collaboration with Dr. Anas M. Abdel Rahman and Dr. Majed Dasouki (King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia).

Sprague-Dawley rats (n = 20) aged between 6 and 8 weeks were housed in the animal facility of the Department of Comparative Medicine (KFSHRC) under standard temperature (20-24°C), humidity (45-50%) and light/dark cycle (12h) conditions, as previously described elsewhere.[3,253] Animals had free access to food and water for the duration of this study. The rats were separated into two groups of 10 animals each. The dexamethasone group received intramuscular injections of 2.5 mg/kg of dexamethasone twice a week for 14 weeks, while the control group was injected with saline solution. Both groups were monitored weekly for weight and blood sugar (glucose level). After week 14, the animals were sacrificed. The brains (5 samples from control animals and 4 samples from dexamethasone-treated rats), hearts (5 control and 5 dexamethasone rats), kidneys (6 control and 5 dexamethasone rats), livers (4 control and 4 dexamethasone rats) and samples of skeletal muscle (6 control and 5 dexamethasone rats) were collected, snap-frozen in liquid nitrogen and stored at -80°C until analysis. The remaining animals for each group were sacrificed for a related study that included radiological examinations and blood work.[253]

7.2.2 Sample preparation

The tissue extraction was performed by a modified Folch procedure that included homogenization followed by liquid-liquid extraction with dichloromethane and methanol.[3] Chemicals and reagents were described in Chapter II (*2.2.1. Chemicals and reagents*, p. 42). Sample preparation and analysis were randomized to minimize batch effects. Reagent and solvent volumes were normalized for the wet mass of tissue to ensure reproducible extractions and reliable comparisons. All samples were extracted in 2 mL polypropylene microcentrifuge tubes obtained from the same manufacturer with identical lot numbers to prevent contamination bias. A detailed

discussion on sources of contamination for lipidomic analysis is available in Chapter II (*2.3.3. Evaluation of contamination sources*, p. 63).

The wet masses of each sample were measured between 160 and 320 mg. Since tissues are solid samples that may not have a homogenous distribution of lipids, we opted for not splitting the samples for full experimental replicates. Instead, the biological replicates of each type of tissue were homogenized and extracted altogether (N = 4 to 6). Samples were mixed with the lipid internal standard mixture (Table II-1, 0.140 µL/mg of tissue – wet mass), methanol (1.86 µL/mg) and water (0.425 µL/mg). Then, each sample was immediately homogenized twice in an ice bath with a Bio-Gen PRO200 handheld homogenizer fitted with a Multi-Gen adapter and Multi-Gen 7XL probes (Pro Scientific, Inc., Monroe, CT, USA) for 30 s, with a 30 s interval. The homogenate was vortexed with 2.00 µL of dichloromethane and 1.00 µL of water/mg of tissue. After equilibration for 15 min in an ice bath, samples were centrifuged for 10 min at 4 °C and 12,000 rpm. The aqueous supernatant was used for metabolomics, as described elsewhere.[3] The organic layer was evaporated to dryness on a SpeedVac for 45 min. The extract was resuspended in 0.4 µL/mg of 3:2 mobile phase A (MPA)/mobile phase B (MPB). For positive ionization, samples were further diluted with 10 µL of 9:1 MPA/MPB /mg of tissue, whereas for negative ionization, we used 1 µL of 9:1 MPA/MPB /mg of tissue. After dilution, aliquots from each sample extract were pooled together for a quality control (QC) mixture. The prepared samples were stored in polypropylene inserts placed in amber autosampler vials with PTFE-lined caps at 4 °C for a maximum of 48h before injection.

7.2.3 Lipidome profiling by UHPLC-MS

Chromatographic separation was performed using a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific, Waltham, MA, USA) with a Waters Acquity BEH C18 column (5 cm × 2.1 mm with 1.7 μm particles, Waters Corporation, Milford, MA, USA), using MPA – 10 mM ammonium formate in 50:40:10 acetonitrile/methanol/water (v/v/v), and MPB – 10 mM ammonium formate in 95:5 isopropanol/water (v/v). Comprehensive lipid profiling was achieved with the following conditions: injection volume of 5.0 μL for positive ionization and 9.0 μL for negative ionization; flowrate of 250 μL/ min; column temperature of 40°C; 22 min gradient separation (0 min – 5% MPB; 1.5 min – 5% MPB; 8.5 min – 30% MPB; 18 min – 95% MPB; 22 min – 95% MPB); and 10 min re-equilibration (0 min – 95% MPB; 3 min – 95% MPB; 4 min – 5% MPB; 10 min – 5% MPB). The UHPLC instrument was coupled to a high-resolution Maxis II quadrupole-time-of-flight (QToF) mass spectrometer (Bruker Daltonics, Billerica, MA, USA) with an electrospray ionization (ESI) source (capillary voltage of 4500 V; endplate offset of 500 V; nebulizer gas pressure of 1.0 bar; dry gas flow rate of 8.0 L/min; dry temperature of 230 °C; spectra acquisition rate of 1 Hz; and m/z range of 150 to 1500 Da). Samples and blank extracts were randomly injected in triplicates (*i.e.*, 4 to 6 biological replicates for each condition analyzed with injection triplicates), with one QC injection after every 10 sample injections. Positive and negative ionization were acquired separately, with polarity switching for consecutive groups of 10 sample injections and one QC, totaling 6 injections for each sample extract. A 1.5 min mass re-calibration segment was inserted at the beginning of each chromatogram during which 1.0 mM sodium formate calibrant solution in 1:1 isopropanol/ water (v/v) was infused with a peristaltic pump to ensure high mass accuracy.

7.2.4 Data processing

Chromatograms were aligned on MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA) with a minimum intensity cut-off of 5000 cts for positive ionization and 2000 cts for negative ionization; minimum peak length of 6 spectra; mass recalibration with the sodium formate segment; m/z tolerance of 5.0 mDa; and retention time tolerance of 15 s. The aligned features were filtered by detection in more than 80% of injections for at least one type of tissue (evaluation of tissue composition) or treatment status (dexamethasone or control). Missing values were substituted by recursive extraction for features detected in at least 10% of injections with a minimum peak length of 6 spectra. The positive and negative ionization results were merged with an m/z tolerance of 10 ppm and retention time tolerance of 15 s. The remaining missing values were substituted by the minimum within-group intensity for features detected in more than 50% of injections in each group (type of tissue and treatment status). If not found in at least 50% of injections within the group, missing values were substituted by the global minimum intensity for all samples and QC injections.

7.2.5 Lipid identification

MS/MS spectra were acquired for identification using the pooled QC sample with collision energies between 10 and 80 eV. The MS/MS spectra were matched to the aligned feature list with a precursor m/z tolerance of 5.0 mDa and retention time tolerance of 15 s. Identification was performed with the MS-Dial LipidBlast (https://fiehnlab.ucdavis.edu/projects/LipidBlast), Human Metabolome Database (https://hmdb.ca) and MassBank of North America (https://mona.fiehnlab.ucdavis.edu) LC-MS/MS libraries in combination with MetaboScape 4.0, using precursor m/z tolerance of 5.0 mDa combined with MS/MS score threshold of 500 and

mSigma (isotope pattern match) tolerance of 100; or MS/MS score threshold of 100 and mSigma tolerance of 50. [73,78,79,81–83,109,151]

The unidentified features were searched in the LipidMaps database (http://www.lipidmaps.org) for putative identification by mass-match with m/z tolerance of 5.0 mDa.[75] Isomers were ranked by the filtering and scoring approach described in Chapter II *(2.2.7. Putative identification of lipids*). Isomers that had the lowest total scores, *i.e*., elution within the expected retention time range, most likely adduct, smallest m/z error and even number of carbons in fatty acyl side chains, were selected as the most likely identification for the choice of lipid subclass. Other isomeric and isobaric possibilities that passed the retention time and adduct filters were kept, but not considered for the determination of lipid subclass. The positions of double bonds and the stereospecific configuration of glycerol derivatives were not determined in this study.

The MS/MS or putatively identified lipids were divided into 41 subclasses and 6 main categories (sphingolipids, glycerolipids, glycerophospholipids, fatty acyls, sterol lipids and others, Table II-2, p. 48). The lipid subclasses and categories followed the classification system proposed by the International Lipid Classification and Nomenclature Committee (ILCNC), the LipidMaps database and the Lipidomics Standard Initiative (https://lipidomics-standards-initiative.org). [16–18,20,76]

7.2.6 Normalization and statistics

Each identified lipid was matched to one of the 14 deuterated internal standards (Table II-1, p. 45) for normalization according to lipid subclass, structural similarity and retention time. Normalized intensities, *i.e.*, the peak intensity of each identified lipid divided by the peak intensity of the matched internal standard, were filtered by relative standard deviation smaller than 30% for

QC injections, auto-scaled and normalized by the summed intensity (*i.e.* total ion count) before statistical analysis.[153] The normalization procedure adopted for this work allows the correction of the ion suppression effect caused by the complex lipid samples during electrospray ionization, as well as small differences that may occur during sample handling and lipid intensity variations between different types of samples, ensuring accurate comparisons.[12] The identification and normalization steps followed the guidelines of the Lipidomics Standard Initiative. [76] A detailed discussion on normalization and scaling strategies for lipidomics is available in Chapter I (*1.3.1.6. Ion suppression and normalization approaches* and *1.3.1.7. Statistics applied to lipidomics*, p. 29).

Statistical analysis was performed on MetaboAnalyst 4.0 (https://www.metaboanalyst.ca).[153] Normalized, auto-scaled intensity ratios were examined through univariate and multivariate analysis, including non-parametric Analysis of Variance (ANOVA), Volcano plots (fold change – FC *versus* p adjusted for false discovery rate –p), Principal Component Analysis (PCA), Partial Least Square – Discriminant Analysis (PLS-DA) and Random Forest. For univariate analysis (non-parametric tests for unequal variances), lipids were considered significant for p (adjusted for false-discovery rate) <0.05 and FC ≥ 1.5 or ≤ 0.67.

7.3 Results

The untargeted LC-MS lipidomics approach employed for this work allowed the detection of 6856 features for positive ionization and 2318 for negative ionization. The aligned features were merged into a unique list of 8765 features. Thirteen deuterated internal standards were detected with a maximum m/z error of 7.2 ppm. The glycerophosphoinositol lipid standard PI d7-15:0/18:1 was not detected due to a small concentration in the commercial mixture employed for this study (Avanti Splash Lipidomix Mass Spec Standard, Table II-1, p. 45). A total of 63.1% of the detected

features were identified in 34 lipid subclasses, including 381 by MS/MS spectral match, 5132 by putative mass match and 20 internal standard ions for positive and negative ionization (Supp. Table 27). More than 40% of the identified lipids were glycerophospholipids, while 16.4% were glycerolipids, 17.8% were sphingolipids, 10.3% were fatty acyl species and the remaining 10.4% were sterols (Figure VII-1). Eight features were not identified as lipids, *i.e.*, medications, phthalates and contaminants.

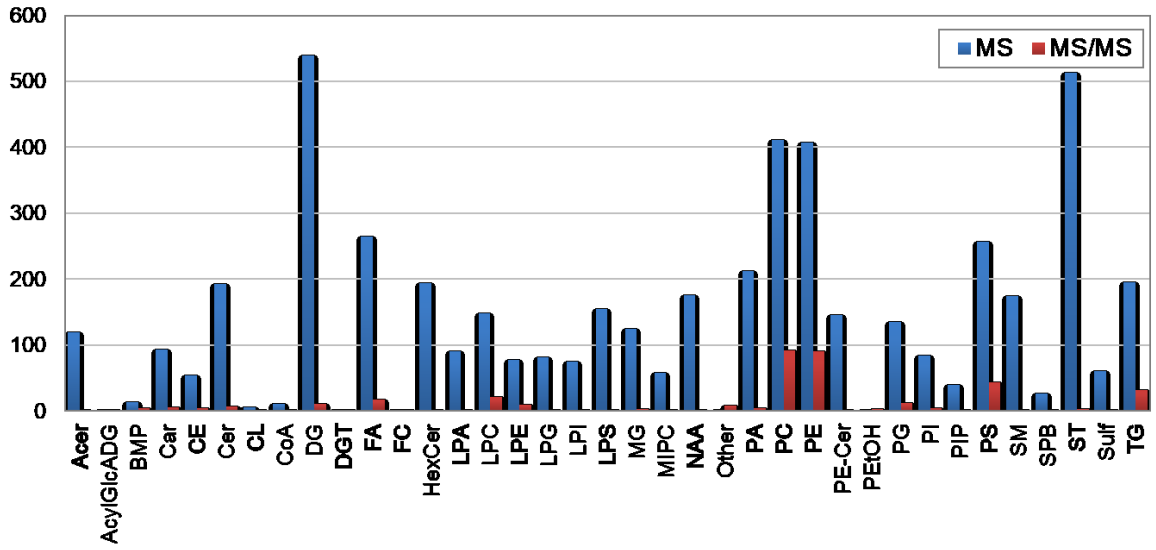

Figure VII-1. Subclass distribution for the lipids identified for liver, kidney, heart, brain and skeletal muscle samples from male rats treated with dexamethasone or normal saline solution (intramuscular injections) for 14 weeks. Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

The normalization procedure using a deuterated internal standard mixture for subclass-match greatly improved the relative standard deviation (RSD) calculated for quality control (QC) replicates. Only 72.0% of features displayed RSD smaller than 50% for the raw data, but the number was improved to 97.4% after normalization. These results show the importance of

adopting appropriate normalization procedures for lipidomics analysis, particularly when very complex samples with potentially different compositions (*e.g.* tissues from different organs) and high ion suppression are compared. A detailed discussion on the necessity of normalization strategies for lipidomics is available in Chapter I (*1.3.1.6. Ion suppression and normalization approaches*, p. 29).

7.3.1. Lipidic composition of different types of tissue

The PCA score plot for the 5 different types of tissues (brain, heart, kidney, liver and muscle) displayed all the QC replicates tightly clustered, indicating good reproducibility for the employed method (Figure VII-2A). Although a complex dataset was used with a simple statistical model, the brain samples were fully separated from the other types of tissue in the first principal component, which shows a dramatically different lipidic composition. The PCA score plot without the brain samples showed that the kidney tissue also has a lipidic composition that causes its separation in the second principal component from heart, liver and muscle (Figure VII-2B). The liver and muscle samples were partially separated in the first and second principal components, while the heart samples were partially overlaid with skeletal muscle.

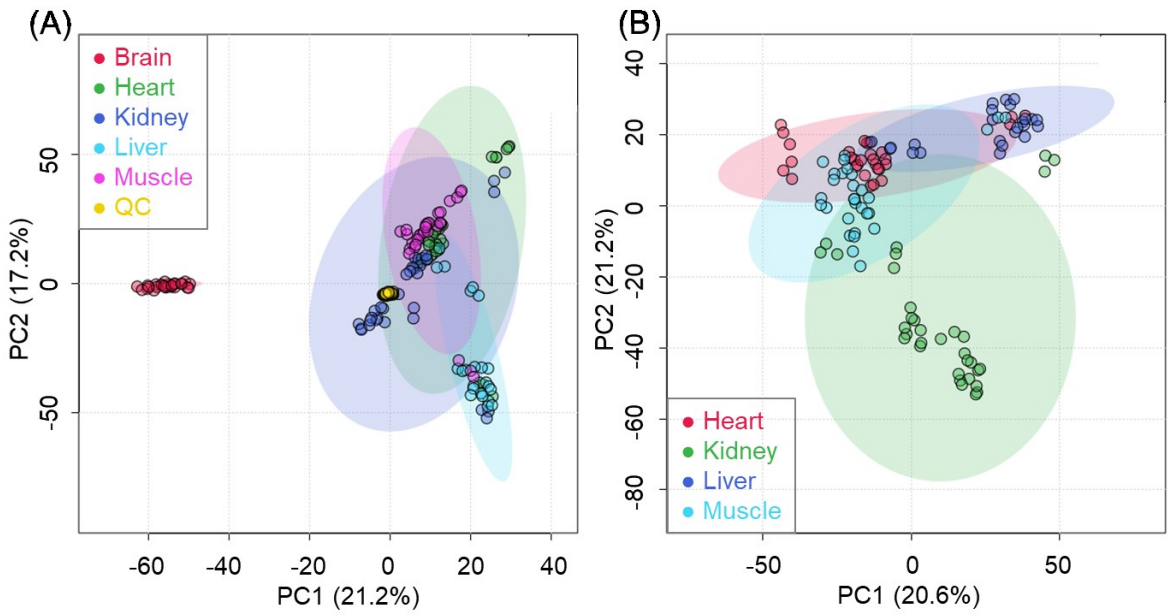

Figure VII-2. PCA score plots for the investigated tissues, regardless of treatment status (dexamethasone or control). (A) All samples and quality control (QC, pooled extracts from all tissues) replicates; (B) Heart, kidney, liver and muscle tissue (brain and QC replicates were removed for better visualization). The 33 QC replicates (yellow) and sample injection triplicates for each biological replicate (N = 4 to 6) were tightly clustered.

The non-parametric ANOVA analysis for the type of tissue (regardless of dexamethasone treatment) resulted in 2937 significant lipids ($p < 0.05$, 99.9% of the compounds employed for statistics, Supp. Table 28). The most significantly altered subclass (smallest p) was sphingomyelin (SM), which presented the highest total normalized intensity (*i.e.*, the sum of normalized intensities for all lipids belonging to the same subclass) for kidney samples (Figure VII-3). Carnitines (Car) were found to be elevated for heart, while hexosylceramides (HexCer) and phosphatidylethanolamines (PE) were more related to brain samples. Fatty acyls (FA), N-acyl amines (NAA) and lysophosphatidic acids (LPA) were elevated for liver tissue. The total

normalized intensities for acyl-coenzyme A (CoA) and monoacylglycerols (MG) were not significantly altered for the five types of tissues by non-parametric ANOVA ($p > 0.05$).

(cont.)

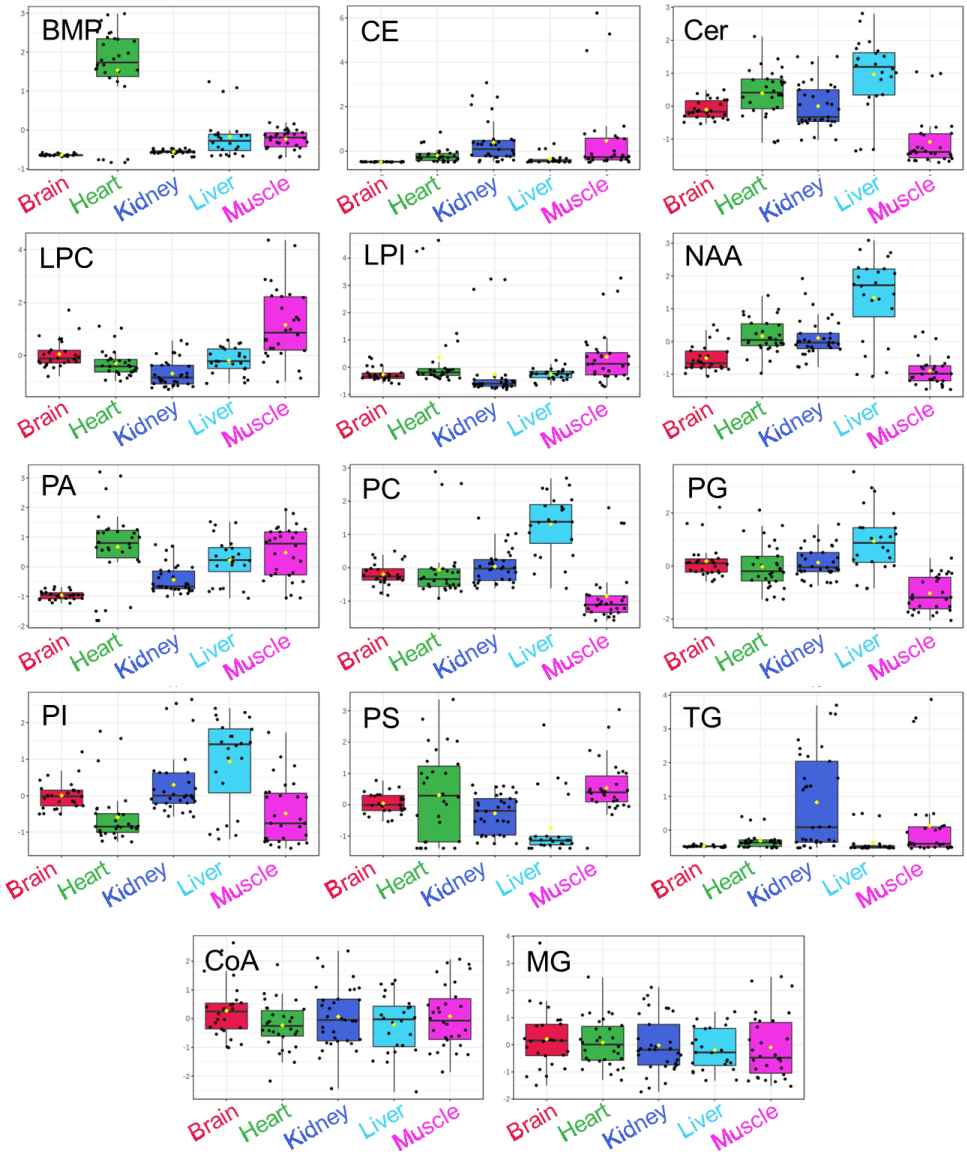

Figure VII-3. Boxplots for the summed normalized intensities for lipids detected in each subclass. Only acyl-coenzyme A (CoA) and monoacylglycerol (MG) species were not significant for non-parametric ANOVA ($p > 0.05$). Lipid subclass abbreviations are defined in Table II-2.

7.3.2. Long-term treatment of dexamethasone for different types of tissues

A Volcano plot analysis for the tissue samples obtained from dexamethasone-treated rats and controls (regardless of tissue type) resulted in 44 significantly altered lipids (fold-change, FC

control/dexamethasone ≥1.5 or ≤0.67 and p adjusted for false-discovery rate, p <0.05, Supp. Table 29), including 14 fatty acyls, 5 phosphatidic acids and 5 sterols (Figure VII-4). The fatty acyl metabolism seems to be deeply affected by long-term dexamethasone treatment for all five types of tissue. The significant alterations for the same lipid species in different tissues indicate a systemic effect of long-term dexamethasone treatment upon the metabolism. The highest significant fold-changes were found for the putatively identified ST 18:0;O4 (or the isomers/isobars ST 18:1;O3, FA 16:1;O, FA 18:4;O and MG O-13:2, with FC control/dexamethasone of 0.17) and PA 38:2 (or PA O-38:3;O and DG 39:3;O2, with FC control/dexamethasone of 0.22), with both displaying elevated intensities for the dexamethasone-treated animals. Amongst the high-confidence MS/MS identified lipids, the highest fold-change was found for the glycerophosphate PA 18:0_18:2 (FC control/dexamethasone of 0.32).

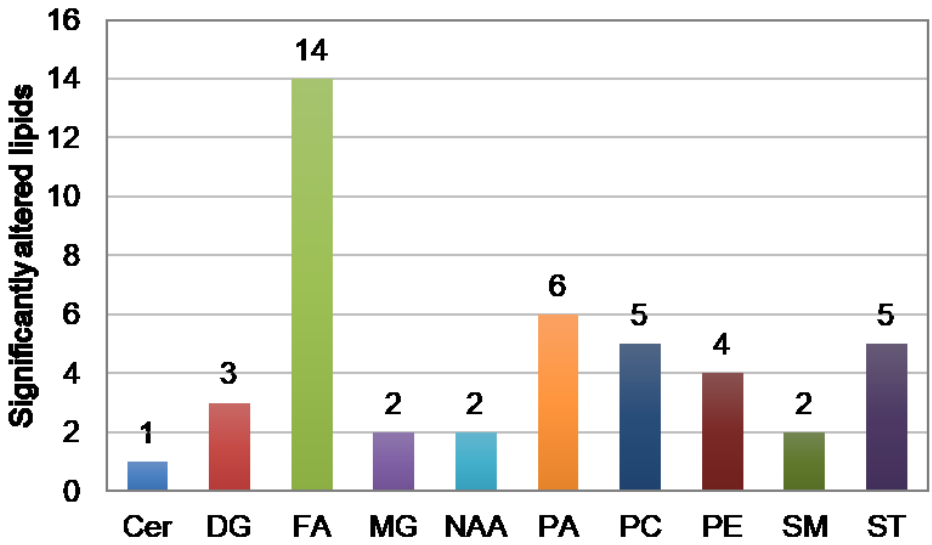

Figure VII-4. Significantly altered lipid classes for the comparison between controls (saline solution) and long-term dexamethasone treatment for brain, heart, kidney, liver and skeletal muscle tissue (fold-change control/dexamethasone ≤0.67 or ≥1.5 and p <0.05).

PLS-DA models for each type of tissue fully separated the dexamethasone-treated animals from controls (injected with saline solution), as shown in Figure VII-5. The Volcano plot analysis emphasized the different effects of long-term dexamethasone treatment upon the studied tissues (Figure VII-6, Supp. Tables 30 to 34). Skeletal muscle presented the highest number of significantly altered lipids (417, corresponding to 14.1% of the lipids employed for statistics), followed by liver (357), heart (325), kidney (200) and brain (47). The Venn diagram in Figure VII-7 compares the significantly altered lipids for the dexamethasone-treated and control animals in all five types of tissue. Muscle and liver shared 30 significantly altered lipids, while heart and kidney shared 25. However, no lipids were significantly altered for all tissues when dexamethasone-treated animals were compared with the controls within each type of tissue.

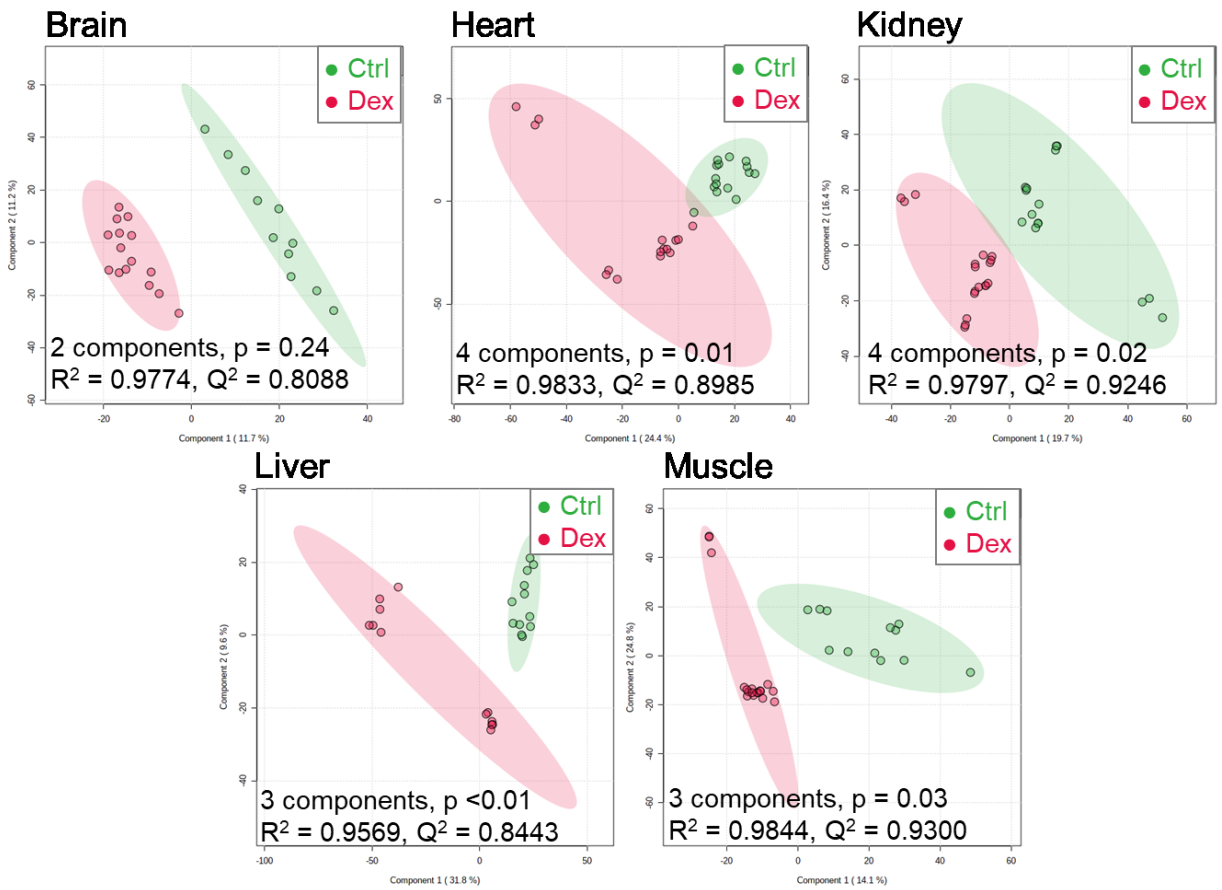

Figure VII-5. PLS-DA score plots obtained for the five types of tissue labeled as control (Ctrl, green) or long-term dexamethasone treatment (Dex, red). The models were cross-validated and passed permutation tests (N = 4 to 6, $p < 0.05$ for 1000 permutations), except for the brain tissue ($p = 0.24$).

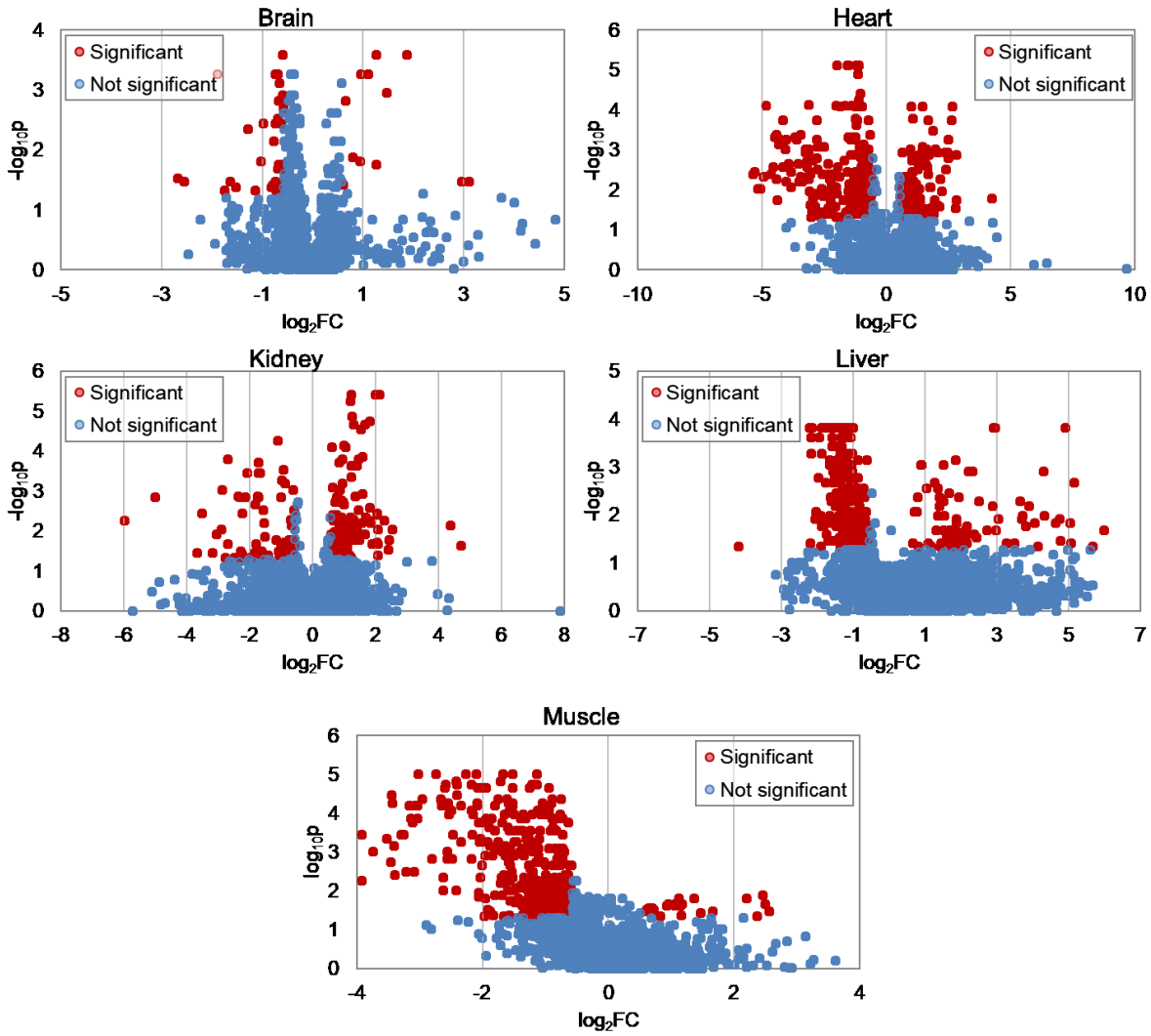

Figure VII-6. Volcano plots for the comparison between tissues obtained from long-term dexamethasone-treated rats and controls (saline solution). Lipids were considered significant for fold-change (FC control/dexamethasone) ≤0.67 or ≥1.5 and $p < 0.05$ (N = 4 to 6, Supp. Tables 30 to 34).

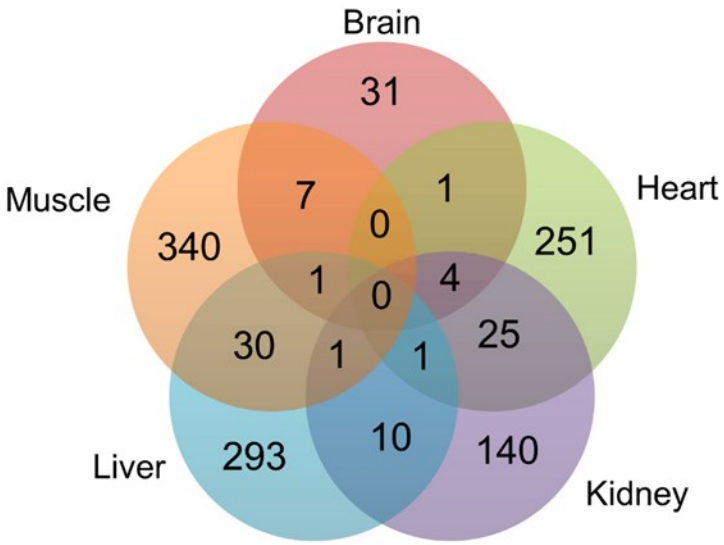

Figure VII-7. Venn diagram for significantly altered lipids in the comparison between dexamethasone-treated rats and controls.

Skeletal muscle was deeply affected by long-term dexamethasone treatment (Figure VII-6). Amongst the 417 significantly altered lipids, 19.9% were identified as diacylglycerols (DG, 83 lipids), 11.8% were phosphatidylethanolamines (PE, 49) and 7.9% were sphingomyelins (SM, 33). In fact, 40.2% of the sphingomyelins employed for statistics were significantly altered, with all of them displaying higher normalized intensities for the dexamethasone group (Figure VII-8). The long-term treatment seems to be related to increased levels of most lipid subclasses in skeletal muscle, as 94.7% of the significantly altered lipids had higher normalized intensities for the dexamethasone group (Figure VII-8, FC control/dexamethasone ≤ 0.67 and $p < 0.05$). The lipids with the most intense fold-changes were identified as SM 42:1;O2 (or PE-Cer 45:1;O2, FC control/dexamethasone of 0.07), PS 26:0_18:1 (MS/MS match, FC control/dexamethasone of 0.07) and PC 22:0_18:1 (MS/MS match, FC control/dexamethasone of 0.07).

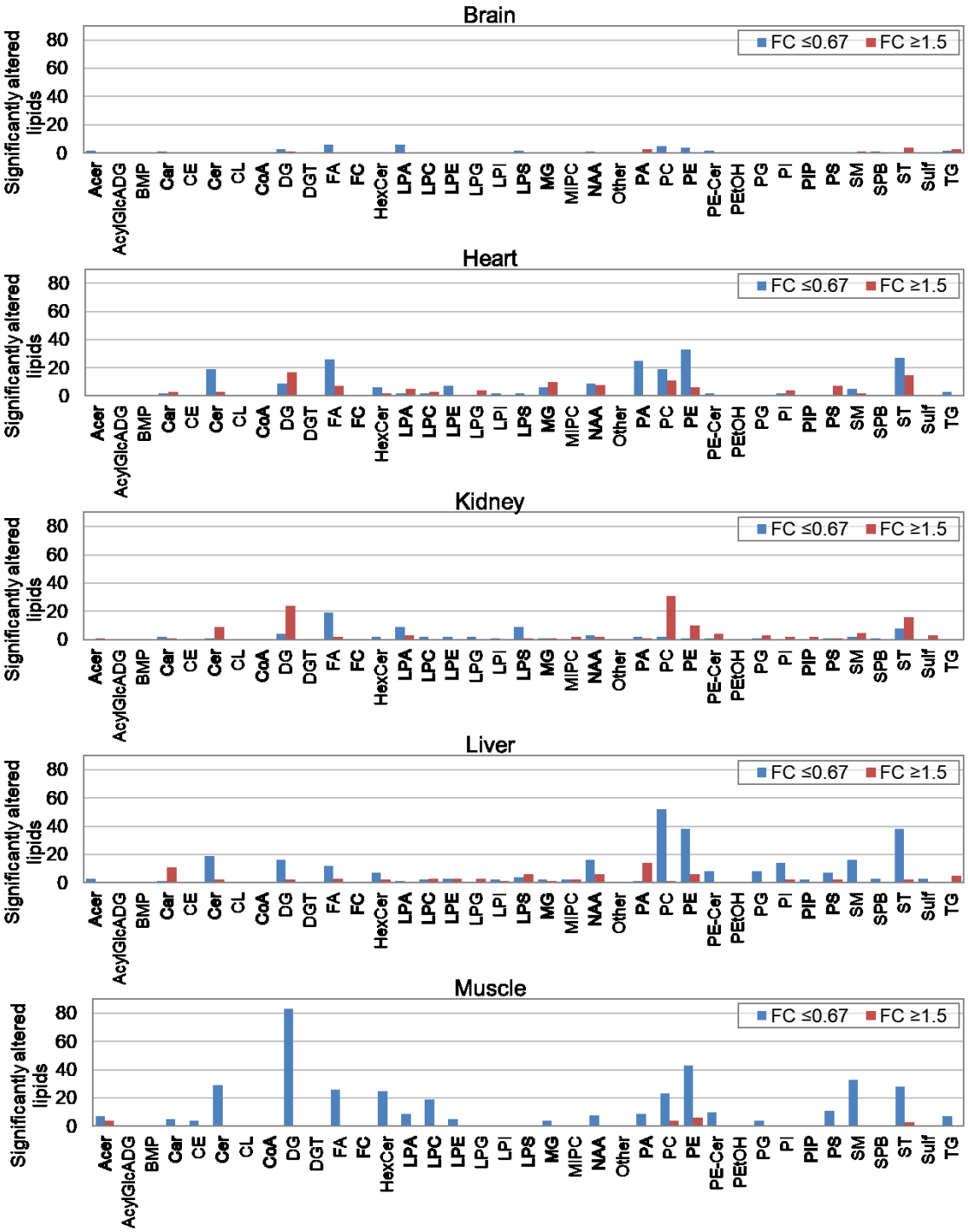

Figure VII-8. Fold-changes (FC) for the significantly altered lipids between brain, heart, kidney, liver and muscle tissue obtained from rats (N = 4 to 6) treated with dexamethasone or saline solution (control). Fold-changes (control/dexamethasone) were calculated by non-parametric Volcano plot analysis for lipids with false-discovery rate adjusted-p <0.05 (Supp. Tables 30 to 34).

The blue bars represent lipids with higher normalized intensities for the dexamethasone group, whereas the red bars represent lipids with lower intensities. Lipid subclass abbreviations are defined in Table II-2 (p. 48).

A similar trend was observed for liver tissue, with 78.4% of the 357 significantly altered lipids showing higher normalized intensities for the dexamethasone-treated group (Figure VII-6). The only subclasses that did not follow the elevated intensity trend for the dexamethasone animals were carnitines (Car), phosphatidic acids (PA) and triacylglycerols (TG), with lower values for 11 (out of 12), 14 (out of 15) and 5 (out of 5) significantly altered lipids, respectively (Figure VII-8). The lipid species with the most intense fold-changes (control/dexamethasone) displayed lower intensities for the dexamethasone group, namely the putatively identified Car 18:2;O (or the isomers/isobars NAE 23:3;O3, FA 25:4;O3, ST 25:1;O5 or MG 22:4;O, FC control/dexamethasone of 63.3), the MS/MS-matched TG 12:0_16:1_18:2 (FC control/dexamethasone of 51.1) and the putatively identified TG 59:7 (or the isomers/isobars TG O-59:8;O, DG 59:8;O or DG O-59:9;O2, FC control/dexamethasone of 35.4). Glycerophospholipids were deeply affected for liver tissue with mostly higher normalized intensities for the dexamethasone group, summing 48.5% of the significant lipids. The most affected subclasses included phosphatidylcholines (PC, 53 lipids, 14.8% of the significantly altered species), phosphatidylethanolamines (PE, 44, 12.3%), sterols (ST, 40, 11.2%), N-acyl amines (NAA, 6.2%) and ceramides (Cer, 21, 5.9%, Figure VII-8).

Out of the 325 significantly altered lipids for heart tissue, 65.5% had higher normalized intensities for the dexamethasone-treated group (Figure VII-6). Glycerophospholipids were also deeply altered (43.1% of the significant lipids), but the most affected subclass was sterols (ST,

12.9% or 42 lipids), followed by phosphatidylethanolamines (PE, 12.0%,39) and fatty acyls (FA, 10.2%, 33). All the 25 significantly altered phosphatidic acids (PA) had higher intensities for the dexamethasone group, along with 84.6% of the phosphatidylethanolamines (33) and 86.4% of the ceramides (Cer, 19, Figure VII-8). The most intense fold-changes were observed for the putatively identified FA 24:4 (or the isomers/isobars NAE 22:3 and ST 24:1;O2, FC control/dexamethasone of 0.02), MG O-17:4 (or the isomers/isobars ST 20:0;O3, FA 18:0;O, FA 20:3;O or MG O-15:1, FC control/dexamethasone of 0.02) and ST 22:0;O3 (or the isomers/isobars FA 20:1, ST 22:1;O2 and FA 22:4, FC control/dexamethasone of 0.03).

Kidney tissue had 200 lipids that were significantly altered for the control group and the dexamethasone-treated rats (Figure VII-6). Most of them (62.0%) showed lower normalized intensities for the dexamethasone group (except for fatty acyls), going against the observations for muscle, liver and heart. More than 90% of the 33 significant phosphatidylcholines (PC) were reduced for the treated group, as well as 90.9% of the 11 phosphatidylethanolamines (PE) and 85.7% of the 28 diacylglycerols (DG, Figure VII-8). The lipid species with the most intense fold-changes were putatively identified as FA 20:3;O2 (or the isomers/isobars Car 10:2;O2, NAE 15:3;O4, MG 17:3 and ST 22:2;O5, with FC control/dexamethasone of 0.02), PA 52:4 (or the isomers/isobars PE 50:3, PA O-52:5;O, PE O-50:4;O and PC 47:3, with FC control/dexamethasone of 0.03) and NAE 16:4;O3 (or the isomers/isobars CAR 11:3;O, FA 21:4;O, MG O-18:5 and ST 23:3;O4, with FC control/dexamethasone of 26.6).

The brain tissue was only mildly affected by long-term treatment with dexamethasone (Figure VII-6). Out of the 47 significantly altered lipids, 6 were identified as fatty acyls (FA), 6 as lysophosphatidic acids (LPA), 5 as phosphatidylcholines (PC) and 5 as triacylglycerols (TG, Figure VII-8). All the significant fatty acyls, lysophosphatidic acids, phosphatidylcholines,

phosphatidylethanolamines (PE) and sterols (ST) had higher normalized intensities for the dexamethasone group (Figure VII-8). The lipid species with the most intense fold-changes were putatively identified as TG 44:1 (or the isomers/isobars DG 44:2;O, TG O-44:2;O and DG O-44:3;O2, FC control/dexamethasone of 8.6), TG 38:1 (or the isomers/isobars DG 38:2;O, DG O-38:3;O2, TG O-38:2;O and Cer 41:2;O5, FC control/dexamethasone of 7.8) and LPS 30:2 (or the isomers/isobars LPG 30:4, LPS O-30:3;O, LPG O-30:5;O or LPS O-28:0;O, with FC control/dexamethasone of 0.2).

7.4 Discussion

The biological mechanisms responsible for the adverse effects of one of the most commonly prescribed class of drugs are not yet well known. We have performed comprehensive, untargeted lipidomics of brain, heart, kidney, liver and skeletal muscle tissue from rats that received high dosages of dexamethasone for 14 weeks *versus* controls treated with normal saline solution. First, we investigated the lipidic composition of each tissue, regardless of treatment. Second, we hypothesized that the long-term treatment with dexamethasone would provoke deep alterations in the tissues' lipidome that could be correlated with common adverse effects. A related study using the same animal model design to perform phenotypic and metabolomic profiling of serum samples was recently published.[253] The authors reported that the dexamethasone-treated rats had an age-dependent weight loss of about 20%, as well as significantly higher blood glucose and total triglycerides. Although the total serum cholesterol was not significantly different, low-density lipoprotein (LDL) was reduced for the dexamethasone group. There were no changes between the kidneys and livers of control and dexamethasone-treated animals by MRI examination, although cysts were observed at the time of sacrifice.[253]

We found that 99.9% of the lipids employed for statistical analysis were significant by ANOVA ($p < 0.05$) between the five types of tissue, regardless of treatment status (dexamethasone or saline). While phosphatidylethanolamines and hexosylceramides were elevated for brain, carnitines and bis(monoacylglycero)phosphates were correlated with heart samples (Figure VII-3). Brain tissue is mostly composed of membrane lipids, which are generally rich in glycerophospholipids, such as phosphatidylethanolamines. However, it depends almost exclusively on glucose for energy, as shown by the low levels of triacylglycerols (Figure VII-3).[255,256] Hexosylceramides are neutral glycosylated N-acyl fatty acid derivatives of sphingosine (Chapter I, *1.2.2. Lipid classification*, p. 3).[16,17] They contain one or more sugar moiety, such as glucose, galactose, N-acetylglucosamine, N-acetylgalactosamine or fucose. The monoglycosylceramides are known as cerebrosides (gluco- and galactocerebrosides), typically found in neural tissue.[15] Galactosylceramide is one of the most important components of the myelin sheet of nerves, acting as insulators for axons of neuronal cells, which correlates with our findings. Furthermore, bis(monoacylglycero)phosphates (BMPs) are a minor constituent of all tissues. The cone-shape and hydrated negatively charged headgroup that is characteristic of BMPs and cardiolipins aid membrane fusion and invaginations, promoting extracellular vesicles. The polyglycerophospholipids are essential for regulation of membrane curvature and stabilization of inner mitochondrial membranes. Interestingly, we found elevated levels of BMPs in heart tissue, but cardiolipins, crucial for ATP production, had higher intensities for skeletal muscle. [257,258] Kidney tissue was particularly related to high levels of sphingolipids, including sphingomyelins, sphingophospholipids and sphingoid bases. The role of sphingolipids in renal function are not yet well known, but sphingomyelins have been previously described as dysregulated in several renal diseases. [259,260] The kidneys displayed high intensities of sphingomyelins, sphingoid bases and

ceramide glycerophospholipids, but liver tissue had elevated levels of lysophosphatidic acids, fatty acyls and N-acyl amines. The particularly high levels of fatty acyls in the liver, regardless of treatment status, are a result of its role in the biosynthesis and metabolism of lipids (Chapter I, *1.2.4.3. Lipogenesis and metabolism*, p. 12).[261,262] Last, skeletal muscle resulted in slightly higher intensities of cardiolipins and lysophosphatidylcholines when compared to the other four types of tissues. Acyl-coenzyme A and monoacylglycerols were the only lipid subclasses the were not significant across the studied samples. Monoacylglycerols are one of the products of triacylglycerol metabolism. They can be hydrolyzed by monoglyceride lipases during lipolysis of adipose tissue to generate fatty acids and glycose. The oxidation of fatty acids to produce energy is dependent upon their activation by an ATP-dependent acylation reaction that results in acyl-CoA. Once activated, the fatty acyl portion of acyl-CoA species are transferred to carnitines by carnitine palmitoyl transferase to allow their transport into the mitochondria, where fatty acids are degraded by β-oxidation. The resulting acetyl-CoA moieties enter the citric acid cycle and oxidative phosphorylation to produce energy. Hence, the two lipid subclasses that were not significantly altered are deeply related.[15]

The long-term dexamethasone treatment caused significant alterations in 44 lipids across all tissues, indicating a systemic effect (Figure VII-4). The glycerophosphate PA 18:0_18:2 was identified by MS/MS match with a significant fold-change of 0.32 for all tissues of control animals compared to dexamethasone-treated rats. Six other phosphatidic acids were putatively identified with significantly higher levels in dexamethasone-treated animals. Heart tissue seems to be particularly affected by the alteration in phosphatidic acid metabolism: all the 25 significantly altered phosphatidic acids showed higher levels in heart tissue of dexamethasone-treated rats, including four species identified by MS/MS-match, namely PA 18:0_18:2 (fold-change for

control/dexamethasone of 0.09), PA 16:0_20:1 (0.10), PA 20:0_20:4 (0.15) and PA 20:0_18:2 (0.22). Phosphatidic acids are important intermediates in the biosynthesis of triacylglycerols and other glycerophospholipids (Chapter I, 1.2.4.3. *Lipogenesis and metabolism*, p. 12). The elevated intensities observed for tissues obtained from the dexamethasone group were not accompanied by significantly higher levels of triacylglycerols, but we noticed 7 elevated phosphatidylcholines and phosphatidylethanolamines, indicating increased biosynthesis rate for glycerophospholipids. Furthermore, phosphatidic acids act as signalling molecules for coagulation, stimulation of membrane fusion, modulation of cardiac muscle contractions and others. Elevated levels of phosphatidic acids have also been previously suggested as markers of cardiac hypertrophy.[263] Inflammatory mediators can activate phospholipases to stimulate the conversion of other glycerophospholipids to phosphatidic acids. Glucocorticoids are used to control exaggerated immune response and inflammations; hence, the systemic elevated levels of phosphatidic acids suggest a regulation of the activity of phospholipases. [264,265]

We also found 14 elevated fatty acyls and 5 sterols for all tissues obtained from the dexamethasone-treated rats. All types of tissues showed elevated levels of fatty acyl lipids for the dexamethasone group compared to control animals. Liver and muscle tissue also had higher levels of most significantly altered diacylglycerols and sterols after dexamethasone treatment. The effect of glucocorticoids upon fatty acid and glycerolipid metabolism is unclear. During normal metabolism conditions, dietary lipids are degraded into fatty acids during digestion, which are then converted to triacylglycerols and packaged into lipoprotein particles (*i.e.*, chylomicrons) along with dietary cholesteryl esters (Chapter I, *1.2.4.1. Digestion and transport*, p. 10). Alternatively, endogenous lipids synthesized in the liver are packaged into very-low density lipoprotein particles (VLDL), which are majorly composed by endogenous triacylglycerols and cholesteryl esters

wrapped in an outer layer of cholesterol, phospholipids and proteins. Lipoproteins containing endogenous (VLDL and others) and exogenous triacylglycerols and sterols (chylomicrons) are transported into tissue cells by low-density lipoprotein (LDL)-receptors. Triacylglycerols are then hydrolyzed by lipoprotein lipases in the capillaries of tissues, leading to free fatty acids and the steady levels of monoacylglycerols observed for all five types of tissues. Hence, lipoprotein lipases regulate serum concentrations of triacylglycerols and lipoprotein particles. The free fatty acids are taken up by the liver, being further activated as fatty acyl-CoAs and esterified with glycerol-3-phosphate (obtained from glucose) to form new triacylglycerol species. The liver regulates the levels of blood glucose and synthesizes or degrades triacylglycerols to generate energy. When the energy demand is high, the liver degrades triacylglycerols into fatty acids, which are then are metabolized into acetyl-CoA by β-oxidation, as previously described (Chapter I, *1.2.4.2. Energy production*, p. 11). When the demand for energy is low, the liver incorporates fatty acids into triacylglycerols, which are excreted to the blood stream as VLDL for further uptake by adipose tissue for future use. In times of starvation and fasting, the stored triacylglycerols are hydrolyzed by adipose triglyceride lipase into diacylglycerols, which are in turn hydrolyzed to monoacylglycerols, glycerol and fatty acids, completing the glycerol and fatty acid metabolism.[15,266]

Glucocorticoids are known to promote increased lipolysis in adipose tissue by inducing the activity of adipose hormone-sensitive triglyceride lipase (promotes the breakdown of triacylglycerols into diacylglycerols and fatty acids); diglyceride lipase (diacylglycerols to monoacylglycerols and fatty acids); and monoglyceride lipase (monoacylglycerols to fatty acids and glycerol). They can also inhibit lipoprotein lipase, which affects serum levels of triacylglycerols.[266] Hence, dexamethasone increases triacylglycerol turnover in adipose tissue and

accumulation of free fatty acids, which are in turn taken up by tissues, explaining the elevated levels observed in the comparisons between controls and dexamethasone-treated rats (Figure VII-8). Furthermore, excessive glucocorticoid consumption or biosynthesis under stress conditions is known to increase *de novo* fatty acid synthesis via acetyl-CoA (Chapter 1 – *1.2.4.3. Lipogenesis and metabolism*, p. 12) by stimulation of the transcription of enzymes such as acetyl-coA carboxylase and fatty acid synthase. The uptake of fatty acyls from the blood stream by the liver is balanced by β-oxidation and secretion of VLDL, which ensures relatively small quantities of triacylglycerols accumulated within the organ. There are reports that glucocorticoids can also supress β-oxidation of fatty acids in the liver.[267] The consequent hepatic lipid deposition upon conversion of free fatty acids to triacylglycerols causes a condition known as steatosis or "fatty liver", one of the adverse effects of long-term dexamethasone treatment.[268] Contrastingly, glucocorticoid treatment seems to increase the activity of lipin-1 in the liver, leading to increased oxidative fatty acid metabolism.[269]

Although we observed increased levels of 12 fatty acyls in liver tissue after long-term dexamethasone treatment, all the five significantly altered triacylglycerols had decreased intensities, which counteracts the expected deposition of triacylglycerols that is a hallmark of hepatic steatosis. Liver tissue displayed increased levels of 78.4% of the significantly altered lipids, including fatty acyls, diacylglycerols, sterols, glycerophospholipids and ceramides, but not triacylglycerols. In fact, the lipid with the second most intense fold-change for liver tissue was identified by high-confidence MS/MS-match as TG 12:0_16:1_18:2 (fold-change control/dexamethasone of 51.1, *i.e.*, significantly decreased levels for liver tissue of dexamethasone-treated rats). TG 12:0_12:0_12:0 was also identified by MS/MS-match with a fold-change (control/dexamethasone) of 17.7. The higher levels of diacylglycerols and fatty acids,

combined with lower levels of triacylglycerols, indicate a higher rate of hydrolysis of complex glycerolipids in the liver. A previous study with the same animal cohort reported higher serum levels of triacylglycerols, which were mainly attributed to the dexamethasone-induced enhanced secretion of lipoprotein lipase, the enzyme that hydrolyzes triacylglycerols from VLDL. [253] Our observations indicate that the excess fatty acid accumulation in tissues may not be counterbalanced by β-oxidation.[253] We found decreased levels of 11 carnitines in liver tissue of dexamethasone-treated rats (out of 12 significantly altered carnitines), including the MS/MS-matched Car 18:2 (fold-change control/dexamethasone of 32.8) and Car 18:1 (24.8). Elevated levels of carnitines were reported for blood serum samples of the same animal cohort.[253] Carnitines are essential for the transportation of fatty acids through mitochondrial membranes for β-oxidation; hence, an alternative compensation mechanism is suggested for the regulation of liver function. The higher levels of diacylglycerols and glycerophospholipids (PC, PE, PI, PS, PG) indicate that the excess of fatty acids may induce higher rates of lipogenesis of membrane lipids instead of energy production through β-oxidation in liver tissue. [270,271]

Contrary to metabolites, brain tissue lipids were remarkedly not affected by the long-term dexamethasone treatment.[3] However, skeletal muscle showed higher levels of multiple lipid subclasses after dexamethasone treatment, including ceramides, diacylglycerols, fatty acyls, hexosylceramides, glycerophospholipids, sphingomyelins, sterols and triacylglycerols. Myopathy, a clinical disorder of skeletal muscle that causes muscle weakness, cramps and fatigue, is a known adverse effect of glucocorticoids that seems to be related to an abnormal accumulation of lipids. An excess of ceramides has been previously related to decreased phosphorylation and activation of protein kinase B, affecting insulin resistance and signalling, which decreases muscle glucose uptake. The lipid subclass may also inhibit IGF-1 protein synthesis and differentiation, as well as

amino acid uptake, affecting muscle development and repair. Oxidative stress and mitochondrial dysfunction are also partially related to upregulation of ceramides.[272] Furthermore, some of the most intense fold-changes for muscle tissue were observed for the MS/MS-matched plasmalogens PE O-18:0_18:3 (fold-change control/dexamethasone of 0.09) and PE O-18:0_18:2 (0.09). The vinyl-ether containing glycerophospholipids act in organization and stability of lipid rafts in membrane regions involved in cell signalling. Although the biological roles of plasmalogens are not yet fully understood, they have been related to antioxidant function. The vinyl-ether bond presents a high reactivity with reactive oxygen species, being preferentially oxidized instead of other types of lipids. Interestingly, the previously published metabolomics study for the same sample cohort pointed to a systemic alteration of the glutathione metabolism, a well-known essential antioxidant against oxidative stress.[3] Decreased levels of oxidized glutathione were found in muscle tissue, indicating a mitigation of oxidative processes. The higher levels of plasmalogens found in this study can be correlated with the lower levels of oxidized glutathione, which shows an inhibition of oxidative stress in muscle. Plasmalogens are synthesized by peroxisomes, small organelles that are also responsible for β-oxidation of long-chain fatty acids. As previously discussed, glucocorticoids can suppress β-oxidation of fatty acids, which were mostly increased for muscle tissue. Hence, there is a suggested effect of dexamethasone upon peroxisomal oxidative processes and redox equilibrium.[273,274]

7.5 Conclusions

Glucocorticoids are widely used for treatment of different pathologies. However, their adverse effects may become increasingly dangerous upon long-term, high-dosage application. The metabolic and phenotypic alterations caused by long-term dexamethasone treatment were

previously reported in related works. Now, we have investigated the lipidic alterations caused by dexamethasone upon brain, heart, kidney, liver and skeletal muscle tissue. The employed LC-MS methodology allowed the reliable comprehensive, untargeted detection of thousands of lipids from different subclasses. Our results show that, although brain lipids are only mildly affected, the synthetic glucocorticoid causes systemic effects upon different organs. The benefits of dexamethasone treatment are undeniable, but this study helps unravel the biochemical implications of long-term use that give rise to the many adverse effects observed by patients and medical professionals.

7.6 Acknowledgments

We thank all collaborators involved in this project, particularly Dr. Anas M. Abdel Rahman (King Faisal Hospital and Research Center, Saudi Arabia) and Xiaohang Wang (University of Alberta, Edmonton, AB, Canada). This work was supported by the King Faisal Hospital and Research Center (Saudi Arabia) and grants from the Natural Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research, the Canada Research Chairs program, the Canadian Foundation for Innovations, Genome Canada and Alberta Innovates.

VIII

Chapter VIII: Lipidomic Alterations Induced by Cystic Fibrosis, Genotypic Mutation and Lung Function Decay [18,19,20]

8.1 Introduction

Cystic fibrosis (CF) is a systemic autosomal recessive syndrome characterized by progressive obstructive lung disease and pancreatic insufficiency. The condition is caused by recessive mutations in the cystic fibrosis transmembrane regulator (*CFTR*) gene, which produces a glycoprotein that controls the transport of chloride and bicarbonate ions through cell membranes, as well as the regulation of transepithelial sodium. [275,276] CFTR is an ATP-binding cassette (ABC) transmembrane protein that contains a cytosolic regulatory domain, two nucleotide-binding domains and two membrane-spanning domains that form a transmembrane channel. The CFTR channel is activated by phosphorylation of the regulatory domain by cAMP-dependent protein kinase A (PKA). The phosphorylated channel is gated by binding of ATP to both nucleotide-binding domains, triggering a conformational change that allows ion passage. ATP hydrolysis disrupts the dimer, which closes the ion channel and prevents ion trafficking. [275,277–279]

The mutated CFTR protein produced by CF patients affects the normal traffic of chloride ions and water though membranes of the respiratory, gastrointestinal and reproductive tracts, causing insufficient hydration of airway surfaces, desiccated mucous secretions and high sodium

[18] A version of this chapter was conditionally accepted for publication at Journal of Proteome Research on Sep/2020.

[19] Supporting figures for this chapter are available in Appendix E. Supporting tables are available with Dr. Liang Li.

[20] The sample collection was performed by Dr. Majed Dasouki and Dr. Anas M. Abdel Rahman's research groups (King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia).

chloride concentrations in sweat. The consequent accumulation of thick mucus is the cause of intestinal obstructions, pancreatic insufficiency, infertility, persistent cough, respiratory infections, and pneumonia. As the CF pathology evolves, the airways become chronically infected with pathogenic bacteria and viruses, leading to exacerbated inflammation. The progressive cycle of chronic infection and inflammation in the respiratory tract culminates in reduced lung respiratory function and permanent lung damage, with progressive chronic obstructive pulmonary disease (COPD) being the cause of death for more than 90% of patients. [280–284] Currently, there is no cure for CF. Although the prognosis may seem dire, the mortality rates and life expectancies of CF patients are improving. When CF was first described in 1938, the life expectancy was less than one year, but the median survival age reached 40.6 years old in the USA and 50.9 in Canada in 2013. [285,286]

While the classical presentation of CF is well known, a definitive CF diagnosis is only achieved by a combination of symptoms, family history, a positive newborn screening test (if available) and laboratory evidence of CFTR malfunction, *e.g.*, positive genetic testing for two CFTR mutations (one in each parental gene) or elevated sweat sodium chloride levels. The newborn screening test with a dried blood spot collected after birth usually includes inspection of the immunoreactive trypsinogen (IRT) protein, a marker of pancreatic insufficiency, which is one of the earliest manifestations of classical CF. However, elevated immunoreactive trypsinogen levels are not conclusive for CF and require confirmation through abnormal sweat tests or identification of two CF-related mutations. [280,281,287–290]

Therapies that correct CFTR function are promising, but still not sufficiently available to patients. The progress of the disease can be slowed down through pancreatic enzyme replacement, mucolytic agents, antibiotics, and mucus-clearing approaches. Lung transplantation is frequently

required in advanced stages of the disease. [280,281,283] However, new mutation-specific drugs are leading towards an essential change in CF management. The most common CFTR mutations are divided into six classes: (I) CFTR is not synthesized (frameshift, splicing or nonsense mutations, resulting in no mRNA or no protein); (II) is inadequately processed, *e.g.* interruption of protein trafficking to the cell membrane, misfolding, abnormal degradation or deletion (lower concentrations); (III) is dysregulated or inactivated, *i.e.*, even though the protein is produced and processed correctly, the chloride channels remain closed in response to stimuli; (IV) has abnormal channel conductance; (V) has reduced synthesis due to the introduction of promoter or splicing abnormalities; and (VI) is unstable, with quick degradation once the protein reaches the cell membrane. Mutations related to the most severe phenotypes (class I, II and III) are associated to loss of CFTR function, leading to early age diagnosis, pancreatic insufficiency and elevated sweat chloride levels; however, patients in classes IV, V and VI have less severe CF symptoms, with sweat chloride levels closer to normal. [280–282,284,288,291,292] Patients may have combinations of two mutations belonging to different classes. The development of targeted therapies and medication based on CFTR gene mutation classes is a promising alternative for CF treatment. For example, Class II mutations may be treated with therapies focusing on the rescue of CFTR protein traffic, whereas class III requires potentiators to restore ion channel activity, such as VX-770 (Ivacaftor).[291] However, the molecular mechanisms associated with each genotypic mutation are not yet fully known.

There is still a need for a better understanding of biological pathways involved in CF pathophysiology and the effect of different mutation classes upon the patient's metabolism. The synthesis, processing and action of the CFTR protein have been the main target of most CF-related studies. Targeted and untargeted metabolomics was applied to blood, sputum, bronchoalveolar

lavage fluid and urine of CF patients, as well as *in vitro* studies. Several metabolites have been related to the disease, *e.g.*, amino acids, lactate, bile acids, nucleotides, carbohydrates and cortisol.[293] Yet, very little is known about CFTR interaction with lipid metabolism and cell membranes. The interaction of the transmembrane CFTR protein with membrane lipids seem to play an important role in protein stability and function, whereas lipid mediators may aid protein folding and traffic.[43,275] There is some evidence of sphingomyelin regulation of CFTR activity and inhibition through catalysis. [294–296] Previous studies targeted to specific lipid categories and subclasses have also shown dysregulation of fatty acids, lysophospholipids and glycerophospholipids related to CF disease severity in plasma and bronchoalveolar lavage, with a significant decrease in levels of phosphatidylcholine species, as well as higher triglyceride levels.[297–299] A more comprehensive characterization of molecular mechanisms involved in CF development, mutation classes and lung function decay may help elucidate the pathophysiology of the condition and improve therapies.

We performed untargeted, comprehensive liquid chromatography-mass spectrometry (LC-MS) lipidomics of blood serum samples from CF adult patients and healthy controls. First, we hypothesized the role of CFTR mutations on the expression of serum lipidome. The lipidome of CF patients was compared with healthy controls using statistical analysis to select significantly altered lipid species and reveal biomarker candidates. Second, we studied lipidic alterations induced by each particular CFTR mutation class. Third, we investigated the correlation between the serum lipidome of patients and their lung function, measured by spirometry tests as the predicted forced expiratory volume in 1 second (FEV1%). FEV1% is an established marker of CF progression, with decreased FEV1% values often related to increased risk of death.[300,301]

8.2 Experimental

8.2.1 Participants[21]

All procedures performed in this study involving human participants followed the ethical standards of the Declaration of Helsinki and the guidelines of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use Good Clinical Practice (ICH-GCP). This study was reviewed and approved by the Institutional Review Board at King Faisal Specialist Hospital and Research Center (Riyadh, Saudi Arabia - approval number 2160 031). Written informed consent was obtained from all participants.

Forty CF patients (17 males and 23 females aged between 12 and 34 years old) were recruited from the CF-Pulmonology clinic at the King Faisal Specialist Hospital and Research Center (Table VIII-1). Patients who had been enrolled in another clinical study in the last 30 days, unable or unwilling to provide informed consent, or diagnosed with conditions other than CF were excluded from this study. The genotypic mutation of each patient was determined as previously described.[302] Patients presented gene mutations belonging to classes II (4 patients), III (13), IV (14), V (4) or VI (1). Spirometry tests were performed for each patient, including forced expiratory volume for the first second of the forced breath (FEV1), forced vital capacity (FVC), the proportion of the lung vital capacity that a person can expire within the first second of forced expiration (FEV1/FVC ratio or FEV1%), and predicted FEV1% (FEV1% divided by the average FEV1% for a population with similar physical characteristics, Table VIII-1). The genotypic class, gene mutation, mutation type, sex, age, FEV1% and predicted FEV1% values for each patient are

[21] The patient recruitment, assessment and sample collection were performed by Mai Abdel Jabar, Imran Nizami, Dr. Majed Dasouki and Dr. Anas M. Abdel Rahman (King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia).

presented in Supp. Table 35. Patients reported having the following CF symptoms: persistent cough (35 patients), lung infections (3), positive septum culture (1), sinusitis (6), bowel irregularities (2), heartburn (1) and flu (1). All patients were taking medications, *i.e.*, tobramycin (32 patients), colistin (1), ciprofloxacin (1), azithromycin (37), tacrolimus (1), dornase alpha (36) or ivacaftor (1). Treatments also included percussive therapy (11), lung transplant (1), pancreatic enzyme supplement (39), vitamin D supplement (2) and liver transplant (1). One patient left the study before collection of genotypic class, mutation type and spirometry results; hence, the patient was included for the comparison between cystic fibrosis patients and healthy controls but excluded for the remaining evaluations. Three patients passed away before spirometry tests were performed and were also excluded from the FEV1% statistical analysis.

Table VIII-1. Patient cohort selected for this study. Detailed information for each patient is available in Supp. Table 35.

	Average (± one SD*)	**Median**	**Minimum**	**Maximum**
Age	21.1 (± 5.2)	20.0	12	34
Sex	17 males / 23 females			
FEV1 (L)	1.6 (± 0.8)	1.4	0.61	3.84
FVC (L)	2.2 (± 0.9)	2.0	0.81	4.92
FEV1% (%)	71.5 (± 13.3)	72.5	48.1	100.0
Predicted FEV1% (%)	53.4 (± 25.9)	47.5	15.0	98.0

*SD: standard deviation

The 20 healthy control samples were obtained from the King Faisal Specialist Hospital and Research Center blood bank. The healthy volunteers were screened to match the age, sex and body

mass index range of patients; however, as samples were collected anonymously from a blood bank, no specific information about each healthy individual is available. All blood samples were collected by vein puncture into plain non- EPIC tubes (Vacutainer, BD Biosciences, San Jose, CA, USA). The serum was separated by centrifugation and immediately frozen at -80°C until analysis.

8.2.2 Sample preparation

The lipid fractions of the blood serum samples were obtained through a modified Folch liquid-liquid extraction with dichloromethane and methanol. The employed chemicals and reagents were previously described in Chapter II (*2.2.1. Chemicals and reagents*, p. 42). Aliquots of 18.0 µL of blood serum were vortexed for 20 s with 4.6 µL of the internal standard mixture (Table II-1, p. 45) and 145.4 µL of methanol. After protein precipitation, the mixture was vortexed again for 20 s with 300 µL of dichloromethane, followed by 10 s vortex with 94.4 µL of water. The mixture was equilibrated at room temperature for 10 min and centrifuged for 10 min at 12,000 rpm and 4°C. The organic phase (240.0 µl) was evaporated to dryness on a SpeedVac for 30 min. The residue was resuspended in 3.6 µL of 1:1 mobile phase A (MPA) / mobile phase B (MPB), vortexed for 30 s and diluted with 32.4 µL of MPA (2× dilution). The extract was not further diluted for negative ionization; however, for positive ionization, an aliquot of 6 µL of the extract was diluted with 2.4 µL of 1:1 MPA/MPB and 21.6 µL of MPA (10× dilution). After extraction, the prepared samples were stored at 4 °C for a maximum of 48h before injection.

Sample preparation was randomized in groups of 6 to 8 samples. A pool containing aliquots of all samples was employed as quality control (QC). One identical QC aliquot was extracted with each group of 6 to 8 samples to control the reproducibility of the extractions and injections. Samples were extracted in polypropylene microcentrifuge tubes and stored in polypropylene

autosampler inserts in amber injection vials capped with polytetrafluoroethylene (PTFE)-lined caps. Blank extracts, using water instead of sample and methanol instead of the internal standard mixture, were prepared with each group of 6-8 samples to control the possibility of contamination from solvents and materials.

8.2.3 LC-MS Analysis

Chromatographic separation was performed using a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific, Waltham, MA, USA) and a Waters Acquity BEH C18 column (5 cm × 2.1 mm with 1.7 μm particles, Waters Corporation, Milford, MA, USA). Mobile phases were composed by 10 mM ammonium formate in 50:40:10 acetonitrile/ methanol/ water (v/v/v, MPA), and 10 mM ammonium formate in 95:5 isopropanol/ water (v/v, MPB). Untargeted, comprehensive lipidomics was achieved with the following conditions: injection volume of 4.0 μL for positive ionization and 5.0 μL for negative ionization; column temperature of 42 °C; 25 min gradient separation (0 min – 2% MPB, 250 μL/min; 3.0 min – 5% MPB, 250 μL/min; 8.0 min – 30% MPB, 250 μL/min; 21 min – 95% MPB, 230 μL/min; 25 min – 95% MPB, 230 μL/min); and 10 min re-equilibration (0 min – 95% MPB, 230 μL/min; 2 min – 5% MPB, 250 μL/min; 9.95 min – 5% MPB, 250 μL/min; 10 min – 5% MPB, 50 μL/min). The UHPLC instrument was coupled to a high-resolution quadrupole-time-of-flight (QToF) mass spectrometer (Maxis II, Bruker Daltonics, Billerica, MA, USA) with electrospray ionization (ESI) source (capillary voltage of 3800 V for positive and 4000 V for negative ionization; endplate offset of 500 V; nebulizer gas pressure of 1.4 bar; dry gas flow rate of 4.0 L/min; dry temperature of 230 °C; spectra acquisition rate of 1.44 Hz; and m/z range of 150 to 1500 Da). MS/MS acquisition was performed for all sample injections in auto-MS/MS mode using the software oToF Control (Bruker Daltonics) for

all sample injections with a cycle time of 1.2 s; active exclusion of precursors detected for more than 3 consecutive spectra within 0.50 min intervals; spectra acquisition rate ramping between 4.0 Hz for low-intensity precursor ions (below 500 cts) and 8.0 Hz for high intensity (> 1,000,000 cts); and collision energy ramping with precursor m/z values (m/z 150 – 8.0 eV; m/z 500 – 12.0 eV; m/z 1000 – 25.0 eV; m/z 1500 – 30.0 eV). Extra MS/MS injections were acquired for the pooled QC with different collision energies (5 to 100 eV), MS spectra acquisition rates (1.0 to 6.0 Hz), cycle time (1.2 to 2.6 s), number of precursor ions in each cycle (3 to 10), precursor m/z ranges (50 to 150 Da intervals), and injection volumes (4.0 to 12.0 µL).

The data acquisition sequence was randomized for samples and blank extracts, with one QC injection after every 6-8 sample injections. Hence, the QC results include extraction and injection replicates of the same pooled mixture. Positive and negative ionization were acquired in separate injections, with polarity switching for consecutive groups of 6-8 sample injections plus one QC. A 1.0 min mass re-calibration segment was inserted between the end of each chromatogram and column re-equilibration, during which 0.10 mM sodium formate calibrant solution in 1:1 isopropanol/ water (v/v) was infused with a peristaltic pump for mass recalibration.

8.2.4 Data processing

Chromatograms were processed on MetaboScape 4.0 (Bruker Daltonics, Billerica, MA, USA) with a minimum intensity cut-off of 7000 cts for positive and 2000 cts for negative ionization, minimum peak length of 6 spectra, m/z recalibration with sodium formate calibrant solution, m/z tolerance of 5.0 mDa and retention time tolerance of 15 s. Missing values were substituted by recursive extraction, *i.e.*, features found for at least 10% of injections were searched again on the raw data without the minimum intensity limit and with a minimum peak length of 6

spectra. The aligned features were filtered by detection in more than 90% of injections in at least one group (CF, control, or QC). The remaining missing values were substituted by the minimum group intensity for features detected in more than 50% of injections in each group (CF, control, or QC); or global minimum intensity (all injections) for features detected in less than 50% of injections in each group.

8.2.5 Lipid identification

A three-tier putative identification routine was applied to the aligned features. Tiers 1 and 2 were based on tandem mass spectrometry (MS/MS) identification, while tier 3 comprised lipids putatively identified by mass match. There was no overlap between the lipids identified in each tier. Lipid identification followed the guidelines established by the Lipidomics Standard Initiative (https://lipidomics-standards-initiative.org). [20,76]

For tiers 1 and 2, tandem mass spectrometry (MS/MS) identification was performed with the MS-Dial LipidBlast library (https://fiehnlab.ucdavis.edu/projects/LipidBlast), the Human Metabolome Database (https://hmdb.ca) and the MassBank of North America LC-MS/MS libraries (https://mona.fiehnlab.ucdavis.edu), combined to MetaboScape 4.0.[73,78,79,81–83,109,151] Tier 1 identification contained lipids with precursor m/z tolerance of 5.0 mDa, MS/MS score threshold of 500 and mSigma (isotope pattern match) threshold of 150, whereas tier 2 employed precursor m/z tolerance of 5.0 mDa, MS/MS score between 100 and 500, and mSigma threshold of 50.

Features not identified for tiers 1 and 2 were searched in the LipidMaps database (http://www.lipidmaps.org) for putative identification by mass-match with m/z tolerance of 5.0 mDa (tier 3).[75] Lipids can have multiple isomers and isobars with different chemical structures; hence, the isomeric or isobaric identifications for each feature were ranked by a five-step filtering

and scoring approach, as described in Chapter II (*2.2.7. Putative identification of lipids*). The positions of double bonds and the stereospecific configuration of glycerol derivatives were not determined in this study.

The MS/MS and putatively identified lipids were divided into 41 subclasses and 6 main categories (Table II-2, p. 48). The lipid subclasses, categories and nomenclature followed the classification system proposed by the International Lipid Classification and Nomenclature Committee (ILCNC), the LipidMaps database (March / 2020) and the Lipidomics Standard Initiative. [16–18,20,76] Internal standard features were found based on retention time and accurate m/z values.

8.2.6 Normalization and statistics

The positively (MS/MS, tiers 1 and 2) or putatively (mass match, tier 3) identified lipids (Supp. Table 36) were matched to one of the 14 deuterated internal standards (Table II-1, p. 45), according to lipid subclass similarity and retention time range. The non-identified features were not used for statistics. Normalized intensities were calculated as the peak intensity of each identified lipid divided by the peak intensity of the matched internal standard, which belonged to the same or the most similar lipid subclass and eluted within similar retention time ranges. Normalized intensities were auto-scaled and normalized for the median intensity value for all samples.

Statistical analysis was performed on MetaboAnalyst 4.0 (https://www.metaboanalyst.ca), which limits the number of features to less than 5000.[153] Hence, lipids with relative standard deviation (RSD) higher than 19% for QCs were excluded before statistical analysis to match the feature limit, as they displayed lower reproducibility for extraction and injection replicates of the

pooled mixture. Statistical analysis included Volcano plots (fold change – FC *versus* p for non-parametric Student's t-test corrected for false discovery rate –p), Principal Component Analysis (PCA), Partial Least Square – Discriminant Analysis (PLS-DA), Random Forest and Receiver Operating Characteristic (ROC) curves for biomarker analysis.[153] For univariate analysis, lipids were considered significantly altered for $p < 0.05$ and $FC \geq 1.5$ or ≤ 0.67 (non-parametric tests for unequal variances). All PLS-DA models were cross-validated and passed permutation tests (1000 permutations). R^2 and Q^2 values for cross-validation, as well as p for permutation tests, are provided in the caption of each PLS-DA score plot.

8.3 Results

A total of 6489 and 4575 features were detected in positive and negative ionization modes, respectively. Representative chromatograms are shown in Figure VIII-1. 1346 lipid species were positively identified by MS/MS, including 879 lipids in tier 1 (precursor m/z error ≤5.0 mDa, MS/MS score ≥500 and mSigma ≤150) and 467 lipids in tier 2 (precursor m/z error ≤5.0 mDa, MS/MS score between 500 and 100, and mSigma ≤50) (Supp. Table 36). The excellent number of MS/MS identifications was a result of the combination of an optimization of the data acquisition parameters employed for all sample injections and the data processing routine (data not shown). The remaining 9718 unidentified features were searched on the LipidMaps database for putative identification (tier 3), and 5600 lipids were identified by accurate mass-match with m/z tolerance of 5.0 mDa (Supp. Table 36). The lipid subclasses of the identified compounds are shown in Figure VIII-2, whereas abbreviations for lipid classes are defined in Table II-2 (p. 48). Identifications included 2160 lipids containing one or more odd-chain fatty acyls (OCFA, 385 identified in tiers 1 or 2) and 4988 lipids with polyunsaturated fatty acyls (PUFA, 1094 identified in tiers 1 and 2).

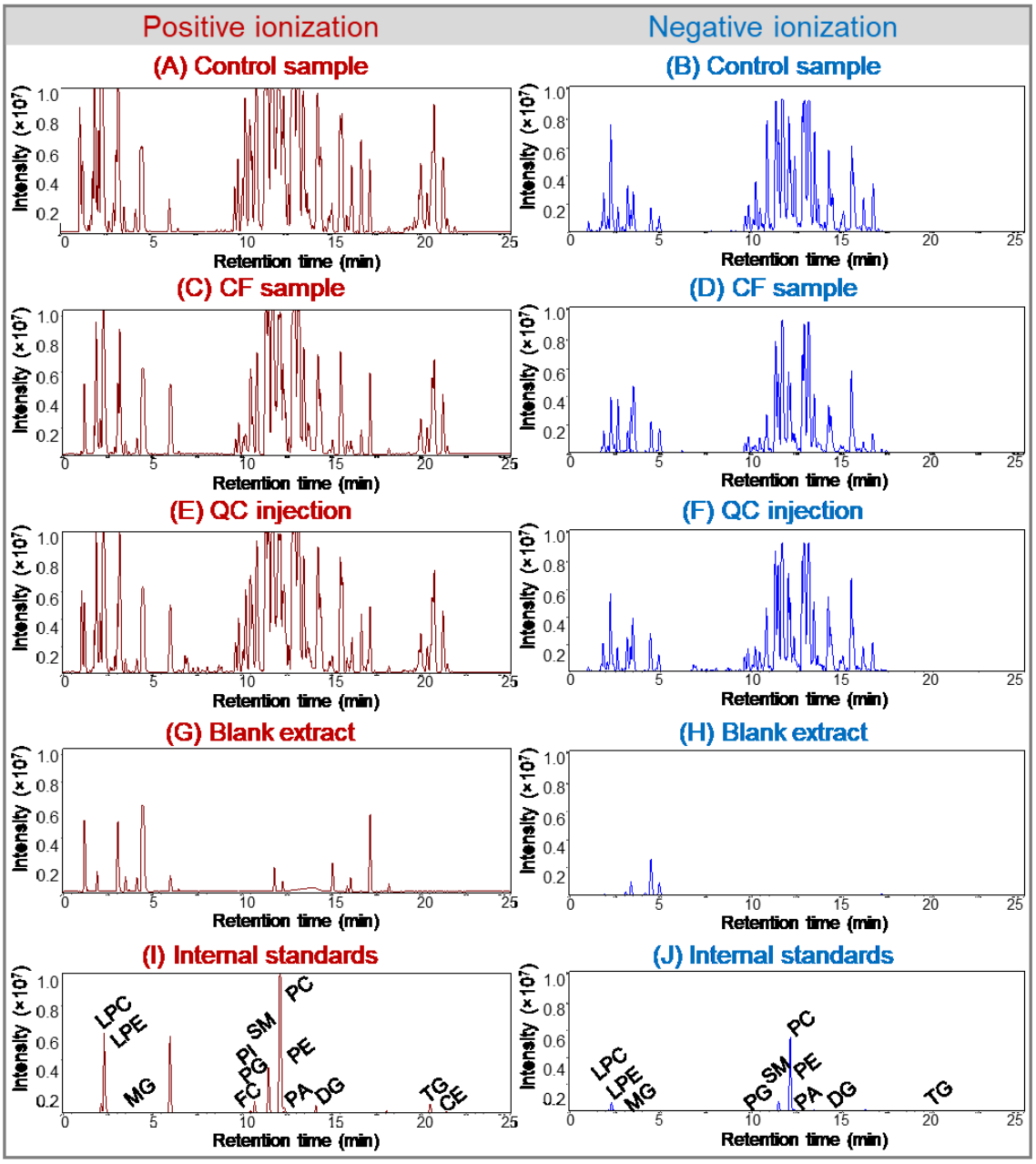

Figure VIII-1. Representative base peak chromatograms obtained under positive (A, C, E, G and I) and negative ionization (B, D, F, H and J) for the evaluation of lipidomics of cystic fibrosis. (A) and (B) a randomly chosen control sample; (C) and (D) a randomly chosen CF sample; (E) and (F) quality control (QC) injections; (G) and (H) a blank extract, *i.e.*, extraction of water instead of the sample; and (I) and (J) the internal standard mixture composed by 14 deuterated lipids (Table II-1). Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

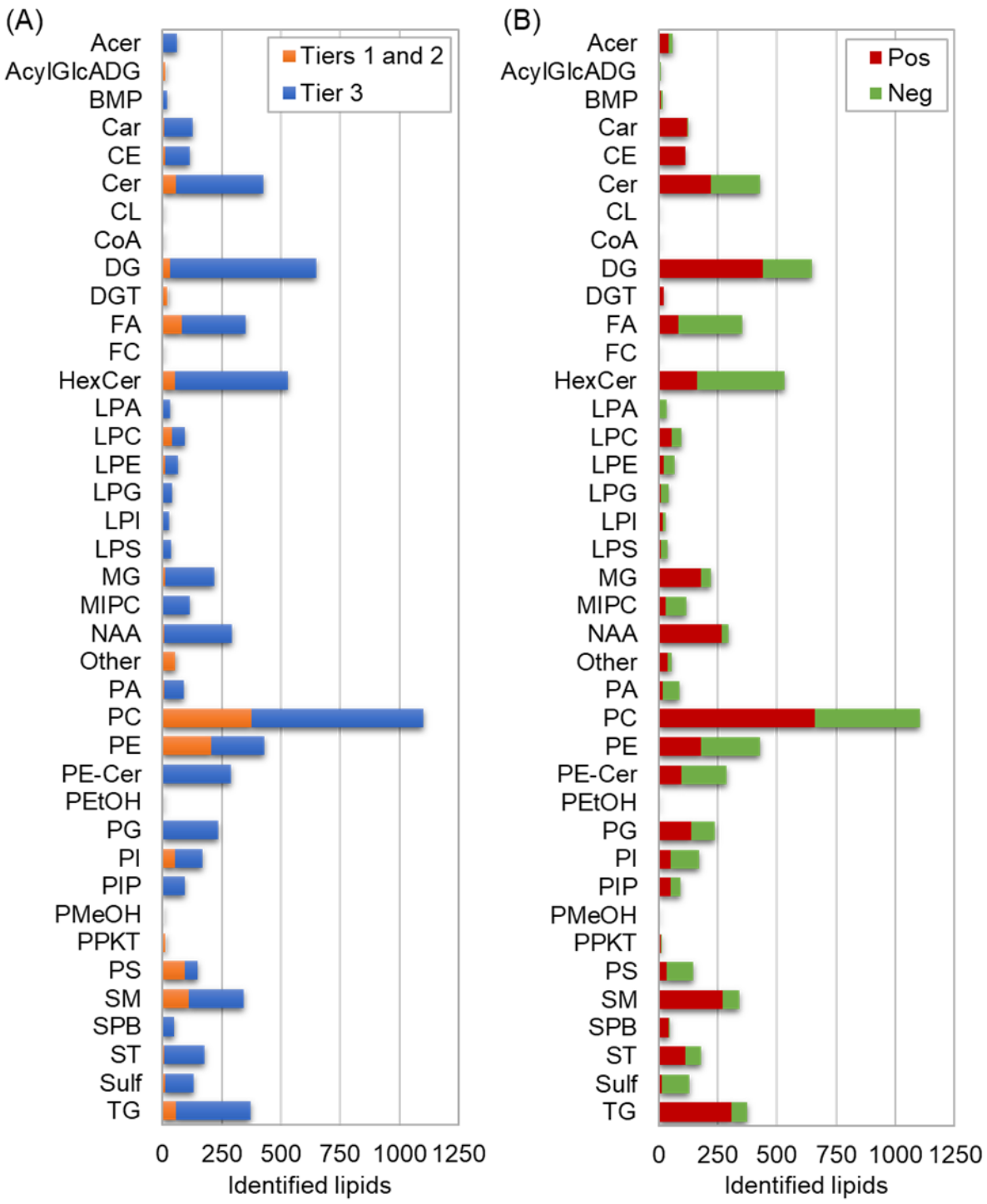

Figure VIII-2. Subclass distribution of identified lipids for lipidomics of cystic fibrosis. The compounds were divided into (A) identification level (tiers 1, 2 or 3) and (B) detection polarity (positive or negative). Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

Thirteen internal standards were identified with a maximum m/z error of 1.4 mDa or 1.7 ppm. The 6946 identified compounds (Supp. Table 36) were matched to one of the internal standards for normalization according to lipid subclass and retention time range. Fifty-four MS/MS-identified features (tier 1 or 2) were excluded from the statistical analysis for not being identified as lipids (medications, other metabolites and contaminants from sample preparation). Thirty internal standard peaks, corresponding to different adducts of the 13 detected internal standards (Table II-1, p. 45), were also removed after normalization.

The QC samples, composed by aliquots of all CF and control samples, were used to gauge the reproducibility of extractions and injections. MetaboAnalyst 4.0 limits the number of features for statistics at a maximum of 5000; hence, 1963 lipids with relative standard deviation (RSD) higher than 19% for QC extraction and injection replicates were excluded before statistical analysis due to low reproducibility. The remaining 4931 lipids with RSD ≤19% were employed for statistics.[153]

8.3.1. CF patients versus *healthy controls*

The PCA score plots in Figure VIII-3A (with QCs) and Appendix D - Figure D - 1 (without QCs) display all the 42 QC and sample experimental replicates tightly clustered, showcasing the reproducibility of the employed methods. The CF and control samples were fully separated by PCA and PLS-DA (Figure VIII-3A and B). A Random Forest model with 7 predictors and 500 trees resulted in correct classification of all samples, with out-of-bag (OOB) error of zero.

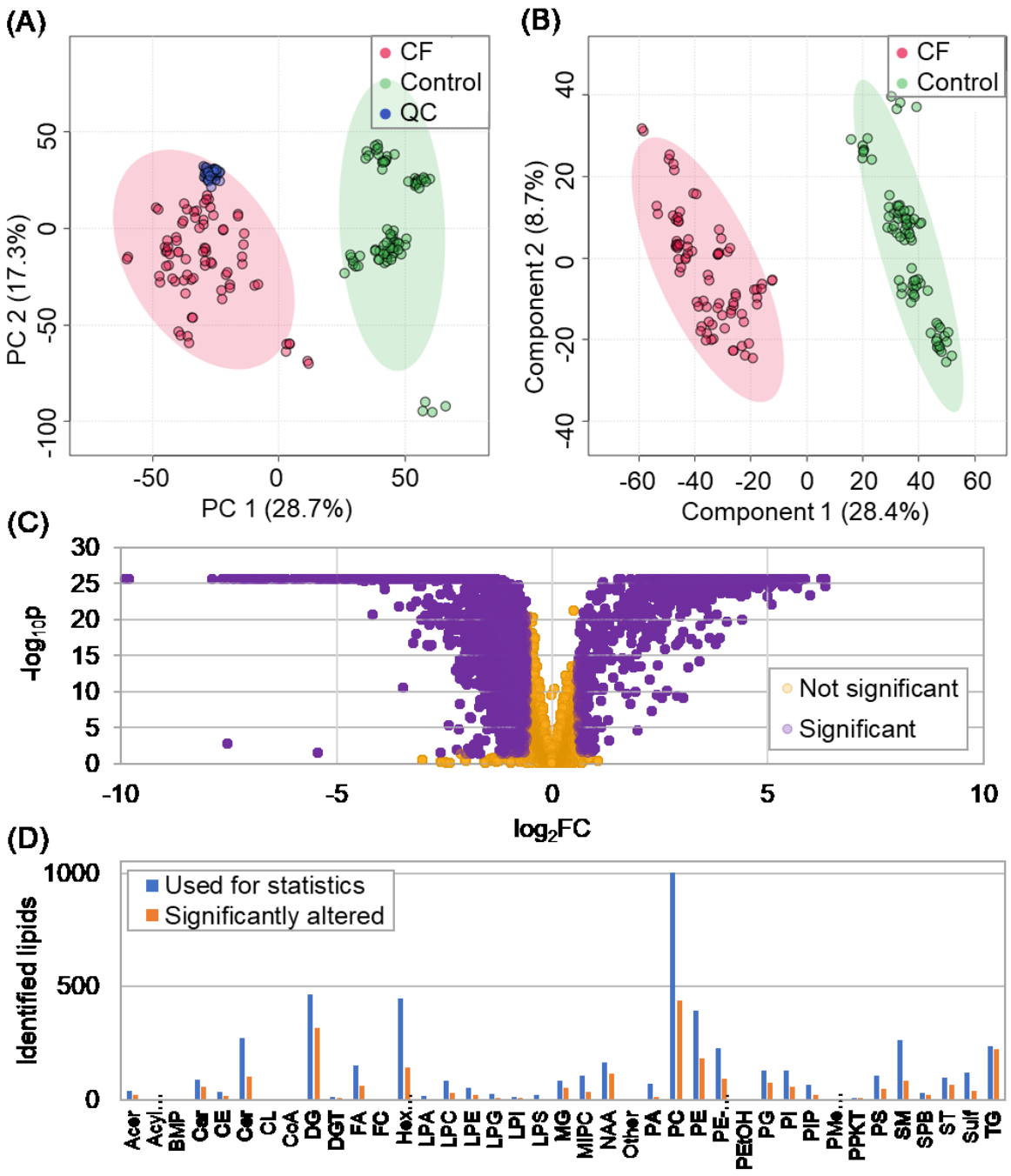

Figure VIII-3. Statistical analysis for cystic fibrosis patients (CF, red) compared to healthy controls (green) using (A) PCA, (B) PLS-DA with 8 components (R^2 of 0.9986, Q^2 of 0.9960 and $p < 0.001$ for 1000 permutations), and (C) Volcano plot, where significant lipids displayed fold-change (FC) ≥ 1.5 or ≤ 0.67 and p value adjusted for false-discovery rate (p) < 0.05. (D) Subclass distribution of all lipids employed for statistics (blue) compared to lipids that were significantly altered when CF

patients were compared to healthy controls (orange). Abbreviations to lipid subclasses are defined in Table II-2 (p. 48). The PCA score plot without quality control (QC) injections (blue) is available in Appendix D - Figure D - 1.

Univariate analysis through a Volcano plot resulted in 2353 significantly altered lipids ($p < 0.05$ and FC ≥ 1.5 or ≤ 0.67), which corresponds to 47.7% of all lipids employed for statistics (Figure VIII-3C and D, Supp. Table 37). The significantly altered lipids included 439 phosphatidylcholines (PC, 43.8% of all PCs used for statistics), 393 ceramides (Cer, 36.3%), 318 diacylglycerols (DG, 68.2%), 222 triacylglycerols (TG, 93.7%) and 181 phosphatidylethanolamines (PE, 46.2%) (Figure VIII-3D). The glycerolipid, sterol and fatty acyl metabolisms were deeply affected by CF: 74.9% of the glycerolipids employed for statistics were significantly altered (598 significant lipids), followed by 60.5% of the employed sterols (78 significant lipids) and 57.9% of fatty acyls (234 significant lipids). The boxplots for significantly altered lipids with the most intense fold-changes are presented in Appendix D - Figure D - 2 (identifications in tiers 1 and 2) and Figure D - 3 (identifications in tier 3).

Overall, 63.1% of the significantly altered lipids resulted in lower normalized intensities for the CF patients when compared to healthy controls (FC CF/control ≤ 0.67, Figure VIII-4). CF seems to be particularly related to lower levels of glycerolipids, sterol and fatty acyls in blood serum. Most acylceramides (Acer, 63.6% of the significantly altered acylceramides), carnitines (Car, 71.9%), N-acyl amines (NAA, 78.1%), monoacylglycerols (MG, 88.0%), diacylglycerols (DG, 62.3%), triacylglycerols (TG, 85.1%), sphingoid bases (SPB, 94.7%) and sterols (ST, 84.1%) displayed lower normalized intensities for the CF patients. Interestingly, 690 lipids identified with one or more odd-chain fatty acyls were also significantly altered, with 69.9% displaying lower

normalized intensities for the CF patients. Only hexosylceramides (HexCer), sulfatides (Sulf), ceramide phosphoinositols (MIPC) and phosphatidylglycerols (PG) had a majority of significantly altered lipids with higher intensity ratios for the CF patients.

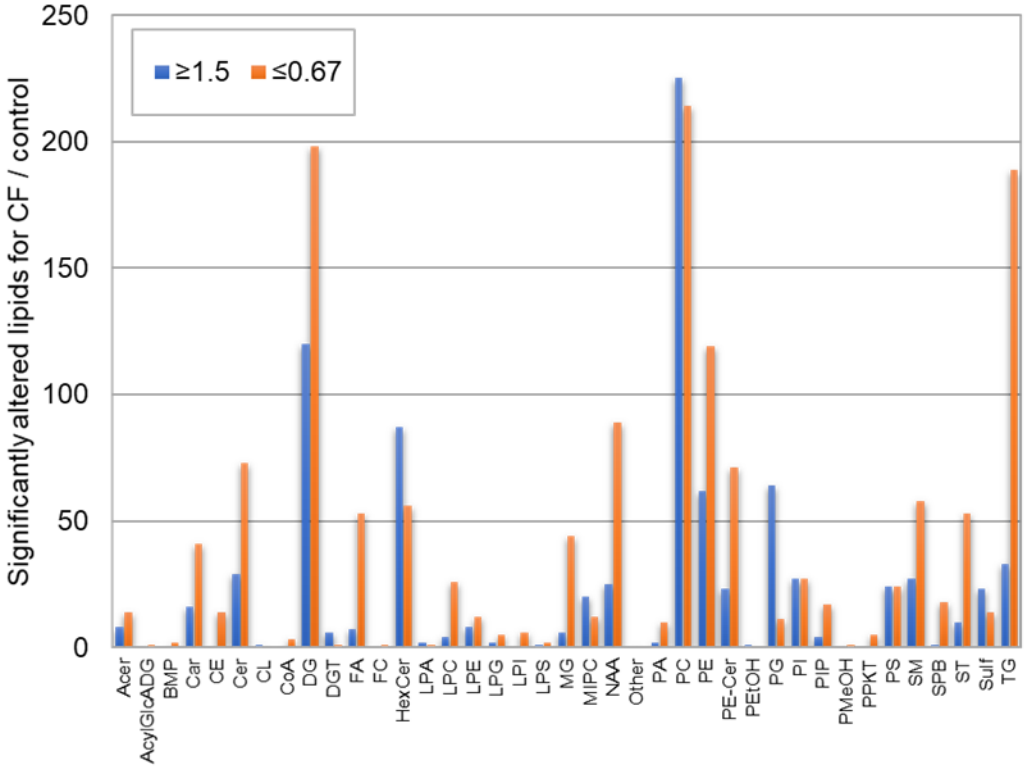

Figure VIII-4. Significantly altered lipids (fold change, FC ≥1.5 or ≤0.67 and p adjusted for false-discovery rate, p <0.05) for the comparison between cystic fibrosis patients (CF) and healthy controls, divided into FC (CF/control) ≥1.5 (blue) or ≤0.67 (orange). Lipids with FC ≥1.5 (blue) displayed higher normalized intensities for the CF patients, while lipids with FC ≤0.67 (orange) had lower normalized intensities for CF patients when compared to the healthy controls. Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

We selected biomarker candidates using the area-under-the-curve (AUC) of Receiver Operating Characteristic (ROC) analysis with sample classification based on Random Forest. The sample-set was first divided into a test set, containing 70% of injections (28 CF patients and 14 healthy controls), and a validation set, with the remaining 30% (12 CF patients and 6 healthy controls). The test set was employed for the selection of biomarkers candidates and construction of the classification model, whereas the validation test was retained for evaluation of the performance of the model. The ROC analysis for the test set resulted in 435 lipids with univariate AUC of 1.0, including 93 phosphatidylcholines, 55 N-acyl amines, 48 ceramides, 27 phosphatidylethanolamines and 27 sterols (Figure D - 4). N-acyl amines, lysoglycerophospholipids and sphingoid bases seem to be particularly affected by CF, as more than 30% of the lipids employed for ROC analysis resulted in the maximum AUC of 1.0. The results indicate that lipid metabolism is highly affected by CF, and a high number of different lipids could be potentially employed as biomarkers for CF diagnosis.

We first selected two lipids with AUC of 1.0 and the highest fold-change values to further investigate as unique biomarker candidates, namely: (1) P373.25848/1.86 (positive ionization, m/z of 373.25848 and retention time of 1.86 min), putatively identified in tier 3 as the eicosanoid FA 20:2;O4 (FC CF/control = 1.05×10^{-3}; p = 3.11×10^{-26}), and (2) P385.2584/2.01, putatively identified in tier 3 as FA 21:3;O4 or FA 19:0;O4 (FC CF/control = 1.12×10^{-3}; p = 3.11×10^{-26}). The extracted ion chromatograms for the selected lipids are shown in Appendix D - Figure D - 5. ROC curves were generated for each of them using Random Forest to classify the test samples as control or CF. Both candidates resulted in AUC of 1.0 (95% confidence interval – CI between 1.0 and 1.0), with sensitivities, specificities and overall accuracy of 100% (Figure VIII-5A and B). The

validation samples were blindly classified as control or CF using each model, but both candidates resulted in perfect specificity, sensitivity and accuracy of 100%.

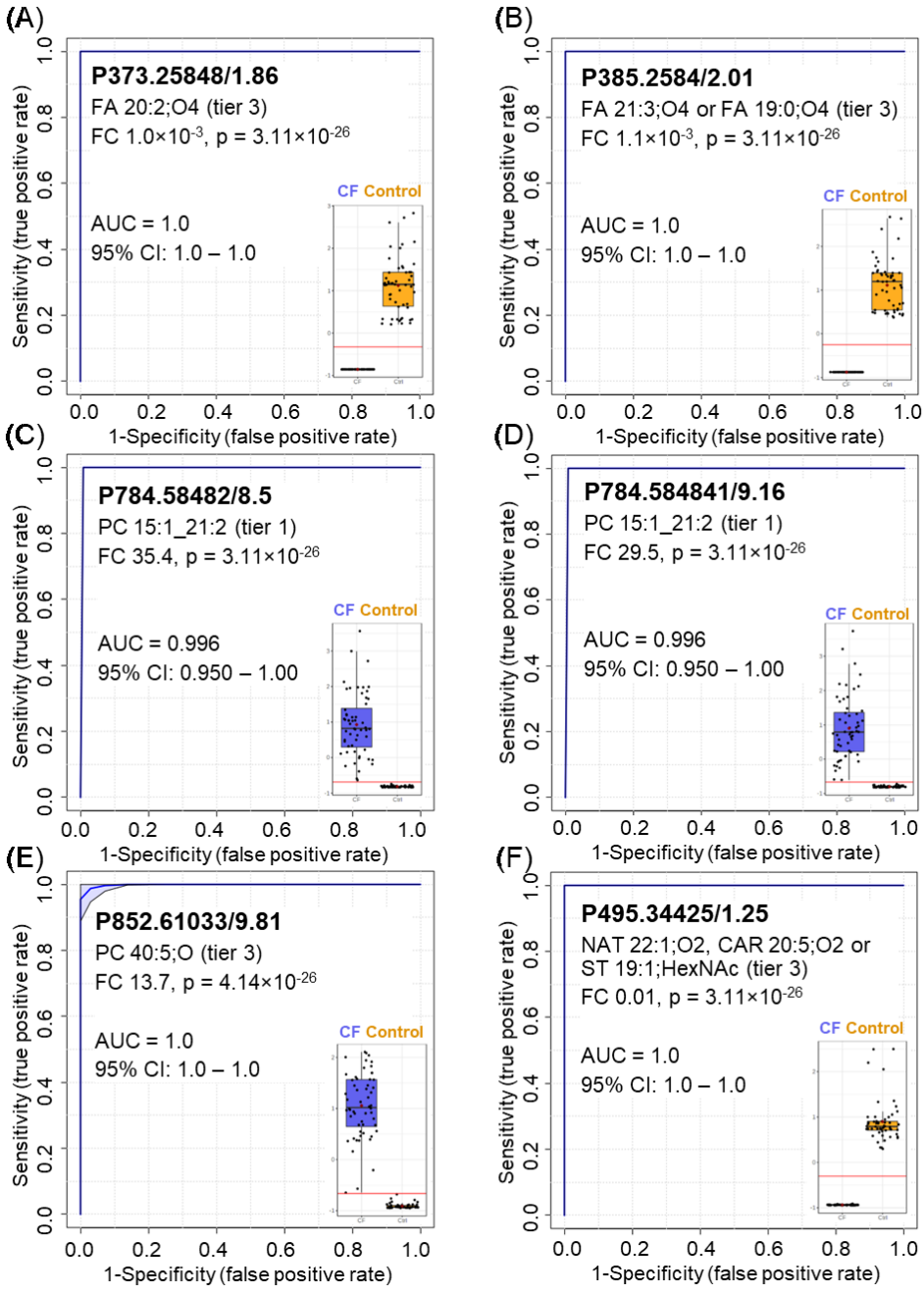

Figure VIII-5. ROC analysis for six distinct lipids selected as potential biomarker candidates for cystic fibrosis. (A) FA 20:2;O4 (tier 3); (B) FA 21:3;O4 | FA 19:0;O4 (tier 3); (C) PC 15:1_21:2

(tier 1); (D) PC 15:1_21:2 (tier 1); (E) PC 40:5;O | PG O-42:7 | PS O-42:5 | PE 43:5;O | CerP 48:6;O5 (tier 3); (F) NAT 22:1;O2 | CAR 20:5;O2 | ST 19:1;O;HexNAc (tier 3). Lipids are named by the ionization polarity (P for positive or N for negative), mass-to-charge ratio (m/z) and retention time (min), as well as putative (tier 3) or MS/MS identification (tiers 1 and 2).

Since the two lipids with the highest FC selected for ROC were only putatively identified in tier 3, we also selected the two lipids positively identified by MS/MS in tier 1 with the highest FC and AUC of 1.0 for further investigation, namely: (1) P784.58482/8.5, identified as PC 15:1_21:2 (FC CF/control = 35.4; $p = 3.11\times10^{-26}$), and (2) P784.584841/9.16, identified as PC 15:1_21:2 (FC CF/control = 29.5; $p = 3.11\times10^{-26}$). Interestingly, both lipids were identified as the same species, differing only in the position of fatty acyls or unsaturation, which were not determined in this study. It is worth noticing that both lipids contain odd-chain fatty acyl groups. The extracted ion chromatograms for the selected lipids are shown in Figure D - 6. Both candidates resulted in AUC of 0.996 (95% CI: 0.950-1.00), with sensitivity of 96.4%, specificity of 100% and overall accuracy of 98.2% for cross-validation of the test set (Figure VIII-5C and D). The validation samples were blindly classified as control or CF using each model. While the first candidate (P784.58482/8.5) resulted in sensitivity of 100%, specificity of 95.8% of accuracy of 97.9% (one misclassified sample), the second (P784.584841/9.16) correctly classified all samples in the validation set (sensitivity, specificity and accuracy of 100%).

Finally, we selected the two lipids with the best classification performance using Random Forest, namely: (1) P852.61033/9.81, putatively identified in tier 3 as PC 40:5;O, PG O-42:7, PS O-42:5, PE 43:5;O or CerP 48:6;O5 (FC CF/control = 13.7; $p = 4.14\times10^{-26}$), and (2) P495.34425/1.25, putatively identified in tier 3 as NAT 22:1;O2, CAR 20:5;O2 or ST

19:1;HexNAc (FC CF/control = 0.01; $p = 3.11 \times 10^{-26}$). The first compound was not matched to the MS/MS libraries employed for identification in tiers 1 and 2, but the characteristic fragment of phosphatidylcholines (m/z 184.07) was observed in its MS/MS spectra, confirming the lipid subclass as PC. The extracted ion chromatograms for the selected lipids are shown in Appendix D - Figure D - 7. Both biomarker candidates resulted in AUC of 1.0 (95% CI: 1.0 – 1.0), with sensitivities, specificities and accuracies of 100% by cross-validation of the test set (Figure VIII-5E and F). The validation samples were blindly classified using each lipid, resulting in sensitivities, specificities and accuracies above 91%. Hence, the six distinct biomarker candidates resulted in excellent classifications for the test set (cross-validation) and the validation set.

8.3.2. CF genotypic class

The sample cohort for this study included patients that belonged to the CF genotypic classes II (N = 4 patients), III (N = 14 patients), IV (N = 14 patients), V (N = 5 patients) and VI (N = 2 patients). The statistical analysis for the genotypic class was performed in two steps. First, all classes were evaluated together. Second, we performed binary comparisons between class III *versus* IV; IV *versus* V; and III *versus* V. Classes II and VI had a small number of patients (4 and 2, respectively); therefore, statistical models for binary comparisons may not be representative and were not considered.

First, samples belonging to patients with class II to VI mutations were evaluated together. The different classes could not be fully separated by PCA (Figure VIII-6A) due to the complexity of the dataset and the simplicity of the statistical model; however, a partial separation can be seen between classes IV and V, while class III falls between them. The PLS-DA model for classes II to VI resulted in a separation between the more severe classes II and III, and the less severe classes

IV, V and VI (Figure VIII-6B). A Random Forest model with 1000 trees and 15 predictors correctly classified all samples from classes III and IV, but one patient from class II, one for class V and one from class VI were misclassified (OOB error of 6.4%), probably due to the smaller number of patients. The non-parametric ANOVA displayed 2820 significant lipids for the five genotypic classes (57.2% of all 4931 lipids employed for statistical analysis), with 608 identified by MS/MS in tiers 1 and 2 (Supp. Table 38). Sterols, glycerophospholipids and sphingolipids were deeply affected by the genotypic class, with more than 56% of the lipids employed for statistics displaying $p < 0.05$ (Figure VIII-6C). Similarly to the control/CF Volcano plot analysis, the subclasses with the highest number of significantly altered lipids included phosphatidylcholines (PC, 593 significant lipids, corresponding to 59.1% of the PCs employed for statistics), ceramides (592 significant, 21.0%), diacylglycerols (DG, 254 significant, 54.5%) and phosphatidylethanolamines (PE, 218 significant, 55.6%) (Figure VIII-6C). It is also worth noticing that all the 68 phosphatidic acids (PA) employed for statistical analysis displayed $p < 0.05$ for ANOVA, as well as more than half of the oxidized lipids (58.0%) and species containing polyunsaturated (PUFA, 58.0%) and odd-chain fatty acyls (57.2%).

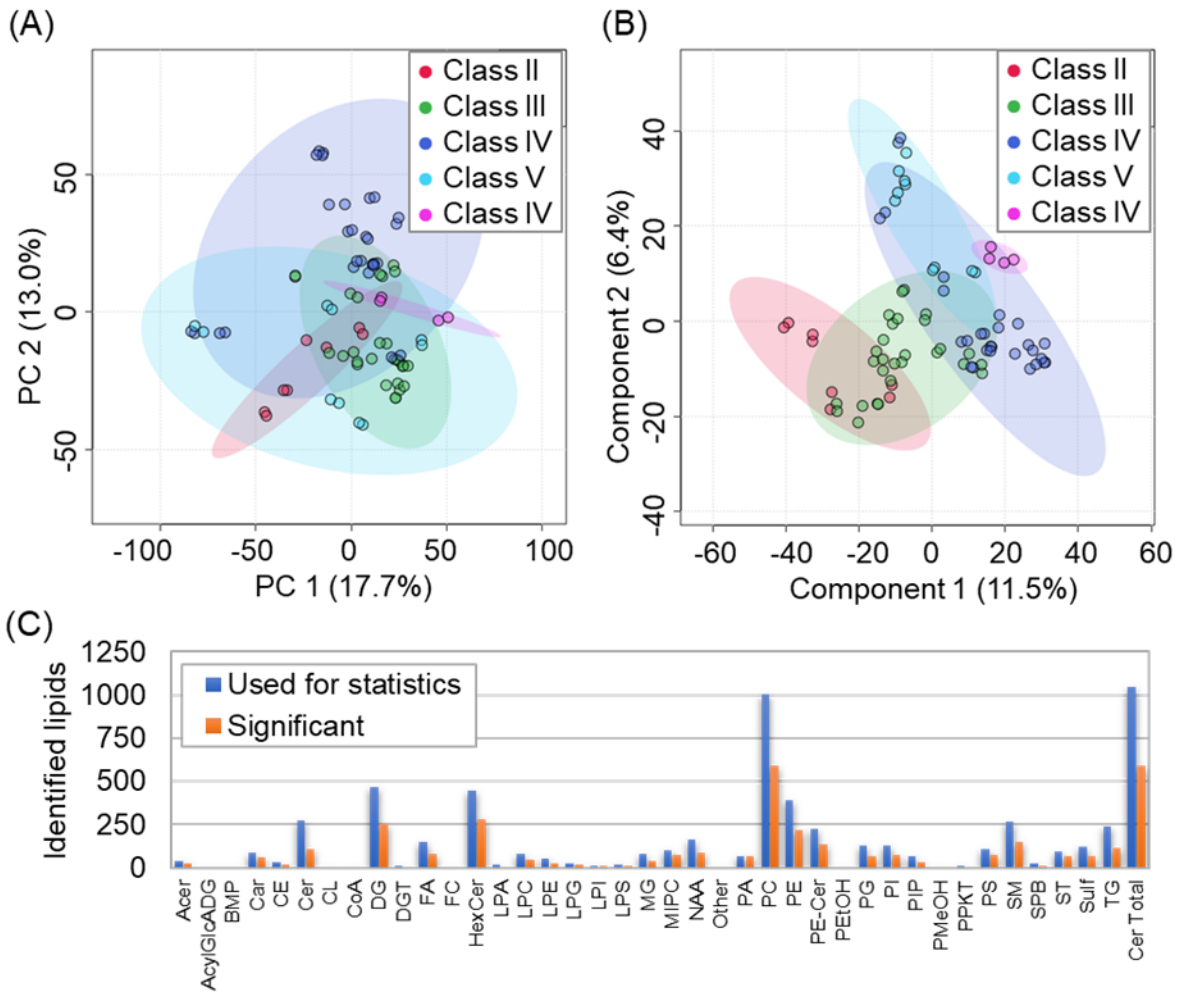

Figure VIII-6. Statistical analysis for all CF genotypic classes (II to VI). (A) PCA score plot; (B) PLS-DA score plot with 8 components (R^2 of 0.9936, Q^2 of 0.8817 and $p < 0.001$ for 1000 permutations); (C) Subclass distribution of all lipids employed for statistics (blue) compared to lipids that were significantly altered ($p < 0.05$) when the different CF genotypic classes were compared by non-parametric ANOVA (orange). Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

Second, we performed binary comparisons for the genotypic classes III, IV and V. Class III is characterized by dysregulated or inactivated CFTR protein (loss of function), while class IV

is related to abnormal conductance of the ion channel and class V, reduced synthesis of CFTR protein. Overall, we observed partial PCA separations (Figure VIII-7A, B and C), while PLS-DA models resulted in better separations (Figure VIII-7D, E and F). Random Forest models for III *versus* IV and IV *versus* V had OOB errors of 0.0% and 2.6%, respectively, but class III *versus* V had a higher error of 5.3%. The Volcano plot comparisons for classes III *versus* IV (603 significantly altered lipids, corresponding to 12.2% of all lipids employed for statistical analysis, Supp. Table 39) and IV *versus* V (696 significantly altered lipids, 18.8%, Supp. Table 40) resulted in high numbers of significantly altered lipids, but classes III *versus* V significantly affected only 130 lipids (3.5% of all lipids employed for statistical analysis, Supp. Table 41) (Figure VIII-7G, H and I). The smaller differences for classes III and V can be related to the characteristic loss of function or reduced synthesis of CFTR, while class IV displays normal CFTR protein, but abnormal ion conductance through the channel.

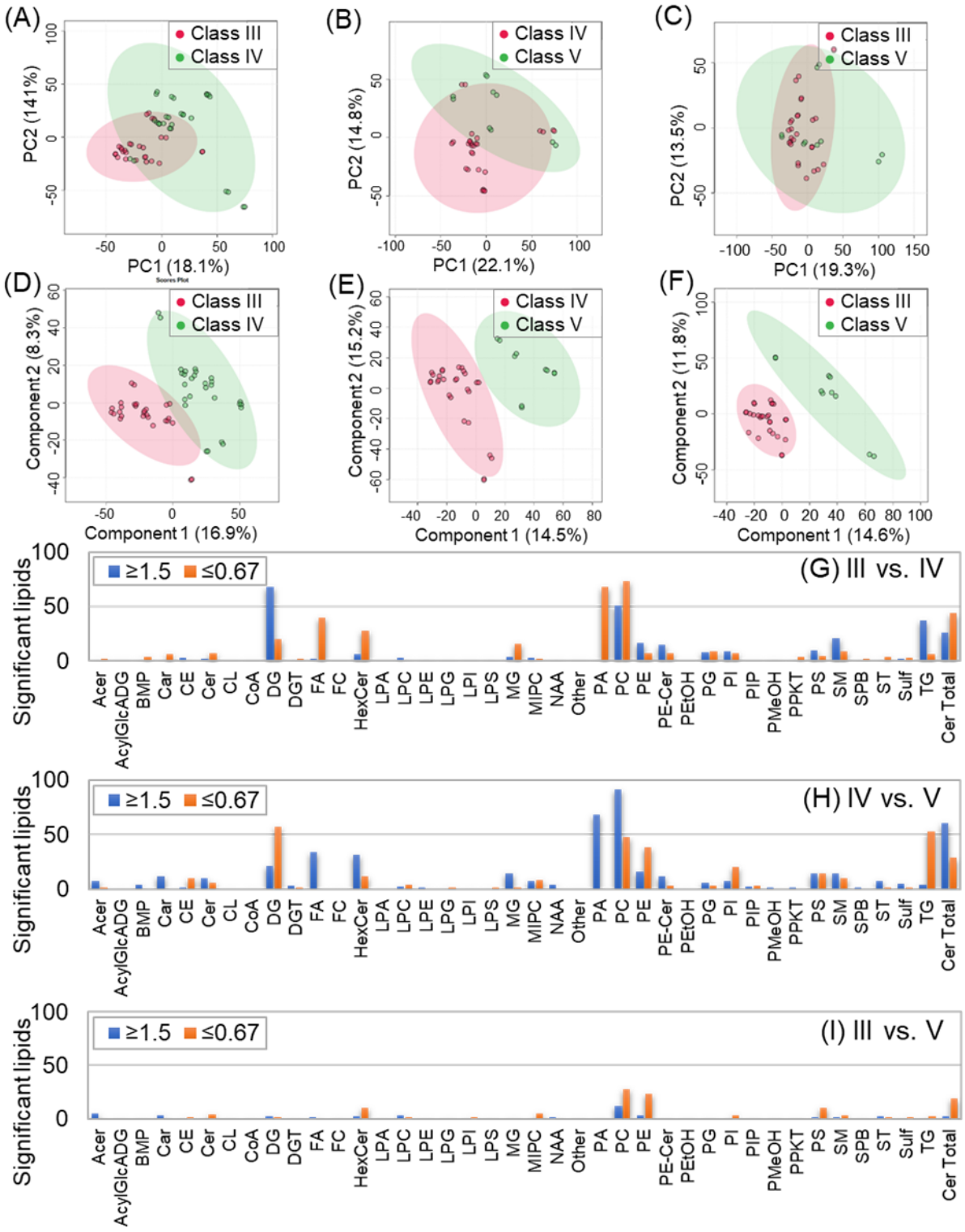

Figure VIII-7. Statistical analysis for binary comparisons between CF genotypic classes III, IV and V. (A) PCA score plot for classes III and IV; (B) PCA score plot for classes IV and V; (C)

PCA score plot for classes III and V; (D) PLS-DA score plot for classes III and IV with 3 components (R^2 of 0.8988, Q^2 of 0.7261 and $p < 0.001$ for 1000 permutations); (E) PLS-DA score plot for classes IV and V with 3 components (R^2 of 0.9417, Q^2 of 0.8040 and $p < 0.001$ for 1000 permutations); (F) PLS-DA score plot for classes III and V with 2 components (R^2 of 0.8820, Q^2 of 0.6749 and $p = 0.001$ for 1000 permutations); (G) lipid subclass distribution of significantly altered lipids in the Volcano plot analysis for the comparison between classes III and IV; (H) lipid subclass distribution of significantly altered lipids in the Volcano plot analysis for the comparison between classes IV and V; (I) lipid subclass distribution of significantly altered lipids in the Volcano plot analysis for the comparison between classes III and V. Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

The abnormal conductance of CFTR in class IV seems to be related to dyslipidemia of complex glycerolipids. We observed lower normalized intensities for diacylglycerols (DG) and triacylglycerols (TG) in class IV, but increased intensities for monoacylglycerols (MG), indicating either an increased breakdown of more complex glycerolipids for class IV or reduction for classes III and V. Furthermore, class IV seems to be closely related to higher normalized intensities of phosphatidic acids (PA) and fatty acyls (FA). All the 68 detected phosphatidic acids were significantly increased for class IV compared to both classes III and IV, but not altered for the direct comparison between classes III and V. Similarly, 47 out of 50 significantly altered fatty acyls were increased for class IV compared to III and V.

Class III and IV were further investigated by ROC analysis. Samples were randomly divided into a test set (10 class III and 10 class IV samples) and a validation set (4 class III and 4 class IV samples). The test set was employed for the selection of biomarker candidates and

building the Random Forest classification model, while the validation set was used to verify the model performance with independent samples. 889 lipids resulted in univariate AUC higher than 0.750, including 211 ceramides (74.9% of the ceramides employed for statistics), 165 phosphatidylcholines (28.5%), 73 diacylglycerols (15.9%) and 64 glycerophosphates (80.0%). The significantly altered lipids were ranked according to their distinct AUC values and importance for Random Forest classification of the test set. A combination of four lipids was selected for a biomarker panel, namely P635.52219/10.62 (identified in tier 3 as HexCer 28:0;O2, TG O-34:0;O, TG 36:2 or TG O-36:3;O), N453.39492/8.94 (identified in tier 3 as FA 28:1;O2 or FA 27:1), N761.54461/9.18 (identified in tier 3 as SM 34:2;O3 or SM 36:2;O5, with the lipid subclass confirmed as SM by the characteristic neutral loss of 60.02 Da and the fragment with m/z 168.04), and P843.69493/13.26 (identified in tier 3 as PE-Cer 46:2;O3, CerP 48:3;O3, PC O-40:3, SM 43:2;O3 or PE O-43:3). The 4-lipid biomarker panel generated an ROC curve with AUC of 0.987 (95% CI of 0.908-1.00), with sensitivity of 90.0%, specificity of 95.0% and overall accuracy of 90.1% (100 cross-validations) (Figure VIII-8). The 4-lipid biomarker panel was validated with the independent validation set, resulting in excellent sensitivity, specificity and accuracy of 100.0%.

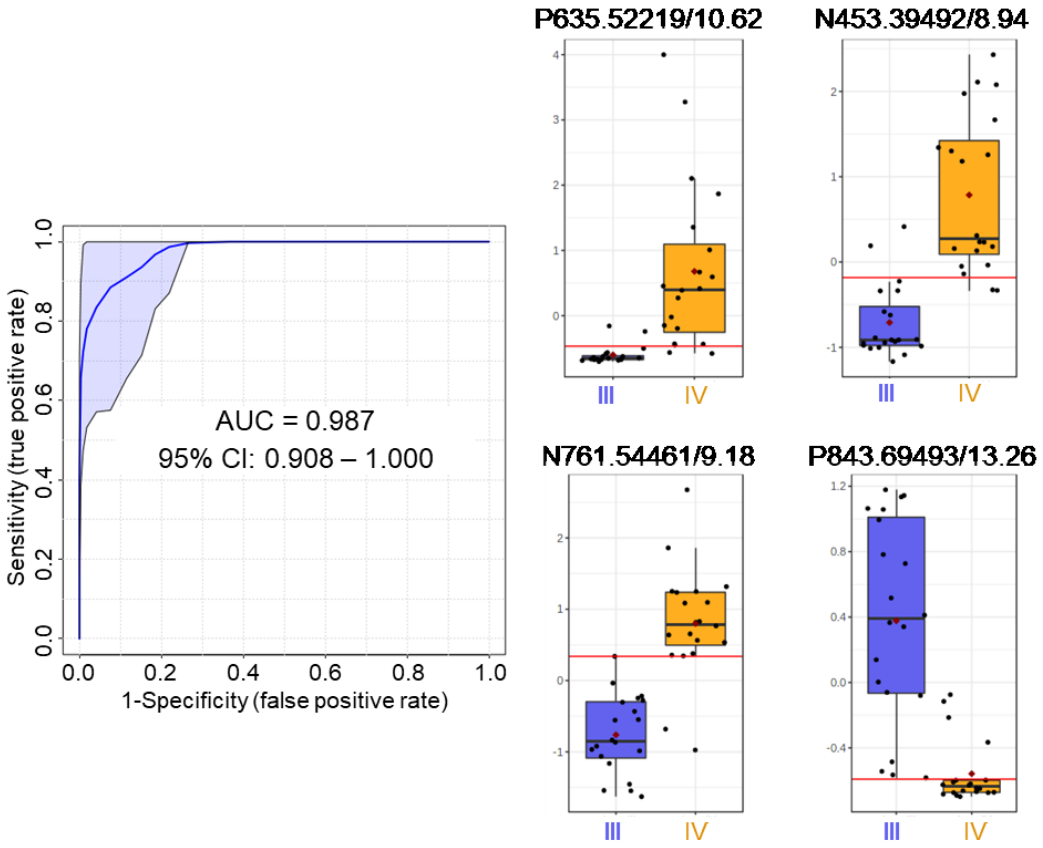

Figure VIII-8. ROC analysis for CF genotypic classes III and IV using a 4-lipid biomarker panel. The boxplots for the biomarker candidates are also shown. Lipids are named by the ionization polarity (P for positive and N for negative), m/z and retention time (min).

8.3.3. Lung function (FEV1, FVC and FEV1%)

The lung function of CF patients can be evaluated through FEV1, the volume of breath forcefully exhaled within 1 second, and FVC, the full amount of air that can be forcefully exhaled in a full breath. The FEV1% or FEV1/FVC ratio, also known as the Tiffeneau-Pinelli index, is often used to diagnose and assess obstructive lung diseases. Decreased FVC values with normal FEV1% is usually an indication of restrictive lung conditions, *e.g.,* idiopathic pulmonary fibrosis, pneumonia or pleural effusion; however, reduced FEV1 and FEV1% values are consistent with

obstructive conditions, *e.g*. asthma, emphysema or bronchiolitis. Normal values for the FEV1% ratio depend on age, sex, height and ethnicity, but are approximately 70-75% or higher. Following spirometry tests, patients were separated into subgroups above and below the median of 72.5% for FEV1%. We performed statistical analysis to compare the two median-split subgroups, looking for alterations in the lipid profile that may be related to loss of lung function in CF patients. The PCA score plots displayed poor separation (Figure VIII-9A), but PLS-DA (Figure VIII-9B) and Random Forest models (OOB error of 1.4%) resulted in good classification performances, indicating differences in the lipid profile that can be related to loss of lung function.

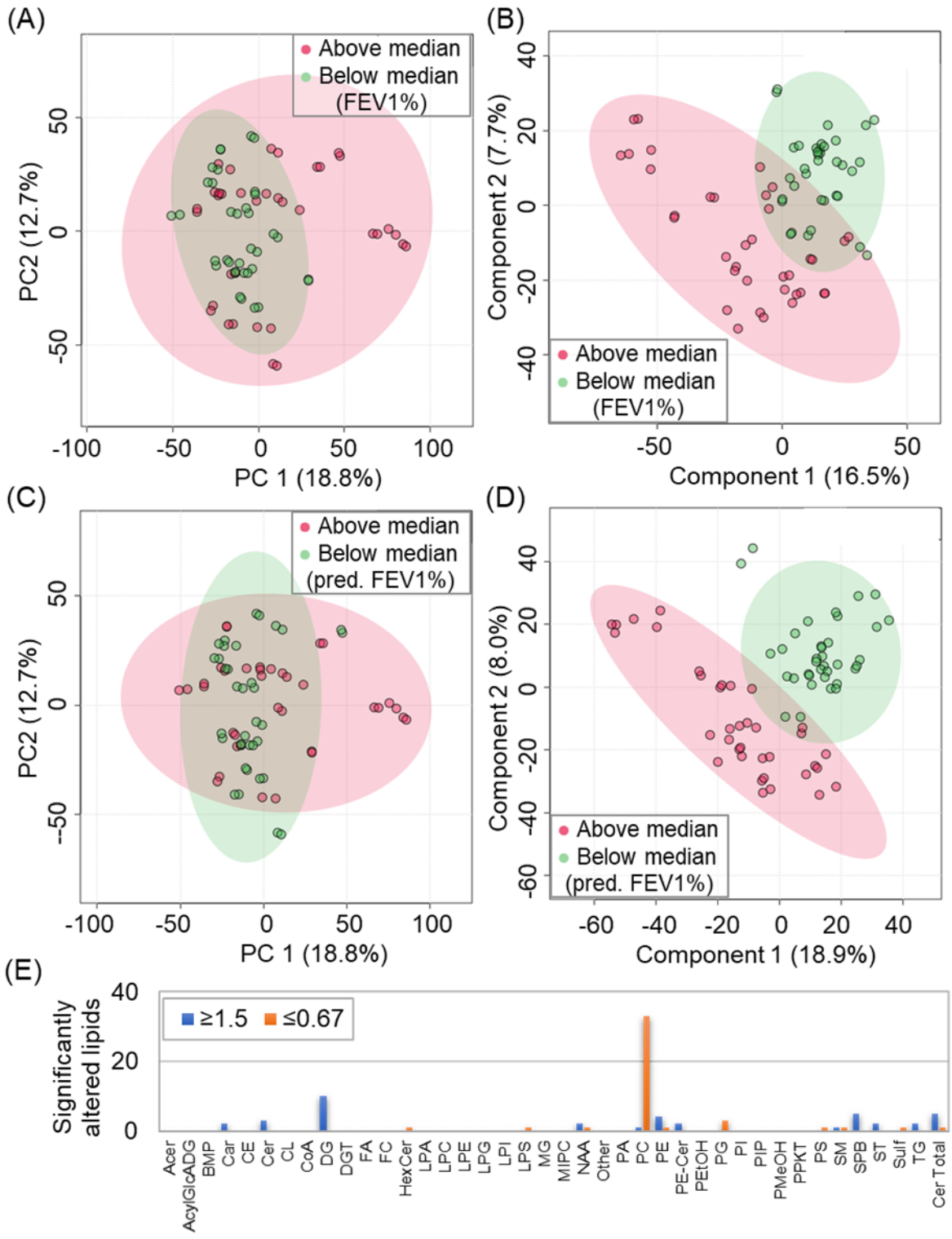

Figure VIII-9. Statistical analysis for patients above (red) and below (green) the median value of FEV1% (A, B) and predicted FEV1% (C, D and E). (A) PCA score plot for FEV1%; (B) PLS-DA

score plot for FEV1% (R^2 of 0.9586, Q^2 of 0.8126 and p=0.04 for 1000 permutations); (C) PCA score plot for predicted FEV1%; (D) PLS-DA score plot for predicted FEV1% (R^2 of 0.9670, Q^2 of 0.8371 and p=0.04 for 1000 permutations); (E) Lipid subclass distribution for significantly altered lipids for the Volcano plot comparison between CF patients with predicted FEV1% values above and below the median of 47.5%, divided into fold-change (FC) ≤0.67 (orange) or ≥1.5 (p adjusted for false-discovery rate <0.05). Lipids with FC ≤0.67 (orange) displayed higher normalized intensities for the patients with reduced lung function, whereas lipids with FC ≥1.5 (blue) had lower normalized intensities. Abbreviations to lipid subclasses are defined in Table II-2 (p. 48).

The predicted FEV1% is defined as the percentage difference between the FEV1% of each patient and the average FEV1% of a population with similar age, sex, height, weight and ethnicity. The European Respiratory Society defines the predicted FEV1% as the main criteria to diagnose chronic obstructive pulmonary disease (COPD, <88% for men or <89% for women). The Global Initiative for Chronic Obstructive Lung Disease (GOLD) sets a limit of predicted FEV1% higher than 80% to diagnose mild COPD, whereas values between 50 and 80% are moderate, 30 to 50% are severe and less than 30% are considered very severe. [303] To evaluate lipidome alterations related to the predicted FEV1%, we separated the patients into subgroups above and below the median of 47.5%. Similarly to the observations found for FEV1%, the PCA score plots displayed poor separation (Figure VIII-9C) due to the simplicity of the model combined to the high complexity of the dataset, but PLS-DA (Figure VIII-9D) and Random Forest models (OOB error of 2.8%) resulted in good classification performances. However, only 77 lipids were significantly affected for the median-split predicted FEV1% (1.6%) (Figure VIII-9E, Supp. Table 42). The most

affected lipid subclasses included glycerophospholipids (44 significantly altered, with 88.6% displaying higher normalized intensities for patients below the median value, *i.e.*, reduced lung function) and sphingolipids (14 significantly altered, with 78.6% displaying lower normalized intensities for patients below the median). Interestingly, all the significantly altered glycerolipids (12 lipids) had lower normalized intensities for the patients with reduced lung function.

We divided the sample set into a test set (13 above the median and 13 below the median patients for FEV1%) and a validation set (5 above the median and 5 below the median samples) for biomarker analysis. The test set was used to select biomarker candidates and build the Random Forest classification model, while the validation set was employed for the independent evaluation of the model performance. The ROC analysis for the median-split groups of predicted FEV1% in the test set resulted in 103 lipids with univariate AUC ≥0.750, including 22 ceramides, 31 phosphatidylcholines and 11 phosphatidylethanolamines. These observations emphasize the effect of reduced lung function upon the ceramide and glycerophospholipid metabolism. The highest univariate AUC was achieved for PE O-18:2_22:0 (positively identified in tier 2); however, the lipid was not significantly altered in the Volcano plot (FC above/ below the median of 1.43, p = 0.001). A panel composed of five lipids was selected amongst the significantly altered lipids based on the highest univariate AUC values and importance for classification of the test set as above or below the median by Random Forest (Figure VIII-10). The panel was composed by N369.17414/1.08 (identified in tier 3 as the sterol sulfate ST 19:1;O2;S), P852.60997/9.04 (identified in tier 3 as PC 40:5;O, with lipid subclass confirmed by the characteristic fragment with m/z 184.07), P779.60475/10.69 (identified in tier 3 as SM 40:5;O2, PE-Cer 43:5;O2, CerP 45:6;O2, PE-Cer 41:2;O2 or SM 38:2;O2), P796.54881/6.52 (identified in tier 3 as PC 36:5;O or PC 34:2;O, with lipid subclass confirmed as PC by the characteristic fragment at m/z 184.07 and

neutral loss of choline), and N731.53423/9.72 (identified in tier 3 as the SM 33:2;O2 or SM 35:2;O4, with the lipid subclass confirmed by the characteristic neutral loss of 60.01 and the fragment with m/z 168.04). The 5-lipid biomarker panel resulted in a multivariate AUC of 0.932 (95% CI of 0.815 – 1.000) for the test set, with an overall accuracy of 83.7% (cross-validation), sensitivity of 80.8% and specificity 88.5%. The validation set resulted in sensitivity, specificity and accuracy of 90.0%. Even though the ROC validation results are inferior to the observed values for the control/CF diagnosis, the selected biomarker panel emphasizes the extensive effect of lung function decay on the lipidome.

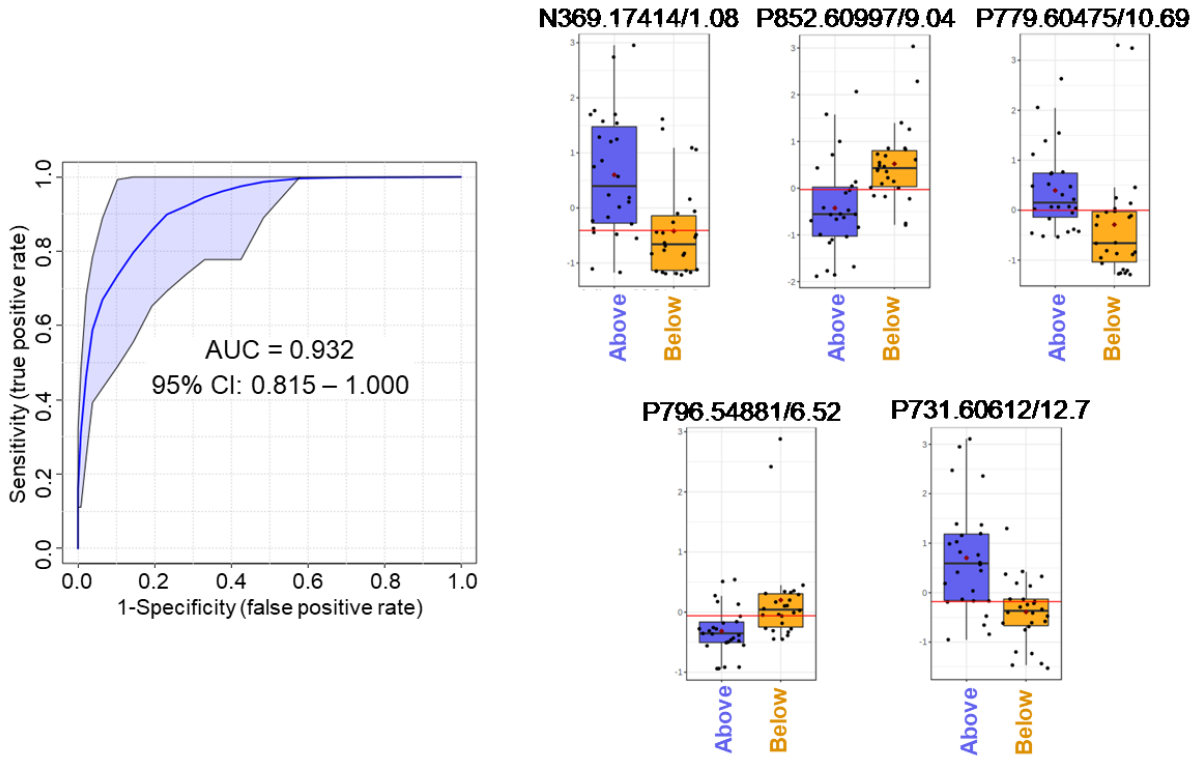

Figure VIII-10. ROC analysis for CF patients with predicted FEV1% values above and below the median of 47.5% using the proposed 5-lipid biomarker panel. Boxplots for the biomarker

candidates are also shown. Lipids are named by the ionization polarity (P for positive and N for negative), m/z and retention time (min).

8.3.4. Evaluation of age and sex as possible confounders

We evaluated sex and age for potential confounding effects upon the observed results. The CF patients were first divided at the median value for age (20 years old). Only 27 lipids were significantly altered in a Volcano plot analysis between the patients above and below the median age, including 10 ceramides, and 6 diacylglycerols (Supp. Table 43). Amongst the 27 significant lipids, 25 were also significantly altered for the comparison between CF patients and healthy controls. For sex, 90 lipids were significantly altered for males and females CF patients, including 29 triacylglycerols, 21 diacylglycerols and 11 phosphatidylcholines (Supp. Table 44). The lipids that were significantly altered for age and sex were not considered as potential biomarkers due to the possibility of influence of the confounders upon the observed results.

8.4 Discussion

We have performed untargeted, comprehensive lipidomics of blood serum to explore the potential of lipidomics for studying the pathogenesis of cystic fibrosis (CF). First, we compared the lipid profile of CF patients with healthy controls to evaluate the changes caused by the disease, regardless of the mutation type. There was a dramatic change in the serum lipidome of CF patients characterized by dyslipidemia for most lipid subclasses. We selected six potential independent biomarkers that were able to classify samples as CF or control with high sensitivity, specificity and accuracy. Second, we employed the same techniques to compare CF patients with different genotypic classes (II to VI). We observed a correlation between class IV and increased intensities

of phosphatidic acids and fatty acyl lipids. Classes III and IV were selected for biomarker analysis, and a panel of four lipids, including two putatively identified ceramides, one sphingomyelin and one fatty acid, correctly classified all samples retained for the validation set. Third, we have explored the correlation between the lipidome and lung function in terms of predicted FEV1%. A 5-lipid biomarker panel was correlated with reduced lung function with an accuracy of 90% for the validation set.

Our approach included novel aspects when compared to previous studies. Untargeted, comprehensive lipidomics of CF has been rarely described in the literature. Traditional lipidomics usually focuses on one or a few selected lipid classes to reduce the high complexity of the lipidome for biological samples. However, the LC-MS methodology employed for this study allowed untargeted, comprehensive lipidomics with high sensitivity for the detection of more than seven thousand lipids belonging to 39 different lipid subclasses using less than 20 µL of blood serum. Also, fully untargeted lipidomics approaches based on LC-QToF-MS usually lack on the identification of lipids; however, our optimized techniques identified 1348 lipids by MS/MS spectral match. When allied to the data processing routine, the method provided reliable results for the statistical analysis of CF lipidome. We were able to pinpoint extensive alterations caused by the disease, as well as select potential biomarkers that may be used for future studies.

The dyslipidemia observed for most lipid subclasses when CF patients were compared to healthy controls is a known effect of this condition. [304] We observed significantly lower levels of serum cholesterol in CF patients (FC CF/control of 0.53), which is corroborated by the poor nutrient absorption that is characteristic of the pathology, as well as higher intracellular cholesterol levels.[299,305,306] However, we also noticed higher levels of 7-ketocholesterol (FC CF/control of 1.95), possibly indicating higher rates of oxidation. Other studies have also reported indications of

increased oxidative stress in CF patients. [307,308] The higher oxidation rate noticed herein is also indicated by the lipid that displayed the highest fold-change between CF patients and healthy controls, identified in tier 3 as FA 20:2;O4 (FC CF/control of 1.0×10^{-3}; putatively identified as prostaglandin 13,14-dihydro-19R-hydroxyPGE1, 19R-hydroxy-PGF1α or TXB1), which is related to inflammation and tissue damage. [26,41,309]

The comparison between CF patients and healthy controls also showcased the importance of odd-chain fatty acids (OCFA), rarely investigated in human samples. The four positively identified lipids (tiers 1 and 2) with the highest significant fold-changes for the CF/control comparison contained at least one OCFA moiety: PC 15:1_21:2, PC 19:2_19:2, PS 22:0_17:2 and DGT 25:0_24:4. In fact, we found 139 significantly altered OCFA lipids that were identified by MS/MS in tiers 1 and 2. This represents the first report of elevated levels of OCFA in the blood serum of CF patients. Lipids with OCFA were also affected for different genotypic mutation classes, indicating that the alterations are present even for patients on controlled diets. While only 45 OCFA lipids were altered for the comparison between classes III and V, over 170 were affected for class III versus IV and IV versus V. Such molecules are more commonly found in ruminant fat and milk, as well as in some plants and bacteria metabolism. The majority of fatty acid studies in human samples focus on the more concentrated even-chain fatty acids; hence, many researchers have used saturated OCFA lipids as low-cost internal standards for GC and LC-MS analysis. Unfortunately, the procedure prevented their incorporation into statistical analysis, limiting our current understanding of OCFA metabolism and its relationship with human diseases.[298,310] Although OCFAs have small concentrations in human fluids and tissues, they can be detected by high sensitivity methods, such as the one adopted for this work. Large-scale epidemiological studies have shown that plasma OCFA levels may be associated with type 2 diabetes and coronary

heart disease; now, our results indicate that OCFA-containing lipids may also play an important role for CF. [311–313]

Previous studies have shown that lipids containing C15:0 fatty acyls may be correlated with diet, but the same relationship was not observed for C17:0 fatty acids, indicating that non-dietary OCFA lipids can be biologically synthesized or modified in animals.[314] The source of distinct OCFAs with variable carbon lengths seems to be related to different processes, including diet and biosynthesis. There has been some evidence that OCFAs may be generated by human metabolism through biosynthesis via propionyl-CoA and α-oxidation of even-chain fatty acids, which can be affected by metabolic and genetic pathologies. [310,314] The biosynthesis of C17 fatty acyls is substantially affected by 2-hydroxyphytanoyl-CoA lyase (*Hacl1*), an enzyme involved in the α-oxidation of even-chain fatty acids through the cleavage of a carbon-carbon bond, generating an OCFA.[314] However, OCFAs can also be synthesized by the conversion of propionate into propionyl-CoA, which in turn is condensed with malonyl-CoA, resulting in 3-oxovaleryl-ACP. The compound is then elongated by several consecutive additions of two carbons though the action of fatty acid synthases. The resulting product is an OCFA, which can be esterified to generate different lipids. Furthermore, the oxidation of plasmalogens can produce odd-chain fatty aldehydes, which may be further incorporated into other lipid classes. Alternatively, the degradation of OCFA though β-oxidation yields propionyl-CoA, which is converted into methylmalonyl-CoA by the action of propionyl-CoA carboxylase. The product is further converted into succinyl-CoA for entry into the citric acid cycle, finalizing the degradation pathway of OCFA, which results in ATP production. Propionyl-CoA is also obtained by the β-oxidation of amino acids and cholesterol.

It is yet unclear if the increased levels of OCFAs observed for CF patients in this work are related to increased synthesis via propionyl-CoA, higher rates of α-oxidation or reduced rate of β-oxidation. We observed decreased levels of most MS/MS-identified plasmalogens, indicating higher rates of oxidative processes, including α and β-oxidation of lipids. We have also previously discussed the indications of oxidative stress for cholesterol, resulting in increased production of propionyl-CoA as a degradation product, which could be related to the observed higher rates of OCFA biosynthesis. Further studies are required to confirm the hypotheses. Nevertheless, higher levels of OCFAs have also been reported for patients with propionic acidemia, a genetic condition characterized by mutations in the propionyl-CoA carboxylase subunit alpha (*PCCA*) and propionyl-CoA carboxylase subunit beta (*PCCB*) genes, causing absence or dysfunction of propionyl-CoA carboxylase and accumulation of propionyl-CoA in the organism.[315] The accumulation can cause impairment, mental disability, lethargy and other issues due to an abnormal buildup of organic acids. Higher levels of propionyl-CoA may induce higher rates of OCFA biosynthesis, as well as inhibition of OCFA degradation, which could cause the previously reported higher levels of OCFA for propionic acidemia patients. Similar alterations could explain the observed increased levels of OCFA for CF patients.[315]

Amongst the most significantly altered lipids for the comparison between CF patients and healthy controls, we also found many elevated PUFA oxylipins, such as PI 18:0_20:4-12-HETE (FC CF/control of 17.2), PE O-18:1_20:4-12-HETE (FC CF/control of 13.4) and PE 16:0_20:4-8-HETE (FC CF/control of 12.6). The hydroxy-eicosatetraenoic acids 12-HETE and 8-HETE are majorly produced by the lipoxygenase pathway, closely related to the cyclooxygenase (COX) pathway, which catalyzes the conversion of arachidonic acid into prostaglandins and thromboxanes. [308] Such oxylipins are key inflammatory mediators derived from arachidonic acid,

which also displayed higher normalized intensities for CF patients (FA 20:4, identified in tier 1 with FC CF/control of 1.30 and p of 1.9×10^{-4}). Although we noticed elevated levels of important lipids related to oxidation and inflammation, most lipid subclasses were significantly decreased for CF patients, including total N-acyl amines, fatty acids and steroid sulfates. Most steroid sulfates are precursors for active hormones, but some may act as neurosteroids and ion channels modulators, including γ-aminobutyric acid ($GABA_A$) and N-methyl-D-aspartate (NMDA) receptors. Steroids are sulfated to promote circulation, but sulfatases can hydrolyze the sulfate esters inside cells to allow the activation of steroid function.[316]

Phosphatidic acids (PA) were particularly affected by different CF genotypic classes. All the 68 detected PAs were significantly elevated for class IV when compared to classes III (FC III/IV between 0.20 and 0.61) and V (FC IV/V between 2.2 and 6.1). However, whether that observation was caused by the upregulation of PAs for class IV or downregulation in classes III and V remains to be determined. Only 12 PAs were significantly altered when all CF patients were compared to healthy controls, including 10 with significantly lower normalized intensities for CF patients, which emphasizes the effect of different genotypic classes upon PA metabolism. PAs act as precursors for the biosynthesis of glycerolipids and other glycerophospholipids, cell signaling and activation of lipid-gated ion channels. Even though PAs are essential for lipid metabolism and survival, their concentrations are usually low, as the molecules are rapidly converted into diacylglycerols and other lipids. We observed higher levels of all the detected PAs for class IV compared to III, but lower levels of 77.3% of the significantly altered diacylglycerols and 86.0% of triacylglycerols, indicating alterations in the activity of phosphatidic acid phosphatases. A similar trend was seen for class IV compared to V. CFTR channel gating is mostly regulated by protein kinase A phosphorylation, conformational changes and protein trafficking from the

endoplasmic reticulum to the Golgi apparatus and the cell membrane.[317] Protein kinases and phosphatases are potential regulators of CFTR channel activation. There is also some evidence of the regulatory function of PAs in CFTR transport between the endoplasmic reticulum and the Golgi apparatus. Hashimoto *et al*. have shown that the concentrations of PA must be regulated by phospholipase D to ensure post-Golgi transport of the CFTR protein. [318] However, the genotypic mutation classes III, IV and V are usually characterized by normal CFTR transport, but abnormal channel functioning (III), abnormal conductance (IV), or insufficient amounts (V). We found evidence that CF genotypic mutations have different effects upon PA levels. The elevated levels of PA and reduced levels of diacylglycerols for class IV may be related to a reduced activity of phosphatases, which would potentially affect concentrations of both lipid classes, as well as CFTR channel gating.

The 5-lipid biomarker panel selected for lung function (predicted FEV1%) included two phosphatidylcholines (PC 40:5;O and PC 36:5;O or 34:2;O), two sphingomyelins (SM 33:2;O2 or 35:2;O4 and SM 40:5;O2 or 38:2;O2) and one steroid sulfate (ST 19:1;O2, putatively identified as 5a-dihydrotestosterone sulfate, 5α-androstane-3α-ol-17-one sulfate or etiocholanolone sulfate). Although all five lipids were putatively identified in tier 3 by accurate mass-match, the lipid subclasses of three of them (both phosphatidylcholines and one sphingomyelin) were confirmed by characteristic MS/MS fragmentation and neutral losses. Interestingly, all 5 molecules were putatively identified as oxidized lipids, correlating to our findings for plasmalogens, cholesterol and oxylipins. CF has been previously related to elevated markers of oxidative stress.[319] Chronically elevated oxidation can perturb homeostasis through alterations of signaling molecules and biological pathways, such as glutathione, cysteine, oxidized nicotinamide adenine dinucleotide phosphate ($NADP^+$), prostaglandins and nitrotyrosine. [307,319,320] Hydrolysis products

of phosphatidylcholines have been previously related to lung surfactant dysfunction. [305] The 5-lipid biomarker panel selected for predicted FEV1% emphasizes the importance of redox regulation in lung function.

Ex vivo lipid oxidation is often cited as a limitation of untargeted lipidomics studies. However, the sample storage and preparation employed in this worked aimed to minimize oxidative processes by protecting the samples from light, atmospheric air and heat. We opted for not employing anti-oxidants as the molecules are known to have little, unreliable effect upon different lipid classes, along with the risk of a pro-oxidative action for selected species and concentrations.[321,322] Nevertheless, we were able to detected and identify more than 280 plasmalogens, a lipid class that is known to undergo quick oxidation due to the highly susceptible vinyl-ether bonds, indicating that oxidative processes were controlled during sample handling. We recognize that this work is limited by the small number of individuals and the presence of potential confounders, *e.g.,* medication and treatments. It is also worth noticing that the diet of patients and healthy control subjects was not controlled, which may affect serum lipidome. However, the results described herein confirm the potential of high-quality untargeted lipidomics studies for pathogenesis analyses and biomarker discovery.

8.5 Conclusions

We have applied comprehensive, untargeted LC-MS lipidomics to investigate CF-related alterations upon the lipidome metabolism in human patients. The results described in this work indicate that CF is deeply related to alterations in lipid pathways, particularly for odd-chain and polyunsaturated fatty acyls and phosphatidic acids. The proposed biomarkers may be used for reliable screening of cystic fibrosis, as well as for studying the affected biological pathways and

the efficacy of new drugs and treatments. Furthermore, we have achieved a novel perspective for CF and found new clues for further research, where the role of lipids may be of vital importance.

8.6 Acknowledgments

We thank Dr. Brian Meyer (Chairman of the Department of Genetics, King Faisal Specialist Hospital and Research Centre) and all the volunteers that donated blood for this work. We also appreciate the collaboration with Dr. Anas M. Abdel Rahman (King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia). This project was funded by the Department of Genetics and Research Center at King Faisal Specialist Hospital and Research Center and grants from the Natural Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research, the Canada Research Chairs program, the Canadian Foundation for Innovations, Genome Canada and Alberta Innovates.

IX

Chapter IX: Conclusions

Lipidomics, along with other *omic* sciences, represent a huge step for biomedical research. The detailed description of physiological and pathological processes should significantly improve our understanding of biological and disease-associated mechanisms, facilitating biomarker discovery and the development of new therapies. The importance of lipids to metabolism is undeniable. The hydrophobic molecules are involved in virtually all processes in our body but impose a challenge for analytical approaches. Lipidomics is a highly complex field of research that demands interdisciplinary knowledge and experience.

This work showcased the potential of lipidomics to study human diseases and physiological processes with LC-MS. Although the analytical aspect was the main goal, each chapter described different lipid pathways and mechanisms that may be related to specific pathologies or treatments. Furthermore, excellent results were achieved with small volumes of precious, invasive biological samples. The integration of lipidomics and metabolomics data was successfully achieved in Chapter V and will be further employed in related studies in the near future.

From extremely low volumes of samples on chapters III, IV and V to the MS/MS identification of more than one thousand lipids in Chapter VIII, the analytical and biochemical possibilities of lipidomics are endless. There is still much room for improvements in the lipidomics field. Although this work presented reliable analytical results, the detection of thousands of features is meaningless without accurate identifications. The standardization of nomenclature employed by different databases and confirmation of identification is a manual, laborious work that still requires a lot of human intellect rather than automatized software. Chapter VIII showed that a carefully optimized LC-MS/MS method can result in the identification of over one thousand

lipids in blood serum, but there is still an under-explored potential that requires the expansion and standardization of lipid databases. Furthermore, the standardization of data analysis and reporting, as well as minimal requirements for quality control and normalization, are urgently needed. Scientific research has been discredited by a considerable portion of the population in the last few years; hence, it is vital to ensure that accurate and reliable results are communicated not only to our peers, but also to the population in an appropriate, simple manner.

In summary, the full potential of lipidomics and *omic* sciences has not yet been achieved, but each step takes us closer to understanding the role of lipids in our organism. This work described major advances in lipidomics of biological fluids and can aid to future developments and breakthroughs in this complex but highly attractive area.

References

(1) Sarkar, I.; Zardini Buzatto, A.; Garg, R.; Li, L.; van Drunen Littel-van den Hurk, S. Metabolomic and Immunological Profiling of Respiratory Syncytial Virus Infection after Intranasal Immunization with a Subunit Vaccine Candidate. *J. Proteome Res.* **2019**, *18* (3), 1145–1161.

(2) Zardini Buzatto, A.; Sarkar, I.; van Drunen Littel-van den Hurk, S.; Li, L. Comprehensive Lipidomic and Metabolomic Analysis for Studying Metabolic Changes in Lung Tissue Induced by a Vaccine against Respiratory Syncytial Virus. *ACS Infect. Dis.* **2020**, *6* (8), 2130–2142.

(3) Dahabiyeh, L.; Malkawi, A.; Wang, X.; Colak, D.; Mujamammi, A.; Sabi, E.; Li, L.; Dasouki, M.; Abdel Rahman, A. Dexamethasone-Induced Perturbations in Tissue Metabolomics Revealed by Chemical Isotope Labeling LC-MS Analysis. *Metabolites* **2020**, *10* (2), 42.

(4) Bou Khalil, M.; Hou, W.; Zhou, H.; Elisma, F.; Swayne, L. A.; Blanchard, A. P.; Yao, Z.; Bennett, S. A. L.; Figeys, D. Lipidomics Era: Accomplishments and Challenges. *Mass Spectrom. Rev.* **2010**, *29* (6), 877–929.

(5) Buzatto, A. Z.; de Sousa, A. C.; Guedes, S. F.; Cieslarová, Z.; Simionato, A. V. C. Metabolomic Investigation of Human Diseases Biomarkers by CE and LC Coupled to MS. *Electrophoresis* **2014**, *35* (9), 1285–1307.

(6) Triebl, A.; Hartler, J.; Trötzmüller, M.; C. Köfeler, H. Lipidomics: Prospects from a Technological Perspective. *Biochim. Biophys. Acta* **2017**, *1862* (8), 740–746.

(7) Wolf, C.; Quinn, P. J. Lipidomics: Practical Aspects and Applications. *Prog. Lipid Res.*

2008, *47* (1), 15–36.

(8) Han, X.; Gross, R. W. Global Analyses of Cellular Lipidomes Directly from Crude Extracts of Biological Samples by ESI Mass Spectrometry. *J. Lipid Res.* **2003**, *44* (6), 1071–1079.

(9) Mouritsen, O. G. *Life - as a Matter of Fat: The Emerging Science of Lipidomics*, 1st ed.; Springer-Verlag: Heidelberg, Germany, 2005.

(10) Navas-Iglesias, N.; Carrasco-Pancorbo, A.; Cuadros-Rodríguez, L. From Lipids Analysis towards Lipidomics, a New Challenge for the Analytical Chemistry of the 21st Century. Part II: Analytical Lipidomics. *TrAC Trends Anal. Chem.* **2009**, *28* (4), 393–403.

(11) Hyötyläinen, T.; Orešič, M. Systems Biology Strategies to Study Lipidomes in Health and Disease. *Prog. Lipid Res.* **2014**, *55* (1), 43–60.

(12) Rustam, Y. H.; Reid, G. E. Analytical Challenges and Recent Advances in Mass Spectrometry Based Lipidomics. *Anal. Chem.* **2018**, *90* (1), 374–397.

(13) Gross, R. W. The Evolution of Lipidomics through Space and Time. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2017**, *1862* (8), 731–739.

(14) Wang, J.; Wang, C.; Han, X. Tutorial on Lipidomics. *Anal. Chim. Acta* **2019**, *1061*, 28–41.

(15) Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentals of Biochemistry*, 1st ed.; John Wiley & Sons, Inc.: New York, NY, USA, 1998.

(16) Fahy, E.; Subramaniam, S.; Brown, H. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Seyama, Y.; Shaw, W.; Shimizu, T.; Spener, F.; van Meer, G.; VanNieuwenhze, M. S.; White, S. H.; Witztum, J. L.; Dennis, E. A. A Comprehensive Classification System for Lipids. *Eur. J. Lipid Sci. Technol.* **2005**, *107* (5), 337–364.

(17) Fahy, E.; Subramaniam, S.; Murphy, R. C.; Nishijima, M.; Raetz, C. R. H.; Shimizu, T.; Spener, F.; van Meer, G.; Wakelam, M. J. O.; Dennis, E. A. Update of the LIPID MAPS Comprehensive Classification System for Lipids. *J. Lipid Res.* **2009**, *50* (Supplement), S9–S14.

(18) Fahy, E.; Cotter, D.; Sud, M.; Subramaniam, S. Lipid Classification, Structures and Tools. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2011**, *1811* (11), 637–647.

(19) Brown, H. A.; Marnett, L. J. Introduction to Lipid Biochemistry, Metabolism, and Signaling. *Chem. Rev.* **2011**, *111* (10), 5817–5820.

(20) Liebisch, G.; Vizcaíno, J. A.; Köfeler, H.; Trötzmüller, M.; Griffiths, W. J.; Schmitz, G.; Spener, F.; Wakelam, M. J. O. Shorthand Notation for Lipid Structures Derived from Mass Spectrometry. *J. Lipid Res.* **2013**, *54* (6), 1523–1530.

(21) *Anatomy and Physiology*, 1st ed.; Rice University, OpenStax, BCcampus Open Textbooks: Houston, TX, USA, 2013.

(22) de Jong, A. J.; Kloppenburg, M.; Toes, R. E. M.; Ioan-Facsinay, A. Fatty Acids, Lipid Mediators, and T-Cell Function. *Front. Immunol.* **2014**, *5* (OCT), 3–9.

(23) Köberlin, M. S.; Snijder, B.; Heinz, L. X.; Baumann, C. L.; Fauster, A.; Vladimer, G. I.; Gavin, A. C.; Superti-Furga, G. A Conserved Circular Network of Coregulated Lipids Modulates Innate Immune Responses. *Cell* **2015**, *162* (1), 170–183.

(24) de Pablo, M. Immune Cell Functions, Lipids and Host Natural Resistance. *FEMS Immunol. Med. Microbiol.* **2000**, *29* (4), 323–328.

(25) Howie, D.; Ten Bokum, A.; Necula, A. S.; Cobbold, S. P.; Waldmann, H. The Role of Lipid Metabolism in T Lymphocyte Differentiation and Survival. *Front. Immunol.* **2018**, *8* (JAN).

(26) Cabral, G. A. Lipids as Bioeffectors in the Immune System. *Life Sci.* **2005**, *77* (14), 1699–1710.

(27) Yaqoob, P. Lipids and the Immune Response: From Molecular Mechanisms to Clinical Applications. *Curr. Opin. Clin. Nutr. Metab. Care* **2003**, *6* (2), 133–150.

(28) García-Sastre, A. Lessons from Lipids in the Fight against Influenza. *Cell* **2013**, *154* (1), 22–23.

(29) D'Avila, H.; Maya-monteiro, C. M.; Bozza, P. T.; Avila, H. D.; Maya-monteiro, C. M.; Bozza, P. T.; D'Avila, H.; Maya-monteiro, C. M.; Bozza, P. T.; Avila, H. D.; Maya-monteiro, C. M.; Bozza, P. T. Lipid Bodies in Innate Immune Response to Bacterial and Parasite Infections. *Int. Immunopharmacol.* **2008**, *8* (10), 1308–1315.

(30) Lee, S. T.; Lee, J. C.; Kim, J. W.; Cho, S. Y.; Seong, J. K.; Moon, M. H. Global Changes in Lipid Profiles of Mouse Cortex, Hippocampus, and Hypothalamus Upon P53 Knockout. *Sci. Rep.* **2016**, *6* (1), 36510.

(31) Schulte, E. C.; Altmaier, E.; Berger, H. S.; Do, K. T.; Kastenmüller, G.; Wahl, S.; Adamski, J.; Peters, A.; Krumsiek, J.; Suhre, K.; Haslinger, B.; Ceballos-Baumann, A.; Gieger, C.; Winkelmann, J. Alterations in Lipid and Inositol Metabolisms in Two Dopaminergic Disorders. *PLoS One* **2016**, *11* (1), e0147129.

(32) Skoumalová, A.; Ivica, J.; Šantorová, P.; Topinková, E.; Wilhelm, J. The Lipid Peroxidation Products as Possible Markers of Alzheimer's Disease in Blood. *Exp. Gerontol.* **2011**, *46* (1), 38–42.

(33) Xicoy, H.; Wieringa, B.; Martens, G. J. M.; Xicoy, H.; Wieringa, B.; Martens, G. J. M. The Role of Lipids in Parkinson's Disease. *Cells* **2019**, *8* (1), 27.

(34) Betsholtz, C. Lipid Transport and Human Brain Development. *Nat. Genet.* **2015**, *47* (7), 699–701.

(35) Müller, C. P.; Reichel, M.; Mühle, C.; Rhein, C.; Gulbins, E.; Kornhuber, J. Brain Membrane Lipids in Major Depression and Anxiety Disorders. *Biochim. Biophys. Acta* **2015**, *1851* (8), 1052–1065.

(36) Galvagnion, C. The Role of Lipids Interacting with α-Synuclein in the Pathogenesis of Parkinson's Disease. *J. Parkinsons. Dis.* **2017**, *7* (3), 433–450.

(37) Fanning, S.; Selkoe, D.; Dettmer, U. Parkinson's Disease: Proteinopathy or Lipidopathy? *npj Park. Dis.* **2020**, *6* (1), 3.

(38) Shichiri, M.; Yoshida, Y.; Niki, E. Unregulated Lipid Peroxidation in Neurological Dysfunction. In *Omega-3 Fatty Acids in Brain and Neurological Health*; Elsevier, 2014; pp 31–55.

(39) Han, X.; Holtzman, D. M.; McKeel, D. W. Plasmalogen Deficiency in Early Alzheimer's Disease Subjects and in Animal Models: Molecular Characterization Using Electrospray Ionization Mass Spectrometry. *J. Neurochem.* **2001**, *77* (4), 1168–1180.

(40) Tam, V. C.; Quehenberger, O.; Oshansky, C. M.; Suen, R.; Armando, A. M.; Treuting, P. M.; Thomas, P. G.; Dennis, E. A.; Aderem, A. Resource Lipidomic Profiling of Influenza Infection Identifies Mediators That Induce and Resolve Inflammation. *Cell* **2013**, *154* (1), 213–227.

(41) McCarthy, M. K.; Weinberg, J. B. Eicosanoids and Respiratory Viral Infection: Coordinators of Inflammation and Potential Therapeutic Targets. *Mediators Inflamm.* **2012**, *2012*, 1–13.

(42) Seo, Y.-J. J.; Alexander, S.; Hahm, B. Does Cytokine Signaling Link Sphingolipid Metabolism to Host Defense and Immunity against Virus Infections? *Cytokine Growth Factor Rev.* **2011**, *22* (1), 55–61.

(43) Worgall, T. S. Lipid Metabolism in Cystic Fibrosis. *Curr. Opin. Clin. Nutr. Metab. Care* **2009**, *12* (2), 105–109.

(44) Horati, H.; Janssens, H. M.; Margaroli, C.; Veltman, M.; Stolarczyk, M.; Kilgore, M. B.; Chou, J.; Peng, L.; Tiddens, H. A. M. W.; Chandler, J. D.; Tirouvanziam, R.; Scholte, B. J. Airway Profile of Bioactive Lipids Predicts Early Progression of Lung Disease in Cystic Fibrosis. *J. Cyst. Fibros.* **2020**, No. xxxx, 1–8.

(45) Tu, J.; Yin, Y.; Xu, M.; Wang, R.; Zhu, Z.-J. Absolute Quantitative Lipidomics Reveals Lipidome-Wide Alterations in Aging Brain. *Metabolomics* **2018**, *14* (1), 5.

(46) Anthonymuthu, T. S.; Kenny, E. M.; Lamade, A. M.; Kagan, V. E.; Bayır, H. Oxidized Phospholipid Signaling in Traumatic Brain Injury. *Free Radic. Biol. Med.* **2018**, *124* (June), 493–503.

(47) Anthonymuthu, T. S.; Kim-Campbell, N.; Bayir, H. Oxidative Lipidomics: Applications in Critical Care. *Curr. Opin. Crit. Care* **2017**, *23* (4), 251–256.

(48) Hall, E. D.; Wang, J. A.; Bosken, J. M.; Singh, I. N. Lipid Peroxidation in Brain or Spinal Cord Mitochondria after Injury. *J. Bioenerg. Biomembr.* **2016**, *48* (2), 169–174.

(49) Niki, E. Lipid Peroxidation: Physiological Levels and Dual Biological Effects. *Free Radic. Biol. Med.* **2009**, *47* (5), 469–484.

(50) Golbraikh, A.; Tropsha, A. Unregulated Lipid Peroxidation in Neurological Dysfunction. *J. Mol. Graph. Model.* **2002**, *20* (4), 269–276.

(51) Rahman, I.; Adcock, I. M. Oxidative Stress and Redox Regulation of Lung Inflammation in COPD. *Eur. Respir. J.* **2006**, *28* (1), 219–242.

(52) MacEyka, M.; Spiegel, S. Sphingolipid Metabolites in Inflammatory Disease. *Nature* **2014**, *510* (7503), 58–67.

(53) Chen, W.-W.; Chao, Y.-J.; Chang, W.-H.; Chan, J.-F.; Hsu, Y.-H. H. Phosphatidylglycerol Incorporates into Cardiolipin to Improve Mitochondrial Activity and Inhibits Inflammation. *Sci. Rep.* **2018**, *8* (1), 4919.

(54) Massey, K. A.; Nicolaou, A. Lipidomics of Oxidized Polyunsaturated Fatty Acids. *Free Radic. Biol. Med.* **2013**, *59*, 45–55.

(55) Lucas, J. H.; Wheeler, D. G.; Guan, Z.; Suntres, Z.; Stokes, B. T. Effect of Glutathione Augmentation on Lipid Peroxidation after Spinal Cord Injury. *J. Neurotrauma* **2002**, *19* (6), 763–775.

(56) Wu, D. Modulation of Immune and Inflammatory Responses by Dietary Lipids. *Curr. Opin. Lipidol.* **2004**, *15* (1), 43–47.

(57) Miller, Y. I.; Shyy, J. Y. J. Context-Dependent Role of Oxidized Lipids and Lipoproteins in Inflammation. *Trends Endocrinol. Metab.* **2017**, *28* (2), 143–152.

(58) Rivera, I.-G.; Gomez-Larrauri, A.; Trueba, M.; Presa, N.; Ordoñez, M.; Gomez-Muñoz, A.; Presa, N.; Gomez-Larrauri, A.; Rivera, I.-G.; Trueba, M.; Ordoñez, M.; Gomez-Larrauri, A.; Trueba, M.; Presa, N.; Ordoñez, M.; Gomez-Muñoz, A.; Presa, N.; Gomez-Larrauri, A.; Rivera, I.-G.; Trueba, M.; Ordoñez, M.; Gomez-Larrauri, A.; Trueba, M.; Presa, N.; Ordoñez, M.; Gomez-Muñoz, A. Control of Inflammatory Responses by Ceramide, Sphingosine 1-Phosphate and Ceramide 1-Phosphate. *Prog. Lipid Res.* **2016**, *61*, 51–62.

(59) Walz, W.; Bhattacharya, S. K.; Walz, W.; Bhattacharya, S. K. *Lipidomics*; Bhattacharya, S. K., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2017; Vol. 1609.

(60) Han, X.; Gross, R. W. Shotgun Lipidomics: Electrospray Ionization Mass Spectrometric Analysis and Quantitation of Cellular Lipidomes Directly from Crude Extracts of Biological Samples. *Mass Spectrom. Rev.* **2005**, *24* (3), 367–412.

(61) Holčapek, M.; Liebisch, G.; Ekroos, K. Lipidomic Analysis. *Anal. Chem.* **2018**, *90* (7), 4249–4257.

(62) Folch, J.; Ascoli, I.; Lees, M.; Meath, J. A.; LeBaron, N. Preparation of Lipide Extracts from Brain Tissue. *J. Biol. Chem.* **1951**, *191* (2), 833–841.

(63) Folch, J.; Lees, M.; Sloane Stanley, G. H. A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues. *J. Biol. Chem.* **1957**, *226* (1), 497–509.

(64) Bligh, E. G.; Dyer, W. J. A Rapid Method of Total Lipid Extraction and Purification. *Can. J. Biochem. Physiol.* **1959**, *37* (8), 911–917.

(65) Cequier-Sánchez, E.; Rodríguez, C.; Ravelo, A. G.; Zárate, R.; Avelo, N. G. R.; Guez, C. O. R. O. D. R. I.; Sa, F.; Cequier-Sánchez, E.; Rodríguez, C.; Ravelo, A. G.; Zárate, R.; Avelo, N. G. R.; Guez, C. O. R. O. D. R. I.; Sa, F.; Cequier-Sánchez, E.; Rodríguez, C.; Ravelo, A. G.; Zárate, R. Dichloromethane as a Solvent for Lipid Extraction and Assessment of Lipid Classes and Fatty Acids from Samples of Different Natures. *J. Agric. Food Chem.* **2008**, *56* (12), 4297–4303.

(66) Chaves, A. L.; Vergara, C. E.; Mayer, J. E. Dichloromethane as an Economic Alternative to Chloroform in the Extraction of DNA from Plant Tissues. *Plant Mol. Biol. Report.* **1995**,

13 (1), 18–25.

(67) Matyash, V.; Liebisch, G.; Kurzchalia, T. V.; Shevchenko, A.; Schwudke, D. Lipid Extraction by Methyl- Tert -Butyl Ether for High-Throughput Lipidomics. *J. Lipid Res.* **2008**, *49* (5), 1137–1146.

(68) Cajka, T.; Fiehn, O. Comprehensive Analysis of Lipids in Biological Systems by Liquid Chromatography-Mass Spectrometry - Supp Info. *TrAC Trends Anal. Chem.* **2014**, *61*, 192–206.

(69) Bruce, S. J.; Tavazzi, I.; Parisod, V.; Rezzi, S.; Kochhar, S.; Guy, P. A. Investigation of Human Blood Plasma Sample Preparation for Performing Metabolomics Using Ultrahigh Performance Liquid Chromatography/Mass Spectrometry. *Anal. Chem.* **2009**, *81* (9), 3285–3296.

(70) Pellegrino, R. M.; Di Veroli, A.; Valeri, A.; Goracci, L.; Cruciani, G. LC/MS Lipid Profiling from Human Serum: A New Method for Global Lipid Extraction. *Anal. Bioanal. Chem.* **2014**, *406* (30), 7937–7948.

(71) Herzog, R.; Schuhmann, K.; Schwudke, D.; Sampaio, J. L.; Bornstein, S. R.; Schroeder, M.; Shevchenko, A. LipidXplorer: A Software for Consensual Cross-Platform Lipidomics. *PLoS One* **2012**, *7* (1), e29851.

(72) Herzog, R.; Schwudke, D.; Shevchenko, A. LipidXplorer: Software for Quantitative Shotgun Lipidomics Compatible with Multiple Mass Spectrometry Platforms. *Curr. Protoc. Bioinforma.* **2013**, *43* (1), 14.12.1-14.12.30.

(73) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. MS-DIAL: Data-Independent MS/MS

Deconvolution for Comprehensive Metabolome Analysis. *Nat. Methods* **2015**, *12* (6), 523–526.

(74) Hartler, J.; Trötzmüller, M.; Chitraju, C.; Spener, F.; Köfeler, H. C.; Thallinger, G. G. Lipid Data Analyzer: Unattended Identification and Quantitation of Lipids in LC-MS Data. *Bioinformatics* **2011**, *27* (4), 572–577.

(75) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H. H.; Russell, D. W.; Subramaniam, S. LMSD: LIPID MAPS Structure Database. *Nucleic Acids Res.* **2007**, *35* (Database), D527–D532.

(76) Liebisch, G.; Ekroos, K.; Hermansson, M.; Ejsing, C. S. Reporting of Lipidomics Data Should Be Standardized. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2017**, *1862* (8), 747–751.

(77) Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. LIPID MAPS Online Tools for Lipid Research. *Nucleic Acids Res.* **2007**, *35* (Web Server), W606–W612.

(78) Kind, T.; Liu, K.-H.; Lee, D. Y.; DeFelice, B.; Meissen, J. K.; Fiehn, O. LipidBlast in Silico Tandem Mass Spectrometry Database for Lipid Identification. *Nat. Methods* **2013**, *10* (8), 755–758.

(79) Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J.; Jia, L.; Cruz, J. A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y.; Clive, D.; Greiner, R.; Nazyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I. HMDB: A Knowledgebase for the Human Metabolome. *Nucleic Acids Res.*

2009, *37* (Database), D603–D610.

(80) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res.* **2012**, *41* (D1), D801–D807.

(81) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M.-A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; MacInnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. HMDB: The Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35* (Database), D521–D526.

(82) Ma, Y.; Kind, T.; Vaniya, A.; Gennity, I.; Fahrmann, J. F.; Fiehn, O. An in Silico MS/MS Library for Automatic Annotation of Novel FAHFA Lipids. *J. Cheminform.* **2015**, *7* (1).

(83) Aoyagi, R.; Ikeda, K.; Isobe, Y.; Arita, M. Comprehensive Analyses of Oxidized Phospholipids Using a Measured MS/MS Spectra Library. *J. Lipid Res.* **2017**, *58* (11), 2229–2237.

(84) Wang, M.; Wang, C.; Han, X. Selection of Internal Standards for Accurate Quantification of Complex Lipid Species in Biological Extracts by Electrospray Ionization Mass Spectrometry-What, How and Why? *Mass Spectrom. Rev.* **2017**, *36* (6), 693–714.

(85) van den Berg, R. A.; Hoefsloot, H. C. J.; Westerhuis, J. A.; Smilde, A. K.; van der Werf,

M. J. Centering, Scaling, and Transformations: Improving the Biological Information Content of Metabolomics Data. *BMC Genomics* **2006**, *7* (1), 142.

(86) Nuzzo, R. How Scientists Fool Themselves – and How They Can Stop. *Nature* **2015**, *526* (7572), 182–185.

(87) Veselkov, K. A.; Vingara, L. K.; Masson, P.; Robinette, S. L.; Want, E.; Li, J. V.; Barton, R. H.; Boursier-Neyret, C.; Walther, B.; Ebbels, T. M.; Pelczer, I.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Optimized Preprocessing of Ultra-Performance Liquid Chromatography/Mass Spectrometry Urinary Metabolic Profiles for Improved Information Recovery. *Anal. Chem.* **2011**, *83* (15), 5864–5872.

(88) Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E. J. J.; van Duijnhoven, J. P. M.; van Dorsten, F. A. Assessment of PLSDA Cross Validation. *Metabolomics* **2008**, *4* (1), 81–89.

(89) Golbraikh, A.; Tropsha, A. Beware of Q2! *J. Mol. Graph. Model.* **2002**, *20* (4), 269–276.

(90) Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.

(91) Biau, G.; Scornet, E. A Random Forest Guided Tour. *TEST* **2016**, *25* (2), 197–227.

(92) Fisher, R. A. *Statistical Methods for Research Workers*, 5th ed.; Crew, F. A. E., Cutler, D. W., Eds.; Oliver and Boyd LTD: London, UK, 1934.

(93) Goodman, S. A Dirty Dozen: Twelve P-Value Misconceptions. *Semin. Hematol.* **2008**, *45* (3), 135–140.

(94) Nuzzo, R. Scientific Method: Statistical Errors. *Nature* **2014**, *506* (7487), 150–152.

(95) Noble, W. S. How Does Multiple Testing Correction Work? *Nat. Biotechnol.* **2009**, *27* (12),

1135–1137.

(96) Godfrey, K. Comparing the Means of Several Groups. *N. Engl. J. Med.* **1985**, *313* (23), 1450–1456.

(97) Vaux, D. L. Know When Your Numbers Are Significant. *Nature* **2012**, *492* (7428), 180–181.

(98) Halsey, L. G.; Curran-Everett, D.; Vowler, S. L.; Drummond, G. B. The Fickle P Value Generates Irreproducible Results. *Nat. Methods* **2015**, *12* (3), 179–185.

(99) Han, X. Lipidomics for Studying Metabolism. *Nature Reviews Endocrinology*. Nature Publishing Group November 2016, pp 668–679.

(100) Hyötyläinen, T.; Orešič, M. Optimizing the Lipidomics Workflow for Clinical Studies—Practical Considerations. *Anal. Bioanal. Chem.* **2015**, *407* (17), 4973–4993.

(101) Han, X. Lipidomics for Precision Medicine and Metabolism: A Personal View. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2017**, *1862* (8), 804–807.

(102) Bang, D. Y.; Kang, D.; Moon, M. H. Nanoflow Liquid Chromatography–Tandem Mass Spectrometry for the Characterization of Intact Phosphatidylcholines from Soybean, Bovine Brain, and Liver. *J. Chromatogr. A* **2006**, *1104* (1–2), 222–229.

(103) Bang, D. Y.; Ahn, E. jeong; Moon, M. H. Shotgun Analysis of Phospholipids from Mouse Liver and Brain by Nanoflow Liquid Chromatography/Tandem Mass Spectrometry. *J. Chromatogr. B* **2007**, *852* (1–2), 268–277.

(104) Kosicek, M.; Kirsch, S.; Bene, R.; Trkanjec, Z.; Titlic, M.; Bindila, L.; Peter-Katalinic, J.; Hecimovic, S. Nano-HPLC–MS Analysis of Phospholipids in Cerebrospinal Fluid of Alzheimer's Disease Patients—a Pilot Study. *Anal. Bioanal. Chem.* **2010**, *398* (7–8), 2929–

2937.

(105) Park, S. M.; Byeon, S. K.; Lee, H.; Sung, H.; Kim, I. Y.; Seong, J. K.; Moon, M. H. Lipidomic Analysis of Skeletal Muscle Tissues of P53 Knockout Mice by NUPLC-ESI-MS/MS. *Sci. Rep.* **2017**, *7* (1), 3302.

(106) Vasilopoulou, C. G.; Sulek, K.; Brunner, A. D.; Meitei, N. S.; Schweiger-Hufnagel, U.; Meyer, S. W.; Barsch, A.; Mann, M.; Meier, F. Trapped Ion Mobility Spectrometry and PASEF Enable In-Depth Lipidomics from Minimal Sample Amounts. *Nat. Commun.* **2020**, *11* (1), 331.

(107) Lee, G. Bin; Lee, J. C.; Moon, M. H. Plasma Lipid Profile Comparison of Five Different Cancers by Nanoflow Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectrometry. *Anal. Chim. Acta* **2019**, *1063*, 117–126.

(108) Danne-rasche, N.; Coman, C.; Ahrends, R. Nano-LC/NSI MS Refines Lipidomics by Enhancing Lipid Coverage, Measurement Sensitivity, and Linear Dynamic Range - Supp Info. *Anal. Chem.* **2018**, *90* (13), 8093–8101.

(109) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. HMDB 4.0: The Human Metabolome Database for 2018. *Nucleic Acids Res.* **2018**, *46* (D1), D608–D617.

(110) Yang, M.; Fazio, S.; Munch, D.; Drumm, P. Impact of Methanol and Acetonitrile on Separations Based on π–π Interactions with a Reversed-Phase Phenyl Column. *J.*

Chromatogr. A **2005**, *1097* (1–2), 124–129.

(111) Berthelette, K. D.; Swann, T.; Fairchild, J. Waters Application Note: Optimizing Selectivity through Intelligent Solvent Selection Using CORTECS Phenyl Columns. Waters Corporation: Milford, MA 2016.

(112) Pöhö, P.; Oresic, M.; Hyötyläinen, T. MS-Based Lipidomics. In *Comprehensive Analytical Chemistry*; García-Cañas, V., Cifuentes, A., Simó, C., Eds.; Elsevier: Kidlington, Oxford, 2014; Vol. 64, pp 375–393.

(113) Cajka, T.; Fiehn, O. Increasing Lipidomic Coverage by Selecting Optimal Mobile-Phase Modifiers in LC–MS of Blood Plasma. *Metabolomics* **2016**, *12* (2), 34.

(114) Sandra, K.; Pereira, A. dos S.; Vanhoenacker, G.; David, F.; Sandra, P. Comprehensive Blood Plasma Lipidomics by Liquid Chromatography/Quadrupole Time-of-Flight Mass Spectrometry. *J. Chromatogr. A* **2010**, *1217* (25), 4087–4099.

(115) Jung, H. R.; Sylvänne, T.; Koistinen, K. M.; Tarasov, K.; Kauhanen, D.; Ekroos, K. High Throughput Quantitative Molecular Lipidomics. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2011**, *1811* (11), 925–934.

(116) Gregory, K. E.; Bird, S. S.; Gross, V. S.; Marur, V. R.; Lazarev, A. V.; Walker, W. A.; Kristal, B. S. Method Development for Fecal Lipidomics Profiling. *Anal. Chem.* **2013**, *85* (2), 1114–1123.

(117) Yamada, T.; Uchikata, T.; Sakamoto, S.; Yokoi, Y.; Fukusaki, E.; Bamba, T. Development of a Lipid Profiling System Using Reverse-Phase Liquid Chromatography Coupled to High-Resolution Mass Spectrometry with Rapid Polarity Switching and an Automated Lipid Identification Software. *J. Chromatogr. A* **2013**, *1292*, 211–218.

(118) Damen, C. W. N.; Isaac, G.; Langridge, J.; Hankemeier, T.; Vreeken, R. J. Enhanced Lipid Isomer Separation in Human Plasma Using Reversed-Phase UPLC with Ion-Mobility/High-Resolution MS Detection. *J. Lipid Res.* **2014**, *55* (8), 1772–1783.

(119) Bird, S. S.; Marur, V. R.; Sniatynski, M. J.; Greenberg, H. K.; Kristal, B. S. Lipidomics Profiling by High-Resolution LC−MS and High-Energy Collisional Dissociation Fragmentation: Focus on Characterization of Mitochondrial Cardiolipins and Monolysocardiolipins. *Anal. Chem.* **2011**, *83* (3), 940–949.

(120) Zhang, J.; Wang, Q.; Kleintop, B.; Raglione, T. Suppression of Peak Tailing of Phosphate Prodrugs in Reversed-Phase Liquid Chromatography. *J. Pharm. Biomed. Anal.* **2014**, *98*, 247–252.

(121) Kaspar, S.; Ledertheil, T.; Hartmer, R.; Nugent, K.; Hagedorn, T.; Baessmann, C.; Nugent, K. *Bruker Technical Note #TN-44: Increasing Peptide Identification Rates for Proteomics Samples by Controlling Peptide Charge States Using CaptiveSpray NanoBooster*; Bruker Daltonics Technical Note TN-44: Bremen ,Germany, Germany, 2013.

(122) Chen, I. S.; Shen, C. S. J.; Sheppard, A. J. Comparison of Methylene Chloride and Chloroform for the Extraction of Fats from Food Products. *J. Am. Oil Chem. Soc.* **1981**, *58* (5), 599–601.

(123) Patterson, R. E.; Ducrocq, A. J.; McDougall, D. J.; Garrett, T. J.; Yost, R. A. Comparison of Blood Plasma Sample Preparation Methods for Combined LC–MS Lipidomics and Metabolomics. *J. Chromatogr. B* **2015**, *1002*, 260–266.

(124) Noga, M.; Sucharski, F.; Suder, P.; Silberring, J. A Practical Guide to Nano-LC Troubleshooting. *J. Sep. Sci.* **2007**, *30* (14), 2179–2189.

(125) Jones, D. R.; Wu, Z.; Chauhan, D.; Anderson, K. C.; Peng, J. A Nano Ultra-Performance Liquid Chromatography–High Resolution Mass Spectrometry Approach for Global Metabolomic Profiling and Case Study on Drug-Resistant Multiple Myeloma. *Anal. Chem.* **2014**, *86* (7), 3667–3675.

(126) Sakamaki, H.; Uchida, T.; Wah, L.; Takeuchi, T. Evaluation of Column Hardware on Liquid Chromatography – Mass Spectrometry of Phosphorylated Compounds. *J. Chromatogr. A* **2015**, *1381*, 125–131.

(127) Ogiso, H.; Suzuki, T.; Taguchi, R. Development of a Reverse-Phase Liquid Chromatography Electrospray Ionization Mass Spectrometry Method for Lipidomics, Improving Detection of Phosphatidic Acid and Phosphatidylserine. *Anal. Biochem.* **2008**, *375* (1), 124–131.

(128) Knittelfelder, O. L.; Weberhofer, B. P.; Eichmann, T. O.; Kohlwein, S. D.; Rechberger, G. N. A Versatile Ultra-High Performance LC-MS Method for Lipid Profiling. *J. Chromatogr. B* **2014**, *951–952* (1), 119–128.

(129) Asakawa, Y.; Tokida, N.; Ozawa, C.; Ishiba, M.; Tagaya, O.; Asakawa, N. Suppression Effects of Carbonate on the Interaction between Stainless Steel and Phosphate Groups of Phosphate Compounds in High-Performance Liquid Chromatography and Electrospray Ionization Mass Spectrometry. *J. Chromatogr. A* **2008**, *1198–1199* (1–2), 80–86.

(130) *Spinal Cord Medicine*, 3rd ed.; Kishblum, S., Lin, V. W., Eds.; Springer Publishing Company: New York, NY, USA, 2019.

(131) *International Perspectives on Spinal Cord Injury, World Health Organization (WHO) and The International Spinal Cord Society (ISCOS)*; Geneva, Switzerland, 2013.

(132) Ahuja, C. S.; Wilson, J. R.; Nori, S.; Kotter, M. R. N.; Druschel, C.; Curt, A.; Fehlings, M. G. Traumatic Spinal Cord Injury. *Nat. Rev. Dis. Prim.* **2017**, *3*, 17018.

(133) World Health Organization Fact Sheet - Spinal Cord Injury.

(134) Okon, E. B.; Streijger, F.; Lee, J. H. T.; Anderson, L. M.; Russell, A. K.; Kwon, B. K. Intraparenchymal Microdialysis after Acute Spinal Cord Injury Reveals Differential Metabolic Responses to Contusive versus Compressive Mechanisms of Injury. *J. Neurotrauma* **2013**, *30* (18), 1564–1576.

(135) Ahuja, C. S.; Nori, S.; Tetreault, L.; Wilson, J.; Kwon, B. K.; Harrop, J.; Choi, D.; Fehlings, M. G. Traumatic Spinal Cord Injury—Repair and Regeneration. *Neurosurgery* **2017**, *80* (3S), S9–S22.

(136) Lubieniecka, J. M.; Streijger, F.; Lee, J. H. T.; Stoynov, N.; Liu, J.; Mottus, R.; Pfeifer, T.; Kwon, B. K.; Coorssen, J. R.; Foster, L. J.; Grigliatti, T. A.; Tetzlaff, W. Biomarkers for Severity of Spinal Cord Injury in the Cerebrospinal Fluid of Rats. *PLoS One* **2011**, *6* (4), e19247.

(137) Evaniew, N.; Noonan, V. K.; Fallah, N.; Kwon, B. K.; Rivers, C. S.; Ahn, H.; Bailey, C. S.; Christie, S. D.; Fourney, D. R.; Hurlbert, R. J.; Linassi, A. G.; Fehlings, M. G.; Dvorak, M. F. Methylprednisolone for the Treatment of Patients with Acute Spinal Cord Injuries: A Propensity Score-Matched Cohort Study from a Canadian Multi-Center Spinal Cord Injury Registry. *J. Neurotrauma* **2015**, *32* (21), 1674–1683.

(138) Badhiwala, J. H.; Wilson, J. R.; Kwon, B. K.; Casha, S.; Fehlings, M. G. A Review of Clinical Trials in Spinal Cord Injury Including Biomarkers. *J. Neurotrauma* **2018**, *35* (16), 1906–1917.

(139) Sekhon, L. H.; Fehlings, M. G. Epidemiology, Demographics, and Pathophysiology of Acute Spinal Cord Injury. *Spine (Phila. Pa. 1976).* **2001**, *26* (24 Suppl), S2-12.

(140) Fehlings, M. G.; Tetreault, L. A.; Wilson, J. R.; Kwon, B. K.; Burns, A. S.; Martin, A. R.; Hawryluk, G.; Harrop, J. S. A Clinical Practice Guideline for the Management of Acute Spinal Cord Injury: Introduction, Rationale, and Scope. *Glob. Spine J.* **2017**, *7* (3_suppl), 84S-94S.

(141) Kwon, B. K.; Hillyer, J.; Tetzlaff, W. Translational Research in Spinal Cord Injury: A Survey of Opinion from the SCI Community. *J. Neurotrauma* **2010**, *27* (1), 21–33.

(142) Kwon, B. K.; Casha, S.; Hurlbert, R. J.; Yong, V. W. Inflammatory and Structural Biomarkers in Acute Traumatic Spinal Cord Injury. *Clin. Chem. Lab. Med.* **2011**, *49* (3), 425–433.

(143) Yokobori, S.; Zhang, Z.; Moghieb, A.; Mondello, S.; Gajavelli, S.; Dietrich, W. D.; Bramlett, H.; Hayes, R. L.; Wang, M.; Wang, K. K. W.; Bullock, M. R. Acute Diagnostic Biomarkers for Spinal Cord Injury: Review of the Literature and Preliminary Research Report. *World Neurosurg.* **2015**, *83* (5), 867–878.

(144) Streijger, F.; Skinnider, M. A.; Rogalski, J. C.; Balshaw, R.; Shannon, C. P.; Prudova, A.; Belanger, L. M.; Ritchie, L.; Tsang, A.; Christie, S.; Parent, S.; Mac-Thiong, J.-M.; Bailey, C. S.; Urquhart, J.; Ailon, T.; Paquette, S.; Boyd, M.; Street, J.; Fisher, C. G.; Dvorak, M. F.; Borchers, C. H.; Foster, L. J.; Kwon, B. K. A Targeted Proteomics Analysis of Cerebrospinal Fluid after Acute Human Spinal Cord Injury. *J. Neurotrauma* **2017**, *34* (12), 2054–2068.

(145) Lee, J. H. T.; Jones, C. F.; Okon, E. B.; Anderson, L.; Tigchelaar, S.; Kooner, P.; Godbey,

T.; Chua, B.; Gray, G.; Hildebrandt, R.; Cripton, P.; Tetzlaff, W.; Kwon, B. K. A Novel Porcine Model of Traumatic Thoracic Spinal Cord Injury. *J. Neurotrauma* **2013**, *30* (3), 142–159.

(146) Shank, C. D.; Walters, B. C.; Hadley, M. N. Current Topics in the Management of Acute Traumatic Spinal Cord Injury. *Neurocrit. Care* **2018**.

(147) Kwon, B. K.; Stammers, A. M. T.; Belanger, L. M.; Bernardo, A.; Chan, D.; Bishop, C. M.; Slobogean, G. P.; Zhang, H.; Umedaly, H.; Giffin, M.; Street, J.; Boyd, M. C.; Paquette, S. J.; Fisher, C. G.; Dvorak, M. F. Cerebrospinal Fluid Inflammatory Cytokines and Biomarkers of Injury Severity in Acute Human Spinal Cord Injury. *J. Neurotrauma* **2010**, *27* (4), 669–682.

(148) Tigchelaar, S.; Streijger, F.; Sinha, S.; Flibotte, S.; Manouchehri, N.; So, K.; Shortt, K.; Okon, E.; Rizzuto, M. A.; Malenica, I.; Courtright-Lim, A.; Eisen, A.; Keuren-Jensen, K. Van; Nislow, C.; Kwon, B. K. Serum MicroRNAs Reflect Injury Severity in a Large Animal Model of Thoracic Spinal Cord Injury. *Sci. Rep.* **2017**, *7* (1), 1376.

(149) Kwon, B. K.; Streijger, F.; Fallah, N.; Noonan, V. K.; Bélanger, L. M.; Ritchie, L.; Paquette, S. J.; Ailon, T.; Boyd, M. C.; Street, J.; Fisher, C. G.; Dvorak, M. F. Cerebrospinal Fluid Biomarkers To Stratify Injury Severity and Predict Outcome in Human Traumatic Spinal Cord Injury. *J. Neurotrauma* **2017**, *34* (3), 567–580.

(150) Jones, C. F.; Newell, R. S.; Lee, J. H. T.; Cripton, P. A.; Kwon, B. K. The Pressure Distribution of Cerebrospinal Fluid Responds to Residual Compression and Decompression in an Animal Model of Acute Spinal Cord Injury. *Spine (Phila. Pa. 1976).* **2012**, *37* (23), E1422–E1431.

(151) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Res.* **2013**, *41* (D1), 801–807.

(152) Drotleff, B.; Lämmerhofer, M. Guidelines for Selection of Internal Standard-Based Normalization Strategies in Untargeted Lipidomic Profiling by LC-HR-MS/MS. *Anal. Chem.* **2019**, *91* (15), 9836–9843.

(153) Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D. S.; Xia, J. MetaboAnalyst 4.0: Towards More Transparent and Integrative Metabolomics Analysis. *Nucleic Acids Res.* **2018**, *46* (W1), W486–W494.

(154) Pleasure, D.; Kim, S. U. Sterol Synthesis by Myelinating Cultures of Mouse Spinal Cord. *Brain Res.* **1976**, *103* (1), 117–126.

(155) Siegenthaler, M. M.; Tu, M. K.; Keirstead, H. S. The Extent of Myelin Pathology Differs Following Contusion and Transection Spinal Cord Injury. *J. Neurotrauma* **2007**, *24* (10), 1631–1646.

(156) Piper, R. C.; Katzmann, D. J. Biogenesis and Function of MVBs. *Annu. Rev. Cell Dev. Biol.* **2010**, *23*, 519–547.

(157) Barile, L.; Vassalli, G. Exosomes: Therapy Delivery Tools and Biomarkers of Diseases. *Pharmacol. Ther.* **2017**, *174*, 63–78.

(158) Jeppesen, D. K.; Fenix, A. M.; Franklin, J. L.; Higginbotham, J. N.; Zhang, Q.; Zimmerman, L. J.; Liebler, D. C.; Ping, J.; Liu, Q.; Evans, R.; Fissell, W. H.; Patton, J. G.; Rome, L. H.;

Burnette, D. T.; Coffey, R. J. Reassessment of Exosome Composition. *Cell* **2019**, *177* (2), 428-445.e18.

(159) Zhang, Y.; Liu, Y.; Liu, H.; Tang, W. H. Exosomes: Biogenesis, Biologic Function and Clinical Potential. *Cell Biosci.* **2019**, *9* (1), 19.

(160) Edgar, J. R. Q & A: What Are Exosomes, Exactly? *BMC Biol.* **2016**, *14* (1), 1–7.

(161) Xu, J.; Chen, G.; Yang, Y. Exosomes: A Rising Star in Failing Hearts. *Front. Physiol.* **2017**, *8* (July).

(162) Skotland, T.; Sandvig, K.; Llorente, A. Lipids in Exosomes: Current Knowledge and the Way Forward. *Prog. Lipid Res.* **2017**, *66*, 30–41.

(163) Pegtel, D. M.; Gould, S. J. Exosomes. *Annu. Rev. Biochem.* **2019**, *88* (1), 487–514.

(164) Lydic, T. A.; Townsend, S.; Adda, C. G.; Collins, C.; Mathivanan, S.; Reid, G. E. Rapid and Comprehensive 'Shotgun' Lipidome Profiling of Colorectal Cancer Cell Derived Exosomes. *Methods* **2015**, *87* (1), 83–95.

(165) Luo, X.; An, M.; Cuneo, K. C.; Lubman, D. M.; Li, L. High-Performance Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry for Exosome Metabolomics. *Anal. Chem.* **2018**, *90* (14), 8314–8319.

(166) An, M.; Wu, J.; Zhu, J.; Lubman, D. M. Comparison of an Optimized Ultracentrifugation Method versus Size-Exclusion Chromatography for Isolation of Exosomes from Human Serum. *J. Proteome Res.* **2018**, *17* (10), 3599–3605.

(167) Record, M.; Carayon, K.; Poirot, M.; Silvente-Poirot, S. Exosomes as New Vesicular Lipid Transporters Involved in Cell-Cell Communication and Various Pathophysiologies. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2014**, *1841* (1), 108–120.

(168) Luquain-Costaz, C.; Lefai, E.; Arnal-Levron, M.; Markina, D.; Sakaï, S.; Euthine, V.; Makino, A.; Guichardant, M.; Yamashita, S.; Kobayashi, T.; Lagarde, M.; Moulin, P.; Delton-Vandenbroucke, I. Bis(Monoacylglycero)Phosphate Accumulation in Macrophages Induces Intracellular Cholesterol Redistribution, Attenuates Liver-X Receptor/ATP-Binding Cassette Transporter A1/ATP-Binding Cassette Ttransporter G1 Pathway, and Impairs Cholesterol Efflux. *Arterioscler. Thromb. Vasc. Biol.* **2013**, *33* (8), 1803–1811.

(169) Higgins, D.; Trujillo, C.; Keech, C. Advances in RSV Vaccine Research and Development – A Global Agenda. *Vaccine* **2016**, *34* (26), 2870–2875.

(170) Stein, R. T.; Bont, L. J.; Zar, H.; Polack, F. P.; Park, C.; Claxton, A.; Borok, G.; Butylkova, Y.; Wegzyn, C. Respiratory Syncytial Virus Hospitalization and Mortality: Systematic Review and Meta-Analysis. *Pediatr. Pulmonol.* **2017**, *52* (4), 556–569.

(171) Mazur, N. I.; Martinón-torres, F.; Baraldi, E.; Fauroux, B.; Greenough, A.; Heikkinen, T.; Manzoni, P.; Mejias, A.; Nair, H.; Papadopoulos, N. G.; Polack, F. P.; Ramilo, O.; Sharland, M.; Stein, R.; Madhi, S. A.; Bont, L. Lower Respiratory Tract Infection Caused by Respiratory Syncytial Virus : Current Management and New Therapeutics. *Lancet Respir. Med.* **2015**, *3* (November 2015), 888–900.

(172) Acosta, P. L.; Caballero, M. T.; Polack, F. P. Brief History and Characterization of Enhanced Respiratory Syncytial Virus Disease. *Clin. Vaccine Immunol.* **2016**, *23* (3), 189–195.

(173) Garg, R.; Latimer, L.; Gerdts, V.; Potter, A.; van Drunen Littel-van den Hurk, S. Intranasal Immunization with a Single Dose of the Fusion Protein Formulated with a Combination Adjuvant Induces Long-Term Protective Immunity against Respiratory Syncytial Virus.

Hum. Vaccin. Immunother. **2017**, *13* (12), 2894–2901.

(174) Rezaee, F.; Linfield, D. T.; Harford, T. J.; Piedimonte, G. Ongoing Developments in RSV Prophylaxis: A Clinician's Analysis. *Curr. Opin. Virol.* **2017**, *24*, 70–78.

(175) Garg, R.; Latimer, L.; Wang, Y.; Simko, E.; Gerdts, V.; Potter, A.; van Drunen Littel-van den Hurk, S.; Hurk, S. V. D. L. Den. Maternal Immunization with Respiratory Syncytial Virus Fusion Protein Formulated with a Novel Combination Adjuvant Provides Protection from RSV in Newborn Lambs. *Vaccine* **2016**, *34* (2), 261–269.

(176) Garg, R.; Babiuk, L.; van Drunen Littel-van den Hurk, S.; Gerdts, V.; Drunen, S. Van; Hurk, L. Den; Gerdts, V.; van Drunen Littel-van den Hurk, S.; Gerdts, V. A Novel Combination Adjuvant Platform for Human and Animal Vaccines. *Vaccine* **2017**, *35* (35), 4486–4489.

(177) Sarkar, I.; Garg, R.; van Drunen Littel-van den Hurk, S. Formulation of the Respiratory Syncytial Virus Fusion Protein with a Polymer-Based Combination Adjuvant Promotes Transient and Local Innate Immune Responses and Leads to Improved Adaptive Immunity. *Vaccine* **2016**, *34* (42), 5114–5124.

(178) Sarkar, I.; Garg, R.; van Drunen Littel-van den Hurk, S. The Respiratory Syncytial Virus Fusion Protein Formulated with a Polymer-Based Adjuvant Induces Multiple Signaling Pathways in Macrophages. *Vaccine* **2018**, *36* (17), 2326–2336.

(179) de Pablo, M. A.; Puertollano, M. A.; Alvarez de Cienfuegos, G. Biological and Clinical Significance of Lipids as Modulators of Immune System Functions. *Clin. Vaccine Immunol.* **2002**, *9* (5), 945–950.

(180) Wang, M.; Wang, C.; Han, X. Selection of Internal Standards for Accurate Quantification of Complex Lipid Species in Biological Extracts by Electrospray Ionization Mass

Spectrometry-What, How and Why? *Mass Spectrom. Rev.* **2016**, *47* (12), n/a-n/a.

(181) Wu, Y.; Li, L. Determination of Total Concentration of Chemically Labeled Metabolites as a Means of Metabolome Sample Normalization and Sample Loading Optimization in Mass Spectrometry-Based Metabolomics. *Anal. Chem.* **2012**, *84* (24), 10723–10731.

(182) Guo, K.; Li, L. High-Performance Isotope Labeling for Profiling Carboxylic Acid-Containing Metabolites in Biofluids by Mass Spectrometry. *Anal. Chem.* **2010**, *82* (21), 8789–8793.

(183) Peng, J.; Li, L. Liquid–Liquid Extraction Combined with Differential Isotope Dimethylaminophenacyl Labeling for Improved Metabolomic Profiling of Organic Acids. *Anal. Chim. Acta* **2013**, *803*, 97–105.

(184) Guo, K.; Li, L. Differential 12 C-/ 13 C-Isotope Dansylation Labeling and Fast Liquid Chromatography/Mass Spectrometry for Absolute and Relative Quantification of the Metabolome. *Anal. Chem.* **2009**, *81* (10), 3919–3932.

(185) Wu, Y.; Li, L. Dansylation Metabolite Assay: A Simple and Rapid Method for Sample Amount Normalization in Metabolomics. *Anal. Chem.* **2014**, *86* (19), 9428–9433.

(186) Zhou, R.; Tseng, C.; Huan, T.; Li, L. IsoMS: Automated Processing of LC-MS Data Generated by a Chemical Isotope Labeling Metabolomics Platform. *Anal. Chem.* **2014**, *86* (10), 4675–4679.

(187) Achaintre, D.; Buleté, A.; Cren-Olivé, C.; Li, L.; Rinaldi, S.; Scalbert, A. Differential Isotope Labeling of 38 Dietary Polyphenols and Their Quantification in Urine by Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry. *Anal. Chem.* **2016**, *88* (5), 2637–2644.

(188) Wu, Y.; Li, L. Sample Normalization Methods in Quantitative Metabolomics. *J. Chromatogr. A* **2016**, *1430*, 80–95.

(189) Huan, T.; Li, L. Quantitative Metabolome Analysis Based on Chromatographic Peak Reconstruction in Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry. *Anal. Chem.* **2015**, *87* (14), 7011–7016.

(190) Li, Y.; Li, L. Retention Time Shift Analysis and Correction in Chemical Isotope Labeling Liquid Chromatography/Mass Spectrometry for Metabolome Analysis. *Rapid Commun. Mass Spectrom.* **2020**, *34* (S1).

(191) Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; Lin, G. MyCompoundID: Using an Evidence-Based Metabolome Library for Metabolite Identification. *Anal. Chem.* **2013**, *85* (6), 3401–3408.

(192) Huan, T.; Tang, C.; Li, R.; Shi, Y.; Lin, G.; Li, L. MyCompoundID MS/MS Search: Metabolite Identification Using a Library of Predicted Fragment-Ion-Spectra of 383,830 Possible Human Metabolites. *Anal. Chem.* **2015**, *87* (20), 10619–10626.

(193) Frolkis, A.; Knox, C.; Lim, E.; Jewison, T.; Law, V.; Hau, D. D.; Liu, P.; Gautam, B.; Ly, S.; Guo, A. C.; Xia, J.; Liang, Y.; Shrivastava, S.; Wishart, D. S. SMPDB: The Small Molecule Pathway Database. *Nucleic Acids Res.* **2010**, *38*, D480–D487.

(194) Jewison, T.; Su, Y.; Disfany, F. M.; Liang, Y.; Knox, C.; Maciejewski, A.; Poelzer, J.; Huynh, J.; Zhou, Y.; Arndt, D.; Djoumbou, Y.; Liu, Y.; Deng, L.; Guo, A. C.; Han, B.; Pon, A.; Wilson, M.; Rafatnia, S.; Liu, P.; Wishart, D. S. SMPDB 2.0: Big Improvements to the Small Molecule Pathway Database. *Nucleic Acids Res.* **2014**, *42* (D1), D478–D484.

(195) Xia, J.; Sinelnikov, I. V.; Han, B.; Wishart, D. S. MetaboAnalyst 3.0—Making

Metabolomics More Meaningful. *Nucleic Acids Res.* **2015**, *43* (W1), W251–W257.

(196) Wright, J. R. Pulmonary Surfactant: A Front Line of Lung Host Defense. *J. Clin. Invest.* **2003**, *111* (10), 1453–1455.

(197) Han, S.; Mallampalli, R. K. The Role of Surfactant in Lung Disease and Host Defense against Pulmonary Infections. *Ann. Am. Thorac. Soc.* **2015**, *12* (5), 765–774.

(198) Agassandian, M.; Mallampalli, R. K. Surfactant Phospholipid Metabolism. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2013**, *1831* (3), 612–625.

(199) Barreira, E. R.; Precioso, A. R.; Bousso, A. Pulmonary Surfactant in Respiratory Syncytial Virus Bronchiolitis: The Role in Pathogenesis and Clinical Implications. *Pediatr. Pulmonol.* **2010**, *46* (5), 415–420.

(200) Xia, J.; Wishart, D. S. MetPA: A Web-Based Metabolomics Tool for Pathway Analysis and Visualization. *Bioinformatics* **2010**, *26* (18), 2342–2344.

(201) Du, L.; Xie, T.; Xu, J.; Kang, A.; Di, L.; Shan, J.; Wang, S. A Metabolomics Approach to Studying the Effects of Jinxin Oral Liquid on RSV-Infected Mice Using UPLC/LTQ-Orbitrap Mass Spectrometry. *J. Ethnopharmacol.* **2015**, *174*, 25–36.

(202) Calder, P. C. Metabolism of Polyunsaturated Fatty Acids by Cells of the Immune System. In *Polyunsaturated Fatty Acid Metabolism*; Elsevier, 2018; pp 135–155.

(203) Gray, L. R.; Tompkins, S. C.; Taylor, E. B. Regulation of Pyruvate Metabolism and Human Disease. *Cell. Mol. Life Sci.* **2014**, *71* (14), 2577–2604.

(204) Sauer, A. V.; Brigida, I.; Carriglio, N.; Aiuti, A. Autoimmune Dysregulation and Purine Metabolism in Adenosine Deaminase Deficiency. *Front. Immunol.* **2012**, *3* (AUG), 1–19.

(205) O'Sullivan, D.; van der Windt, G. J. W.; Huang, S. C.-C.; Curtis, J. D.; Chang, C.-H.; Buck, M. D.; Qiu, J.; Smith, A. M.; Lam, W. Y.; DiPlato, L. M.; Hsu, F.-F.; Birnbaum, M. J.; Pearce, E. J.; Pearce, E. L. Memory CD8+ T Cells Use Cell-Intrinsic Lipolysis to Support the Metabolic Programming Necessary for Development. *Immunity* **2014**, *41* (1), 75–88.

(206) Pan, Y.; Tian, T.; Park, C. O.; Lofftus, S. Y.; Mei, S.; Liu, X.; Luo, C.; O'Malley, J. T.; Gehad, A.; Teague, J. E.; Divito, S. J.; Fuhlbrigge, R.; Puigserver, P.; Krueger, J. G.; Hotamisligil, G. S.; Clark, R. A.; Kupper, T. S. Survival of Tissue-Resident Memory T Cells Requires Exogenous Lipid Uptake and Metabolism. *Nature* **2017**, *543* (7644), 252–256.

(207) Bellucci, A.; Mercuri, N. B.; Venneri, A.; Faustini, G.; Longhena, F.; Pizzi, M.; Missale, C.; Spano, P. Review: Parkinson's Disease: From Synaptic Loss to Connectome Dysfunction. *Neuropathol. Appl. Neurobiol.* **2016**, *42* (1), 77–94.

(208) Rocca, W. A. The Future Burden of Parkinson's Disease. *Mov. Disord.* **2018**, *33* (1), 8–9.

(209) Tysnes, O.-B. B.; Storstein, A. Epidemiology of Parkinson's Disease. *J. Neural Transm.* **2017**, *124* (8), 901–905.

(210) Macleod, A. D.; Taylor, K. S. M. M.; Counsell, C. E.; Pringsheim, T.; Jette, N.; Frolkis, A.; Steeves, T. D. L. L.; Macleod, A. D.; Taylor, K. S. M. M.; Counsell, C. E.; Pringsheim, T.; Jette, N.; Frolkis, A.; Steeves, T. D. L. L.; Macleod, A. D.; Taylor, K. S. M. M.; Counsell, C. E.; Pringsheim, T.; Jette, N.; Frolkis, A.; Steeves, T. D. L. L.; Macleod, A. D.; Taylor, K. S. M. M.; Counsell, C. E.; Pringsheim, T.; Jette, N.; Frolkis, A.; Steeves, T. D. L. L. Mortality in Parkinson's Disease: A Systematic Review and Meta-Analysis. *Mov. Disord.* **2014**, *29* (13), 1615–1622.

(211) Naghavi, M.; Abajobir, A. A.; Abbafati, C.; Abbas, K. M.; Abd-Allah, F.; Abera, S. F.;

Aboyans, V.; Adetokunboh, O.; Afshin, A.; Agrawal, A.; Ahmadi, A.; Ahmed, M. B.; Aichour, A. N.; Aichour, M. T. E.; Aichour, I.; Aiyar, S.; Alahdab, F.; Al-Aly, Z.; Alam, K.; Alam, N.; Alam, T.; Alene, K. A.; Al-Eyadhy, A.; Ali, S. D.; Alizadeh-Navaei, R.; Alkaabi, J. M.; Alkerwi, A.; Alla, F.; Allebeck, P.; Allen, C.; Al-Raddadi, R.; Alsharif, U.; Altirkawi, K. A.; Alvis-Guzman, N.; Amare, A. T.; Amini, E.; Ammar, W.; Amoako, Y. A.; Anber, N.; Andersen, H. H.; Andrei, C. L.; Androudi, S.; Ansari, H.; Antonio, C. A. T.; Anwari, P.; Ärnlöv, J.; Arora, M.; Artaman, A.; Aryal, K. K.; Asayesh, H.; Asgedom, S. W.; Atey, T. M.; Avila-Burgos, L.; Avokpaho, E. F. G. A.; Awasthi, A.; Babalola, T. K.; Bacha, U.; Balakrishnan, K.; Barac, A.; Barboza, M. A.; Barker-Collo, S. L.; Barquera, S.; Barregard, L.; Barrero, L. H.; Baune, B. T.; Bedi, N.; Beghi, E.; Béjot, Y.; Bekele, B. B.; Bell, M. L.; Bennett, J. R.; Bensenor, I. M.; Berhane, A.; Bernabé, E.; Betsu, B. D.; Beuran, M.; Bhatt, S.; Biadgilign, S.; Bienhoff, K.; Bikbov, B.; Bisanzio, D.; Bourne, R. R. A.; Breitborde, N. J. K.; Bulto, L. N. B.; Bumgarner, B. R.; Butt, Z. A.; Cahuana-Hurtado, L.; Cameron, E.; Campuzano, J. C.; Car, J.; Cárdenas, R.; Carrero, J. J.; Carter, A.; Casey, D. C.; Castañeda-Orjuela, C. A.; Catalá-López, F.; Charlson, F. J.; Chibueze, C. E.; Chimed-Ochir, O.; Chisumpa, V. H.; Chitheer, A. A.; Christopher, D. J.; Ciobanu, L. G.; Cirillo, M.; Cohen, A. J.; Colombara, D.; Cooper, C.; Cowie, B. C.; Criqui, M. H.; Dandona, L.; Dandona, R.; Dargan, P. I.; das Neves, J.; Davitoiu, D. V.; Davletov, K.; de Courten, B.; Defo, B. K.; Degenhardt, L.; Deiparine, S.; Deribe, K.; Deribew, A.; Dey, S.; Dicker, D.; Ding, E. L.; Djalalinia, S.; Do, H. P.; Doku, D. T.; Douwes-Schultz, D.; Driscoll, T. R.; Dubey, M.; Duncan, B. B.; Echko, M.; El-Khatib, Z. Z.; Ellingsen, C. L.; Enayati, A.; Ermakov, S. P.; Erskine, H. E.; Eskandarieh, S.; Esteghamati, A.; Estep, K.; Farinha, C. S. e S.; Faro, A.; Farzadfar, F.; Feigin, V. L.; Fereshtehnejad, S.-M.; Fernandes, J. C.; Ferrari,

A. J.; Feyissa, T. R.; Filip, I.; Finegold, S.; Fischer, F.; Fitzmaurice, C.; Flaxman, A. D.; Foigt, N.; Frank, T.; Fraser, M.; Fullman, N.; Fürst, T.; Furtado, J. M.; Gakidou, E.; Garcia-Basteiro, A. L.; Gebre, T.; Gebregergs, G. B.; Gebrehiwot, T. T.; Gebremichael, D. Y.; Geleijnse, J. M.; Genova-Maleras, R.; Gesesew, H. A.; Gething, P. W.; Gillum, R. F.; Giref, A. Z.; Giroud, M.; Giussani, G.; Godwin, W. W.; Gold, A. L.; Goldberg, E. M.; Gona, P. N.; Gopalani, S. V.; Gouda, H. N.; Goulart, A. C.; Griswold, M.; Gupta, R.; Gupta, T.; Gupta, V.; Gupta, P. C.; Haagsma, J. A.; Hafezi-Nejad, N.; Hailu, A. D.; Hailu, G. B.; Hamadeh, R. R.; Hambisa, M. T.; Hamidi, S.; Hammami, M.; Hancock, J.; Handal, A. J.; Hankey, G. J.; Hao, Y.; Harb, H. L.; Hareri, H. A.; Hassanvand, M. S.; Havmoeller, R.; Hay, S. I.; He, F.; Hedayati, M. T.; Henry, N. J.; Heredia-Pi, I. B.; Herteliu, C.; Hoek, H. W.; Horino, M.; Horita, N.; Hosgood, H. D.; Hostiuc, S.; Hotez, P. J.; Hoy, D. G.; Huynh, C.; Iburg, K. M.; Ikeda, C.; Ileanu, B. V.; Irenso, A. A.; Irvine, C. M. S.; Islam, S. M. S.; Jacobsen, K. H.; Jahanmehr, N.; Jakovljevic, M. B.; Javanbakht, M.; Jayaraman, S. P.; Jeemon, P.; Jha, V.; John, D.; Johnson, C. O.; Johnson, S. C.; Jonas, J. B.; Jürisson, M.; Kabir, Z.; Kadel, R.; Kahsay, A.; Kamal, R.; Karch, A.; Karimi, S. M.; Karimkhani, C.; Kasaeian, A.; Kassaw, N. A.; Kassebaum, N. J.; Katikireddi, S. V.; Kawakami, N.; Keiyoro, P. N.; Kemmer, L.; Kesavachandran, C. N.; Khader, Y. S.; Khan, E. A.; Khang, Y.-H.; Khoja, A. T. A.; Khosravi, M. H.; Khosravi, A.; Khubchandani, J.; Kiadaliri, A. A.; Kieling, C.; Kievlan, D.; Kim, Y. J.; Kim, D.; Kimokoti, R. W.; Kinfu, Y.; Kissoon, N.; Kivimaki, M.; Knudsen, A. K.; Kopec, J. A.; Kosen, S.; Koul, P. A.; Koyanagi, A.; Kulikoff, X. R.; Kumar, G. A.; Kumar, P.; Kutz, M.; Kyu, H. H.; Lal, D. K.; Lalloo, R.; Lambert, T. L. N.; Lan, Q.; Lansingh, V. C.; Larsson, A.; Lee, P. H.; Leigh, J.; Leung, J.; Levi, M.; Li, Y.; Li Kappe, D.; Liang, X.; Liben, M. L.; Lim, S. S.; Liu, P. Y.; Liu, A.; Liu, Y.; Lodha, R.;

Logroscino, G.; Lorkowski, S.; Lotufo, P. A.; Lozano, R.; Lucas, T. C. D. D.; Ma, S.; Macarayan, E. R. K.; Maddison, E. R.; Magdy Abd El Razek, M.; Majdan, M.; Majdzadeh, R.; Majeed, A.; Malekzadeh, R.; Malhotra, R.; Malta, D. C.; Manguerra, H.; Manyazewal, T.; Mapoma, C. C.; Marczak, L. B.; Markos, D.; Martinez-Raga, J.; Martins-Melo, F. R.; Martopullo, I.; McAlinden, C.; McGaughey, M.; McGrath, J. J.; Mehata, S.; Meier, T.; Meles, K. G.; Memiah, P.; Memish, Z. A.; Mengesha, M. M.; Mengistu, D. T.; Menota, B. G.; Mensah, G. A.; Meretoja, T. J.; Meretoja, A.; Millear, A.; Miller, T. R.; Minnig, S.; Mirarefin, M.; Mirrakhimov, E. M.; Misganaw, A.; Mishra, S. R.; Mohamed, I. A.; Mohammad, K. A.; Mohammadi, A.; Mohammed, S.; Mokdad, A. H.; Mola, G. L. D.; Mollenkopf, S. K.; Molokhia, M.; Monasta, L.; Montañez, J. C.; Montico, M.; Mooney, M. D.; Moradi-Lakeh, M.; Moraga, P.; Morawska, L.; Morozoff, C.; Morrison, S. D.; Mountjoy-Venning, C.; Mruts, K. B.; Muller, K.; Murthy, G. V. S.; Musa, K. I.; Nachega, J. B.; Naheed, A.; Naldi, L.; Nangia, V.; Nascimento, B. R.; Nasher, J. T.; Natarajan, G.; Negoi, I.; Ngunjiri, J. W.; Nguyen, C. T.; Nguyen, Q. Le; Nguyen, T. H.; Nguyen, G.; Nguyen, M.; Nichols, E.; Ningrum, D. N. A.; Nong, V. M.; Noubiap, J. J. N.; Ogbo, F. A.; Oh, I.-H.; Okoro, A.; Olagunju, A. T.; Olsen, H. E.; Olusanya, B. O.; Olusanya, J. O.; Ong, K.; Opio, J. N.; Oren, E.; Ortiz, A.; Osman, M.; Ota, E.; PA, M.; Pacella, R. E.; Pakhale, S.; Pana, A.; Panda, B. K.; Panda-Jonas, S.; Papachristou, C.; Park, E.-K.; Patten, S. B.; Patton, G. C.; Paudel, D.; Paulson, K.; Pereira, D. M.; Perez-Ruiz, F.; Perico, N.; Pervaiz, A.; Petzold, M.; Phillips, M. R.; Pigott, D. M.; Pinho, C.; Plass, D.; Pletcher, M. A.; Polinder, S.; Postma, M. J.; Pourmalek, F.; Purcell, C.; Qorbani, M.; Quintanilla, B. P. A.; Radfar, A.; Rafay, A.; Rahimi-Movaghar, V.; Rahman, M. H. U.; Rahman, M.; Rai, R. K.; Ranabhat, C. L.; Rankin, Z.; Rao, P. C.; Rath, G. K.; Rawaf, S.; Ray, S. E.; Rehm, J.; Reiner,

R. C.; Reitsma, M. B.; Remuzzi, G.; Rezaei, S.; Rezai, M. S.; Rokni, M. B.; Ronfani, L.; Roshandel, G.; Roth, G. A.; Rothenbacher, D.; Ruhago, G. M.; SA, R.; Saadat, S.; Sachdev, P. S.; Sadat, N.; Safdarian, M.; Safi, S.; Safiri, S.; Sagar, R.; Sahathevan, R.; Salama, J.; Salamati, P.; Salomon, J. A.; Samy, A. M.; Sanabria, J. R.; Sanchez-Niño, M. D.; Santomauro, D.; Santos, I. S.; Santric Milicevic, M. M.; Sartorius, B.; Satpathy, M.; Schmidt, M. I.; Schneider, I. J. C. C.; Schulhofer-Wohl, S.; Schutte, A. E.; Schwebel, D. C.; Schwendicke, F.; Sepanlou, S. G.; Servan-Mori, E. E.; Shackelford, K. A.; Shahraz, S.; Shaikh, M. A.; Shamsipour, M.; Shamsizadeh, M.; Sharma, J.; Sharma, R.; She, J.; Sheikhbahaei, S.; Shey, M.; Shi, P.; Shields, C.; Shigematsu, M.; Shiri, R.; Shirude, S.; Shiue, I.; Shoman, H.; Shrime, M. G.; Sigfusdottir, I. D.; Silpakit, N.; Silva, J. P.; Singh, J. A.; Singh, A.; Skiadaresi, E.; Sligar, A.; Smith, D. L.; Smith, A.; Smith, M.; Sobaih, B. H. A. A.; Soneji, S.; Sorensen, R. J. D. D.; Soriano, J. B.; Sreeramareddy, C. T.; Srinivasan, V.; Stanaway, J. D.; Stathopoulou, V.; Steel, N.; Stein, D. J.; Steiner, C.; Steinke, S.; Stokes, M. A.; Strong, M.; Strub, B.; Subart, M.; Sufiyan, M. B.; Sunguya, B. F.; Sur, P. J.; Swaminathan, S.; Sykes, B. L.; Tabarés-Seisdedos, R.; Tadakamadla, S. K.; Takahashi, K.; Takala, J. S.; Talongwa, R. T.; Tarawneh, M. R.; Tavakkoli, M.; Taveira, N.; Tegegne, T. K.; Tehrani-Banihashemi, A.; Temsah, M.-H.; Terkawi, A. S.; Thakur, J. S.; Thamsuwan, O.; Thankappan, K. R.; Thomas, K. E.; Thompson, A. H.; Thomson, A. J.; Thrift, A. G.; Tobe-Gai, R.; Topor-Madry, R.; Torre, A.; Tortajada, M.; Towbin, J. A.; Tran, B. X.; Troeger, C.; Truelsen, T.; Tsoi, D.; Tuzcu, E. M.; Tyrovolas, S.; Ukwaja, K. N.; Undurraga, E. A.; Updike, R.; Uthman, O. A.; Uzochukwu, B. S. C. C.; van Boven, J. F. M. M.; Vasankari, T.; Venketasubramanian, N.; Violante, F. S.; Vlassov, V. V.; Vollset, S. E.; Vos, T.; Wakayo, T.; Wallin, M. T.; Wang, Y.-P.; Weiderpass, E.; Weintraub, R. G.; Weiss, D.

J.; Werdecker, A.; Westerman, R.; Whetter, B.; Whiteford, H. A.; Wijeratne, T.; Wiysonge, C. S.; Woldeyes, B. G.; Wolfe, C. D. A. A.; Woodbrook, R.; Workicho, A.; Xavier, D.; Xiao, Q.; Xu, G.; Yaghoubi, M.; Yakob, B.; Yano, Y.; Yaseri, M.; Yimam, H. H.; Yonemoto, N.; Yoon, S.-J.; Yotebieng, M.; Younis, M. Z.; Zaidi, Z.; Zaki, M. E. S.; Zegeye, E. A.; Zenebe, Z. M.; Zerfu, T. A.; Zhang, A. L.; Zhang, X.; Zipkin, B.; Zodpey, S.; Lopez, A. D.; Murray, C. J. L. Global, Regional, and National Age-Sex Specific Mortality for 264 Causes of Death, 1980–2016: A Systematic Analysis for the Global Burden of Disease Study 2016. *Lancet* **2017**, *390* (10100), 1151–1210.

(212) Mantri, S.; Morley, J. F.; Siderowf, A. D. The Importance of Preclinical Diagnostics in Parkinson Disease. *Park. Relat. Disord.* **2018**, *64* (May), 0–1.

(213) Adler, C. H.; Beach, T. G.; Hentz, J. G.; Shill, H. A.; Caviness, J. N.; Driver-Dunckley, E.; Sabbagh, M. N.; Sue, L. I.; Jacobson, S. A.; Belden, C. M.; Dugger, B. N.; Hentz, J. G.; Shill, H. A.; Caviness, J. N.; Driver-Dunckley, E.; Sabbagh, M. N.; Sue, L. I.; Jacobson, S. A.; Belden, C. M.; Dugger, B. N. Low Clinical Diagnostic Accuracy of Early vs Advanced Parkinson Disease: Clinicopathologic Study. *Neurology* **2014**, *83* (5), 406–412.

(214) Rizzo, G.; Copetti, M.; Arcuti, S.; Martino, D.; Fontana, A.; Logroscino, G. Accuracy of Clinical Diagnosis of Parkinson Disease. *Neurology* **2016**, *86* (6), 566–576.

(215) Wang, M.; Wang, C.; Han, R. H.; Han, X. Novel Advances in Shotgun Lipidomics for Biology and Medicine. *Prog. Lipid Res.* **2016**, *61*, 83–108.

(216) de Farias, C. C.; Maes, M.; Bonifácio, K. L.; Bortolasci, C. C.; de Souza Nogueira, A.; Brinholi, F. F.; Matsumoto, A. K.; do Nascimento, M. A.; de Melo, L. B.; Nixdorf, S. L.; Lavado, E. L.; Moreira, E. G.; Barbosa, D. S. Highly Specific Changes in Antioxidant

Levels and Lipid Peroxidation in Parkinson's Disease and Its Progression: Disease and Staging Biomarkers and New Drug Targets. *Neurosci. Lett.* **2016**, *617*, 66–71.

(217) Bhat, S.; Acharya, U. R.; Hagiwara, Y.; Dadmehr, N.; Adeli, H. Parkinson's Disease: Cause Factors, Measurable Indicators, and Early Diagnosis. *Comput. Biol. Med.* **2018**, *102* (July), 234–241.

(218) Pchelina, S.; Baydakova, G.; Nikolaev, M.; Senkevich, K.; Emelyanov, A.; Kopytova, A.; Miliukhina, I.; Yakimovskii, A.; Timofeeva, A.; Berkovich, O.; Fedotova, E.; Illarioshkin, S.; Zakharova, E. Blood Lysosphingolipids Accumulation in Patients with Parkinson's Disease with Glucocerebrosidase 1 Mutations. *Mov. Disord.* **2018**, *33* (8), 1325–1330.

(219) Migdalska-Richards, A.; Schapira, A. H. V. The Relationship between Glucocerebrosidase Mutations and Parkinson Disease. *J. Neurochem.* **2016**, *139*, 77–90.

(220) Do, C. B.; Tung, J. Y.; Dorfman, E.; Kiefer, A. K.; Drabant, E. M.; Francke, U.; Mountain, J. L.; Goldman, S. M.; Tanner, C. M.; Langston, J. W.; Wojcicki, A.; Eriksson, N. Web-Based Genome-Wide Association Study Identifies Two Novel Loci and a Substantial Genetic Component for Parkinson's Disease. *PLoS Genet.* **2011**, *7* (6), e1002141.

(221) Pankratz, N.; Wilk, J. B.; Latourelle, J. C.; DeStefano, A. L.; Halter, C.; Pugh, E. W.; Doheny, K. F.; Gusella, J. F.; Nichols, W. C.; Foroud, T.; Myers, R. H.; PSG-PROGENI; GenePD. Genomewide Association Study for Susceptibility Genes Contributing to Familial Parkinson Disease. *Hum. Genet.* **2010**, *124* (6), 593–605.

(222) Robak, L. A.; Jansen, I. E.; Rooij, J. van; Uitterlinden, A. G.; Kraaij, R.; Jankovic, J.; Heutink, P.; Shulman, J. M.; Nalls, M. A.; Plagnol, V.; Hernandez, D. G.; Sharma, M.; Sheerin, U.-M. M.; Saad, M.; Simón-Sánchez, J.; Schulte, C.; Lesage, S.; Sveinbjörnsdóttir,

S.; Arepalli, S.; Barker, R.; Ben, Y.; Berendse, H. W.; Berg, D.; Bhatia, K.; de Bie, R. M. A. A.; Biffi, A.; Bloem, B.; Bochdanovits, Z.; Bonin, M.; Bras, J. M.; Brockmann, K.; Brooks, J.; Burn, D. J.; Majounie, E.; Charlesworth, G.; Lungu, C.; Chen, H.; Chinnery, P. F.; Chong, S.; Clarke, C. E.; Cookson, M. R.; Cooper, J. M.; Corvol, J. C.; Counsell, C.; Damier, P.; Dartigues, J.-F. F.; Deloukas, P.; Deuschl, G.; Dexter, D. T.; van Dijk, K. D.; Dillman, A.; Durif, F.; Dürr, A.; Edkins, S.; Evans, J. R.; Foltynie, T.; Dong, J.; Gardner, M.; Gibbs, J. R.; Goate, A.; Gray, E.; Guerreiro, R.; Harris, C.; van Hilten, J. J.; Hofman, A.; Hollenbeck, A.; Holton, J.; Hu, M.; Huang, X.; Wurster, I.; Mätzler, W.; Hudson, G.; Hunt, S. E.; Huttenlocher, J.; Illig, T.; Jónsson, P. V.; Lambert, J.-C. C.; Langford, C.; Lees, A.; Lichtner, P.; Limousin, P.; Lopez, G.; Lorenz, D.; Lungu, C.; McNeill, A.; Moorby, C.; Moore, M.; Morris, H. R.; Morrison, K. E.; Escott-Price, V.; Mudanohwo, E.; O'Sullivan, S. S.; Pearson, J.; Perlmutter, J. S.; Pétursson, H.; Pollak, P.; Post, B.; Potter, S.; Ravina, B.; Revesz, T.; Riess, O.; Rivadeneira, F.; Rizzu, P.; Ryten, M.; Sawcer, S.; Schapira, A.; Scheffer, H.; Shaw, K.; Shoulson, I.; Shulman, J. M.; Sidransky, E.; Smith, C.; Spencer, C. C. A. A.; Stefánsson, H.; Bettella, F.; Stockton, J. D.; Strange, A.; Talbot, K.; Tanner, C. M.; Tashakkori-Ghanbaria, A.; Tison, F.; Trabzuni, D.; Traynor, B. J.; Uitterlinden, A. G.; Velseboer, D.; Vidailhet, M.; Walker, R.; Warrenburg, B. van de; Wickremaratchi, M.; Williams, N.; Williams-Gray, C. H.; Winder-Rhodes, S.; Stefánsson, K.; Martinez, M.; Wood, N. W.; Hardy, J.; Heutink, P.; Brice, A.; Gasser, T.; Singleton, A. B.; van Rooij, J.; Uitterlinden, A. G.; Kraaij, R.; Jankovic, J.; Heutink, P.; Shulman, J. M.; Nalls, M. A.; Plagnol, V.; Hernandez, D. G.; Sharma, M.; Sheerin, U.-M. M.; Saad, M.; Simón-Sánchez, J.; Schulte, C.; Lesage, S.; Sveinbjörnsdóttir, S.; Arepalli, S.; Barker, R.; Ben-, Y.; Berendse, H. W.; Berg, D.; Bhatia, K.; de Bie, R. M. A. A.; Biffi, A.; Bloem, B.;

Bochdanovits, Z.; Bonin, M.; Bras, J. M.; Brockmann, K.; Brooks, J.; Burn, D. J.; Majounie, E.; Charlesworth, G.; Lungu, C.; Chen, H.; Chinnery, P. F.; Chong, S.; Clarke, C. E.; Cookson, M. R.; Mark Cooper, J.; Corvol, J. C.; Counsell, C.; Damier, P.; Dartigues, J.-F. F.; Deloukas, P.; Deuschl, G.; Dexter, D. T.; van Dijk, K. D.; Dillman, A.; Durif, F.; Dürr, A.; Edkins, S.; Evans, J. R.; Foltynie, T.; Dong, J.; Gardner, M.; Raphael Gibbs, J.; Goate, A.; Gray, E.; Guerreiro, R.; Harris, C.; van Hilten, J. J.; Hofman, A.; Hollenbeck, A.; Holton, J.; Hu, M.; Huang, X.; Wurster, I.; Mätzler, W.; Hudson, G.; Hunt, S. E.; Huttenlocher, J.; Illig, T.; Jónsson, P. V.; Lambert, J.-C. C.; Langford, C.; Lees, A.; Lichtner, P.; Limousin, P.; Lopez, G.; Lorenz, D.; Lungu, C.; McNeill, A.; Moorby, C.; Moore, M.; Morris, H. R.; Morrison, K. E.; Escott-Price, V.; Mudanohwo, E.; O'Sullivan, S. S.; Pearson, J.; Perlmutter, J. S.; Pétursson, H.; Pollak, P.; Post, B.; Potter, S.; Ravina, B.; Revesz, T.; Riess, O.; Rivadeneira, F.; Rizzu, P.; Ryten, M.; Sawcer, S.; Schapira, A.; Scheffer, H.; Shaw, K.; Shoulson, I.; Shulman, J. M.; Sidransky, E.; Smith, C.; Spencer, C. C. A. A.; Stefánsson, H.; Bettella, F.; Stockton, J. D.; Strange, A.; Talbot, K.; Tanner, C. M.; Tashakkori-Ghanbaria, A.; Tison, F.; Trabzuni, D.; Traynor, B. J.; Uitterlinden, A. G.; Velseboer, D.; Vidailhet, M.; Walker, R.; van de Warrenburg, B.; Wickremaratchi, M.; Williams, N.; Williams-Gray, C. H.; Winder-Rhodes, S.; Stefánsson, K.; Martinez, M.; Wood, N. W.; Hardy, J.; Heutink, P.; Brice, A.; Gasser, T.; Singleton, A. B.; Rooij, J. van; Uitterlinden, A. G.; Kraaij, R.; Jankovic, J.; Heutink, P.; Shulman, J. M.; Nalls, M. A.; Plagnol, V.; Hernandez, D. G.; Sharma, M.; Sheerin, U.-M. M.; Saad, M.; Simón-Sánchez, J.; Schulte, C.; Lesage, S.; Sveinbjörnsdóttir, S.; Arepalli, S.; Barker, R.; Ben, Y.; Berendse, H. W.; Berg, D.; Bhatia, K.; de Bie, R. M. A. A.; Biffi, A.; Bloem, B.; Bochdanovits, Z.; Bonin, M.; Bras, J. M.; Brockmann, K.; Brooks, J.; Burn, D. J.; Majounie, E.; Charlesworth,

G.; Lungu, C.; Chen, H.; Chinnery, P. F.; Chong, S.; Clarke, C. E.; Cookson, M. R.; Cooper, J. M.; Corvol, J. C.; Counsell, C.; Damier, P.; Dartigues, J.-F. F.; Deloukas, P.; Deuschl, G.; Dexter, D. T.; van Dijk, K. D.; Dillman, A.; Durif, F.; Dürr, A.; Edkins, S.; Evans, J. R.; Foltynie, T.; Dong, J.; Gardner, M.; Gibbs, J. R.; Goate, A.; Gray, E.; Guerreiro, R.; Harris, C.; van Hilten, J. J.; Hofman, A.; Hollenbeck, A.; Holton, J.; Hu, M.; Huang, X.; Wurster, I.; Mätzler, W.; Hudson, G.; Hunt, S. E.; Huttenlocher, J.; Illig, T.; Jónsson, P. V.; Lambert, J.-C. C.; Langford, C.; Lees, A.; Lichtner, P.; Limousin, P.; Lopez, G.; Lorenz, D.; Lungu, C.; McNeill, A.; Moorby, C.; Moore, M.; Morris, H. R.; Morrison, K. E.; Escott-Price, V.; Mudanohwo, E.; O'Sullivan, S. S.; Pearson, J.; Perlmutter, J. S.; Pétursson, H.; Pollak, P.; Post, B.; Potter, S.; Ravina, B.; Revesz, T.; Riess, O.; Rivadeneira, F.; Rizzu, P.; Ryten, M.; Sawcer, S.; Schapira, A.; Scheffer, H.; Shaw, K.; Shoulson, I.; Shulman, J. M.; Sidransky, E.; Smith, C.; Spencer, C. C. A. A.; Stefánsson, H.; Bettella, F.; Stockton, J. D.; Strange, A.; Talbot, K.; Tanner, C. M.; Tashakkori-Ghanbaria, A.; Tison, F.; Trabzuni, D.; Traynor, B. J.; Uitterlinden, A. G.; Velseboer, D.; Vidailhet, M.; Walker, R.; Warrenburg, B. van de; Wickremaratchi, M.; Williams, N.; Williams-Gray, C. H.; Winder-Rhodes, S.; Stefánsson, K.; Martinez, M.; Wood, N. W.; Hardy, J.; Heutink, P.; Brice, A.; Gasser, T.; Singleton, A. B. Excessive Burden of Lysosomal Storage Disorder Gene Variants in Parkinson's Disease. *Brain* **2017**, *140* (12), 3191–3203.

(223) Farmer, K.; Smith, C. A.; Hayley, S.; Smith, J. Major Alterations of Phosphatidylcholine and Lysophosphotidylcholine Lipids in the Substantia Nigra Using an Early Stage Model of Parkinson's Disease. *Int. J. Mol. Sci.* **2015**, *16* (8), 18865–18877.

(224) Boutin, M.; Sun, Y.; Shacka, J. J.; Auray-Blais, C. Tandem Mass Spectrometry Multiplex Analysis of Glucosylceramide and Galactosylceramide Isoforms in Brain Tissues at

Different Stages of Parkinson Disease. *Anal. Chem.* **2016**, *88* (3), 1856–1863.

(225) Ferrazza, R.; Cogo, S.; Melrose, H.; Bubacco, L.; Greggio, E.; Guella, G.; Civiero, L.; Plotegher, N. LRRK2 Deficiency Impacts Ceramide Metabolism in Brain. *Biochem. Biophys. Res. Commun.* **2016**, *478* (3), 1141–1146.

(226) Xing, Y.; Tang, Y.; Zhao, L.; Wang, Q.; Qin, W.; Ji, X.; Zhang, J.; Jia, J. Associations between Plasma Ceramides and Cognitive and Neuropsychiatric Manifestations in Parkinson's Disease Dementia. *J. Neurol. Sci.* **2016**, *370*, 82–87.

(227) Chan, R. B.; Perotte, A. J.; Zhou, B.; Liong, C.; Shorr, E. J.; Marder, K. S.; Kang, U. J.; Waters, C. H.; Levy, O. A.; Xu, Y.; Shim, H. Bin; Pe'er, I.; Di Paolo, G.; Alcalay, R. N. Elevated GM3 Plasma Concentration in Idiopathic Parkinson's Disease: A Lipidomic Analysis. *PLoS One* **2017**, *12* (2), e0172348.

(228) Guedes, L. C.; Chan, R. B.; Gomes, M. A.; Conceição, V. A.; Machado, R. B.; Soares, T.; Xu, Y.; Gaspar, P.; Carriço, J. A.; Alcalay, R. N.; Ferreira, J. J.; Outeiro, T. F.; Miltenberger-Miltenyi, G. Serum Lipid Alterations in GBA-Associated Parkinson's Disease. *Park. Relat. Disord.* **2017**, *44*, 58–65.

(229) Han, W.; Sapkota, S.; Camicioli, R.; Dixon, R. A.; Li, L. Profiling Novel Metabolic Biomarkers for Parkinson's Disease Using in-Depth Metabolomic Analysis. *Mov. Disord.* **2017**, *32* (12), 1720–1728.

(230) Acharya, H. J.; Bouchard, T. P.; Emery, D. J.; Camicioli, R. M. Axial Signs and Magnetic Resonance Imaging Correlates in Parkinson's Disease. *Can. J. Neurol. Sci. / J. Can. des Sci. Neurol.* **2007**, *34* (1), 56–61.

(231) Camicioli, R.; Sabino, J.; Gee, M.; Bouchard, T.; Fisher, N.; Hanstock, C.; Emery, D.;

Martin, W. R. R. W. Ventricular Dilatation and Brain Atrophy in Patients with Parkinson's Disease with Incipient Dementia. *Mov. Disord.* **2011**, *26* (8), 1443–1450.

(232) McDermott, K. L.; Fisher, N.; Bradford, S.; Camicioli, R. Parkinson's Disease Mild Cognitive Impairment Classifications and Neurobehavioral Symptoms. *Int. Psychogeriatrics* **2018**, *30* (2), 253–260.

(233) Miller, I. N.; Cronin-Golomb, A. Gender Differences in Parkinson's Disease: Clinical Characteristics and Cognition. *Mov. Disord.* **2010**, *25* (16), 2695–2703.

(234) Gegg, M. E.; Schapira, A. H. V. The Role of Glucocerebrosidase in Parkinson Disease Pathogenesis. *FEBS J.* **2018**, *285* (19), 3591–3603.

(235) Savica, R.; Murray, M. E.; Persson, X.-M.; Kantarci, K.; Parisi, J. E.; Dickson, D. W.; Petersen, R. C.; Ferman, T. J.; Boeve, B. F.; Mielke, M. M. Plasma Sphingolipid Changes with Autopsy-confirmed Lewy Body or Alzheimer's Pathology. *Alzheimer's Dement. Diagnosis, Assess. Dis. Monit.* **2016**, *3* (1), 43–50.

(236) Nelson, M. P.; Boutin, M.; Tse, T. E.; Lu, H.; Haley, E. D.; Ouyang, X.; Zhang, J.; Auray-Blais, C.; Shacka, J. J. The Lysosomal Enzyme Alpha-Galactosidase A Is Deficient in Parkinson's Disease Brain in Association with the Pathologic Accumulation of Alpha-Synuclein. *Neurobiol. Dis.* **2018**, *110* (December 2017), 68–81.

(237) Di Pardo, A.; Maglione, V. Sphingolipid Metabolism: A New Therapeutic Opportunity for Brain Degenerative Disorders. *Front. Neurosci.* **2018**, *12* (APR), 58–65.

(238) Jana, A.; Hogan, E. L.; Pahan, K. Ceramide and Neurodegeneration: Susceptibility of Neurons and Oligodendrocytes to Cell Damage and Death. *J. Neurol. Sci.* **2009**, *278* (1–2), 5–15.

(239) Alza, N. P.; Iglesias González, P. A.; Conde, M. A.; Uranga, R. M.; Salvador, G. A. Lipids at the Crossroad of α-Synuclein Function and Dysfunction: Biological and Pathological Implications. *Front. Cell. Neurosci.* **2019**, *13* (May), 1–17.

(240) Stoessel, D.; Schulte, C.; Teixeira dos Santos, M. C.; Scheller, D.; Rebollo-Mesa, I.; Deuschle, C.; Walther, D.; Schauer, N.; Berg, D.; Nogueira da Costa, A.; Maetzler, W. Promising Metabolite Profiles in the Plasma and CSF of Early Clinical Parkinson's Disease. *Front. Aging Neurosci.* **2018**, *10* (March), 1–14.

(241) Lin, G.; Wang, L.; Marcogliese, P. C.; Bellen, H. J. Sphingolipids in the Pathogenesis of Parkinson's Disease and Parkinsonism. *Trends Endocrinol. Metab.* **2019**, *30* (2), 106–117.

(242) Wood, P. L.; Tippireddy, S.; Feriante, J.; Woltjer, R. L. Augmented Frontal Cortex Diacylglycerol Levels in Parkinson's Disease and Lewy Body Disease. *PLoS One* **2018**, *13* (3), e0191815.

(243) Cain, D. W.; Cidlowski, J. A. Immune Regulation by Glucocorticoids. *Nat. Rev. Immunol.* **2017**, *17* (4), 233–247.

(244) Oray, M.; Abu Samra, K.; Ebrahimiadib, N.; Meese, H.; Foster, C. S. Long-Term Side Effects of Glucocorticoids. *Expert Opin. Drug Saf.* **2016**, *15* (4), 457–465.

(245) Walker, B. R. Glucocorticoids and Cardiovascular Disease. *Eur. J. Endocrinol.* **2007**, *157* (5), 545–559.

(246) Ye, Q.; He, X.-O.; D'Urzo, A. A Review on the Safety and Efficacy of Inhaled Corticosteroids in the Management of Asthma. *Pulm. Ther.* **2017**, *3* (1), 1–18.

(247) Horby, P.; Lim, W. S.; Emberson, J.; Mafham, M.; Bell, J.; Linsell, L.; Staplin, N.; Brightling, C.; Ustianowski, A.; Elmahi, E.; Prudon, B.; Green, C.; Felton, T.; Chadwick,

D.; Rege, K.; Fegan, C.; Chappell, L. C.; Faust, S. N.; Jaki, T.; Jeffery, K.; Montgomery, A.; Rowan, K.; Juszczak, E.; Baillie, J. K.; Haynes, R.; Landray, M. J.; Group, R. C. Effect of Dexamethasone in Hospitalized Patients with COVID-19: Preliminary Report. *medRxiv* **2020**, 2020.06.22.20137273.

(248) Johnson, R. M.; Vinetz, J. M. Dexamethasone in the Management of Covid -19. *BMJ* **2020**, *370*, m2648.

(249) World Health Organization Model List of Essential Medicines, 21st List. World Health Organization (WHO): Geneva, Switzerland, Switzerland 2019, pp 119–134.

(250) Buchman, A. L. Side Effects of Corticosteroid Therapy. *J. Clin. Gastroenterol.* **2001**, *33* (4), 289–294.

(251) Mangos, G. J.; Whitworth, J. A.; Williamson, P. M.; Kelly, J. J. Glucocorticoids and the Kidney. *Nephrology* **2003**, *8* (6), 267–273.

(252) Malkawi, A. K.; Masood, A.; Shinwari, Z.; Jacob, M.; Benabdelkamel, H.; Matic, G.; Almuhanna, F.; Dasouki, M.; Alaiya, A. A.; Rahman, A. M. A. Proteomic Analysis of Morphologically Changed Tissues after Prolonged Dexamethasone Treatment. *Int. J. Mol. Sci.* **2019**, *20* (13), 3122.

(253) Malkawi, A. K.; Alzoubi, K. H.; Jacob, M.; Matic, G.; Ali, A.; Al Faraj, A.; Almuhanna, F.; Dasouki, M.; Abdel Rahman, A. M. Metabolomics Based Profiling of Dexamethasone Side Effects in Rats. *Front. Pharmacol.* **2018**, *9* (FEB), 1–14.

(254) Lv, Z.-P.; Peng, Y.-Z.; Zhang, B.-B.; Fan, H.; Liu, D.; Guo, Y.-M. Glucose and Lipid Metabolism Disorders in the Chickens with Dexamethasone-Induced Oxidative Stress. *J. Anim. Physiol. Anim. Nutr. (Berl).* **2018**, *102* (2), e706–e717.

(255) Trayssac, M.; Hannun, Y. A.; Obeid, L. M. Role of Sphingolipids in Senescence: Implication in Aging and Age-Related Diseases. *J. Clin. Invest.* **2018**, *128* (7), 2702–2712.

(256) Bianca Velasco, A.; Tan, Z. S. Fatty Acids and the Aging Brain. In *Omega-3 Fatty Acids in Brain and Neurological Health*; Elsevier, 2014; pp 201–219.

(257) Schulze, P. C.; Drosatos, K.; Goldberg, I. J. Lipid Use and Misuse by the Heart. *Circ. Res.* **2016**, *118* (11), 1736–1751.

(258) Scherer, M.; Schmitz, G. Metabolism, Function and Mass Spectrometric Analysis of Bis(Monoacylglycero)Phosphate and Cardiolipin. *Chem. Phys. Lipids* **2011**, *164* (6), 556–562.

(259) Bhat, O. M.; Yuan, X.; Li, G.; Lee, R.; Li, P.-L. Sphingolipids and Redox Signaling in Renal Regulation and Chronic Kidney Diseases. *Antioxid. Redox Signal.* **2018**, *28* (10), 1008–1026.

(260) Harrison-Bernard, L. M. Sphingolipids, New Kids on the Block, Promoting Glomerular Fibrosis in the Diabetic Kidney. *American Journal of Physiology - Renal Physiology*. American Physiological Society 2015, pp F685–F686.

(261) Frayn, K. N.; Arner, P.; Yki-Järvinen, H. Fatty Acid Metabolism in Adipose Tissue, Muscle and Liver in Health and Disease. *Essays Biochem.* **2006**, *42*, 89–103.

(262) Atshaves, B. P.; Martin, G. G.; Hostetler, H. A.; McIntosh, A. L.; Kier, A. B.; Schroeder, F. Liver Fatty Acid-Binding Protein and Obesity. *Journal of Nutritional Biochemistry*. Elsevier November 1, 2010, pp 1015–1032.

(263) Dhalla, N. S.; Xu, Y. J.; Sheu, S. S.; Tappia, P. S.; Panagia, V. Phosphatidic Acid: A Potential Signal Transducer for Cardiac Hypertrophy. *Journal of Molecular and Cellular*

Cardiology. Academic Press November 1, 1997, pp 2865–2871.

(264) Piltch, A.; Sun, L.; Fava, R. A.; Hayashi, J. Lipocortin-Independent Effect of Dexamethasone on Phospholipase Activity in a Thymic Epithelial Cell Line. *Biochem. J.* **1989**, *261* (2), 395–400.

(265) Zawalich, W. S.; Tesz, G. J.; Yamazaki, H.; Zawalich, K. C.; Philbrick, W. Dexamethasone Suppresses Phospholipase C Activation and Insulin Secretion from Isolated Rat Islets. *Metabolism* **2006**, *55* (1), 35–42.

(266) Wang, J. *Glucocorticoid Signaling*; Wang, J.-C., Harris, C., Eds.; Advances in Experimental Medicine and Biology; Springer New York: New York, NY, 2015; Vol. 872.

(267) Rahimi, L.; Rajpal, A.; Ismail-Beigi, F. Glucocorticoid-Induced Fatty Liver Disease. *Diabetes, Metab. Syndr. Obes. Targets Ther.* **2020**, *Volume 13*, 1133–1145.

(268) Kawano, Y.; Cohen, D. E. Mechanisms of Hepatic Triglyceride Accumulation in Non-Alcoholic Fatty Liver Disease. *J. Gastroenterol.* **2013**, *48* (4), 434–441.

(269) Alves-Bezerra, M.; Cohen, D. E. Triglyceride Metabolism in the Liver. In *Comprehensive Physiology*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2017; Vol. 8, pp 1–22.

(270) Makrecka-Kuka, M.; Sevostjanovs, E.; Vilks, K.; Volska, K.; Antone, U.; Kuka, J.; Makarova, E.; Pugovics, O.; Dambrova, M.; Liepinsh, E. Plasma Acylcarnitine Concentrations Reflect the Acylcarnitine Profile in Cardiac Tissues. *Sci. Rep.* **2017**, *7* (1), 17528.

(271) Schooneman, M. G.; Vaz, F. M.; Houten, S. M.; Soeters, M. R. Acylcarnitines. *Diabetes* **2013**, *62* (1), 1–8.

(272) Owczarek, J.; Jasińska, M.; Orszulak-Michalak, D. Drug-Induced Myopathies. An

Overview of the Possible Mechanisms. *Pharmacol. Reports* **2005**, *57* (1), 23–34.

(273) da Silva, T. F.; Sousa, V. F.; Malheiro, A. R.; Brites, P. The Importance of Ether-Phospholipids: A View from the Perspective of Mouse Models. *Biochim. Biophys. Acta - Mol. Basis Dis.* **2012**, *1822* (9), 1501–1508.

(274) Brites, P.; Waterham, H. R.; Wanders, R. J. . A. Functions and Biosynthesis of Plasmalogens in Health and Disease. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2004**, *1636* (2–3), 219–231.

(275) Farinha, C. M.; Miller, E.; McCarty, N. Protein and Lipid Interactions – Modulating CFTR Trafficking and Rescue. *J. Cyst. Fibros.* **2018**, *17* (2), S9–S13.

(276) Gulbins, E. Lipids Control Mucus Production in Cystic Fibrosis. *Nat. Med.* **2010**, *16* (3), 267–268.

(277) Zhang, Z.; Liu, F.; Chen, J. Conformational Changes of CFTR upon Phosphorylation and ATP Binding. *Cell* **2017**, *170* (3), 483-491.e8.

(278) Ehrhardt, A.; Chung, W. J.; Pyle, L. C.; Wang, W.; Nowotarski, K.; Mulvihill, C. M.; Ramjeesingh, M.; Hong, J.; Velu, S. E.; Lewis, H. A.; Atwell, S.; Aller, S.; Bear, C. E.; Lukacs, G. L.; Kirk, K. L.; Sorscher, E. J. Channel Gating Regulation by the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) First Cytosolic Loop. *J. Biol. Chem.* **2016**, *291* (4), 1854–1865.

(279) Csanády, L.; Vergani, P.; Gadsby, D. C. Structure, Gating, and Regulation of the CFTR Anion Channel. *Physiol. Rev.* **2019**, *99* (1), 707–738.

(280) Accurso, F. J.; Ratjen, F.; Bell, S. C.; Rowe, S. M.; Goss, C. H.; Quittner, A. L.; Bush, A.; Tullis, E.; Accurso, F. J.; Ratjen, F.; Bell, S. C.; Rowe, S. M.; Goss, C. H.; Quittner, A. L.;

Bush, A.; Walker, J. M.; Naehrig, S.; Chao, C.-M.; Naehrlich, L.; Ratjen, F.; Bell, S. C.; Rowe, S. M.; Goss, C. H.; Quittner, A. L.; Bush, A. Cystic Fibrosis. *Nat. Rev. Dis. Prim.* **2015**, *1* (May), 15010.

(281) Accurso, F. J. Cystic Fibrosis. In *Goldman's Cecil Medicine*; Elsevier, 2012; Vol. 1, pp 544–548.

(282) Cutting, G. R. Cystic Fibrosis. In *International Encyclopedia of Public Health*; Elsevier, 2017; Vol. 2, pp 208–210.

(283) Michelson, P.; Faro, A.; Ferkol, T. Pulmonary Disease in Cystic Fibrosis. In *Kendig's Disorders of the Respiratory Tract in Children*; Wilmott, R., Bush, A., Deterding, R., Ratjen, F., Sly, P., Zar, H., Li, A. P., Eds.; Elsevier Inc., 2019; pp 777–787.

(284) Garcia, A. M.; Dorsey, J. Nonpulmonary Manifestations of Cystic Fibrosis. In *Kendig's Disorders of the Respiratory Tract in Children*; Wilmott, R., Bush, A., Deterding, R., Ratjen, F., Sly, P., Zar, H., Li, A. P., Eds.; Elsevier Inc., 2019; pp 788–799.

(285) MacKenzie, T.; Gifford, A. H.; Sabadosa, K. A.; Quinton, H. B.; Knapp, E. A.; Goss, C. H.; Marshall, B. C. Longevity of Patients With Cystic Fibrosis in 2000 to 2010 and Beyond: Survival Analysis of the Cystic Fibrosis Foundation Patient Registry. *Ann. Intern. Med.* **2014**, *161* (4), 233.

(286) Stephenson, A. L.; Sykes, J.; Stanojevic, S.; Quon, B. S.; Marshall, B. C.; Petren, K.; Ostrenga, J.; Fink, A. K.; Elbert, A.; Goss, C. H. Survival Comparison of Patients With Cystic Fibrosis in Canada and the United States. *Ann. Intern. Med.* **2017**, *166* (8), 537.

(287) Levy, H.; Farrell, P. M. New Challenges in the Diagnosis and Management of Cystic Fibrosis. *J. Pediatr.* **2015**, *166* (6), 1337–1341.

(288) Molina, S. A.; Hunt, W. R. Cystic Fibrosis. In *Lung Epithelial Biology in the Pathogenesis of Pulmonary Disease*; Elsevier, 2017; pp 219–249.

(289) Bienvenu, T.; Nguyen-Khoa, T. Current and Future Diagnosis of Cystic Fibrosis: Performance and Limitations. *Arch. Pédiatrie* **2020**, *27*, eS19–eS24.

(290) Wallis, C. Diagnosis and Presentation of Cystic Fibrosis. In *Kendig's Disorders of the Respiratory Tract in Children*; Wilmott, R., Bush, A., Deterding, R., Ratjen, F., Sly, P., Zar, H., Li, A. P., Eds.; Elsevier Inc., 2019; pp 769–776.

(291) Marson, F. A. L.; Bertuzzo, C. S.; Ribeiro, J. D. Classification of CFTR Mutation Classes. *Lancet Respir. Med.* **2016**, *4* (8), e37–e38.

(292) Goetzinger, K. R. Cystic Fibrosis. In *Obstetric Imaging: Fetal Diagnosis and Care*; Elsevier, 2018; pp 579-581.e1.

(293) Muhlebach, M. S.; Sha, W. Lessons Learned from Metabolomics in Cystic Fibrosis. *Mol. Cell. Pediatr.* **2015**, *2* (1), 9.

(294) Ramu, Y.; Xu, Y.; Lu, Z. Inhibition of CFTR Cl- Channel Function Caused by Enzymatic Hydrolysis of Sphingomyelin. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (15), 6448–6453.

(295) Becker, K. A.; Riethmüller, J.; Lüth, A.; Döring, G.; Kleuser, B.; Gulbins, E. Acid Sphingomyelinase Inhibitors Normalize Pulmonary Ceramide and Inflammation in Cystic Fibrosis. *Am. J. Respir. Cell Mol. Biol.* **2010**, *42* (6), 716–724.

(296) Stauffer, B. B.; Cui, G.; Cottrill, K. A.; Infield, D. T.; McCarty, N. A. Bacterial Sphingomyelinase Is a State-Dependent Inhibitor of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR). *Sci. Rep.* **2017**, *7* (1), 1–14.

(297) Guerrera, I. C.; Astarita, G.; Jais, J.-P.; Sands, D.; Nowakowska, A.; Colas, J.; Sermet-

Gaudelus, I.; Schuerenberg, M.; Piomelli, D.; Edelman, A.; Ollero, M. A Novel Lipidomic Strategy Reveals Plasma Phospholipid Signatures Associated with Respiratory Disease Severity in Cystic Fibrosis Patients. *PLoS One* **2009**, *4* (11), e7735.

(298) Ollero, M.; Astarita, G.; Guerrera, I. C.; Sermet-Gaudelus, I.; Trudel, S.; Piomelli, D.; Edelman, A. Plasma Lipidomics Reveals Potential Prognostic Signatures within a Cohort of Cystic Fi Brosis Patients. *J. Lipid Res.* **2011**, *52* (5), 1011–1022.

(299) Seidl, E.; Kiermeier, H.; Liebisch, G.; Ballmann, M.; Hesse, S.; Paul-Buck, K.; Ratjen, F.; Rietschel, E.; Griese, M. Lavage Lipidomics Signatures in Children with Cystic Fibrosis and Protracted Bacterial Bronchitis. *J. Cyst. Fibros.* **2019**, *18* (6), 790–795.

(300) Corriveau, S.; Sykes, J.; Stephenson, A. L. Cystic Fibrosis Survival. *Curr. Opin. Pulm. Med.* **2018**, *24* (6), 574–578.

(301) Szczesniak, R.; Heltshe, S. L.; Stanojevic, S.; Mayer-Hamblett, N. Use of FEV1 in Cystic Fibrosis Epidemiologic Studies and Clinical Trials: A Statistical Perspective for the Clinical Researcher. *J. Cyst. Fibros.* **2017**, *16* (3), 318–326.

(302) Banjar, H.; Al-Mogarri, I.; Nizami, I.; Al-Haider, S.; AlMaghamsi, T.; Alkaf, S.; Al-Enazi, A.; Moghrabi, N. Geographic Distribution of Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Gene Mutations in Saudi Arabia (Article in Press). *Int. J. Pediatr. Adolesc. Med.* **2019**.

(303) GOLD 2017. *Global Initiative for Chronic Obstructive Lung Disease : Pocket Guide To COPD Diagnosis, Management, and Prevention, A Guide for Health Care Professionals (Https://Goldcopd.Org)*; 2017.

(304) Rhodes, B.; Nash, E. F.; Tullis, E.; Pencharz, P. B.; Brotherwood, M.; Dupuis, A.;

Stephenson, A. Prevalence of Dyslipidemia in Adults with Cystic Fibrosis. *J. Cyst. Fibros.* **2010**, *9* (1), 24–28.

(305) Gunasekara, L.; Al-Saiedy, M.; Green, F.; Pratt, R.; Bjornson, C.; Yang, A.; Michael Schoel, W.; Mitchell, I.; Brindle, M.; Montgomery, M.; Keys, E.; Dennis, J.; Shrestha, G.; Amrein, M. Pulmonary Surfactant Dysfunction in Pediatric Cystic Fibrosis: Mechanisms and Reversal with a Lipid-Sequestering Drug. *J. Cyst. Fibros.* **2017**, *16* (5), 565–572.

(306) Woestenenk, J. W.; Schulkes, D. A.; Schipper, H. S.; van der Ent, C. K.; Houwen, R. H. J. Dietary Intake and Lipid Profile in Children and Adolescents with Cystic Fibrosis. *J. Cyst. Fibros.* **2017**, *16* (3), 410–417.

(307) Scholte, B. J.; Horati, H.; Veltman, M.; Vreeken, R. J.; Garratt, L. W.; Tiddens, H. A. W. M.; Janssens, H. M.; Stick, S. M. Oxidative Stress and Abnormal Bioactive Lipids in Early Cystic Fibrosis Lung Disease. *J. Cyst. Fibros.* **2019**, *18* (6), 781–789.

(308) Yang, J.; Eiserich, J. P.; Cross, C. E.; Morrissey, B. M.; Hammock, B. D. Metabolomic Profiling of Regulatory Lipid Mediators in Sputum from Adult Cystic Fibrosis Patients. *Free Radic. Biol. Med.* **2012**, *53* (1), 160–171.

(309) Calder, P.; Grimble, R. Polyunsaturated Fatty Acids, Inflammation and Immunity. *Eur. J. Clin. Nutr.* **2002**, *56* (S3), S14–S19.

(310) Jenkins, B.; West, J. A.; Koulman, A. A Review of Odd-Chain Fatty Acid Metabolism and the Role of Pentadecanoic Acid (C15:0) and Heptadecanoic Acid (C17:0) in Health and Disease. *Molecules* **2015**, *20* (2), 2425–2444.

(311) Khaw, K.-T. T.; Friesen, M. D.; Riboli, E.; Luben, R.; Wareham, N. Plasma Phospholipid Fatty Acid Concentration and Incident Coronary Heart Disease in Men and Women: The

EPIC-Norfolk Prospective Study. *PLoS Med.* **2012**, *9* (7), e1001255.

(312) Jacobs, S.; Schiller, K.; Jansen, E.; Fritsche, A.; Weikert, C.; Di Giuseppe, R.; Boeing, H.; Schulze, M. B.; Kröger, J. Association between Erythrocyte Membrane Fatty Acids and Biomarkers of Dyslipidemia in the EPIC-Potsdam Study. *Eur. J. Clin. Nutr.* **2014**, *68* (4), 517–525.

(313) Forouhi, N. G.; Koulman, A.; Sharp, S. J.; Imamura, F.; Kröger, J.; Schulze, M. B.; Crowe, F. L.; Huerta, J. M.; Guevara, M.; Beulens, J. W. J.; van Woudenbergh, G. J.; Wang, L.; Summerhill, K.; Griffin, J. L.; Feskens, E. J. M.; Amiano, P.; Boeing, H.; Clavel-Chapelon, F.; Dartois, L.; Fagherazzi, G.; Franks, P. W.; Gonzalez, C.; Jakobsen, M. U.; Kaaks, R.; Key, T. J.; Khaw, K.-T. T.; Kühn, T.; Mattiello, A.; Nilsson, P. M.; Overvad, K.; Pala, V.; Palli, D.; Quirós, J. R.; Rolandsson, O.; Roswall, N.; Sacerdote, C.; Sánchez, M.-J. J.; Slimani, N.; Spijkerman, A. M. W.; Tjonneland, A.; Tormo, M.-J. J.; Tumino, R.; van der A, D. L.; van der Schouw, Y. T.; Langenberg, C.; Riboli, E.; Wareham, N. J. Differences in the Prospective Association between Individual Plasma Phospholipid Saturated Fatty Acids and Incident Type 2 Diabetes: The EPIC-InterAct Case-Cohort Study. *Lancet Diabetes Endocrinol.* **2014**, *2* (10), 810–818.

(314) Jenkins, B. J.; Seyssel, K.; Chiu, S.; Pan, P.-H. H.; Lin, S.-Y. Y.; Stanley, E.; Ament, Z.; West, J. A.; Summerhill, K.; Griffin, J. L.; Vetter, W.; Autio, K. J.; Hiltunen, K.; Hazebrouck, S.; Stepankova, R.; Chen, C.-J. J.; Alligier, M.; Laville, M.; Moore, M.; Kraft, G.; Cherrington, A.; King, S.; Krauss, R. M.; de Schryver, E.; Van Veldhoven, P. P.; Ronis, M.; Koulman, A. Odd Chain Fatty Acids; New Insights of the Relationship Between the Gut Microbiota, Dietary Intake, Biosynthesis and Glucose Intolerance. *Sci. Rep.* **2017**, *7* (1), 44845.

(315) Sbaï, D.; Narcy, C.; Thompson, G. N.; Mariotti, A.; Poggi, F.; Saudubray, J. M.; Bresson, J. L. Contribution of Odd-Chain Fatty Acid Oxidation to Propionate Production in Disorders of Propionate Metabolism. *Am. J. Clin. Nutr.* **1994**, *59* (6), 1332–1337.

(316) Mueller, J. W.; Gilligan, L. C.; Idkowiak, J.; Arlt, W.; Foster, P. A. The Regulation of Steroid Action by Sulfation and Desulfation. *Endocr. Rev.* **2015**, *36* (5), 526–563.

(317) Esteves, C. Z.; de Aguiar Dias, L.; de Oliveira Lima, E.; de Oliveira, D. N.; Rodrigues Melo, C. F. O.; Delafiori, J.; Souza Gomez, C. C.; Ribeiro, J. D.; Ribeiro, A. F.; Levy, C. E.; Catharino, R. R. Skin Biomarkers for Cystic Fibrosis: A Potential Non-Invasive Approach for Patient Screening. *Front. Pediatr.* **2018**, *5* (19).

(318) Hashimoto, Y.; Okiyoneda, T.; Harada, K.; Ueno, K.; Sugahara, T.; Yamashita, A.; Shuto, T.; Suico, M. A.; Kai, H. Phosphatidic Acid Metabolism Regulates the Intracellular Trafficking and Retrotranslocation of CFTR. *Biochim. Biophys. Acta - Mol. Cell Res.* **2008**, *1783* (1), 153–162.

(319) Ziady, A. G.; Hansen, J. Redox Balance in Cystic Fibrosis. *Int. J. Biochem. Cell Biol.* **2014**, *52* (2014), 113–123.

(320) Schunemann, H. J.; Muti, P.; Freudenheim, J. L.; Armstrong, D.; Browne, R.; Klocke, R. A.; Trevisan, M. Oxidative Stress and Lung Function. *Am. J. Epidemiol.* **1997**, *146* (11), 939–948.

(321) Jonasdottir, H. S.; Brouwers, H.; Toes, R. E. M.; Ioan-Facsinay, A.; Giera, M. Effects of Anticoagulants and Storage Conditions on Clinical Oxylipid Levels in Human Plasma. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2018**, *1863* (12), 1511–1522.

(322) Magnusardottir, A. R.; Skuladottir, G. V. Effects of Storage Time and Added Antioxidant

on Fatty Acid Composition of Red Blood Cells at −20°C. *Lipids* **2006**, *41* (4), 401–404.

Appendix A

Appendix A: Supplementary Figures for Chapter 2

LIST OF FIGURES

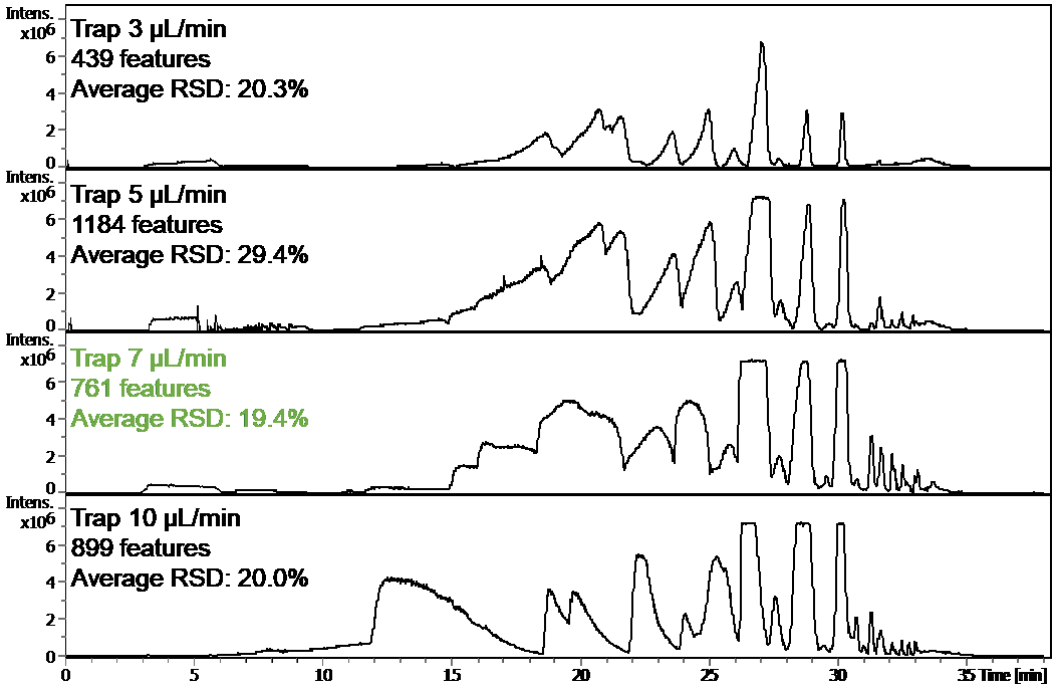

Figure A - 1. Base peak chromatograms (BPC) obtained for the initial optimization of the trapping flowrate using a serum sample from a pig. The parameters chosen as optimized are displayed in green. Average RSD: average relative standard deviation of peak intensities for injection replicates for all detected features, taken a measure of reproducibility. Initial nanoLC method: MPA – 20 mM NH_4COOH, 5 mM formic acid in 45:45:10 methanol/ acetonitrile/ water; MPB - 20 mM NH_4COOH, 5 mM formic acid in 2-propanol; trapping for 1 min at 7 µL/min (95% MPA); 300 nL/min; 45ºC, 38 min gradient (0 min – 5% MPB, 2 min – 5% MPB, 12 min – 30% MPB, 24 min – 90% MPB, 28 min – 1% MPB, 38 min – 1% MPB); 20 min equilibrium (95% MPA); 2 µL injection. Initial MS conditions: electrospray ion source capillary voltage of 1300 V, dry gas flow rate of 3 L/min, source temperature of 200ºC, spectra acquisition rate of 2 Hz and nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar. The initial sample preparation method is described in the Experimental Section.

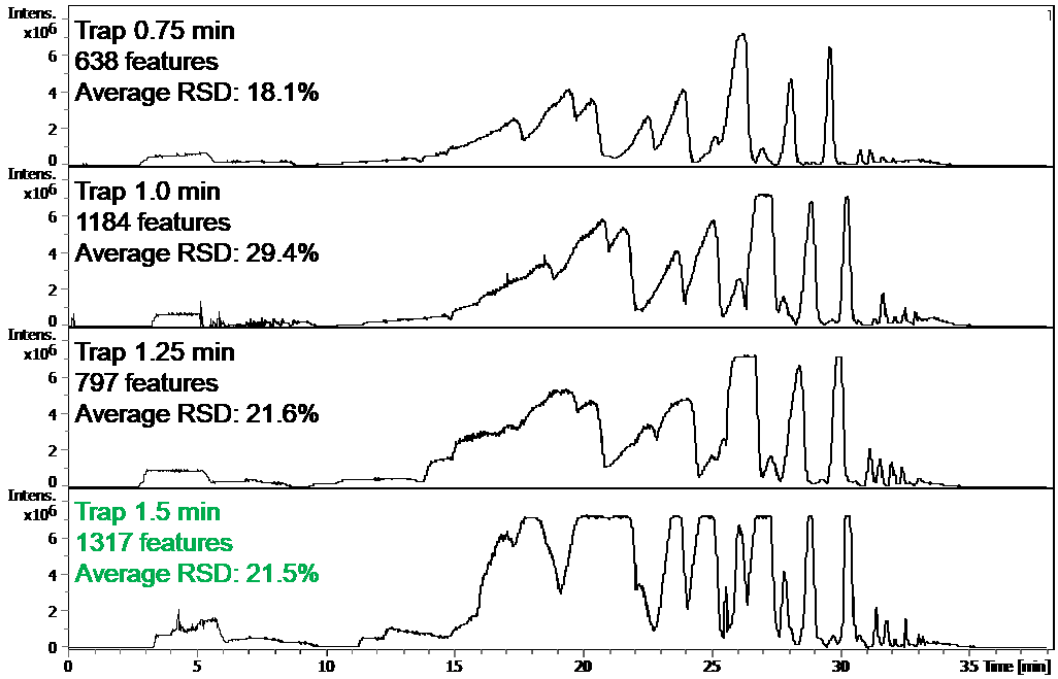

Figure A - 2. BPC obtained for the initial optimization of trapping time using a serum sample from a pig. The nanoLC and MS methods are described in Figure A - 1. The initial sample preparation method is described in the Experimental Section.

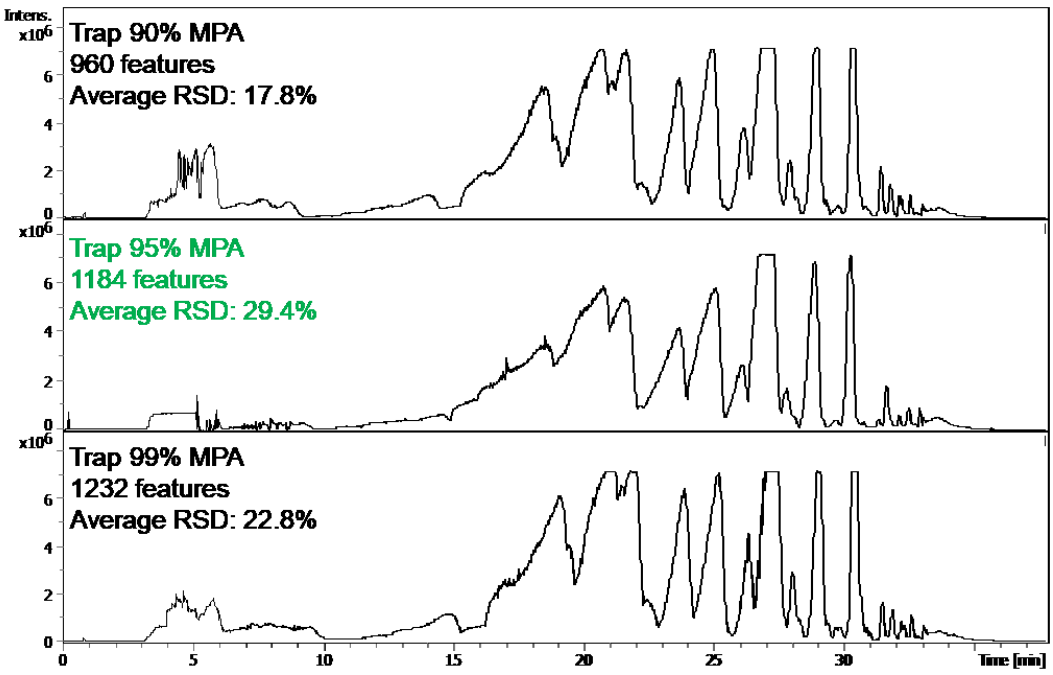

Figure A - 3. BPC obtained for the initial optimization of trapping mobile phase composition using a serum sample from a pig. The nanoLC and MS methods are described in Figure A - 1. The initial sample preparation method is described in the Experimental Section.

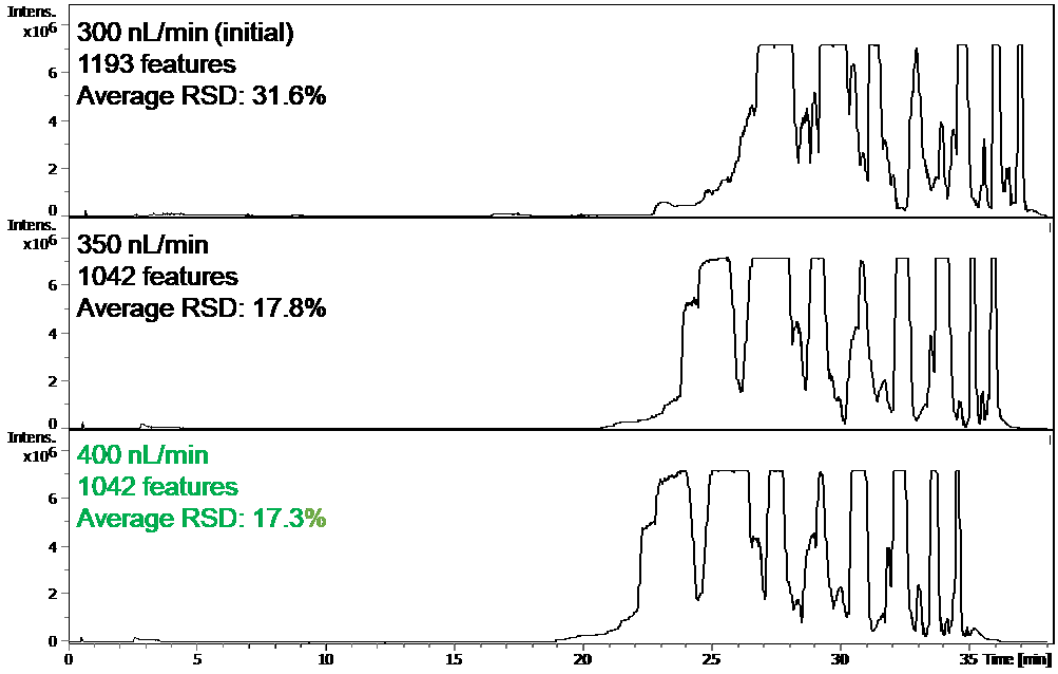

Figure A - 4. Initial optimization of the analytical flowrate. NanoLC method: trapping for 1.50 min at 7 µL/min (95% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water; MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water; 45ºC, 38 min gradient (0 min – 5% MPB, 2 min – 5% MPB, 12 min – 30% MPB, 24 min – 90% MPB, 28 min – 1% MPB, 38 min – 1% MPB); 20 min equilibrium run (95% MPA); 2 µL injection. MS conditions are identical to Figure A - 1.

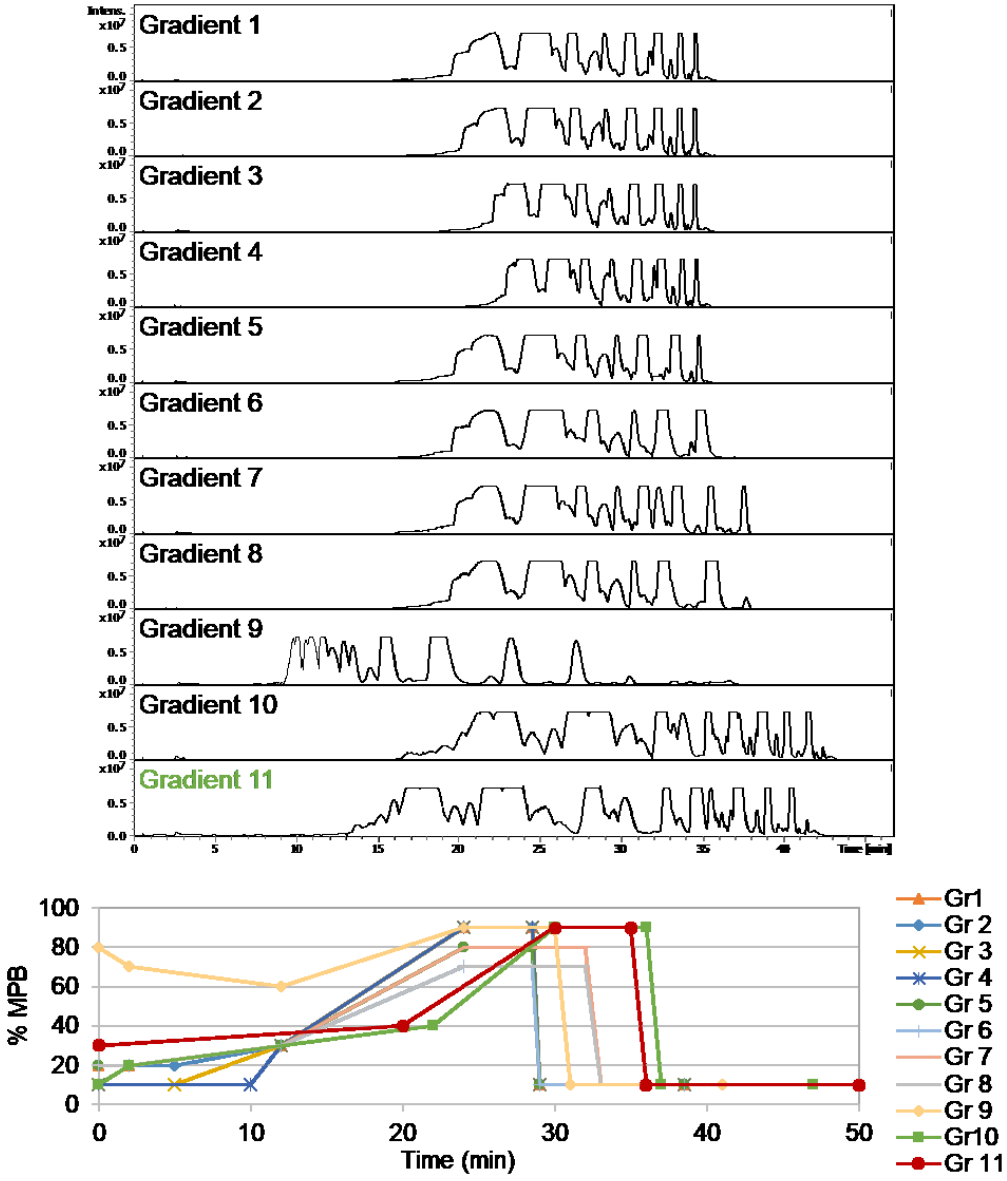

Figure A - 5. Evaluation of the nanoLC gradient (1 to 11) for the first phase of optimization. Gradient 11 was selected as the best option. NanoLC method: trapping for 1.50 min at 7 µL/min (95% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v/); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 400 nL/min; 45ºC; 20 min equilibrium (95% MPA); 2 µL injection. MS conditions are identical to Figure A - 1.

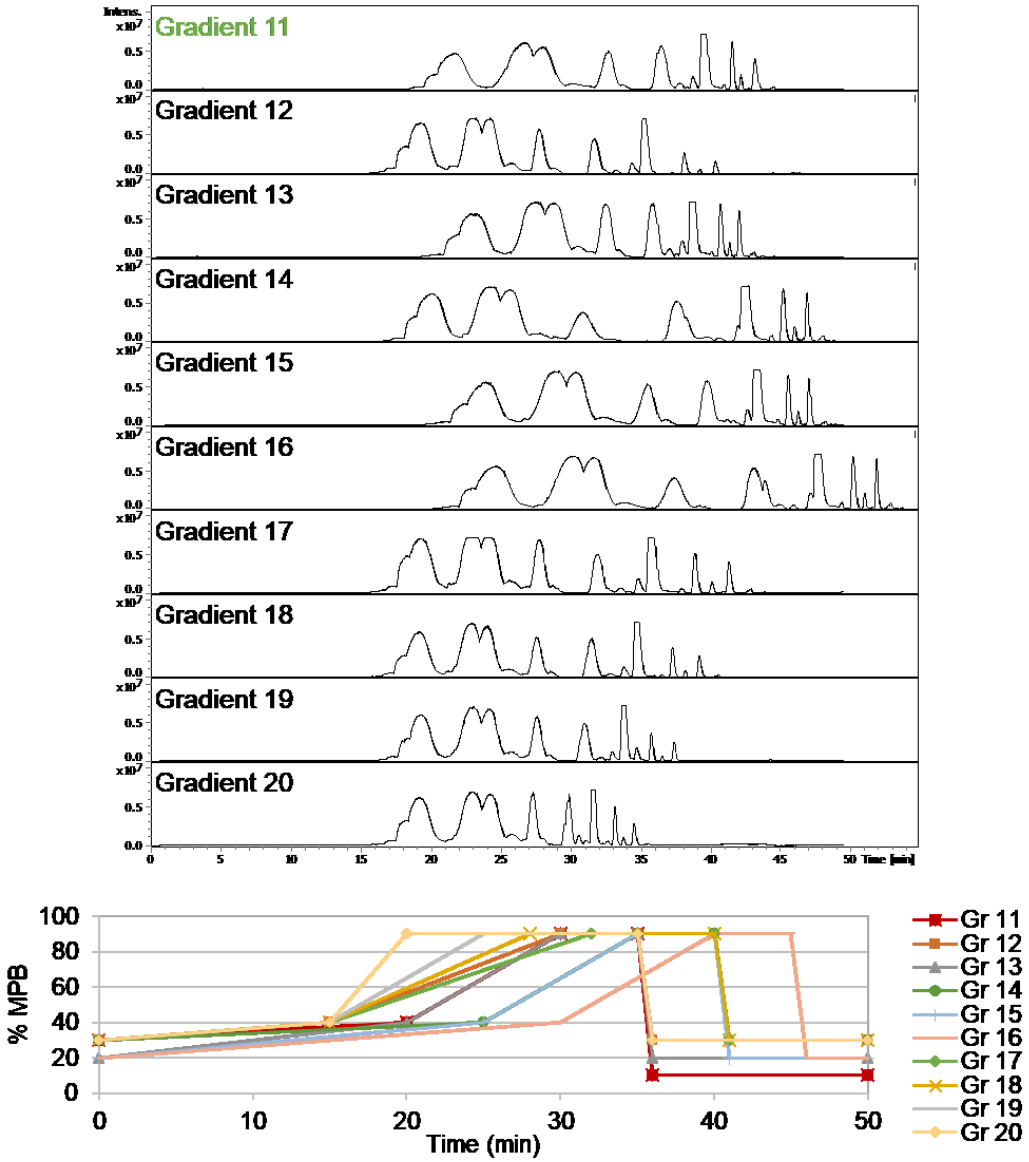

Figure A - 6. Evaluation of the nanoLC gradient (11 to 20) for the first phase of optimization. Gradient 11 was selected as the best option. The nanoLC method is described in Figure A - 7. MS conditions are identical to Figure A - 1.

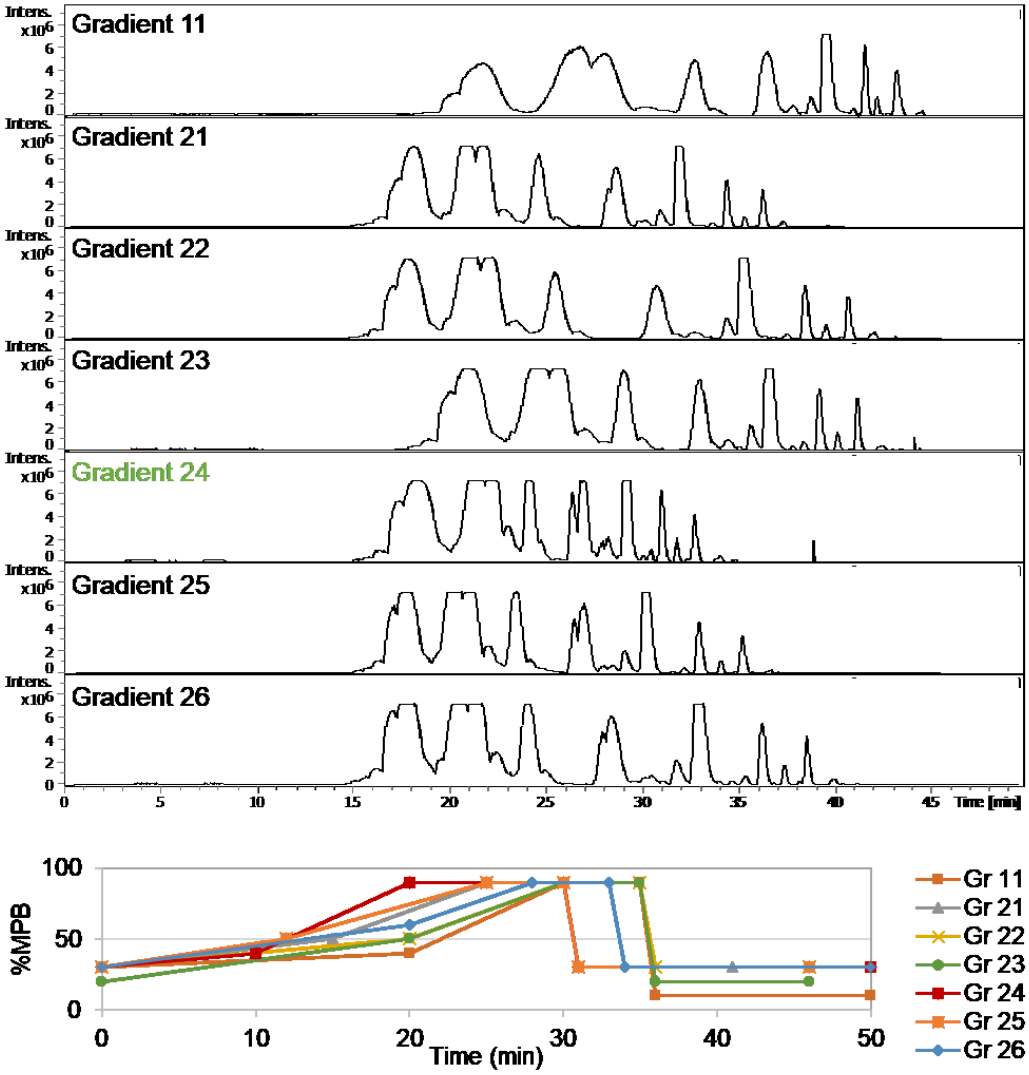

Figure A - 7. Evaluation of the nanoLC gradient (21 to 26) for the first phase of optimization. Gradient 24 was selected as the best option. The nanoLC method is described in Figure A - 5. MS conditions are identical to Figure A - 1.

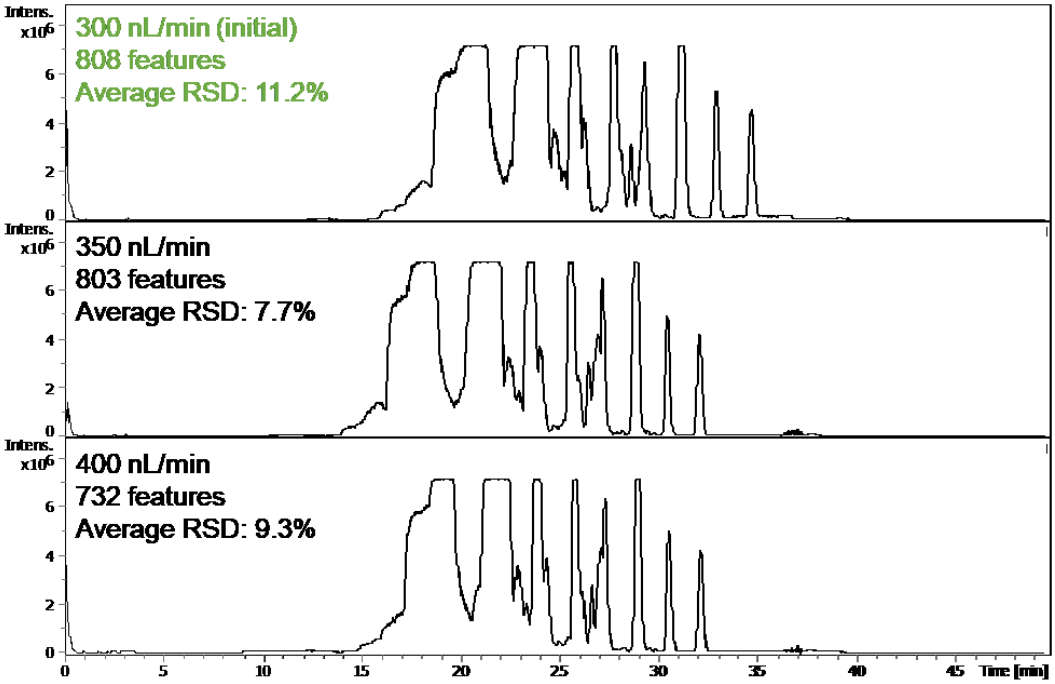

Figure A - 8. Re-evaluation of the flowrate for the optimized gradient (gradient 24, Figure A - 7). NanoLC method: trapping for 1.50 min at 7 µL/min (95% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC; 50 min gradient (0 min – 30% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 30 min – 90% MPB, 31 min – 30% MPB, 50 min – 30% MPB), 20 min equilibrium (95% MPA); 2 µL injection. MS conditions are identical to Figure A - 1.

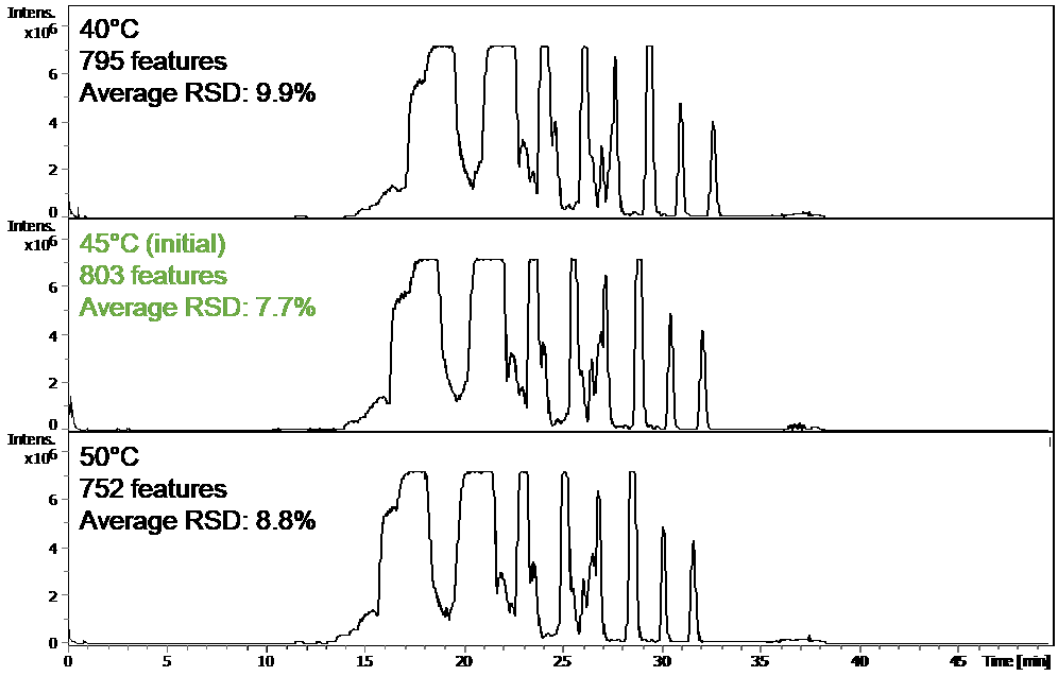

Figure A - 9. Optimization of the column temperature. NanoLC method: trapping for 1.50 min at 7 µL/min (95% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 400 nL/min; 50 min gradient (0 min – 30% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 30 min – 90% MPB, 31 min – 30% MPB, 50 min – 30% MPB), 20 min equilibrium (95% MPA); 2 µL injection. MS conditions are identical to Figure A - 1.

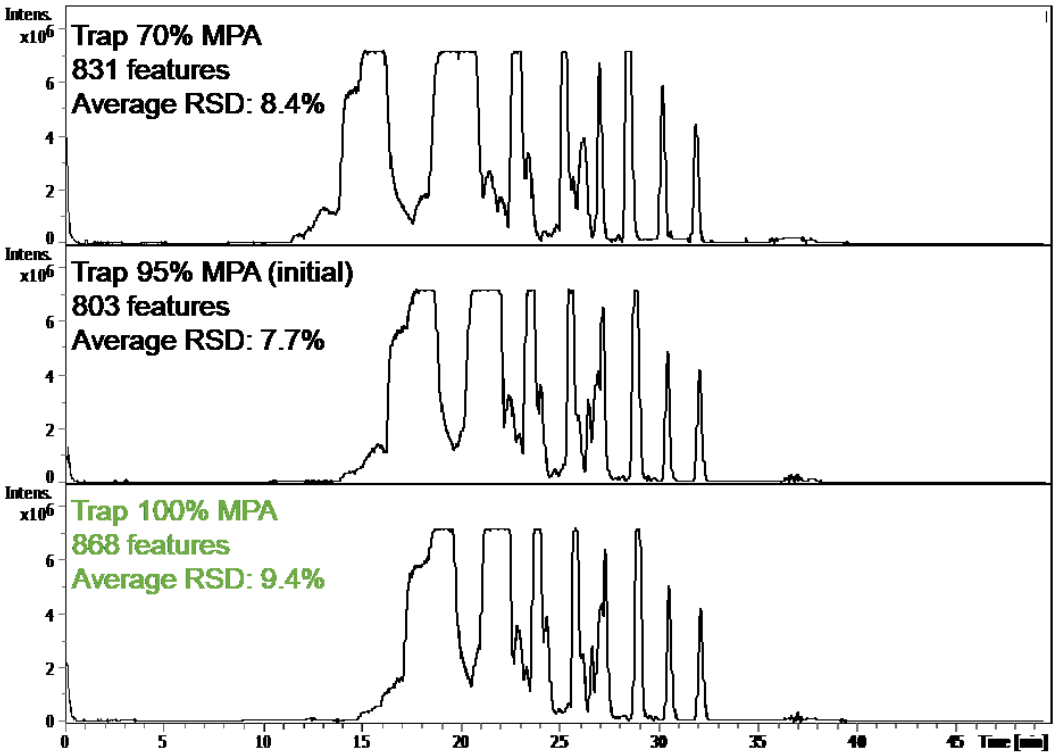

Figure A - 10. Optimization of the trapping mobile phase composition. NanoLC method: MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 50 min gradient (0 min – 30% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 30 min – 90% MPB, 31 min – 30% MPB, 50 min – 30% MPB), 20 min equilibrium (95% MPA); 2 µL injection. MS conditions are identical to Figure A - 1.

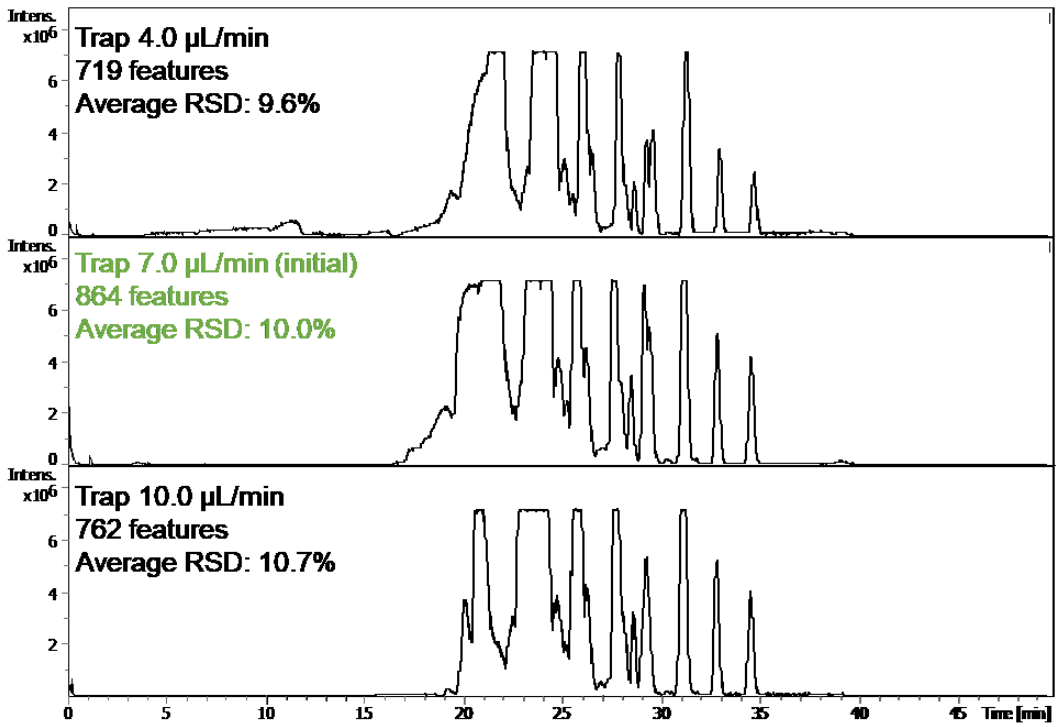

Figure A - 11. Optimization of the trapping flowrate. The nanoLC method is described in Figure A - 10. MS conditions are identical to Figure A - 1.

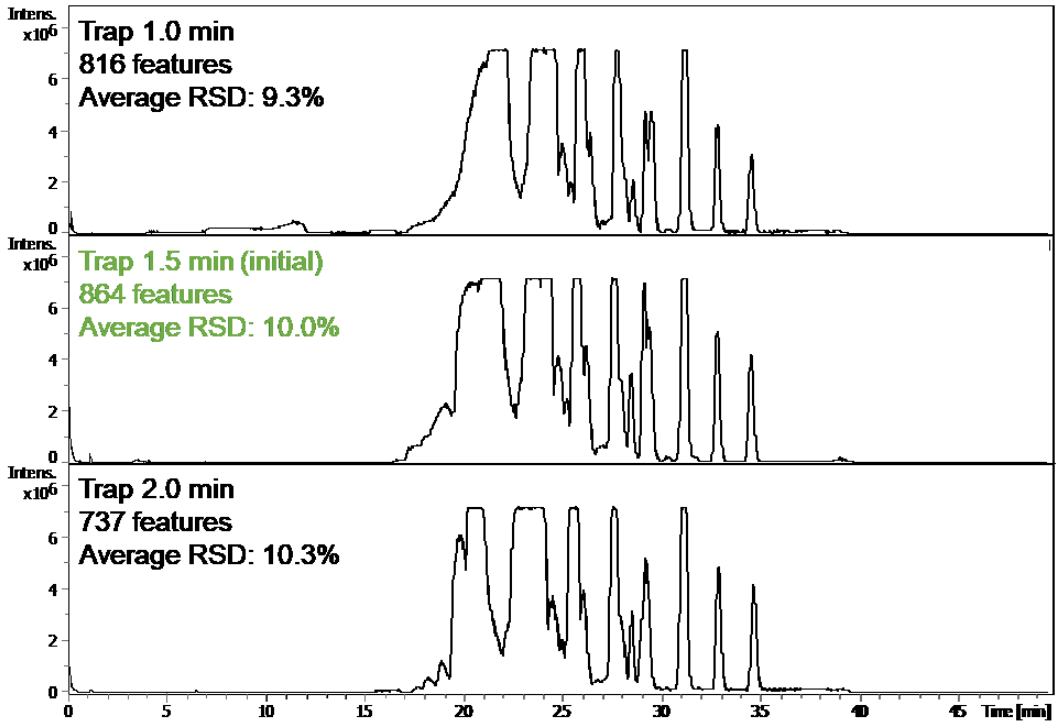

Figure A - 12. Optimization of the trapping time. The nanoLC method is described in Figure A - 10. MS conditions are identical to Figure A - 1.

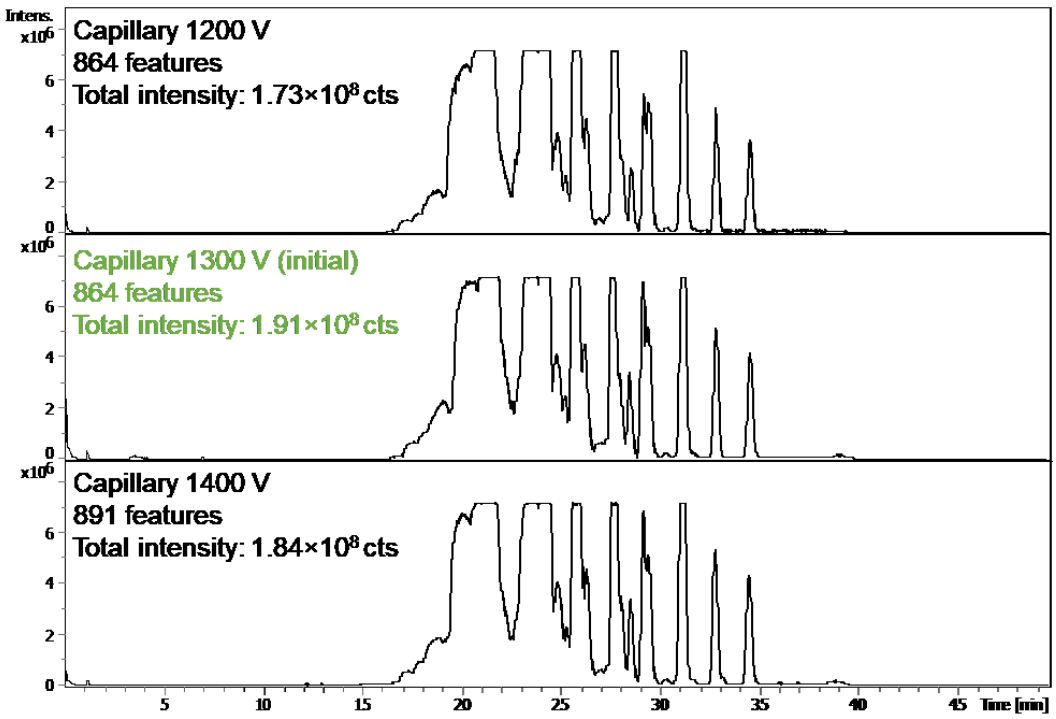

Figure A - 13. Optimization of MS capillary voltage. Total intensity: summed intensity of all detected features for each injection. NanoLC method: trapping for 1.5 min at 7 µL/min (100% MPA), MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 50 min gradient (0 min – 30% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 30 min – 90% MPB, 31 min – 30% MPB, 50 min – 30% MPB), 20 min equilibrium (95% MPA); 2 µL injection. MS conditions: dry gas flowrate of 3 L/min, source temperature of 200ºC, spectra acquisition rate of 2 Hz and nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar.

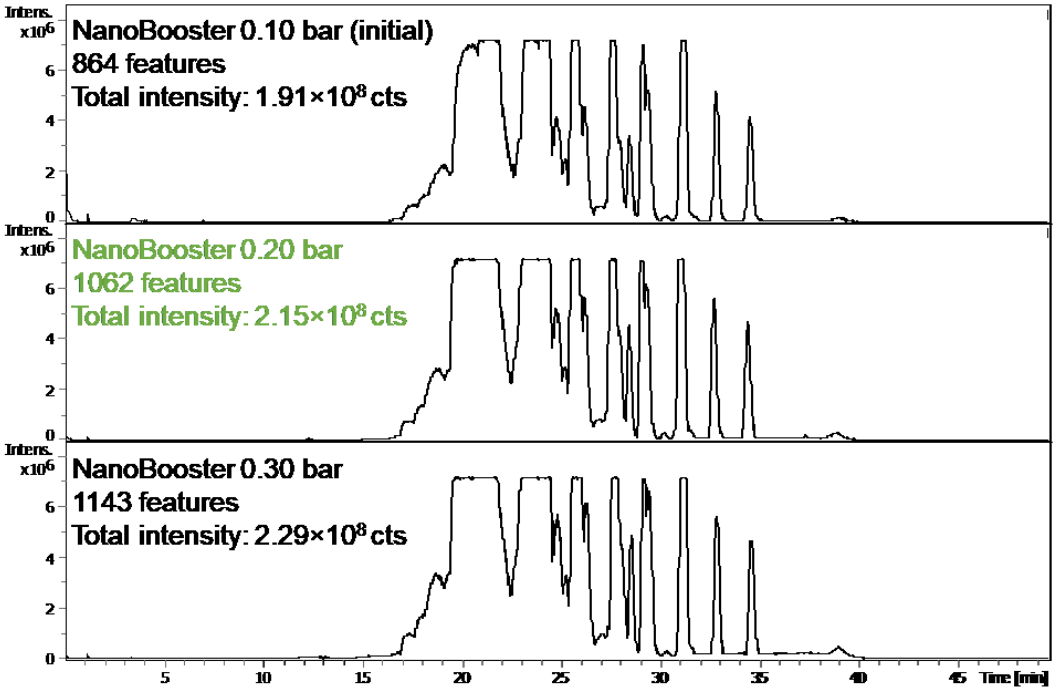

Figure A - 14. Optimization of nanoBooster acetonitrile-enriched nitrogen gas pressure. Although 0.30 bar provided more detected features, the consumption of acetonitrile to enrich the nanoBooster nitrogen gas was unreasonable (over 1 L for each day of analysis). NanoLC method: identical to Figure A - 13. MS conditions: electrospray ion source capillary voltage of 1300 V, dry gas flow rate of 3 L/min, source temperature of 200ºC and spectra acquisition rate of 2 Hz.

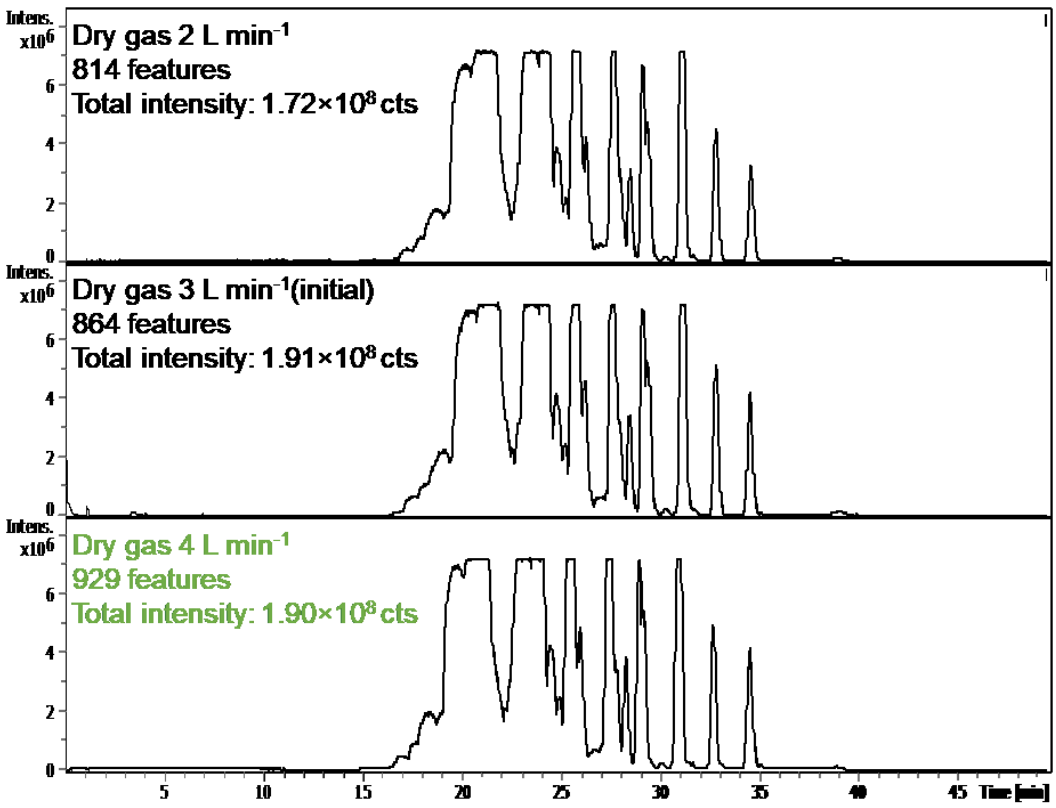

Figure A - 15. Optimization of ion source dry nitrogen gas flowrate. NanoLC method: identical to Figure A - 13. MS conditions: electrospray ion source capillary voltage of 1300 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar, source temperature of 200ºC and spectra acquisition rate of 2 Hz.

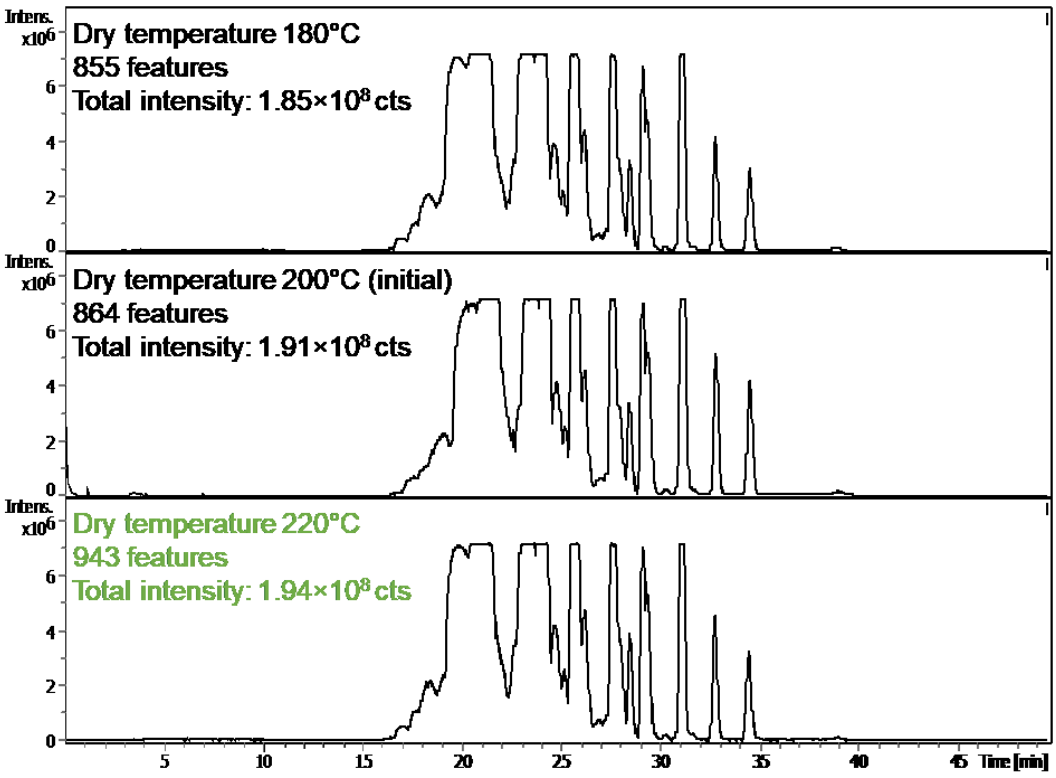

Figure A - 16. Optimization of ion source temperature. NanoLC method: identical to Figure A - 13. MS conditions: electrospray ion source capillary voltage of 1300 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar, dry nitrogen gas flow rate of 3 L/min and spectra acquisition rate of 2 Hz.

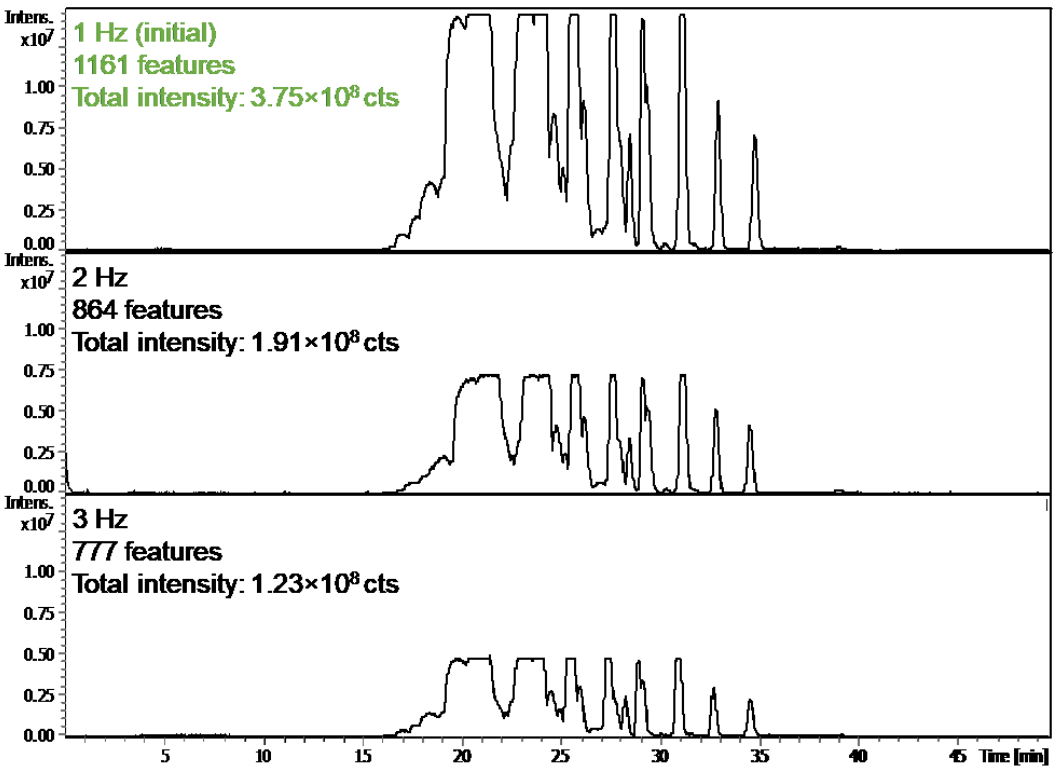

Figure A - 17. Optimization of the spectra acquisition rate. NanoLC method: identical to Figure A - 13. MS conditions: electrospray ion source capillary voltage of 1300 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar, dry nitrogen gas flow rate of 3 L/min and ion source temperature of 200ºC.

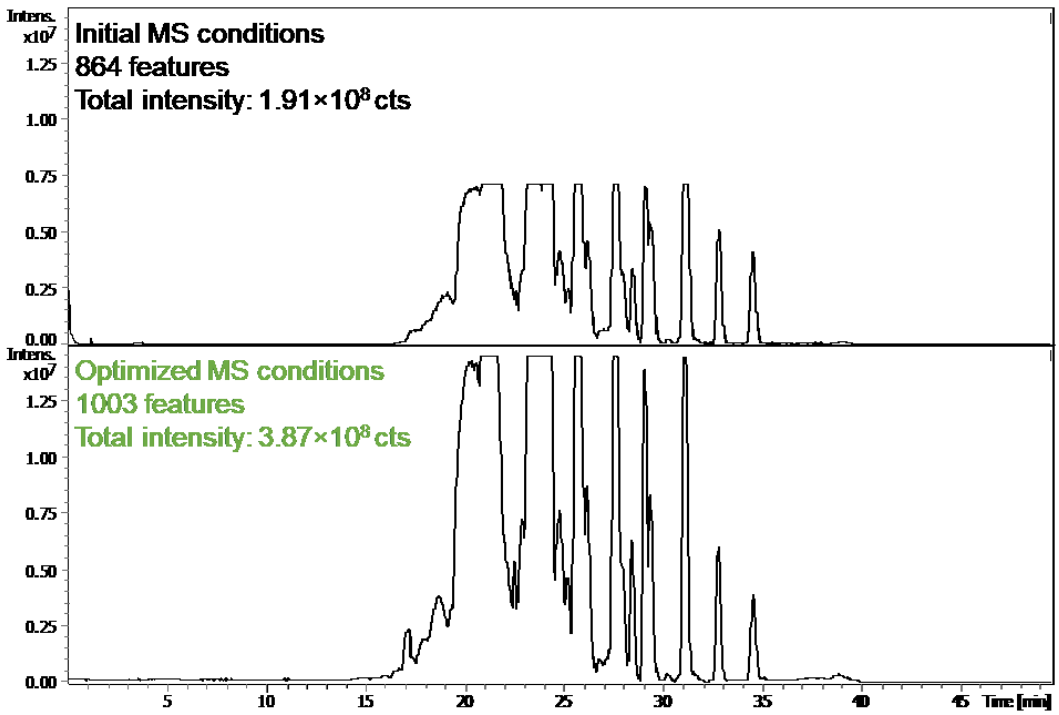

Figure A - 18. Initial versus optimized mass spectrometry methods. NanoLC method: identical to Figure A - 13. Initial MS conditions: electrospray ion source capillary voltage of 1300 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.10 bar, dry nitrogen gas flow rate of 3 L/min, ion source temperature of 200ºC and spectra acquisition rate of 2 Hz. Optimized MS conditions: electrospray ion source capillary voltage of 1300 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.20 bar, dry nitrogen gas flow rate of 4 L/min, ion source temperature of 220ºC and spectra acquisition rate of 1 Hz.

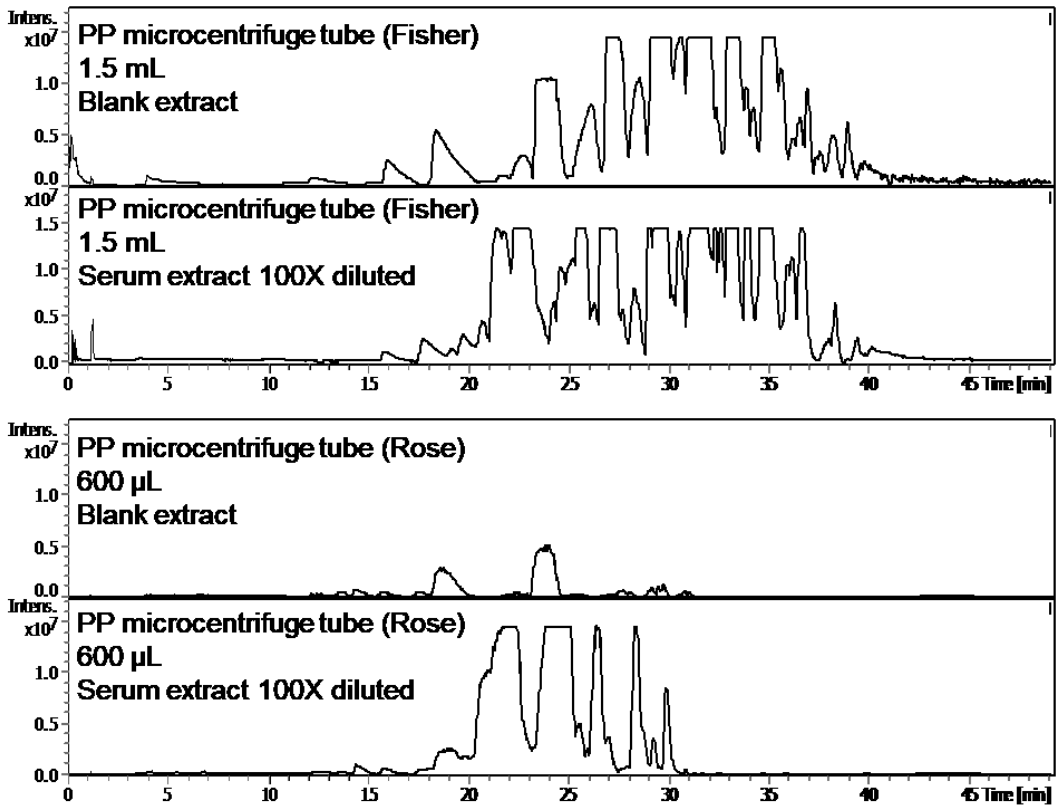

Figure A - 19. Comparison between pig serum extracts and blank extracts (water instead of the sample) when two different brands of polypropylene (PP) microcentrifuge tubes were employed for extraction (1.5 mL tubes from Fisher Scientific and 600 µL tubes from Rose Scientific acquired between 2016 and 2017).

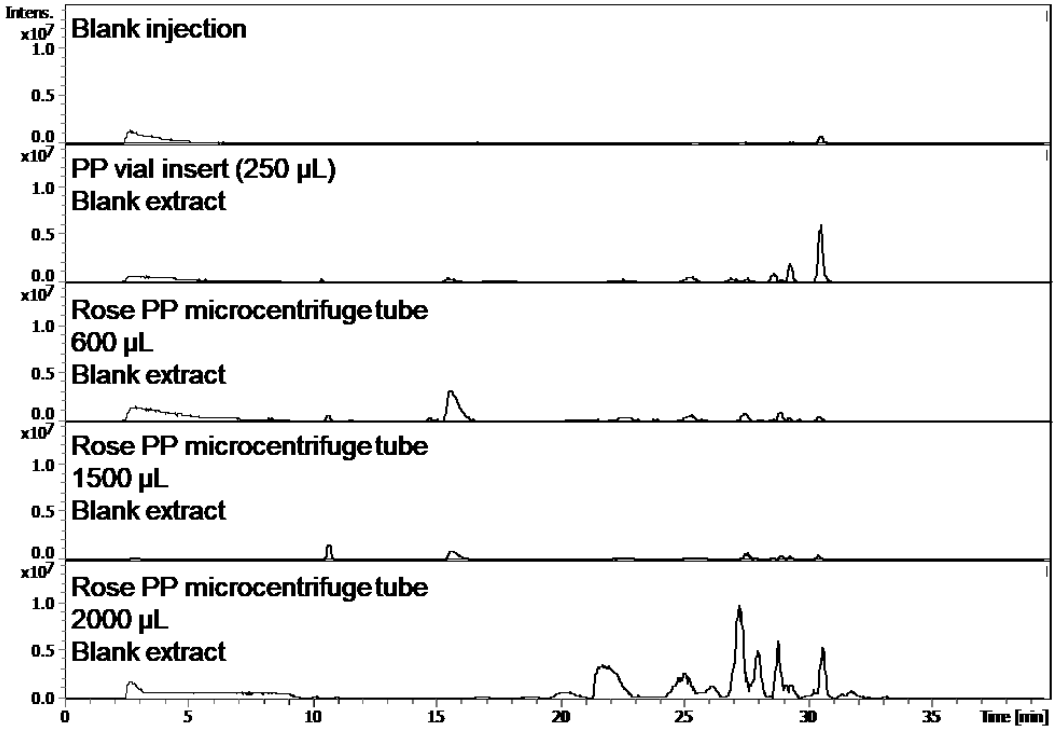

Figure A - 20. Chromatograms obtained for blank extracts (water instead of the sample) when extractions were performed in a PP autosampler insert (Canadian Life Science) and different sizes of PP microcentrifuge tubes (Rose Scientific, acquired between 2016 and 2017), compared to an injection blank (pure mobile phase).

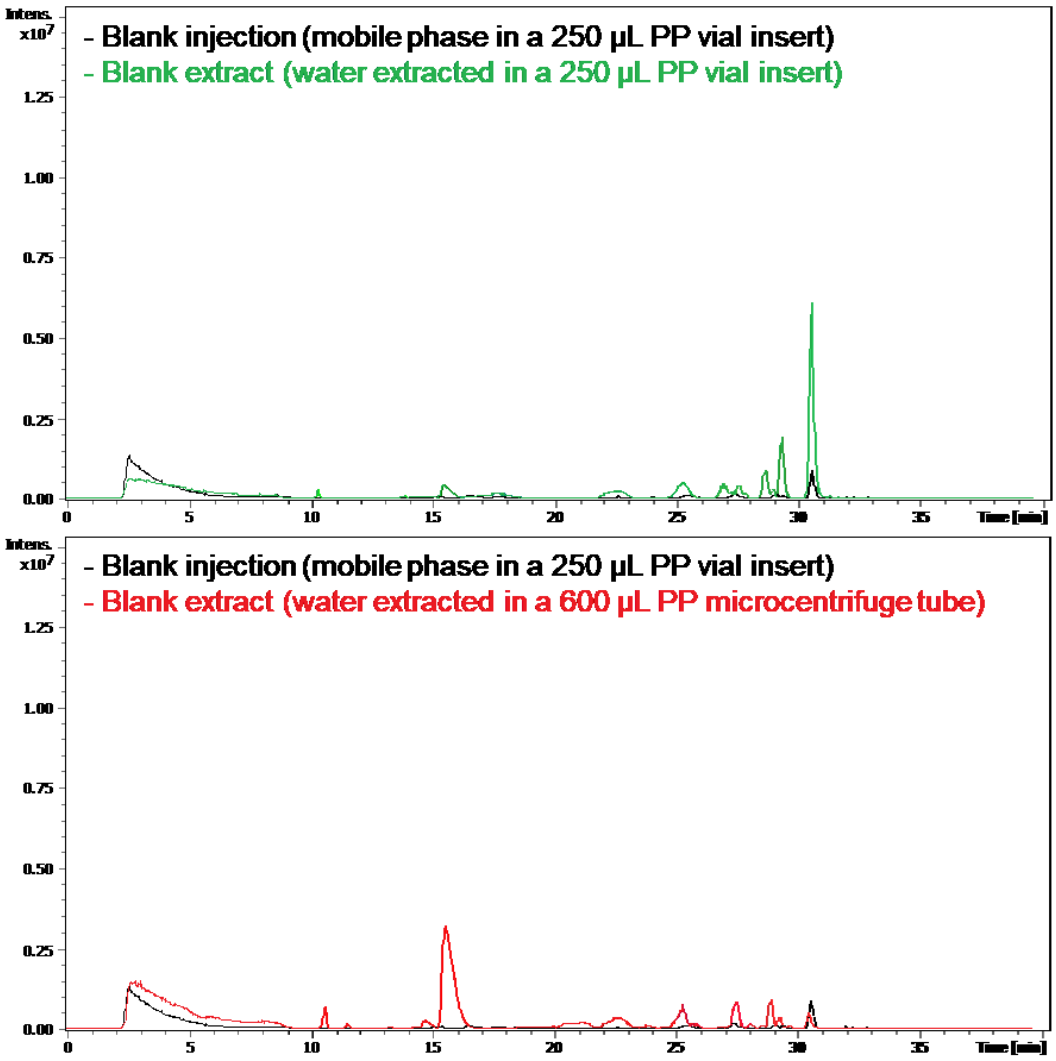

Figure A - 21. Comparison between blank injections (injection of pure mobile phase, black) with blank extracts (extraction of water instead of the sample, green and red) for extractions performed inside a PP autosampler insert (Canadian Life Science, green) and a 600 µL PP microcentrifuge tube (Rose Scientific, red). All samples were transferred to PP autosampler vial inserts for injection.

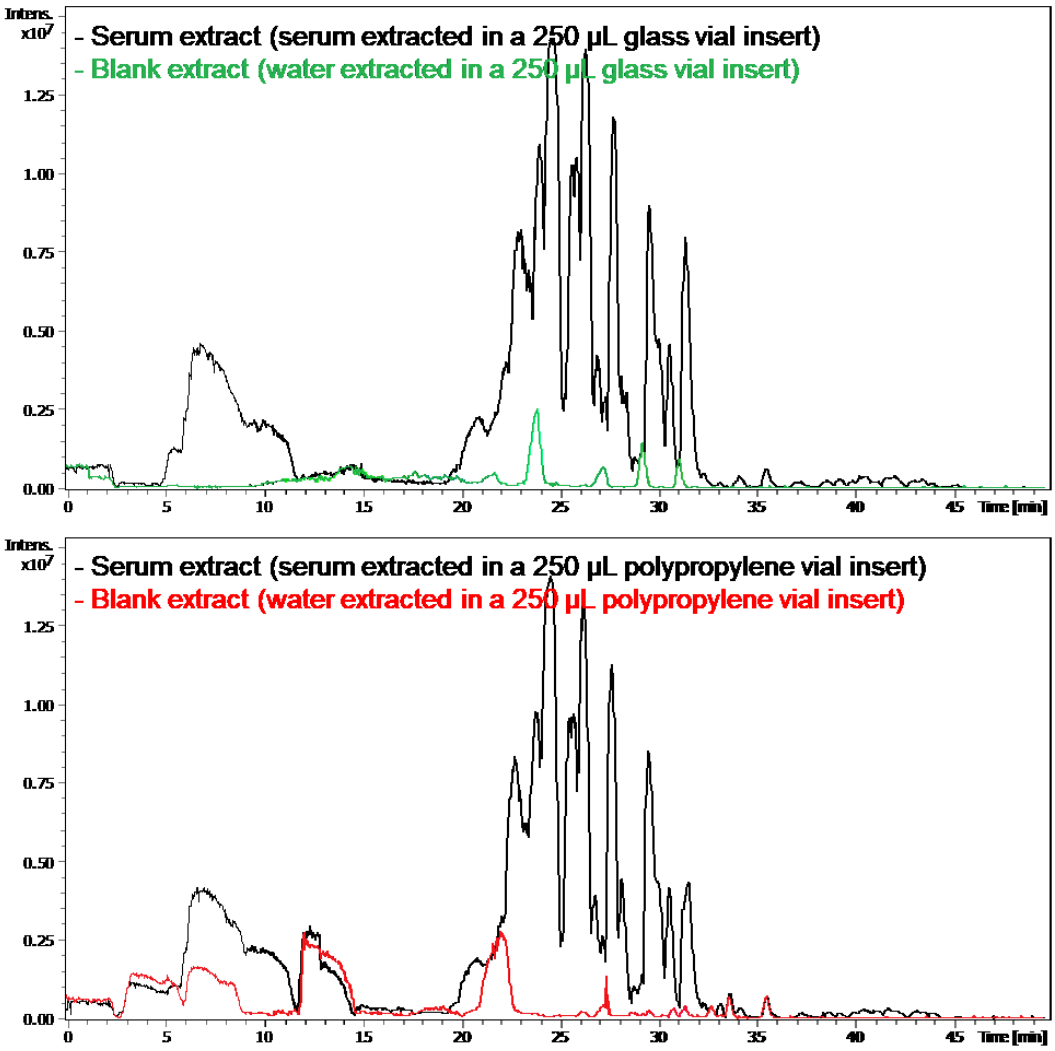

Figure A - 22. Comparison between serum (black) and blank extracts (extraction of water instead of the sample) for extractions performed inside a glass (Agilent Technologies, green) and a PP autosampler insert (Canadian Life Science, red).

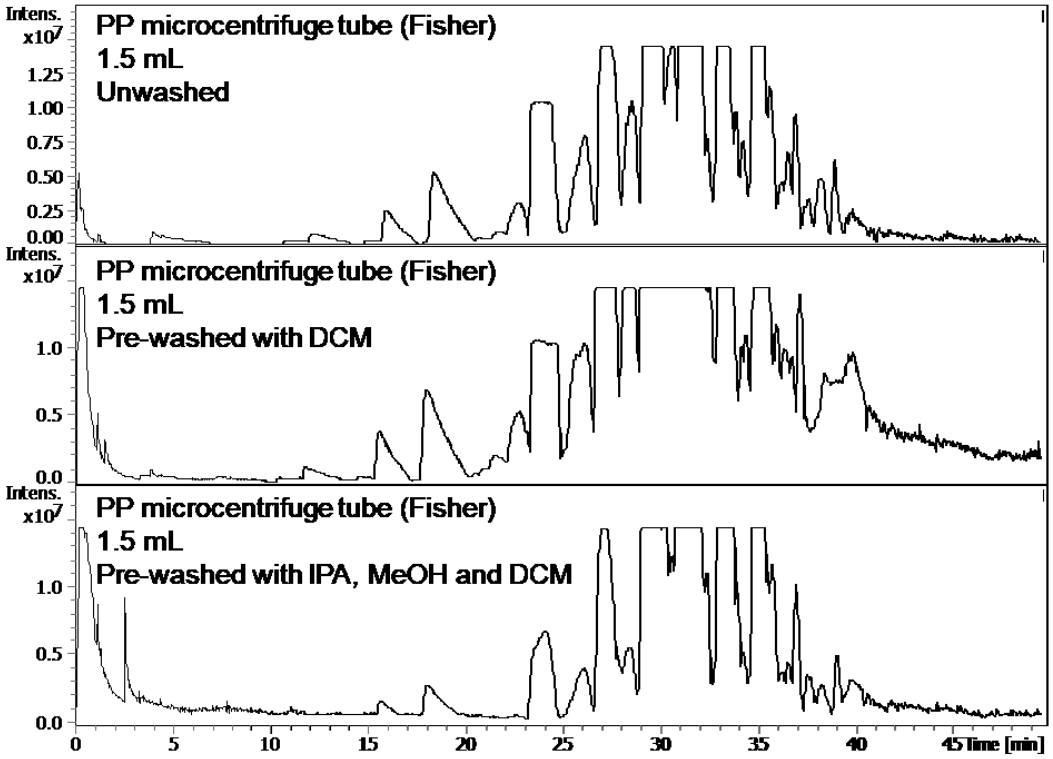

Figure A - 23. Chromatograms obtained for blank extracts (extraction of water instead of the sample) for extractions performed in 1.5 mL PP microcentrifuge tubes (Fisher Scientific, acquired between 2016 and 2017) unwashed, washed with dichloromethane (DCM), and sequentially washed with 2-propanol (IPA), methanol (MeOH) and dichloromethane (DCM). Each tube was rinsed with 1 mL of each solvent and air-dried at room temperature in a fume hood before extraction. The chromatograms obtained for the pure solvents are displayed in Figure A - 25.

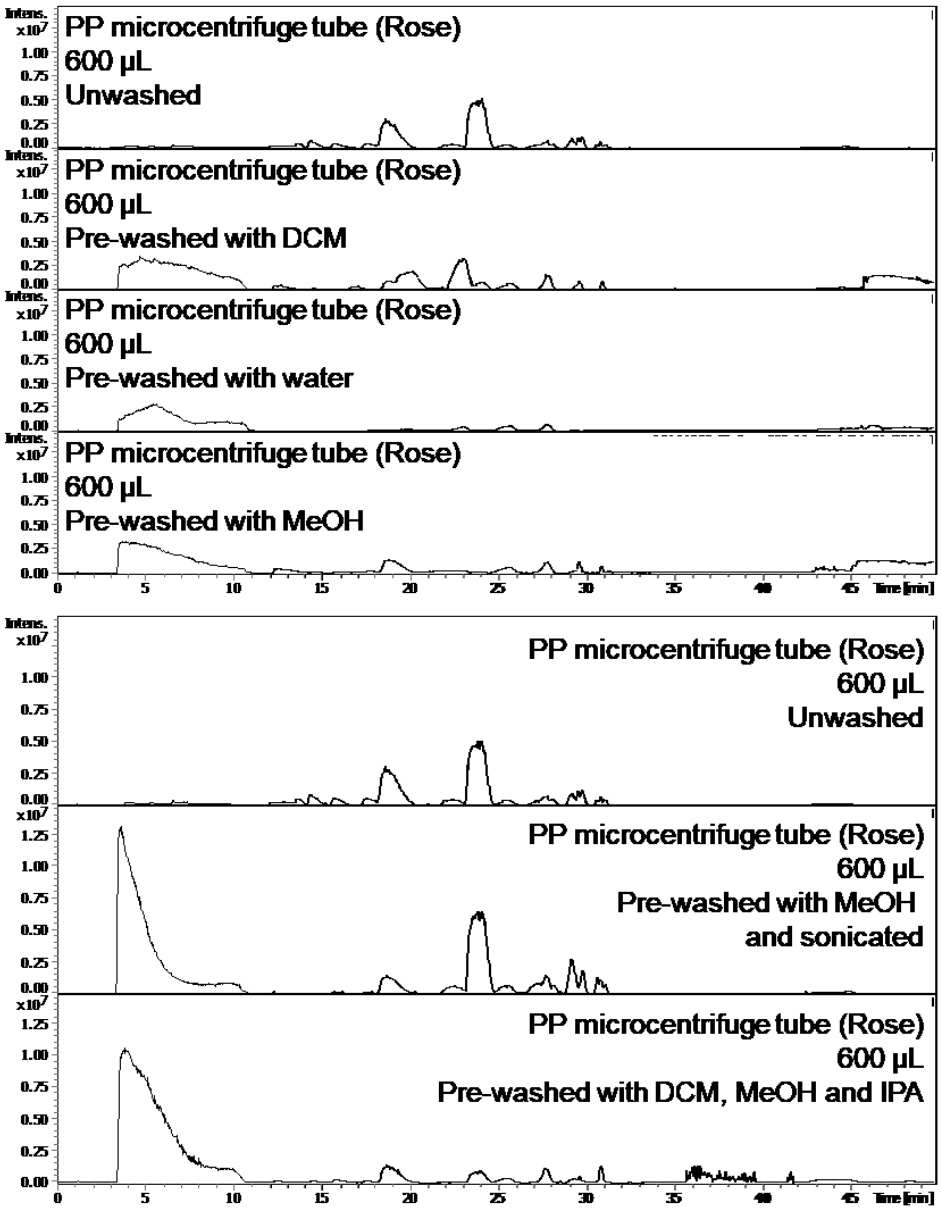

Figure A - 24. Chromatograms obtained for blank extracts (extraction of water instead of the sample) for extractions performed in 600 µL PP microcentrifuge tubes (Rose Scientific, acquired between 2016 and 2017) unwashed, washed with dichloromethane (DCM), washed with water, washed with methanol (MeOH), washed with methanol and sonicated, and sequentially washed with dichloromethane (DCM), methanol and 2-propanol (IPA). Each tube was rinsed with 1 mL of each solvent and air-dried at room temperature in a fume hood before extraction. The chromatograms obtained for the pure solvents are displayed in Figure A - 25.

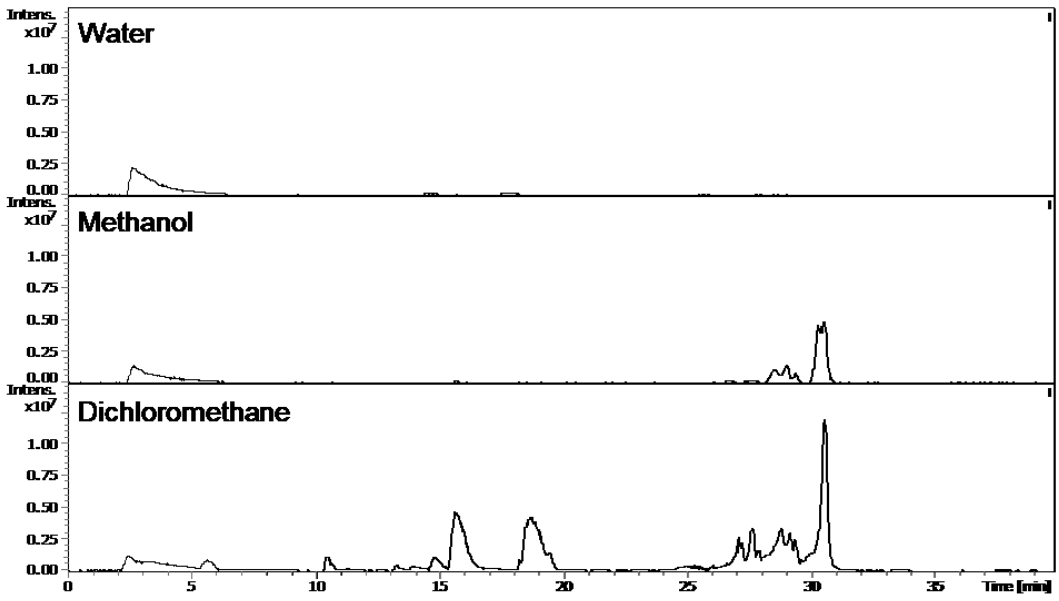

Figure A - 25. Chromatograms obtained for pure LC-MS grade water, methanol and dichloromethane.

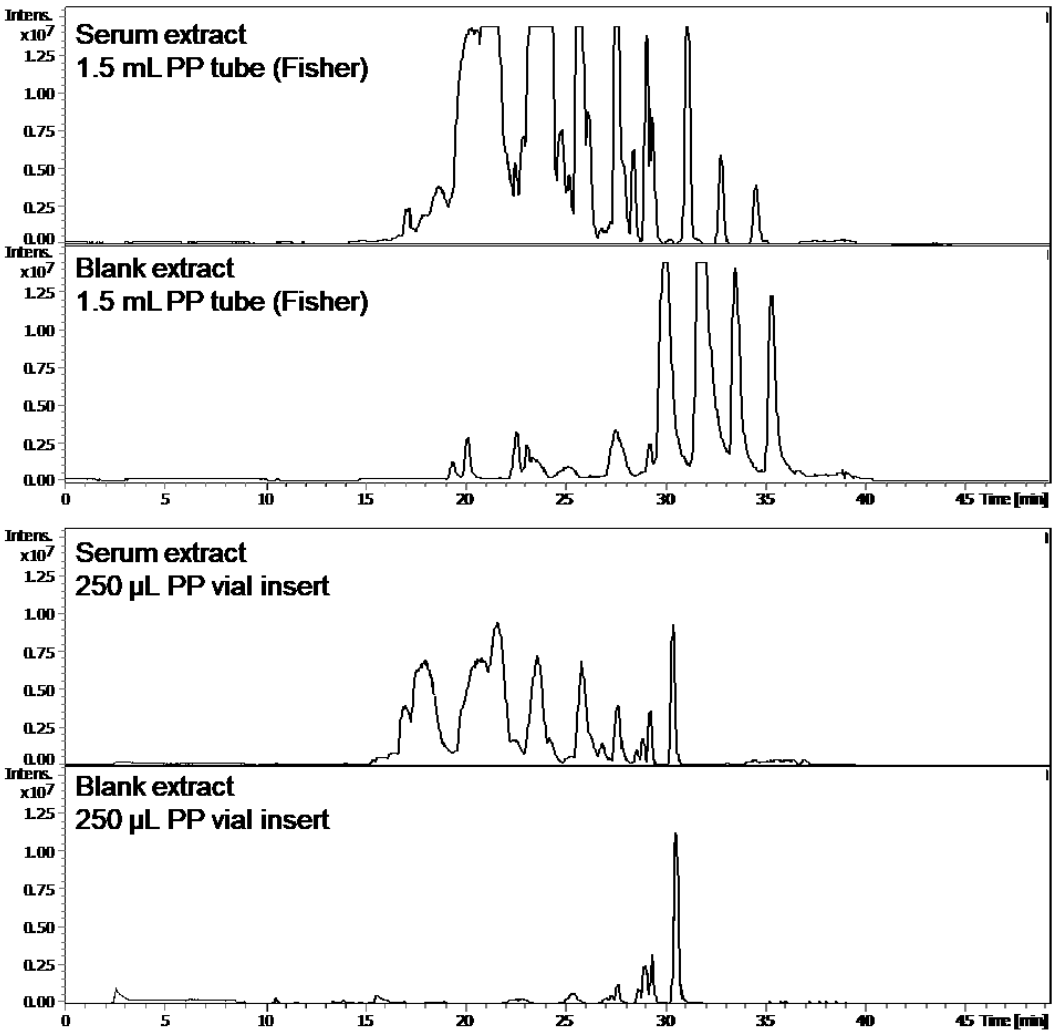

Figure A - 26. Comparison between pig serum extracts and blank extracts (extraction of water instead of the sample) when extractions were performed inside a 1.5 mL polypropylene (PP) microcentrifuge tube (Fisher Scientific, acquired between 2016 and 2017, top) and in a 250 µL PP autosampler vial insert (Canadian Life Science, bottom).

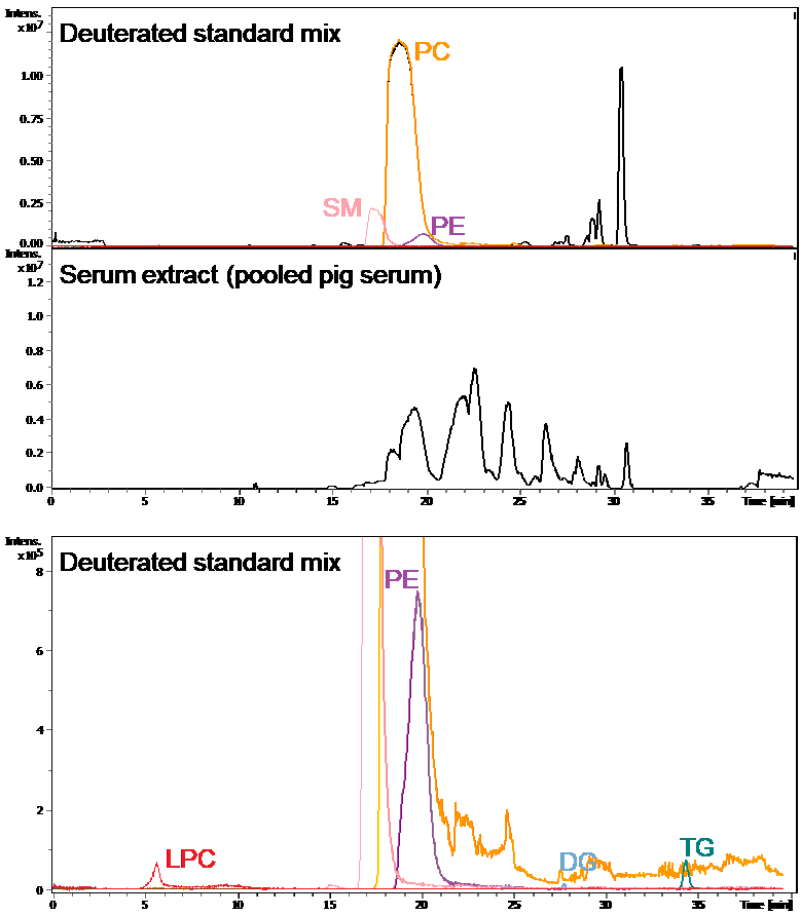

Figure A - 27. Chromatograms obtained for a pool of serum samples from 8 pigs compared to a mixture of 14 deuterated lipid standards, belonging to different lipid subclasses (Table II-1). The top and middle base peak chromatograms (BPC) are shown in full scale, whereas the bottom chromatogram is at a lower intensity range for better visualization of the less intense standards. NanoLC method: trapping for 1.5 min at 7 µL/min (100% MPA), MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 50 min gradient (0 min – 30% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 30 min – 90% MPB, 31 min – 30% MPB, 50 min – 30% MPB), 20 min equilibrium (95% MPA); 2 µL injection. MS conditions: electrospray ion source capillary voltage of 1300 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.20 bar, dry nitrogen gas flow rate of 4 L/min, ion source temperature of 220ºC and spectra acquisition rate of 1 Hz.

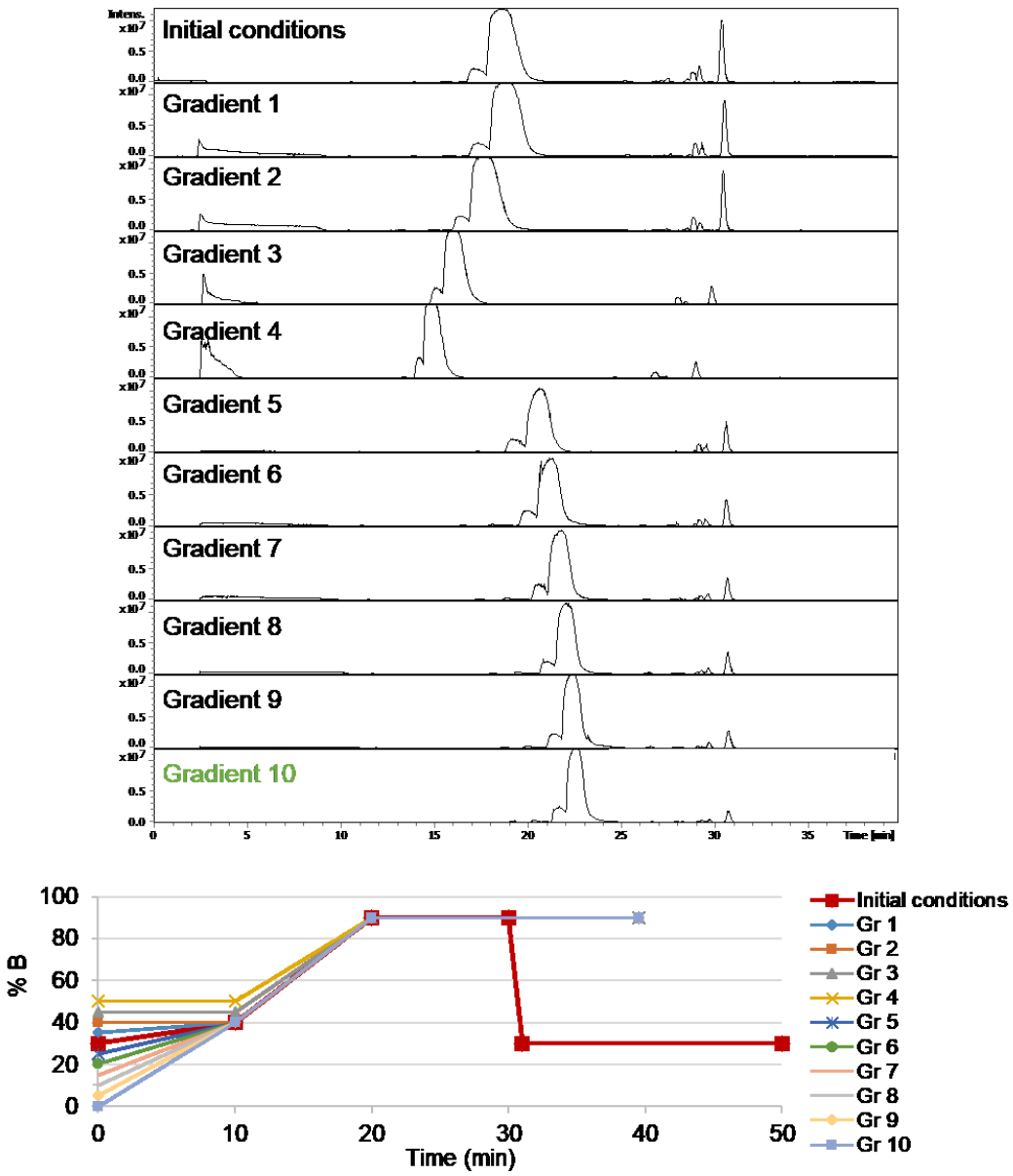

Figure A - 28. Re-evaluation of the separation gradient employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). Gradient 10 was chosen as the best option for this step. The nanoLC-MS method is identical to Figure A - 27, except for gradient.

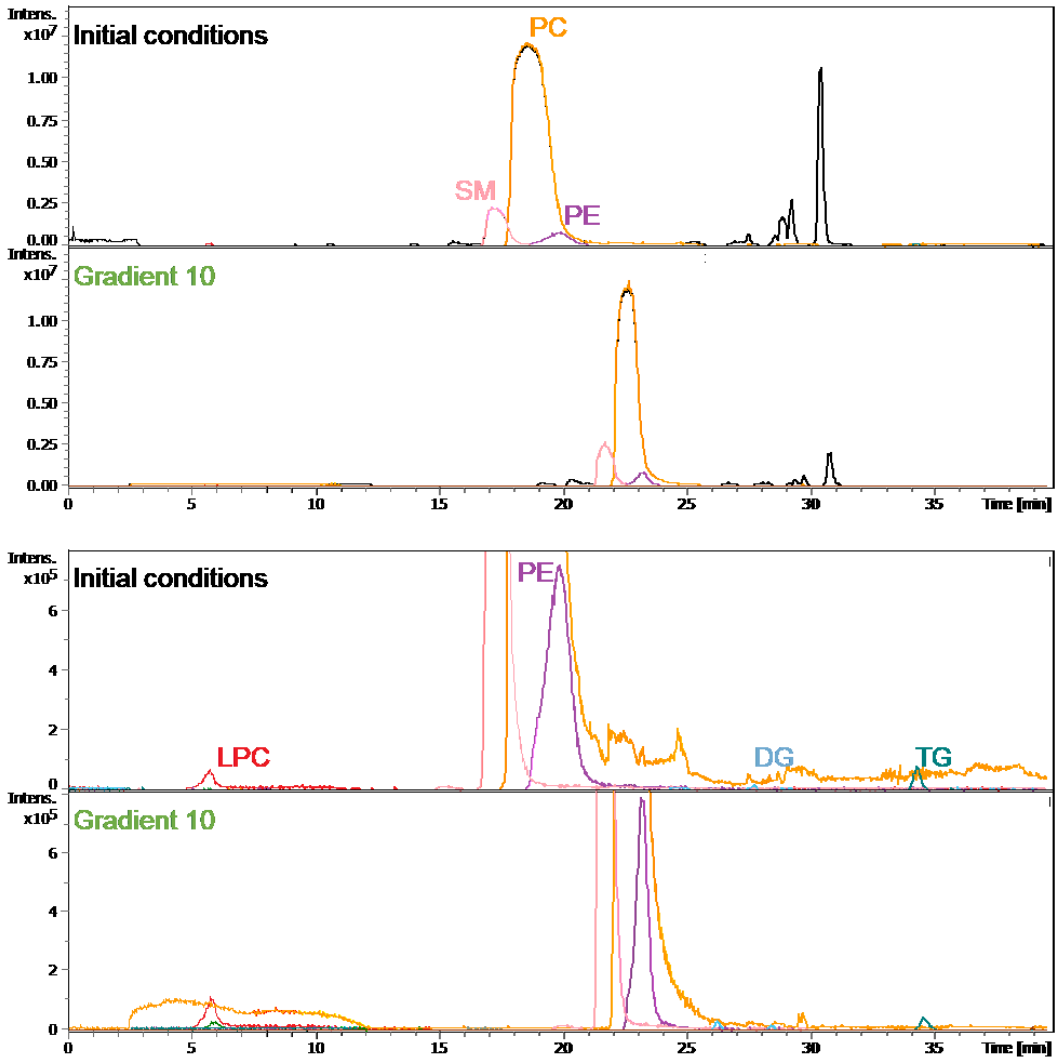

Figure A - 29. Comparison between the gradient optimized during the first phase (initial conditions) and gradient 10 for the 14 deuterated lipid standards (Table II-1). Both gradients are described in Figure A - 28. The nanoLC-MS method is identical to Figure A - 27, except for gradient 10.

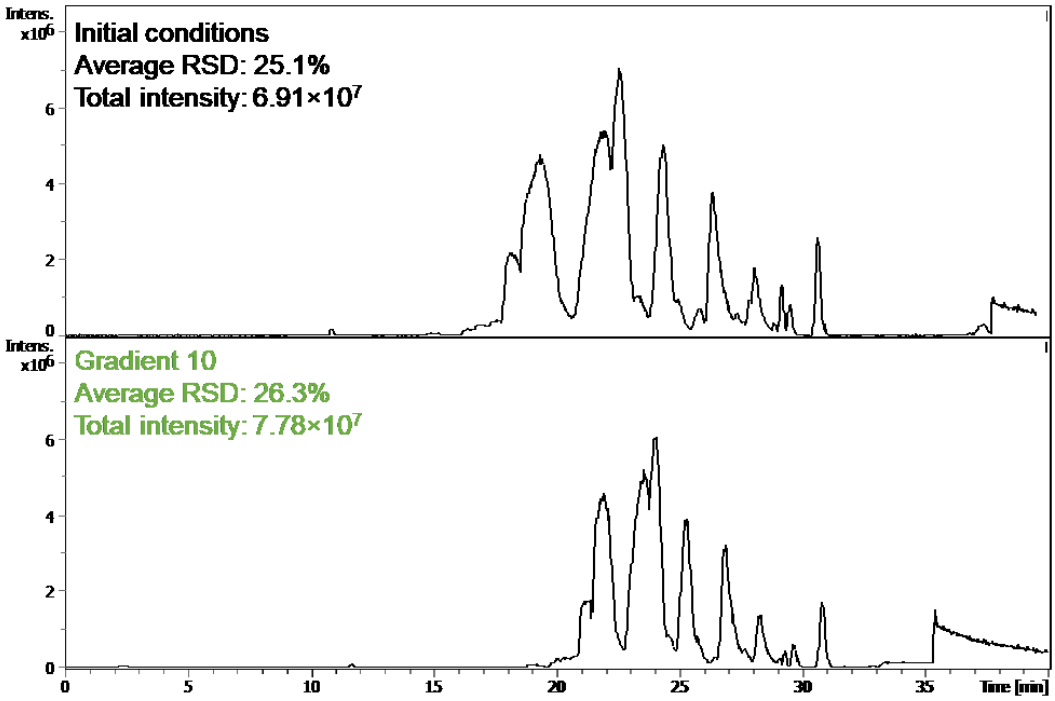

Figure A - 30. Comparison between the gradient optimized during the first phase (initial conditions) and gradient 10 for a pool of serum samples from 8 pigs. Both gradients are described in Figure A - 28. The nanoLC-MS method is identical to Figure A - 27, except for gradient 10.

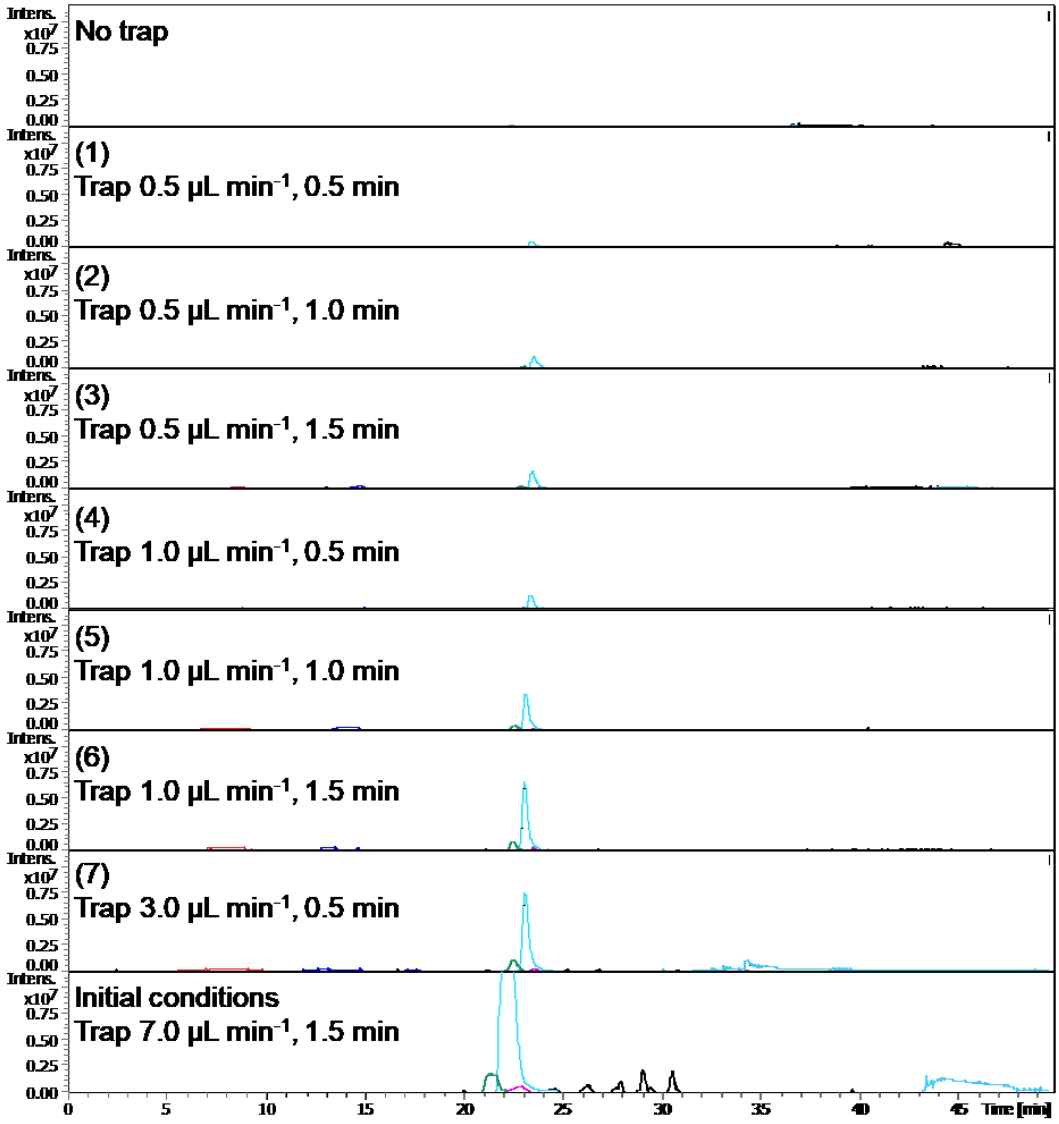

Figure A - 31. Re-optimization of trapping parameters (conditions 1 to 7) employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). Trapping conditions 8 to 14 are shown in Figure A - 32. The nanoLC-MS method is identical to Figure A - 27, except for gradient (gradient 10, Figure A - 28) and trapping.

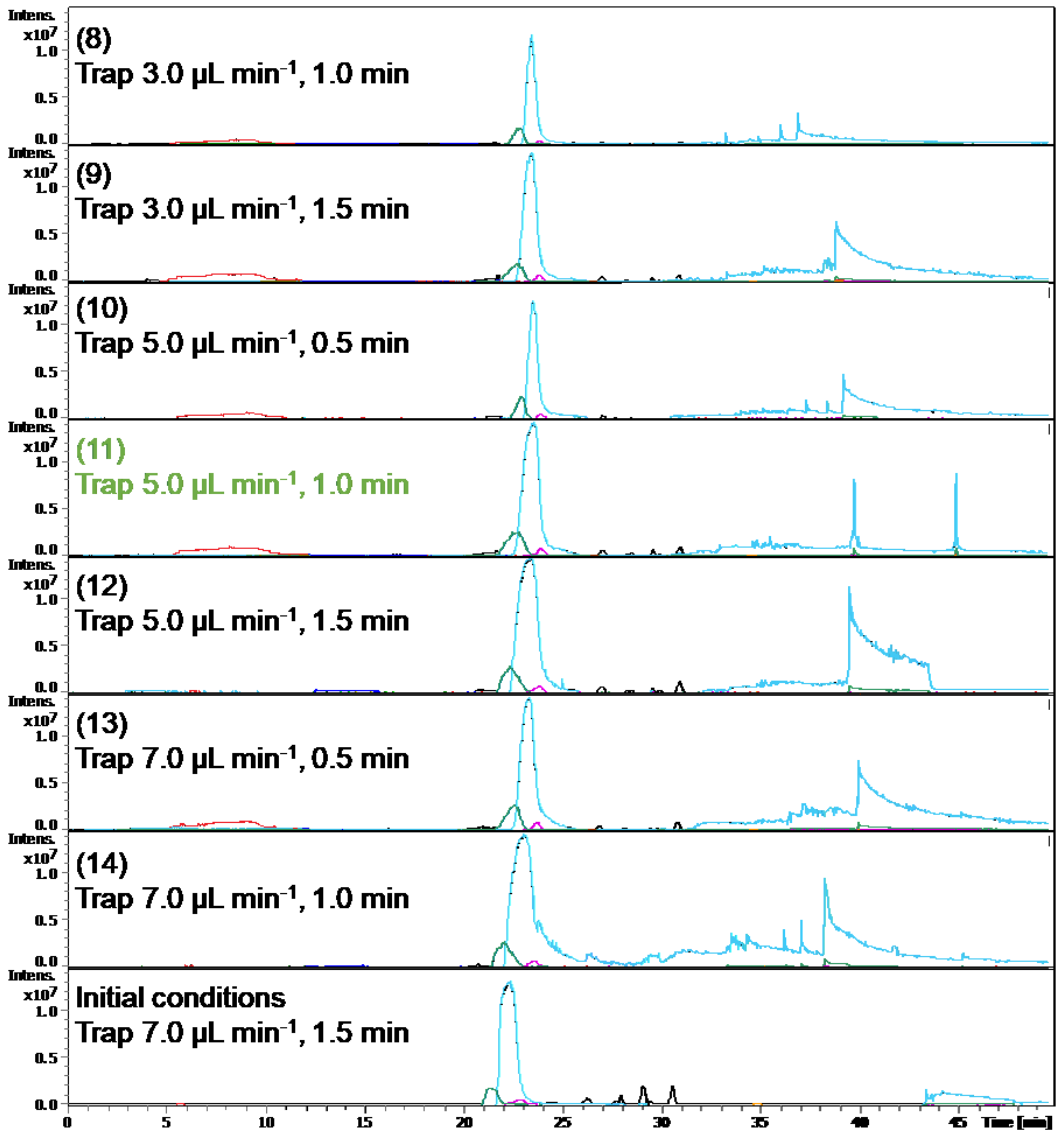

Figure A - 32. Re-optimization of trapping parameters (conditions 8 to 14) employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). Trapping conditions 1 to 7 are shown in Figure A - 31. Trapping 11 was chosen as the best option for this step. The nanoLC-MS method is identical to Figure A - 27, except for gradient (gradient 10, Figure A - 28) and trapping.

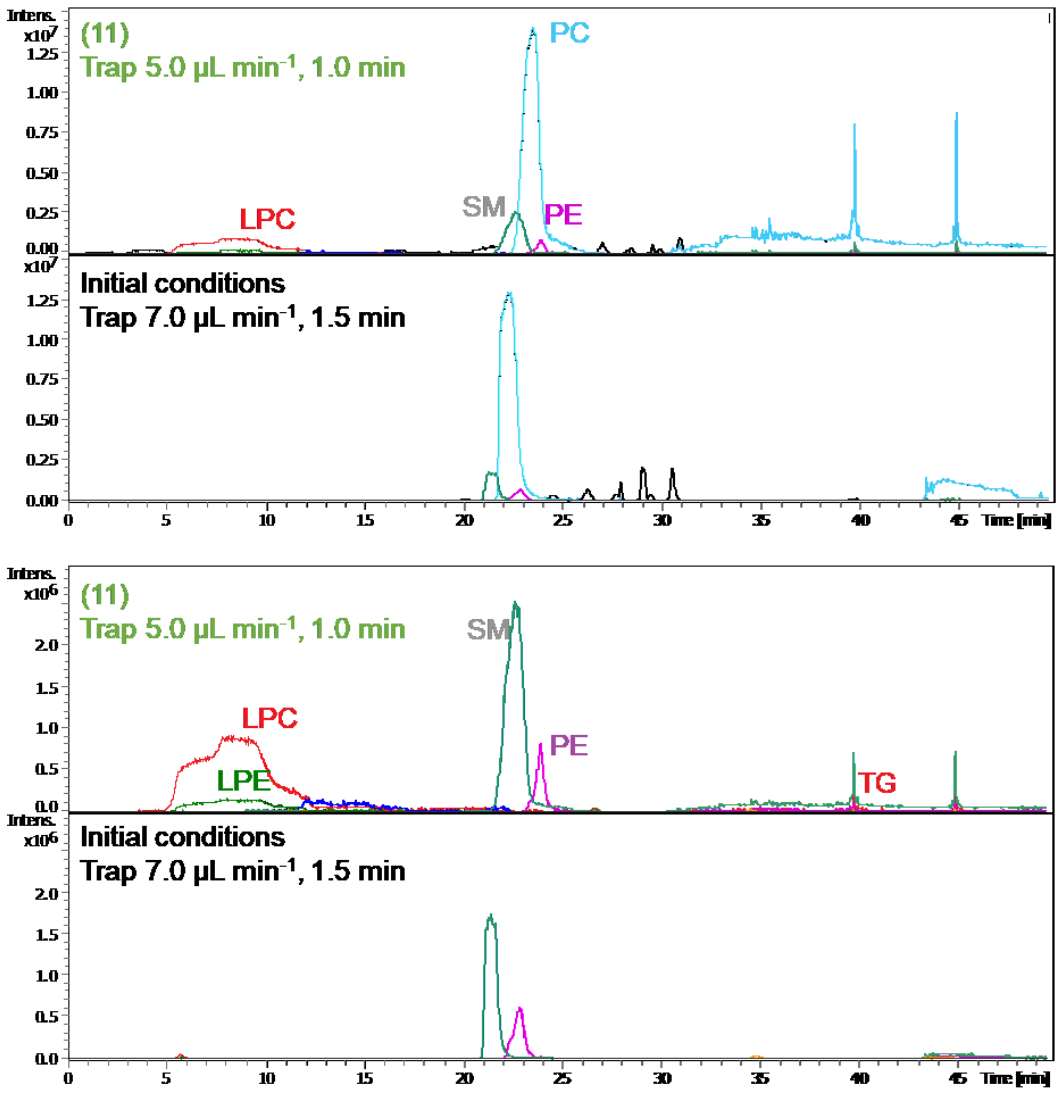

Figure A - 33. Detailed comparison between the previously optimized trapping conditions and the best parameters found for the second phase of optimization, employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). All standards are shown on the top two base peak chromatograms, whereas the most intense standard (PC d7-15:0/18:1) is removed on the bottom two chromatograms for better visualization. The nanoLC-MS method is identical to Figure A - 27, except for gradient (gradient 10, Figure A - 28) and trapping (trapping 11, Figure A - 32).

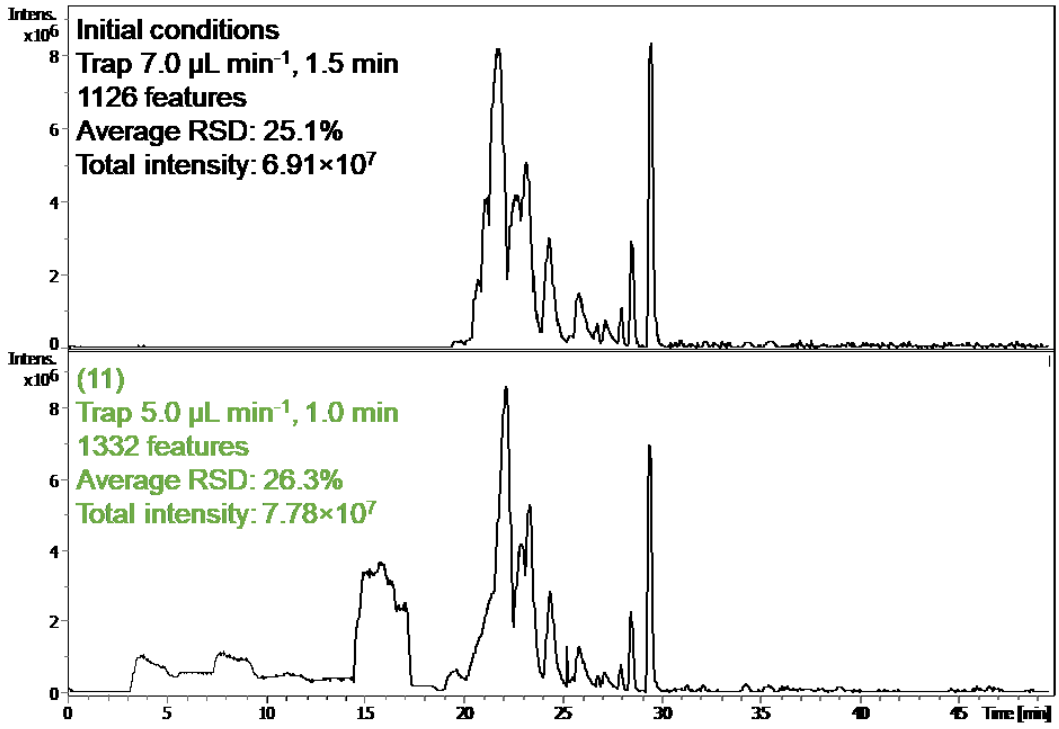

Figure A - 34. Comparison between the trapping conditions optimized during the first phase (initial conditions, top) and the parameters chosen during the second phase of optimization (gradient 10 and trap at 5.0 µL/min for 1.0 min, bottom) for a pool of serum samples from 8 pigs. The nanoLC-MS method is identical to Figure A - 27, except for gradient (gradient 10, Figure A - 28) and trapping (trapping 11, Figure A - 32).

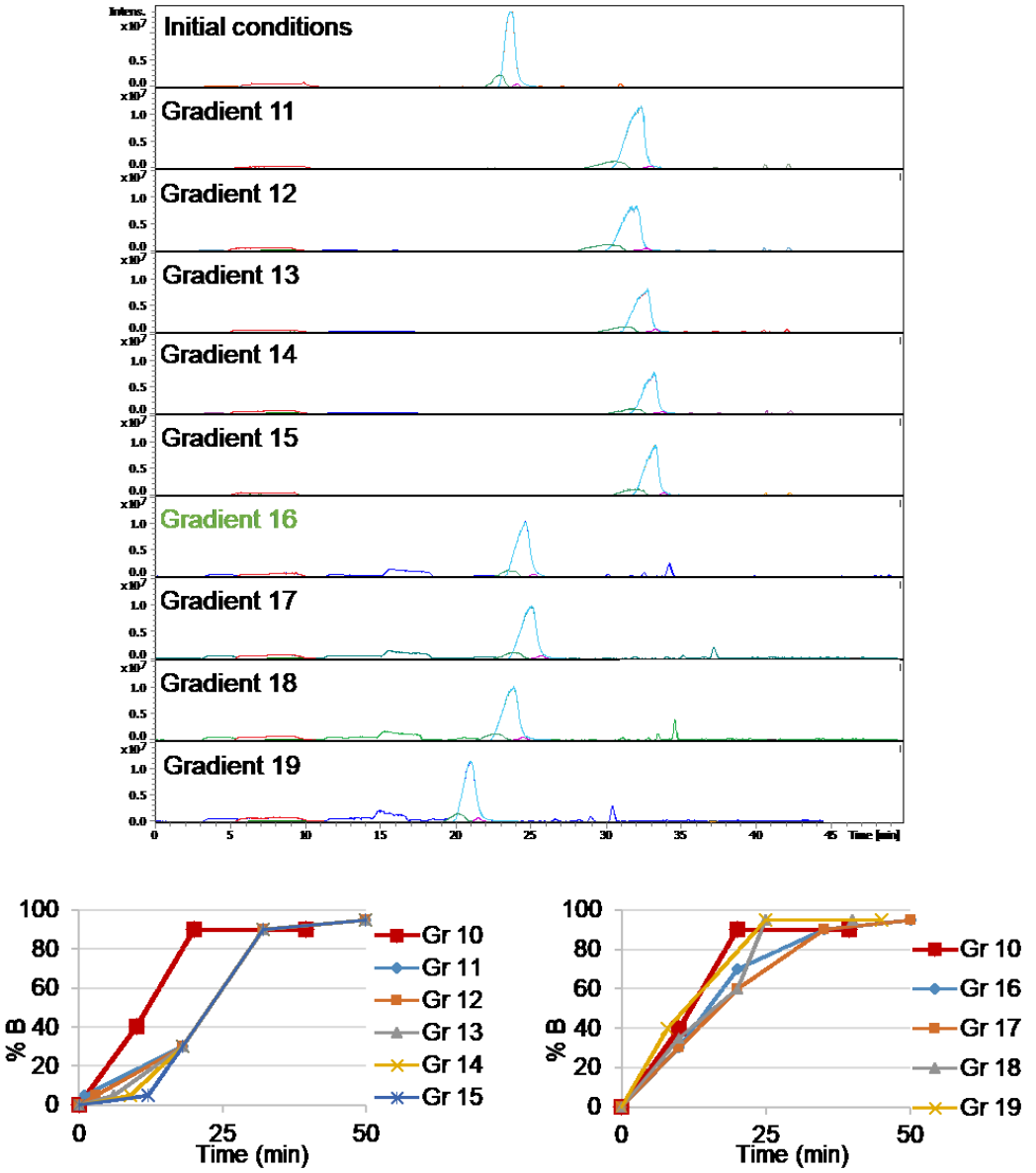

Figure A - 35. Fine adjustment of the separation gradient employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). Gradient 16 was chosen as the best option for this step. The nanoLC-MS method is identical to Figure A - 27, except for gradient and trapping (5 µL/min, 1.0 min).

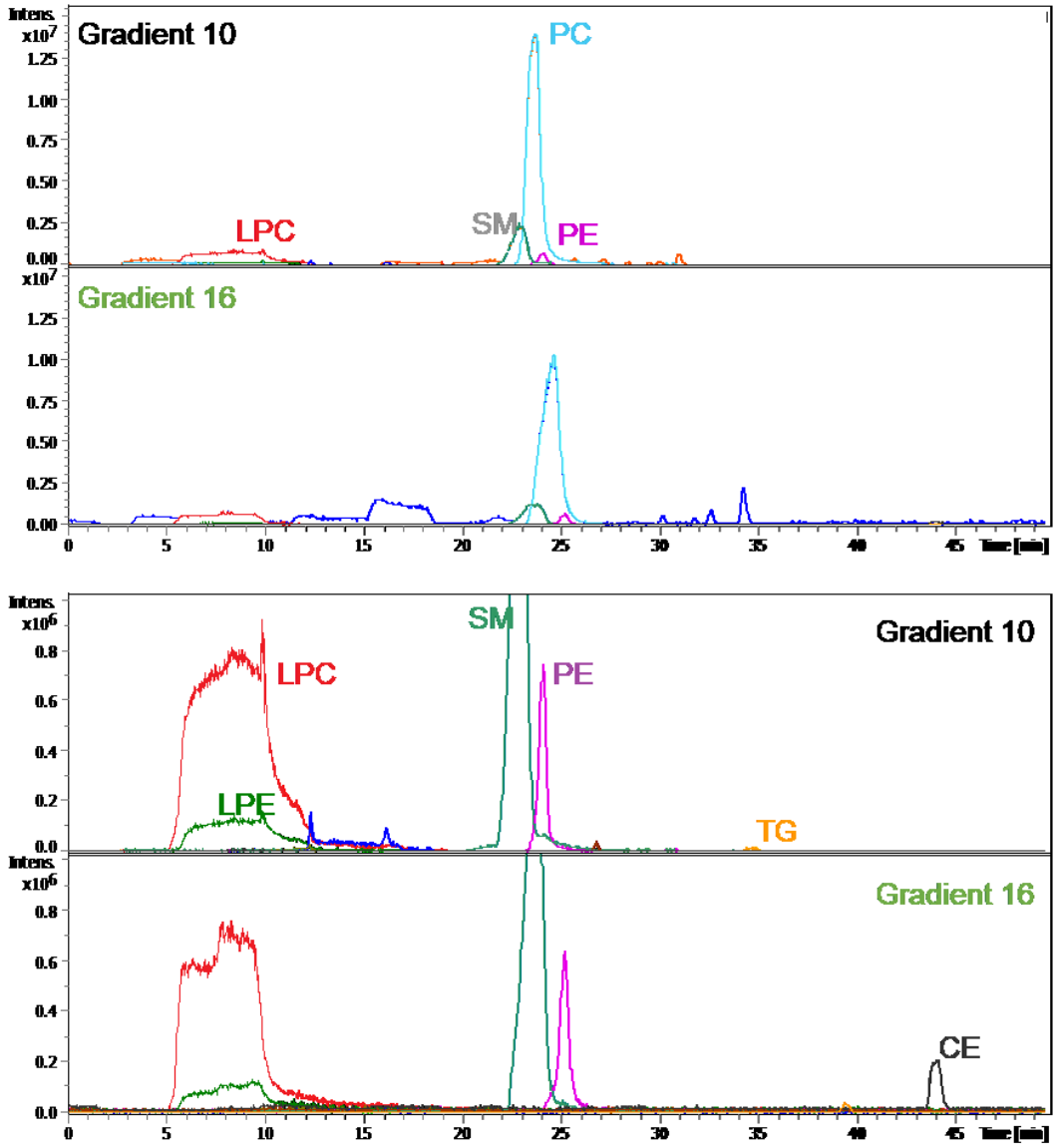

Figure A - 36. Detailed comparison of the previously optimized gradient (gradient 10, Figure A - 28) and the optimized gradient (gradient 16, Figure A - 35) employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). The two top base peak chromatograms display all standards, whereas the most intense peak (PC d7-15:0/18:1) is removed on the bottom chromatograms for better visualization. Both gradients are described in Figure A - 35. The nanoLC-MS method is identical to Figure A - 27, except for gradient and trapping (5 µL/min, 1.0 min).

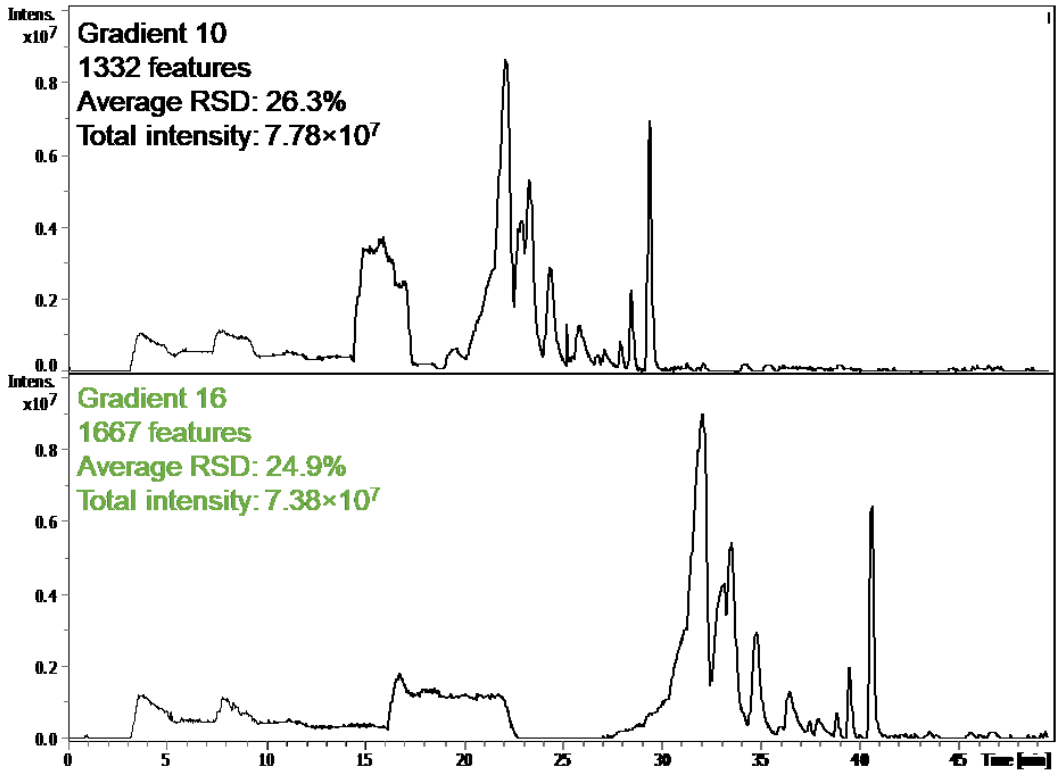

Figure A - 37. Detailed comparison of the previously optimized gradient (gradient 10, Figure A - 28) and the new optimized gradient (gradient 16, Figure A - 35) employing a pool of serum from 8 pigs. Both gradients are described in Figure A - 35. The nanoLC-MS method is identical to Figure A - 27, except for gradient and trapping (5 µL/min, 1.0 min).

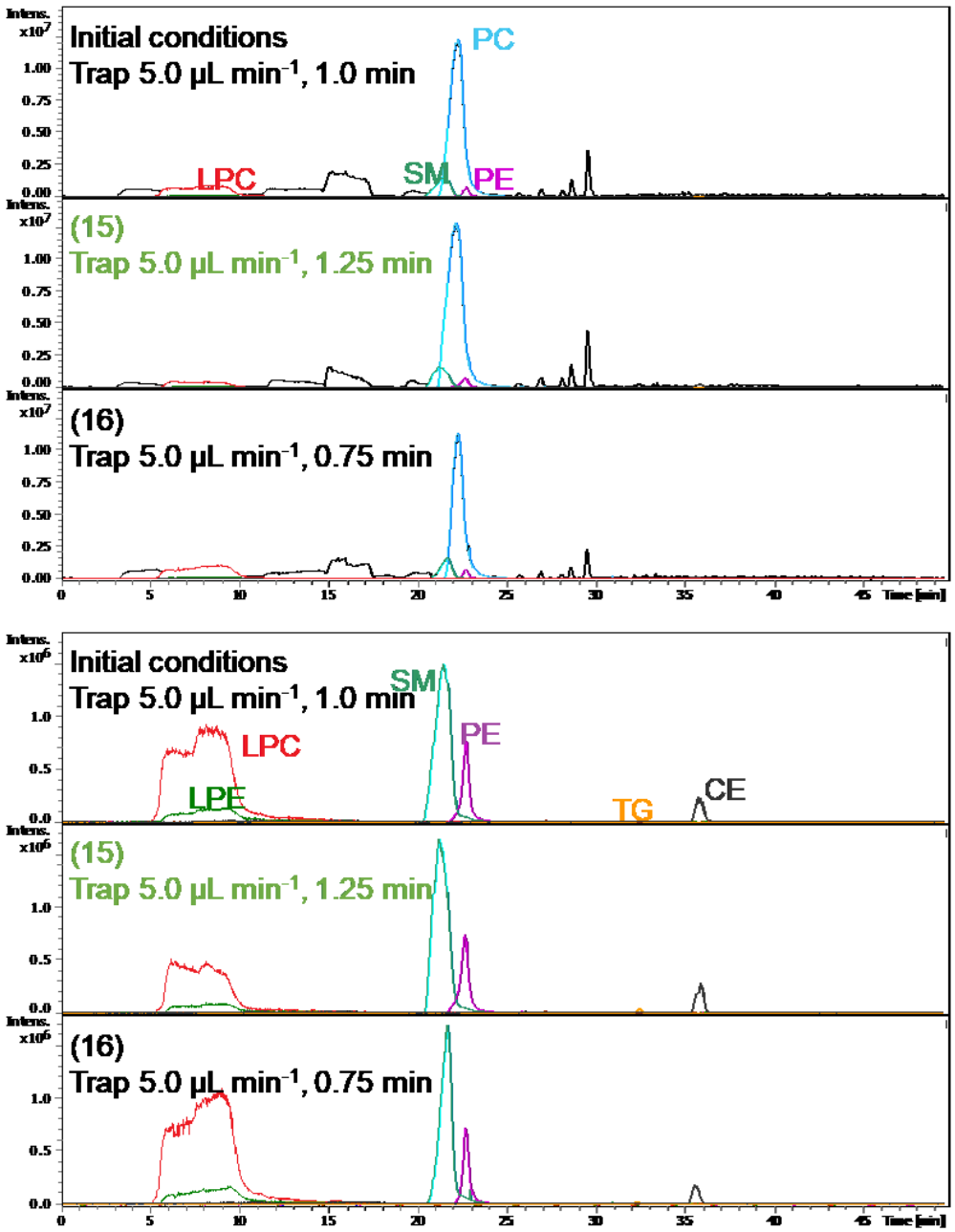

Figure A - 38. Fine adjustment of trapping conditions employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). The top three base peak chromatograms display all standards, whereas the most intense peak (PC d7-15:0/18:1) is removed on the bottom chromatograms for better visualization. Trapping 15 (5.0 µL/min, 1.25 min) was chosen as the best option for this step. The nanoLC-MS method is identical to Figure A - 27, except for gradient (gradient 10 described in Figure A - 35) and trapping.

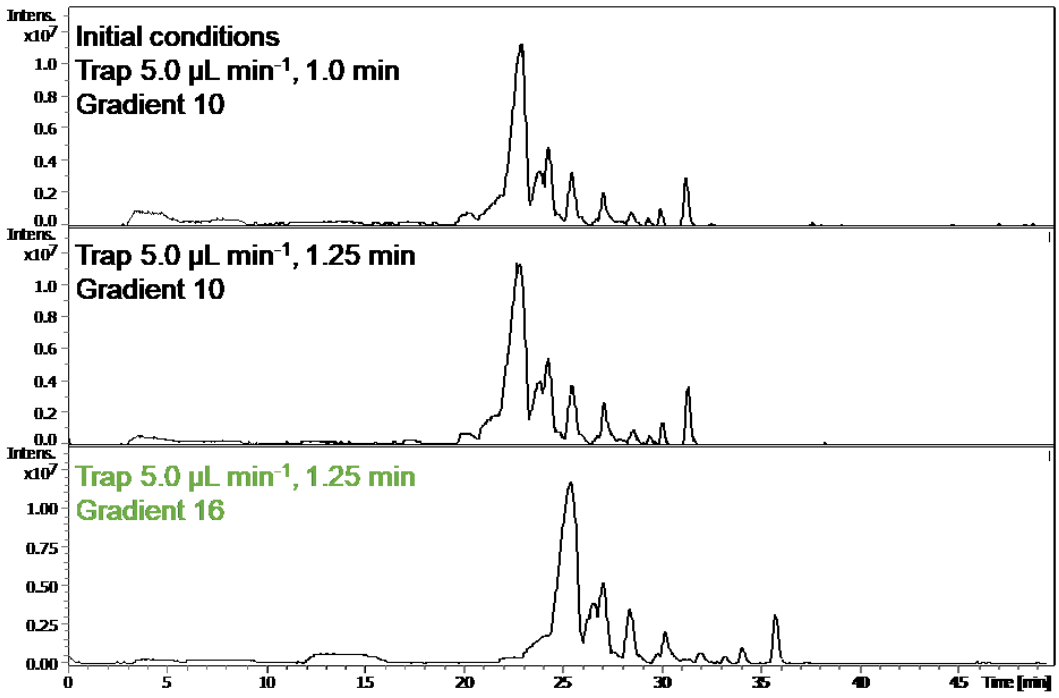

Figure A - 39. Comparison between the previously optimized conditions for nanoLC-MS analysis of lipids obtained from a pool of serum samples from 8 pigs (top) and the fine-tuned parameters for phase 2 (bottom). NanoLC method: MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 20 min equilibrium (95% MPA); 2 µL injection. Trapping conditions are described in each chromatogram, with the best results for 5.0 µL/min for 1.25 min (100% MPA, bottom chromatogram). Gradient 10: 0 min – 0% MPB, 10 min – 40% MPB, 20 min – 90% MPB, 39.5 min – 90% MPB. Gradient 16: 0 min – 0% MPB, 10 min – 30% MPB, 20 min – 70% MPB, 32 min – 95% MPB, 50 min – 95% MPB. The MS method is identical to Figure A - 27.

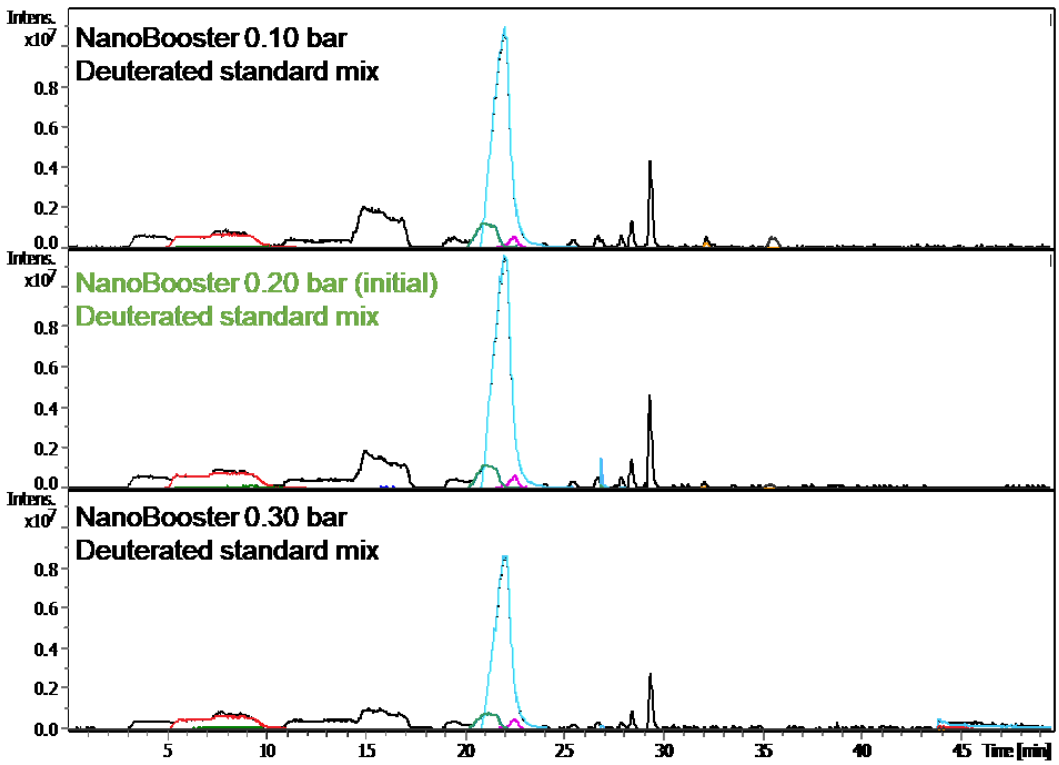

Figure A - 40. Re-evaluation of the nanoBooster acetonitrile-enriched nitrogen gas pressure between 0.10 and 0.30 bar employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). The base peak chromatogram is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors. NanoLC method: trapping at 5.0 µL/min for 1.25 min (100% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 50 min gradient (0 min – 0% MPB, 10 min – 30% MPB, 20 min – 70% MPB, 32 min – 95% MPB, 50 min – 95% MPB), 20 min equilibrium (95% MPA); 2 µL injection. The MS method is identical to Figure A - 27, except for nanoBooster nitrogen gas pressure.

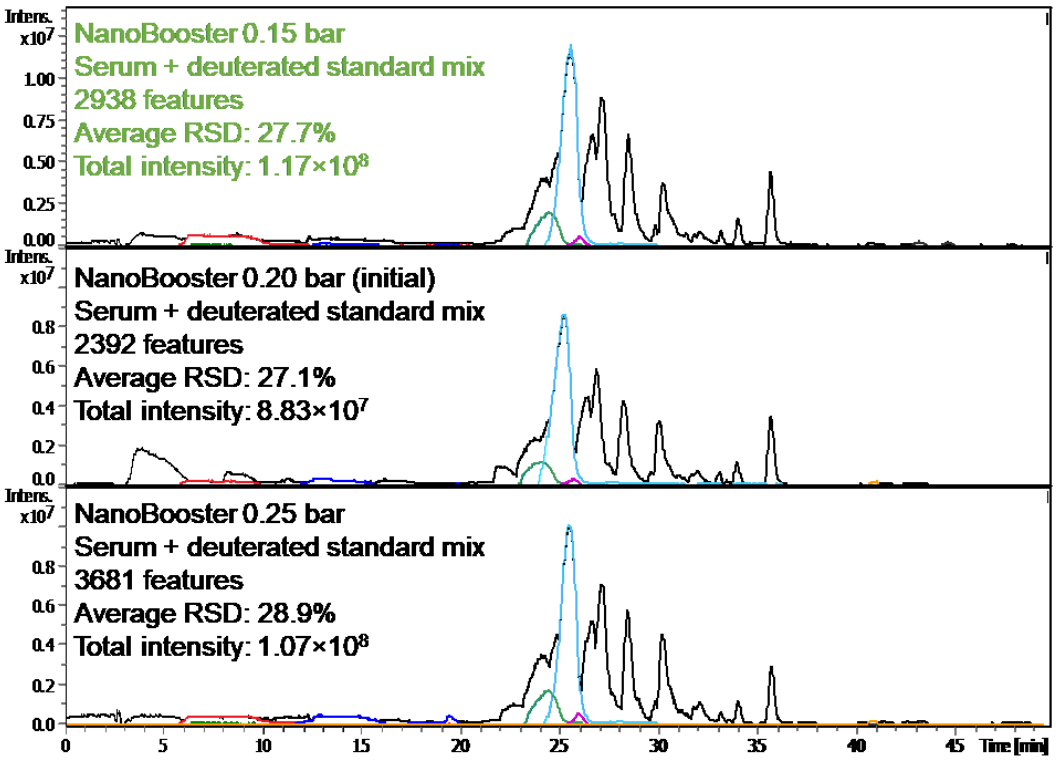

Figure A - 41. Fine-tuning of the nanoBooster acetonitrile-enriched nitrogen gas pressure with a pool of serum samples from 8 pigs. The base peak chromatogram for serum is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors (Table II-1). More features were detected for 0.25 bar, but there was a high consumption of acetonitrile and higher average relative standard deviation (RSD), *i.e.* lower peak intensity reproducibility. The nanoLC method is identical to Figure A - 40. The MS method is identical to Figure A - 27, except for nanoBooster nitrogen gas pressure.

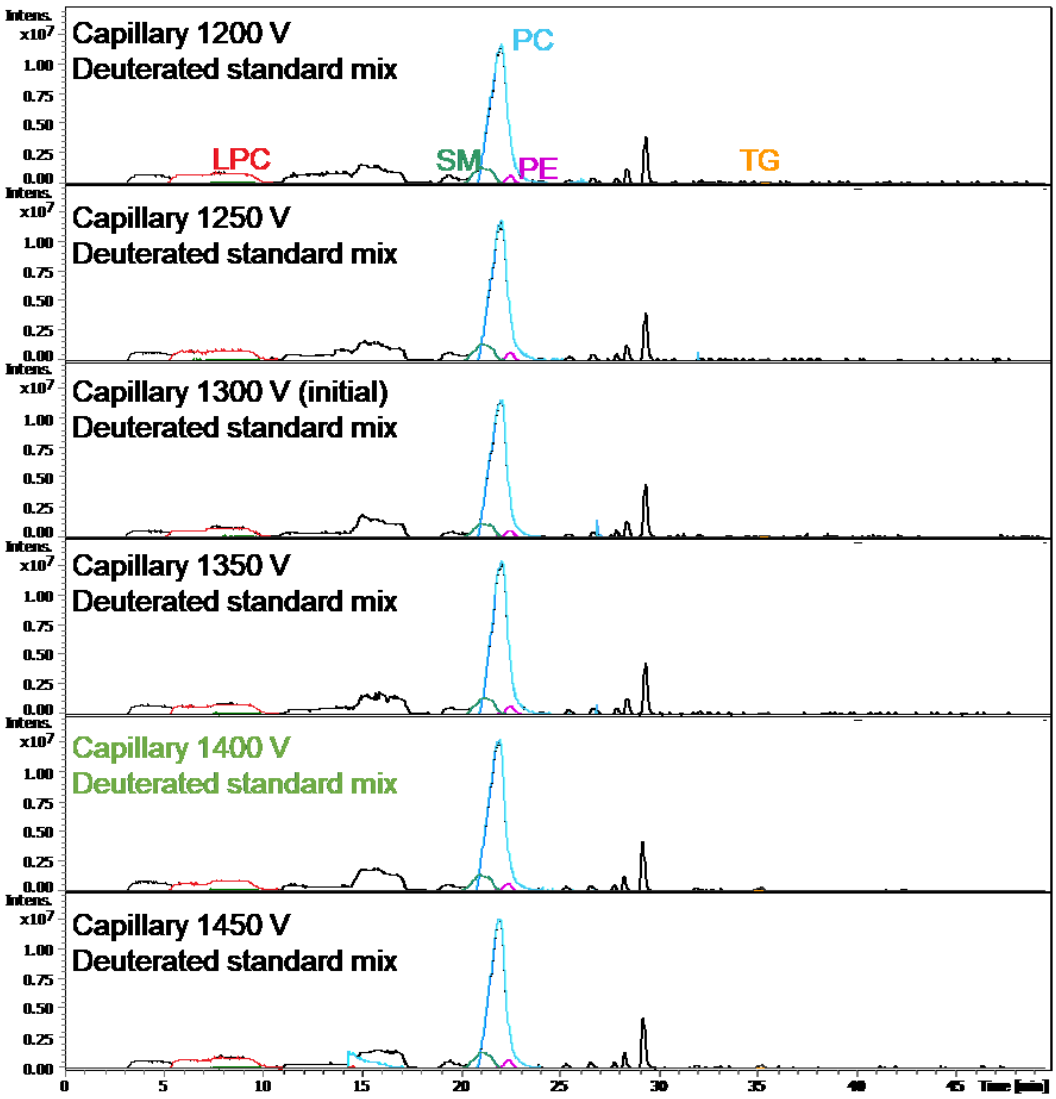

Figure A - 42. Re-evaluation of the capillary voltage for the optimized nanoLC method employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). The base peak chromatogram is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors. The nanoLC method is identical to Figure A - 40. The MS method is identical to Figure A - 27, except for capillary voltage.

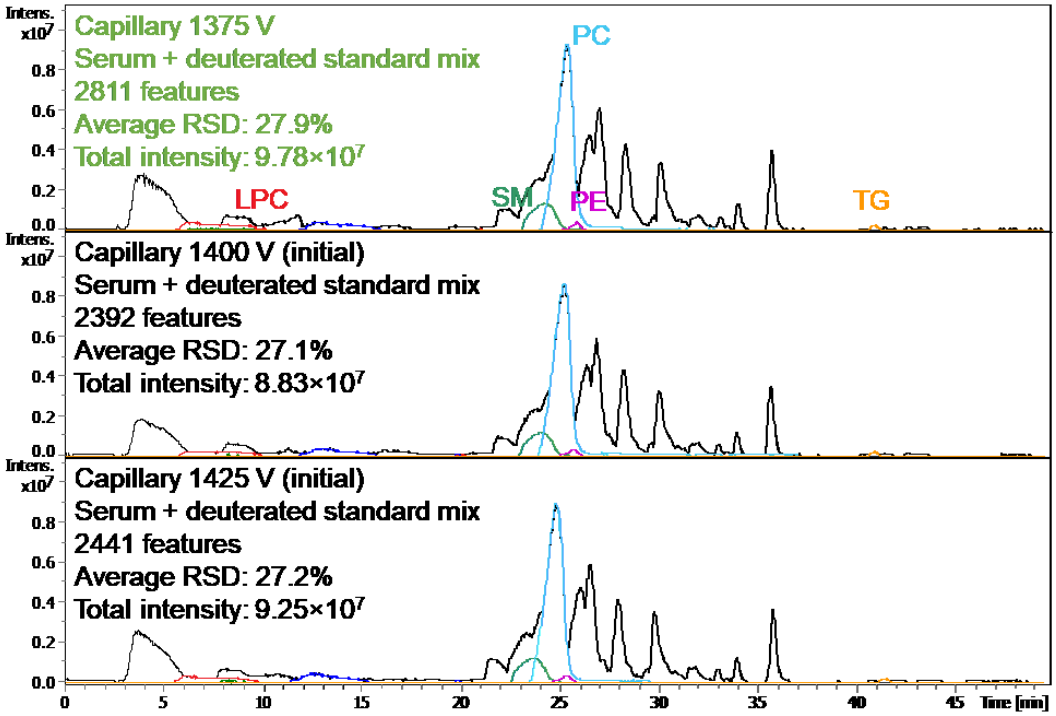

Figure A - 43. Fine-tuning of the capillary voltage for the optimized nanoLC method with a pool of serum samples from 8 pigs. The base peak chromatogram for serum is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors (Table II-1). The nanoLC method is identical to Figure A - 40. The MS method is identical to Figure A - 27, except for capillary voltage.

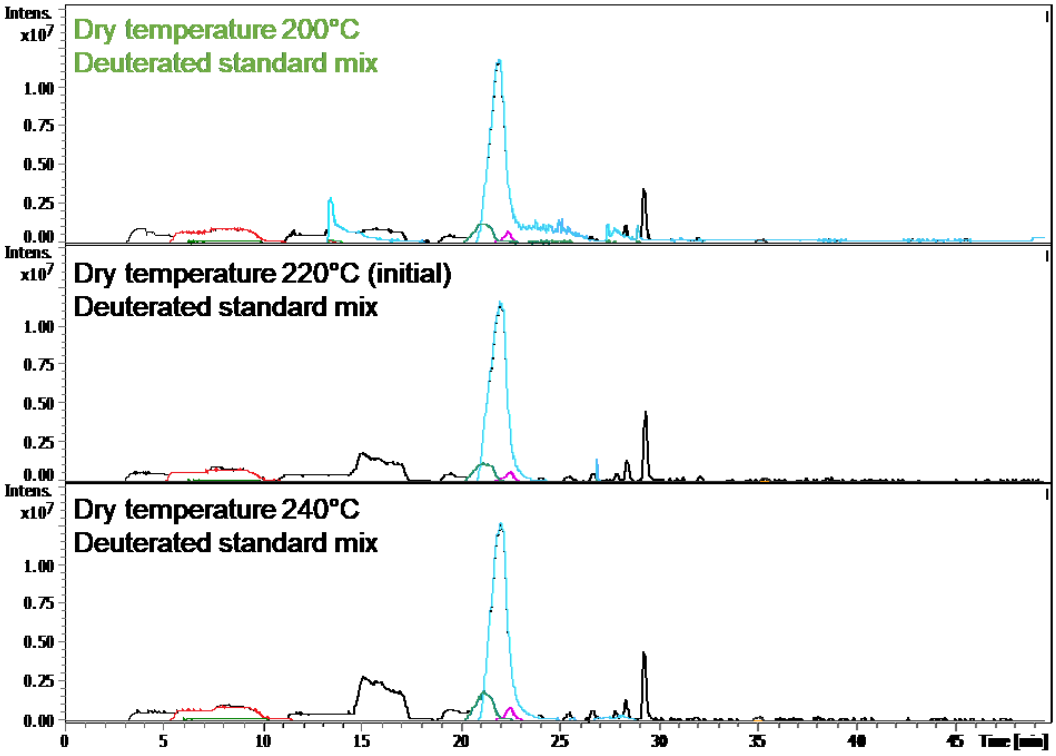

Figure A - 44. Re-evaluation of the ion source temperature for the optimized nanoLC method employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). The base peak chromatogram is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors. The nanoLC method is identical to Figure A - 40. The MS method is identical to Figure A - 27, except for source temperature.

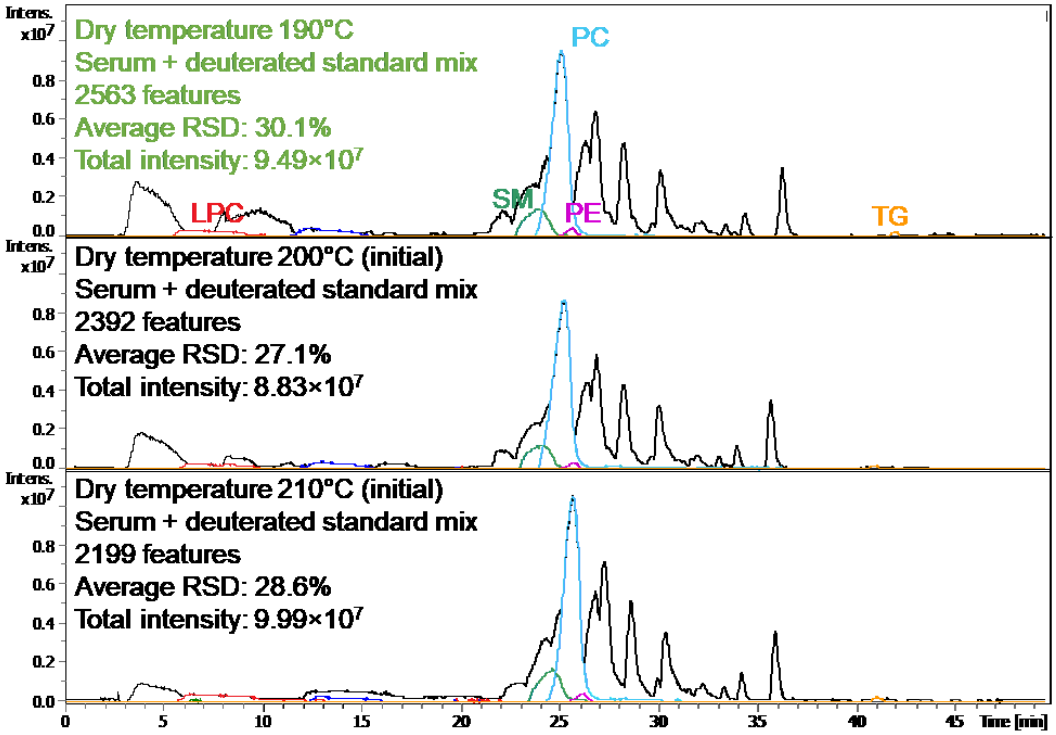

Figure A - 45. Fine-tuning of the ion source temperature for the optimized nanoLC method with a pool of serum samples from 8 pigs. The base peak chromatogram for serum is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors (Table II-1). The nanoLC method is identical to Figure A - 40. The MS method is identical to Figure A - 27, except for source temperature.

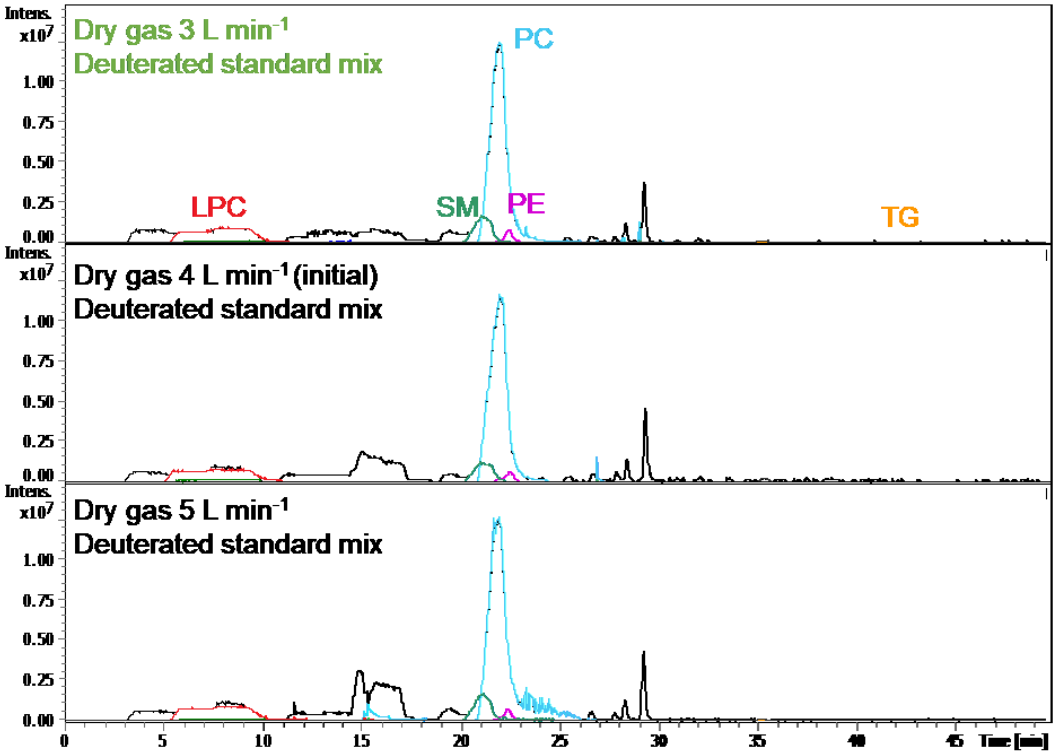

Figure A - 46. Re-evaluation of the dry nitrogen gas flowrate for the optimized nanoLC method employing a mixture of 14 deuterated lipid standards from different lipid subclasses (Table II-1). The base peak chromatogram is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors. The nanoLC method is identical to Figure A - 40. The MS method is identical to Figure A - 27, except for dry nitrogen gas flowrate.

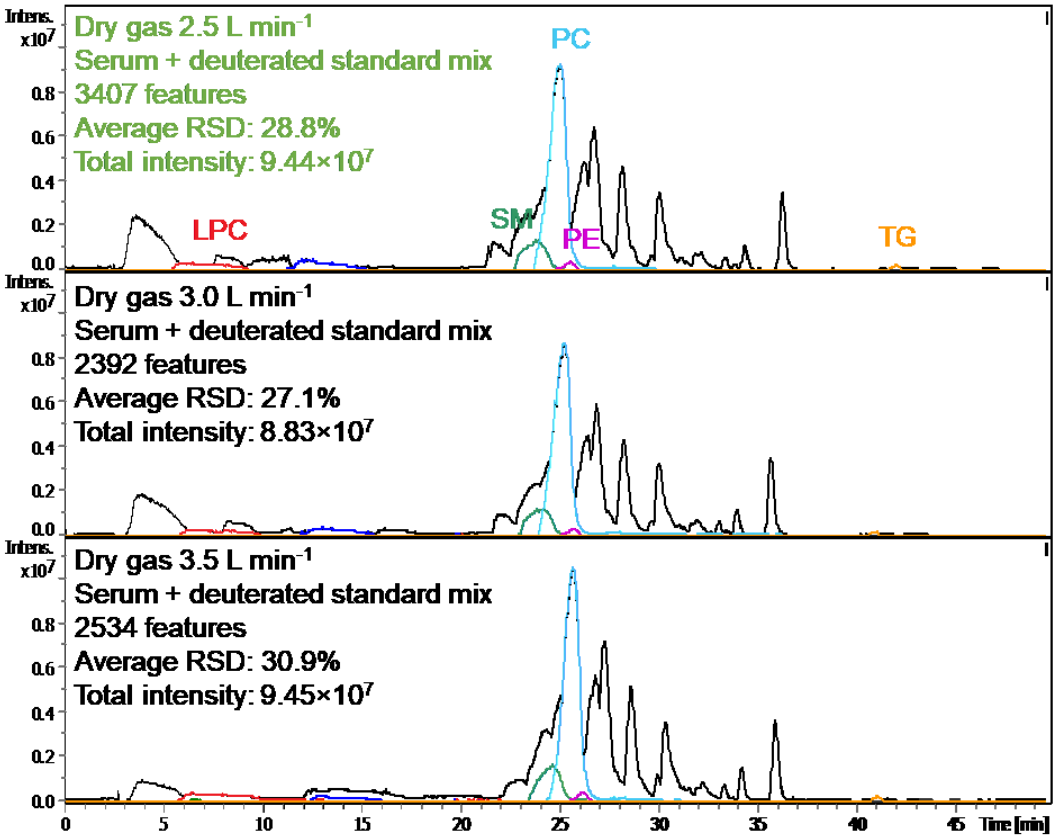

Figure A - 47. Fine-tuning of the dry nitrogen gas flowrate for the optimized nanoLC method with a pool of serum samples from 8 pigs. The base peak chromatogram for serum is shown in black, whereas the extracted ion chromatograms for the lipid deuterated standards are shown in different colors (Table II-1). The nanoLC method is identical to Figure A - 40. The MS method is identical to Figure A - 27, except for dry gas flowrate.

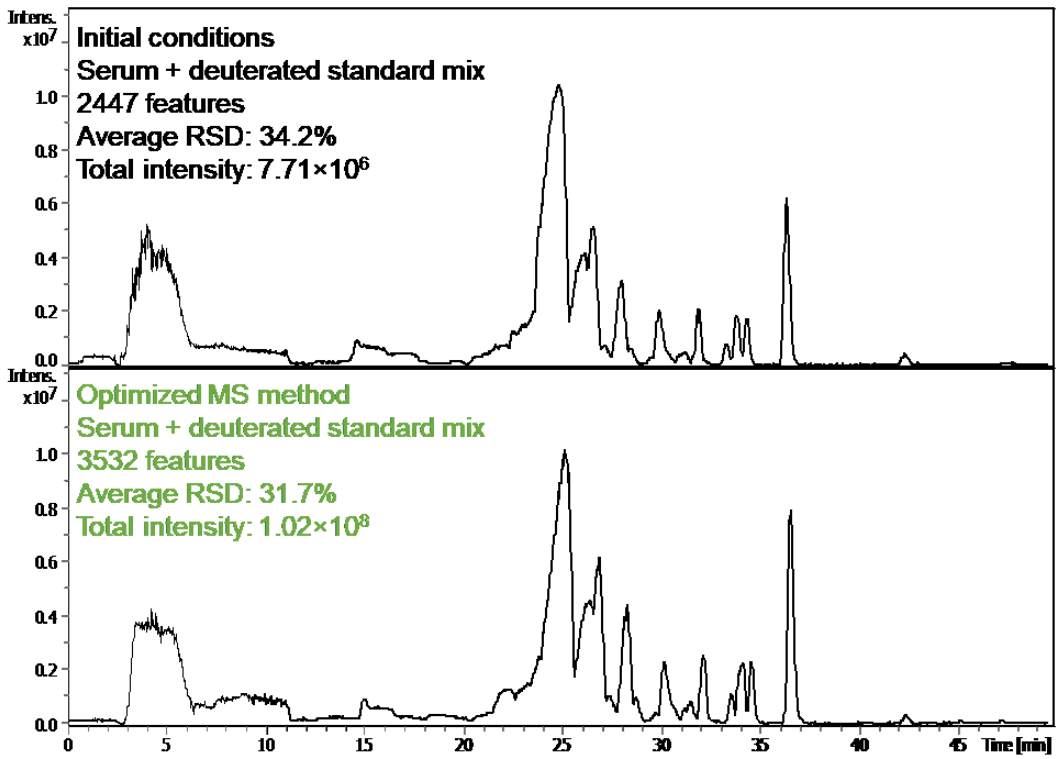

Figure A - 48. Comparison between chromatograms for a pool of serum samples from 8 pigs obtained with the previously optimized MS method (top, electrospray ion source capillary voltage of 1300 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.20 bar, dry nitrogen gas flow rate of 4 L/min, ion source temperature of 220ºC) and the re-evaluated, fine-tuned parameters (bottom, electrospray ion source capillary voltage of 1375 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.15 bar, dry nitrogen gas flow rate of 2.5 L/min, ion source temperature of 190ºC). The spectra acquisition rate was kept at 1 Hz. The nanoLC method is identical to Figure A - 40.

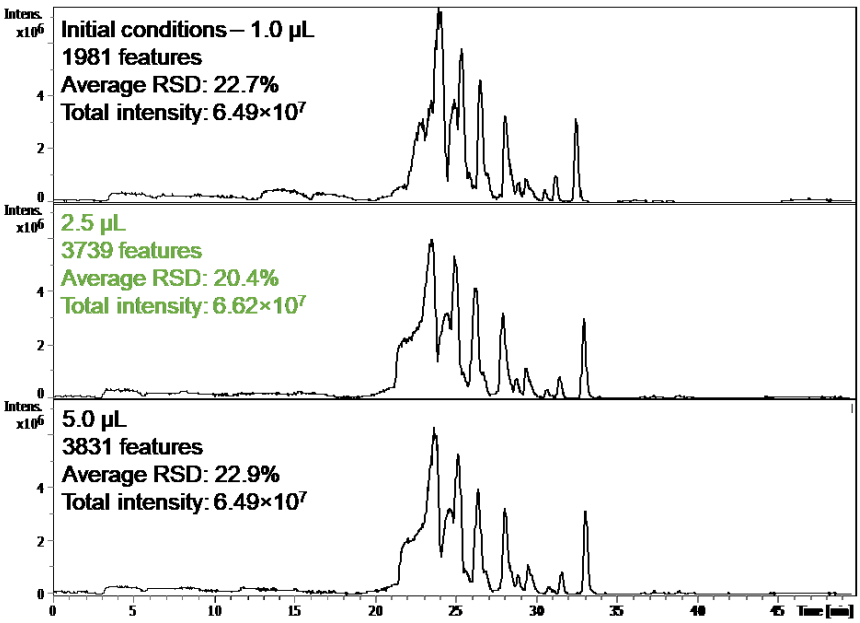

Figure A - 49. Evaluation of the sample volume for extraction. Sample preparation method: 1.0, 2.5 or 5.0 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with equal volumes of the mixture of 14 deuterated lipid standards (Table II-1) and 55.8 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 33.4 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min and resuspended with 10% of 6:4 MPA/MPB and 90% of 9:1 MPA/MPB for 100× dilution (total of 100.0 µL for the 1.0 µL serum aliquots; 250.0 µL for the 2.5 µL serum aliquots; and 500.0 µL for the 5.0 µL serum aliquots). NanoLC method: trapping at 5.0 µL/min for 1.25 min (100% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 50 min gradient (0 min – 0% MPB, 10 min – 30% MPB, 20 min – 70% MPB, 32 min – 95% MPB, 50 min – 95% MPB), 20 min equilibrium (95% MPA); 2 µL injection. MS method: electrospray ion source capillary voltage of 1375 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.15 bar, dry nitrogen gas flow rate of 2.5 L/min, ion source temperature of 190ºC.

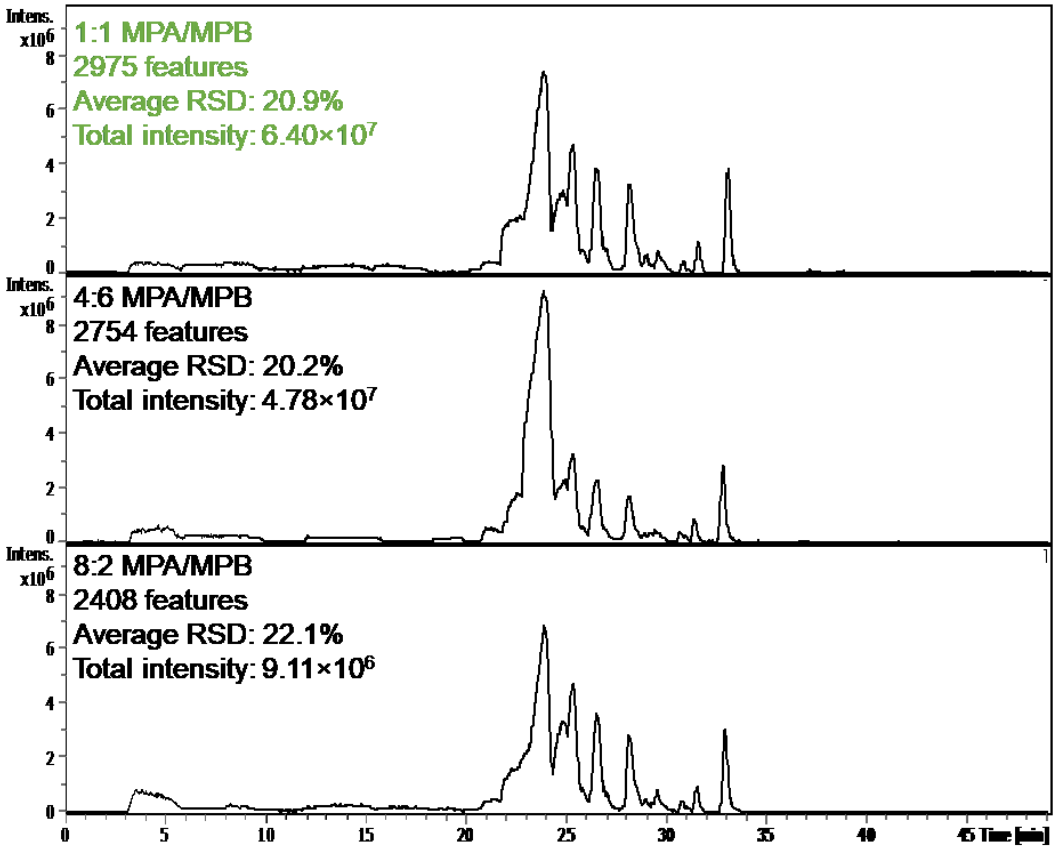

Figure A - 50. Optimization of the resuspension of the dried lipid extract. Sample preparation method: 1.0 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 1.0 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 55.8 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 33.4 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 10.0 µL of the resuspension solvents (as described for each chromatogram) and diluted with 90.0 µL of 9:1 MPA/MPB to 100× dilution. The nanoLC-MS methods are identical to Figure A - 49.

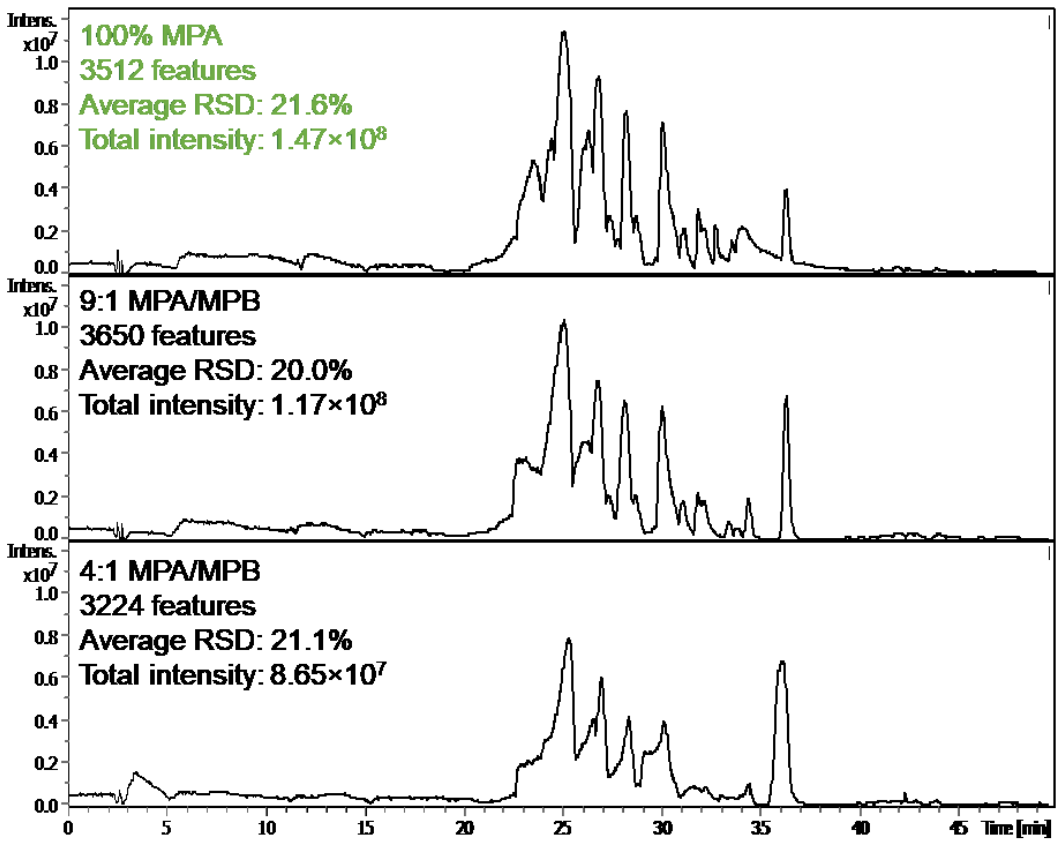

Figure A - 51. Optimization of the composition of mobile phases for sample dilution before injection. Sample preparation method: 2.5 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 55.8 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 33.4 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 7.5 µL of 6:4 MPA/MPB and diluted with 67.5 µL of the mobile phase mixture described for each chromatogram. The nanoLC-MS methods are identical to Figure A - 49.

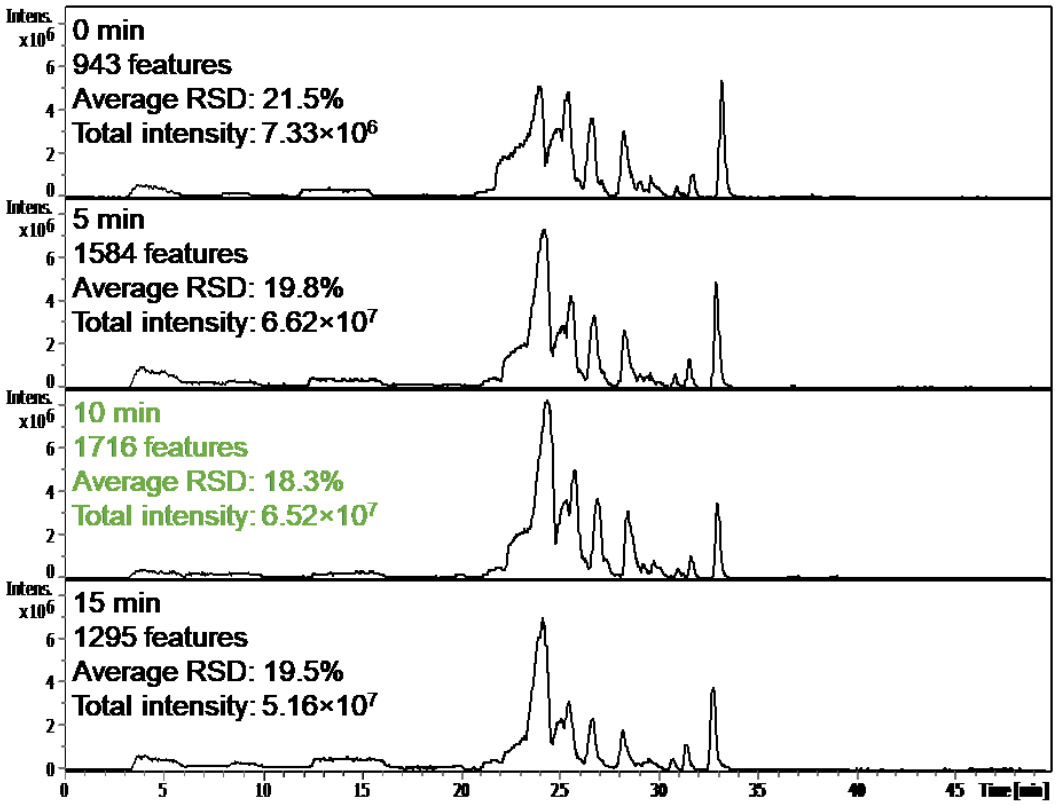

Figure A - 52. Optimization of the equilibrium time before centrifugation of the biphasic mixture obtained by liquid-liquid extraction. Sample preparation method: 1.0 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 1.0 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 55.8 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 33.4 µL of water and vortexed for 10 s; after resting for different times (as described for each chromatogram) at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 10.0 µL of 6:4 MPA/MPB and diluted with 90.0 µL of 9:1 MPA/MPB. The nanoLC-MS methods are identical to Figure A - 49.

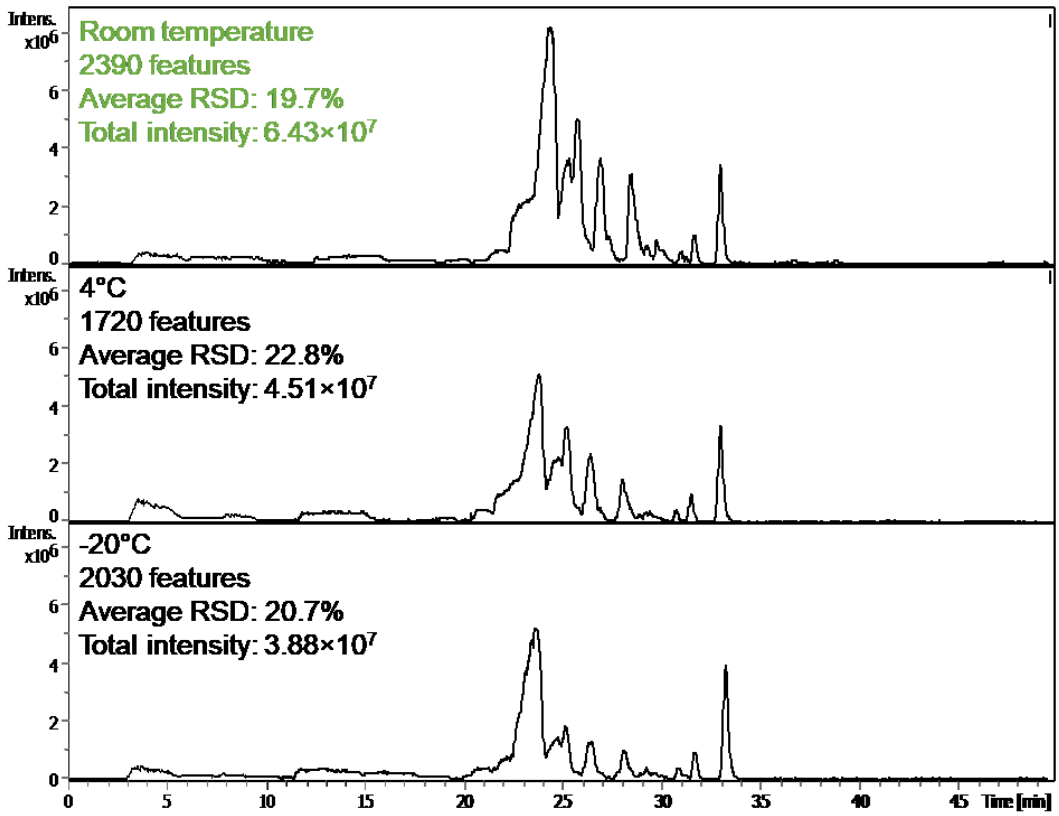

Figure A - 53. Optimization of the temperature for equilibrium before centrifugation of the biphasic mixture obtained by liquid-liquid extraction. Sample preparation method: 1.0 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 1.0 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 55.8 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 33.4 µL of water and vortexed for 10 s; after resting for 10 min at variable temperatures (as described for each chromatogram), the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 10.0 µL of 6:4 MPA/MPB and diluted with 90.0 µL of 9:1 MPA/MPB. The nanoLC-MS methods are identical to Figure A - 49.

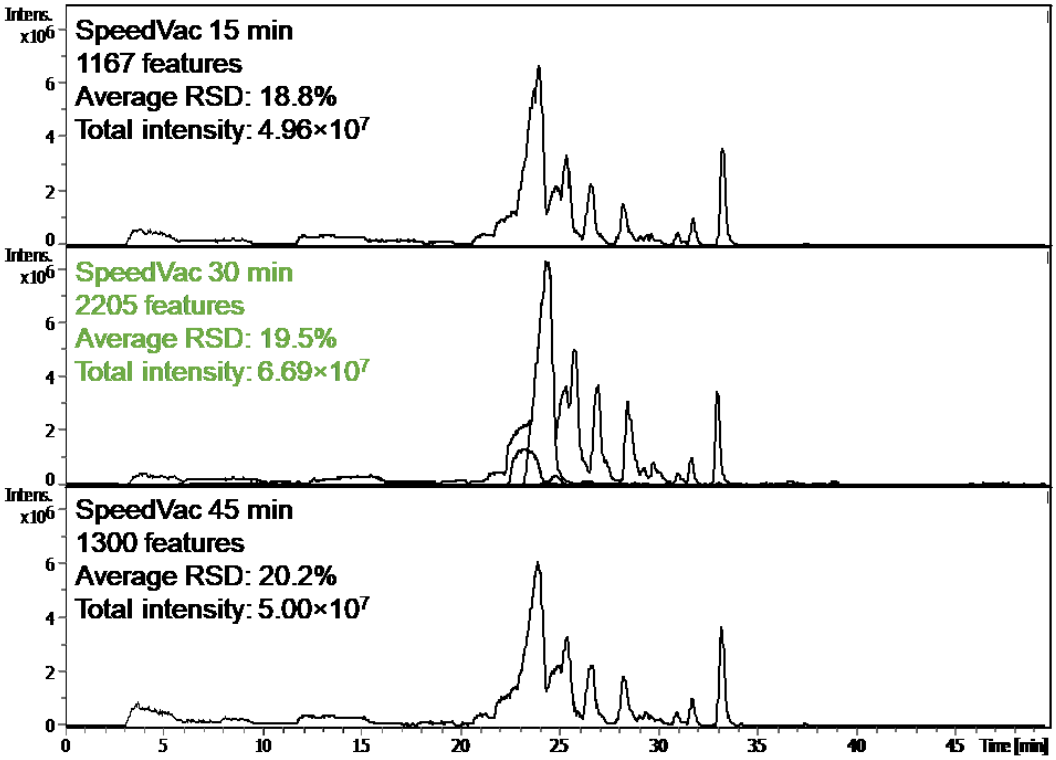

Figure A - 54. Optimization of the SpeedVac drying time for the organic phase obtained after liquid-liquid extraction of serum samples. The three intervals allowed visible drying of the organic solvents, but traces of solvents and water may remain for reduced periods. Sample preparation method: 1.0 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 1.0 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 55.8 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 33.4 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for variable periods (as described for each chromatogram), resuspended with 10.0 µL of 6:4 MPA/MPB and diluted with 90.0 µL of 9:1 MPA/MPB. The nanoLC-MS methods are identical to Figure A - 49.

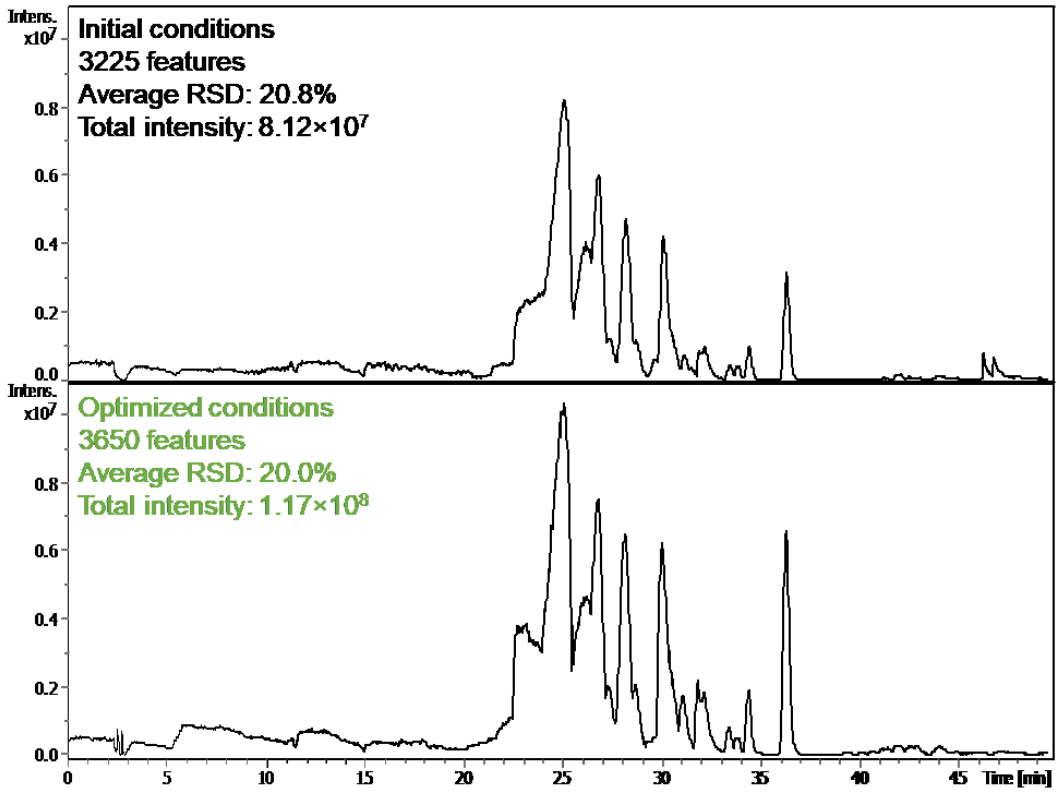

Figure A - 55. Comparison between the initial sample preparation method and the optimized volume of serum, resuspension mobile phase mixture, equilibrium time and temperature, and SpeedVac drying time. The initial sample preparation is described in Figure A - 49 for 1.0 µL of serum. Optimized conditions for sample preparation: 2.5 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 55.8 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 33.4 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 10.0 µL of 1:1 MPA/MPB and diluted with 90.0 µL of 9:1 MPA/MPB. The nanoLC-MS methods are identical to Figure A - 49.

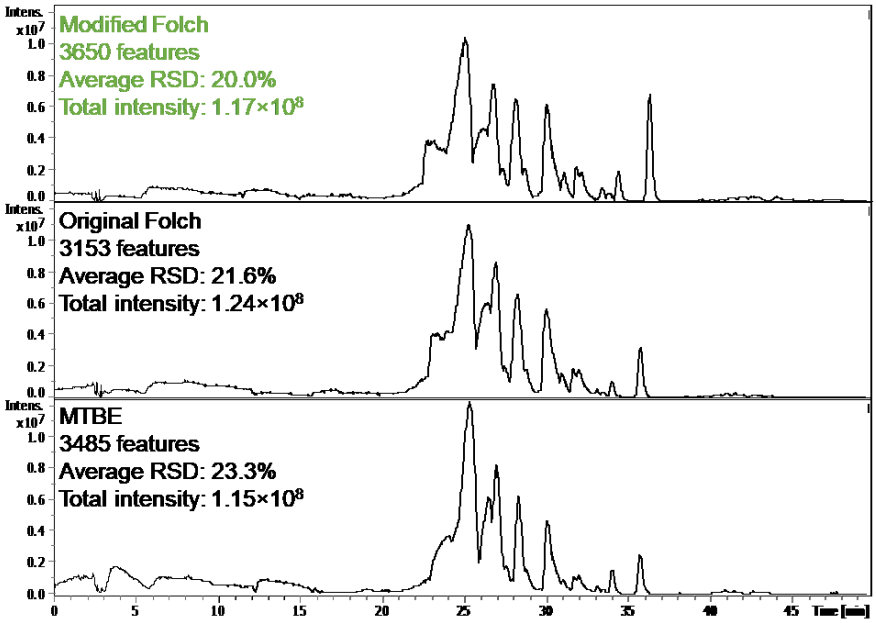

Figure A - 56. Comparison between the modified Folch method, the original Folch method and the MTBE method adapted from Matyash *et al*. for 2.5 µL of serum.[67,70] Modified Folch method: 2.5 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 53.3 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 31.9 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 10.0 µL of 1:1 MPA/MPB and diluted with 90.0 µL of 9:1 MPA/MPB. Folch method: the same procedure was followed, but with 2.5 µL of serum, 2.5 µL of the mixture of internal standards, 14.2 µL of methanol, 33.3 µL of dichloromethane and 8.2 µL of water. Adapted MTBE method: 2.5 µL of serum, 2.5 µL of the mixture of internal standards, 16.2 µL of methanol and 62.5 µL of MTBE were incubated in a shaker for 1 h at room temperature, followed by vortex with 15.6 µL of water; the remaining steps were identical to the modified Folch method, except for drying time of 1h30min. The nanoLC-MS methods are identical to Figure A - 49.

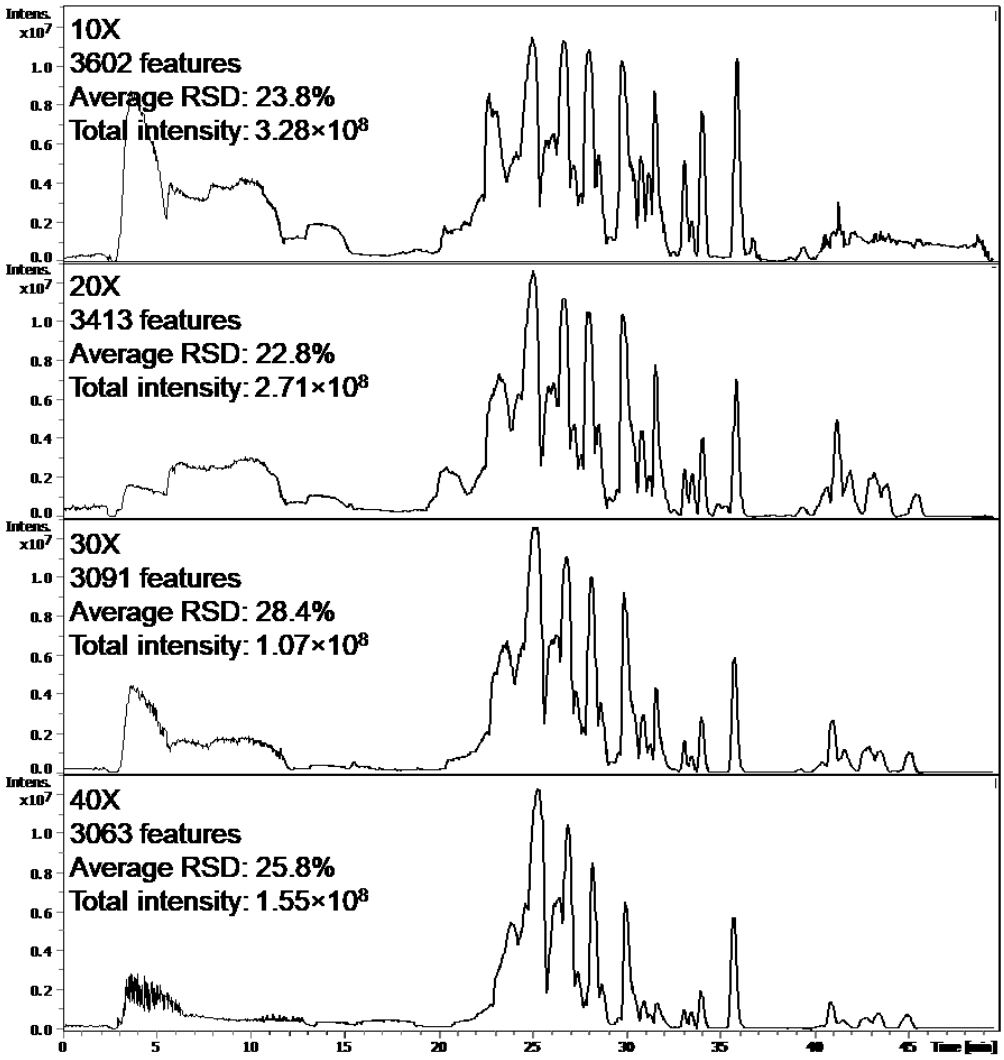

Figure A - 57. Dilution of the sample after extraction. Sample preparation: 2.5 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 53.3 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 31.9 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 10% of 1:1 MPA/MPB and diluted with 90% of 9:1 MPA/MPB, according to the dilutions described for each chromatogram. The nanoLC-MS methods are identical to Figure A - 49.

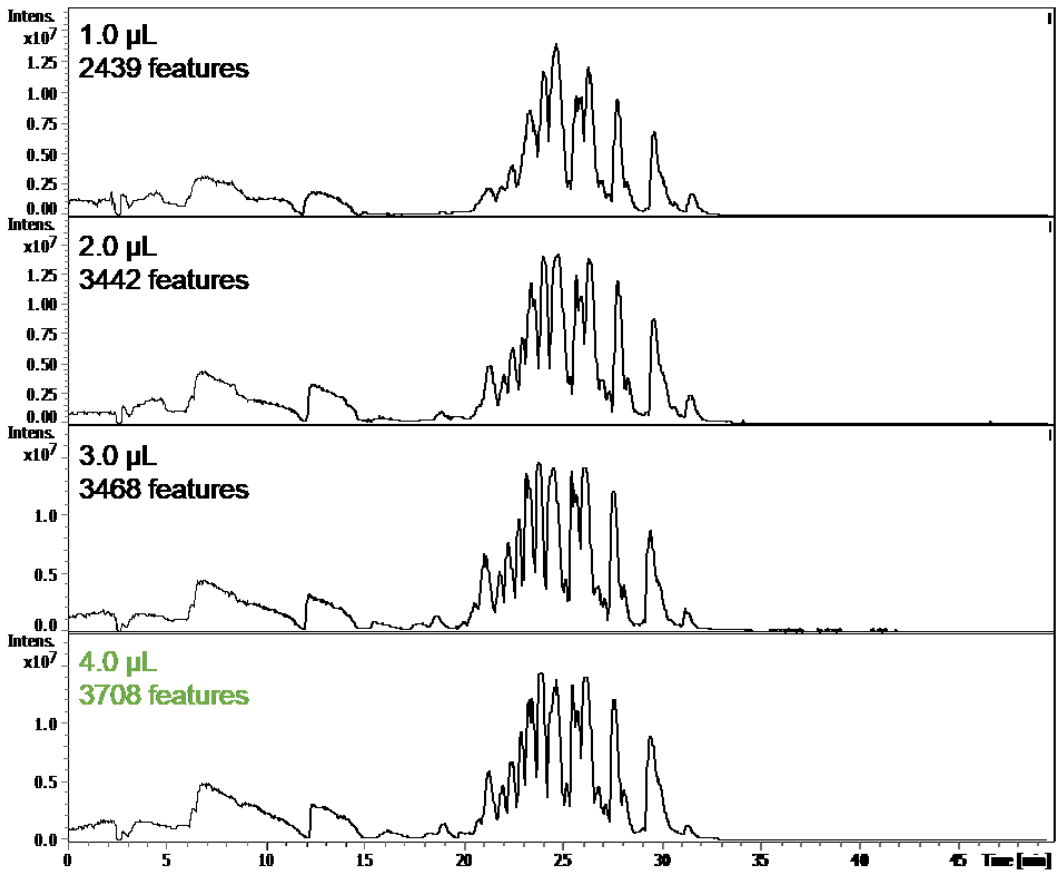

Figure A - 58. Optimization of the sample injection volume for the nanoLC-MS analysis. Optimized sample preparation: 2.5 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 53.3 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 31.9 µL of water and vortexed for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 2.5 µL of 1:1 MPA/MPB and diluted with 22.5 µL of 9:1 MPA/MPB. The nanoLC-MS methods are identical to Figure A - 49, except for the injection volume.

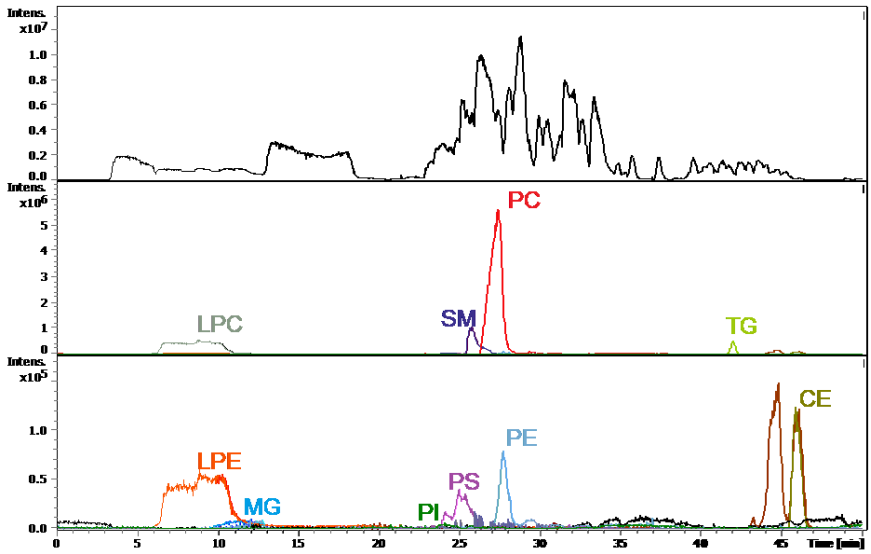

Figure A - 59. Base peak chromatogram for a pool of serum samples from 100 healthy humans and extracted ion chromatograms for the deuterated lipids employed as internal standards. Optimized nanoLC method: trapping at 5.0 µL/min for 1.25 min (100% MPA); MPA – 10 mM NH_4COOH in 50:40:10 methanol/ acetonitrile/ water (v/v/v); MPB – 10 mM NH_4COOH in 95:5 2-propanol/ water (v/v); 45ºC, 300 nL/min; 50 min gradient (0 min – 0% MPB, 10 min – 30% MPB, 20 min – 70% MPB, 25 min – 80% MPB, 30 min – 95% MPB, 40 min – 95% MPB, 42 min –0% MPB, 50 min –0% MPB), 5 min equilibrium (100% MPA); 4 µL injection. Optimized MS method: electrospray ion source capillary voltage of 1375 V, nanoBooster acetonitrile-enriched nitrogen gas pressure of 0.15 bar, dry nitrogen gas flow rate of 2.5 L/min, ion source temperature of 190ºC. Optimized sample preparation method: 2.5 µL of a pool of serum samples from 8 pigs were vortexed for 20 s with 2.5 µL of the mixture of 14 deuterated lipid standards (Table II-1) and 53.3 µL of methanol; 110.8 µL of dichloromethane was added, followed by vortex for 20 s; the mixture was washed with 31.9 µL of water and vortex for 10 s; after resting for 10 min at room temperature, the mixture was centrifuged for 10 min at 12,000 rpm and 4ºC; the bottom organic layer was evaporated to dryness on a SpeedVac for 30 min, resuspended with 2.5 µL of 1:1 MPA/MPB and diluted with 22.5 µL of 9:1 MPA/MPB.

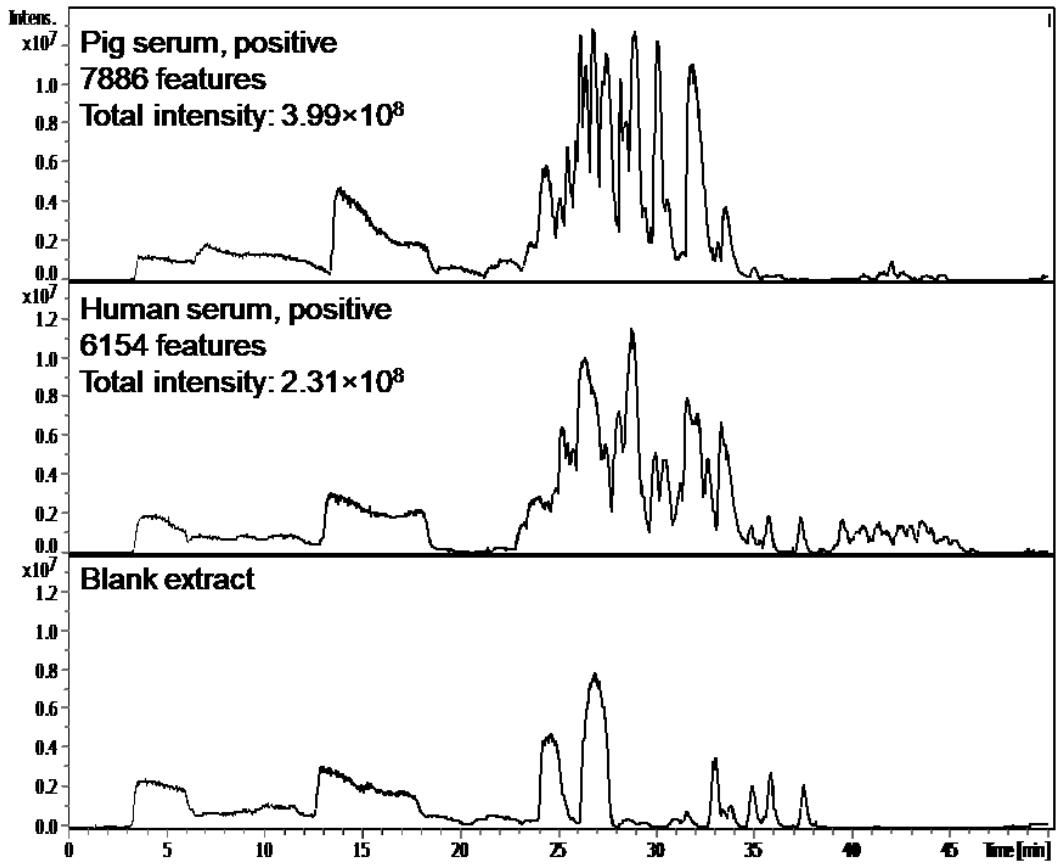

Figure A - 60. Comparison between chromatograms obtained for a pool of serum samples from 8 pigs (top); serum from 100 healthy humans (mid); and a blank extract (extract of water instead of the sample, bottom chromatogram) for positive ionization. The nanoLC-MS and sample preparation methods are described in Figure A - 59.

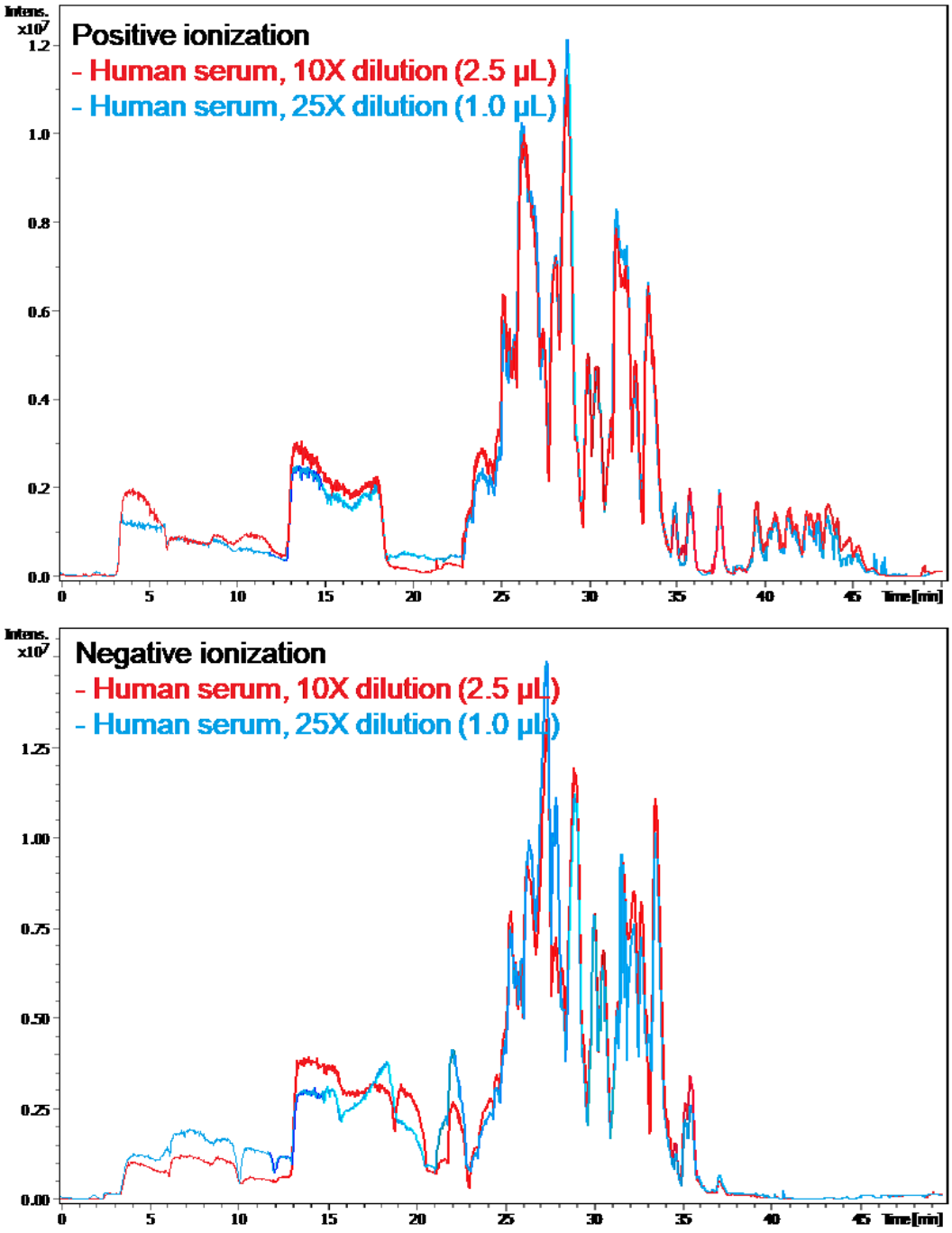

Figure A - 61. Comparison of base peak chromatograms obtained for 1.0 µL (25× dilution) and 2.5 µL (10× dilution) of a pool of serum samples obtained from 100 healthy humans with positive and negative ionization. The nanoLC-MS and sample preparation methods are described in Figure A - 59.

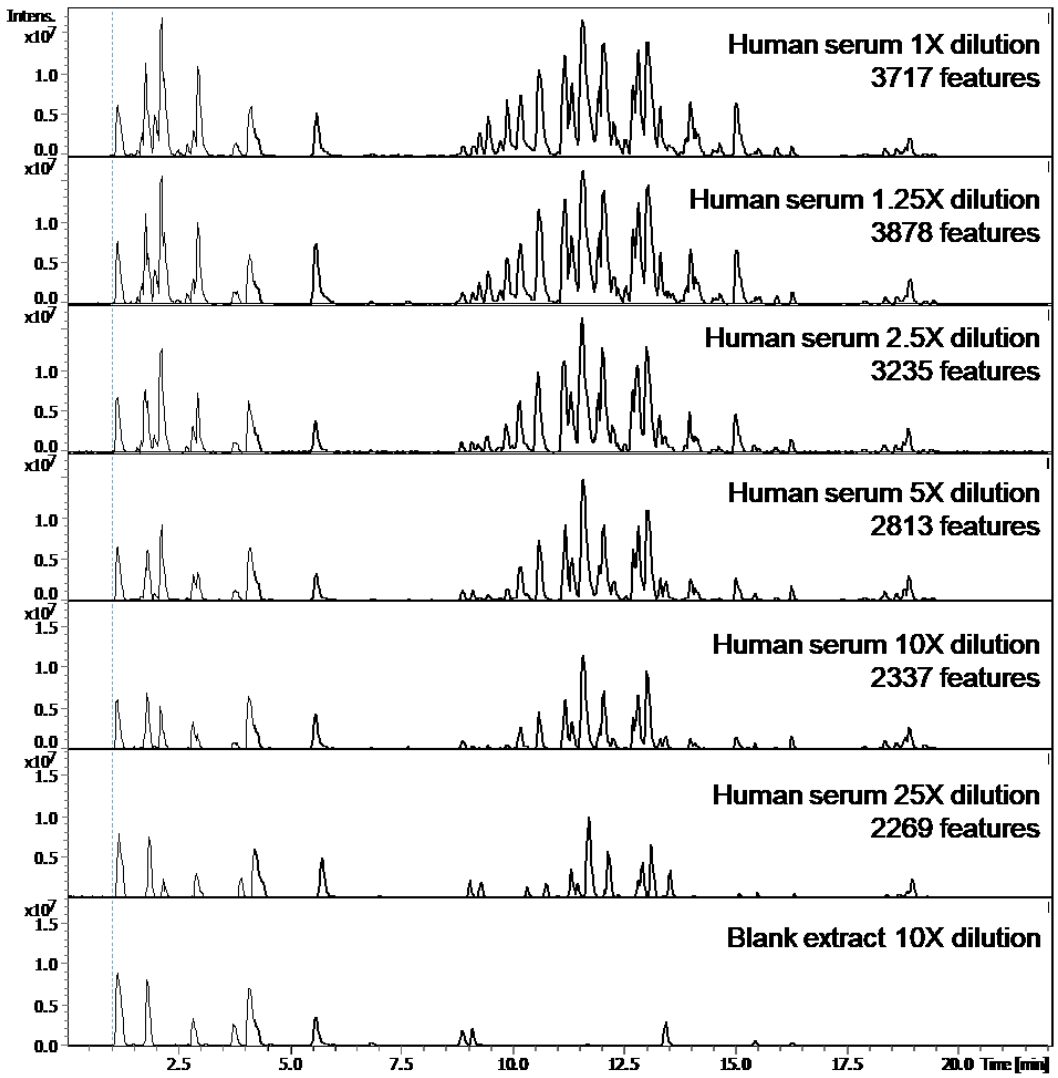

Figure A - 62. Base peak chromatograms obtained for a pool of serum samples from 100 healthy humans with UHPLC-MS (positive ionization). The UHPLC-MS method is described in the Experimental Section.

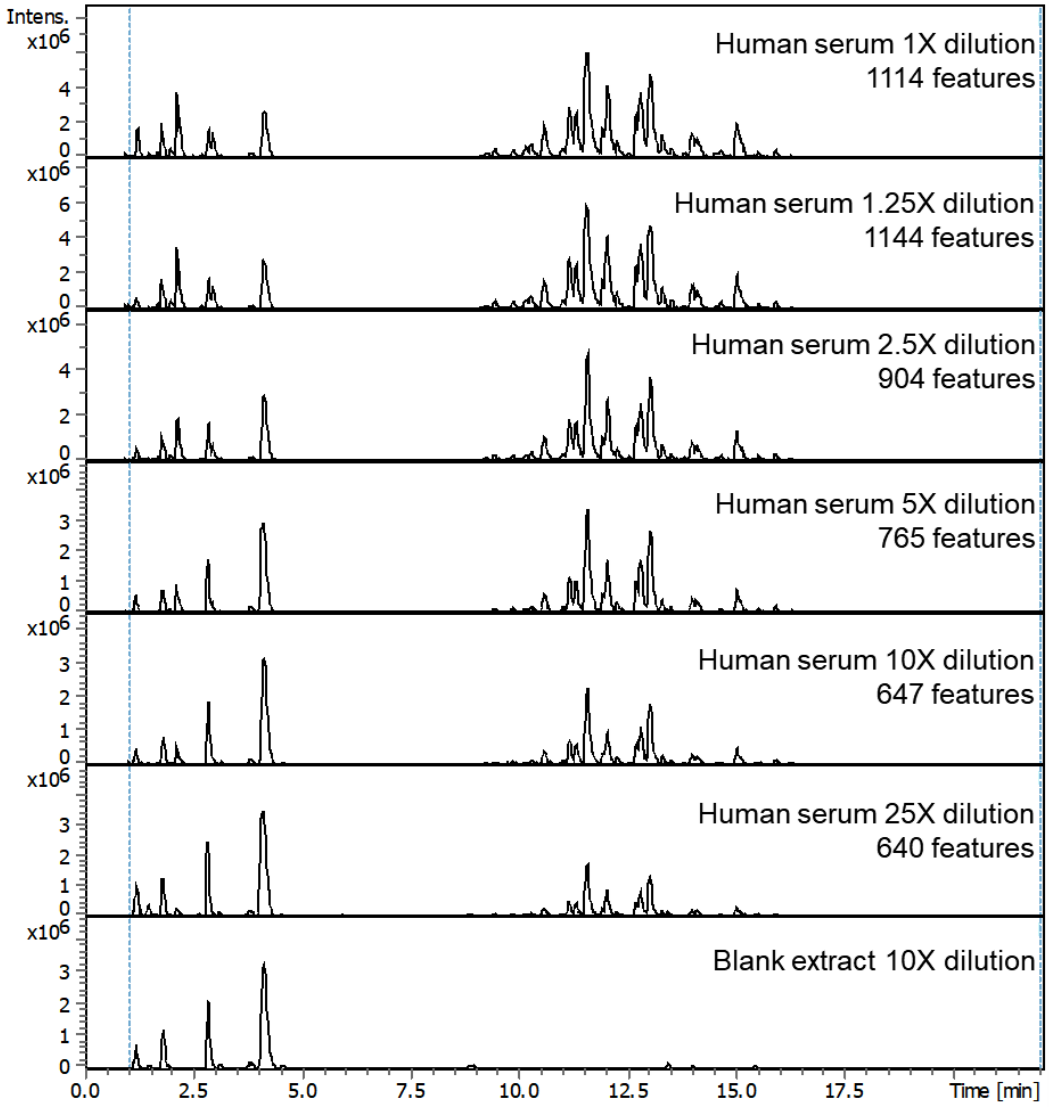

Figure A - 63. Base peak chromatograms obtained for a pool of serum samples from 100 healthy humans with UHPLC-MS (negative ionization). The UHPLC-MS method is described in the Experimental Section.

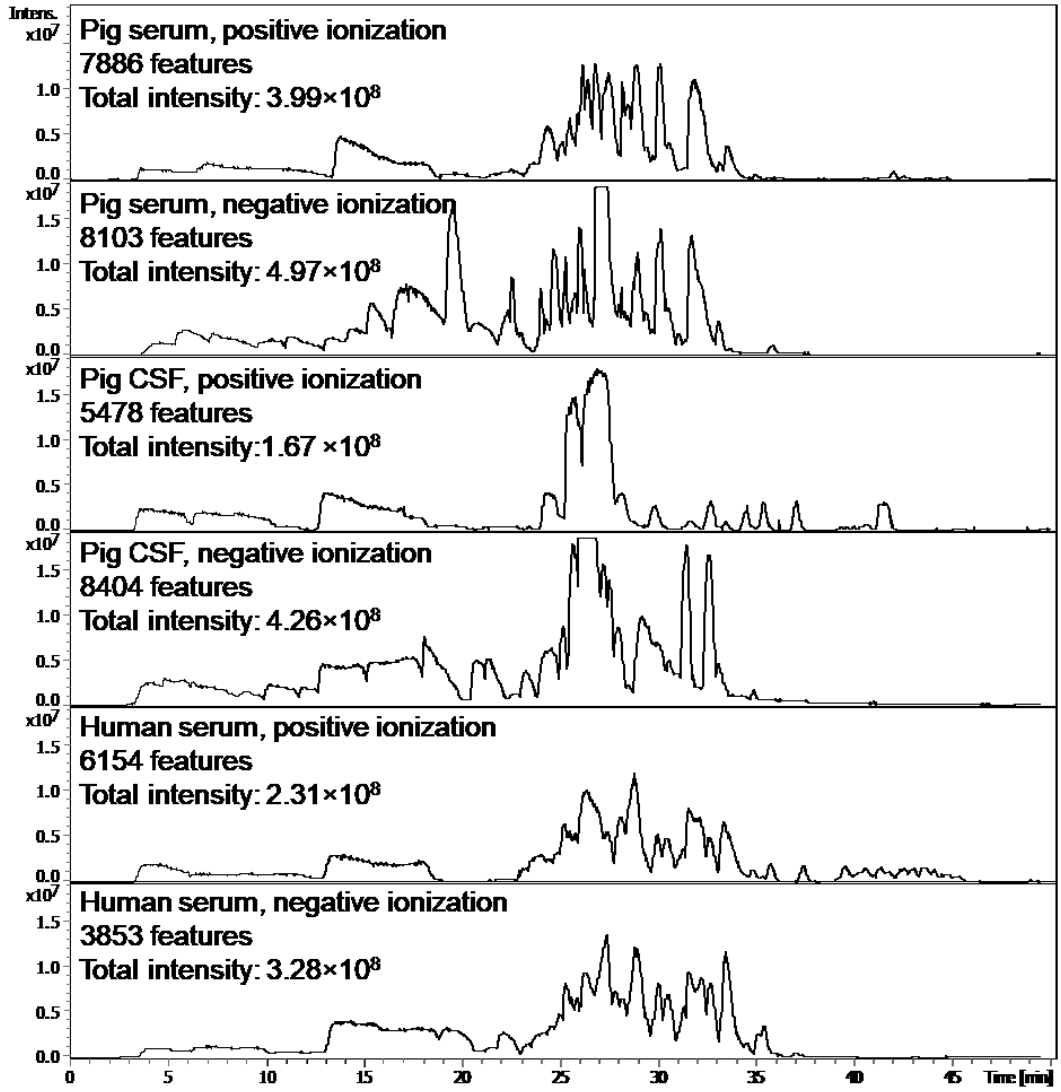

Figure A - 64. Base peak chromatograms obtained for a pool of serum or cerebrospinal fluid samples from 8 pigs, and a pool of serum from 100 healthy humans with positive and negative ionization. The nanoLC-MS and sample preparation methods are described in Figure A - 59.

Appendix B

Appendix B: Supplementary Figures for Chapter 5

LIST OF FIGURES

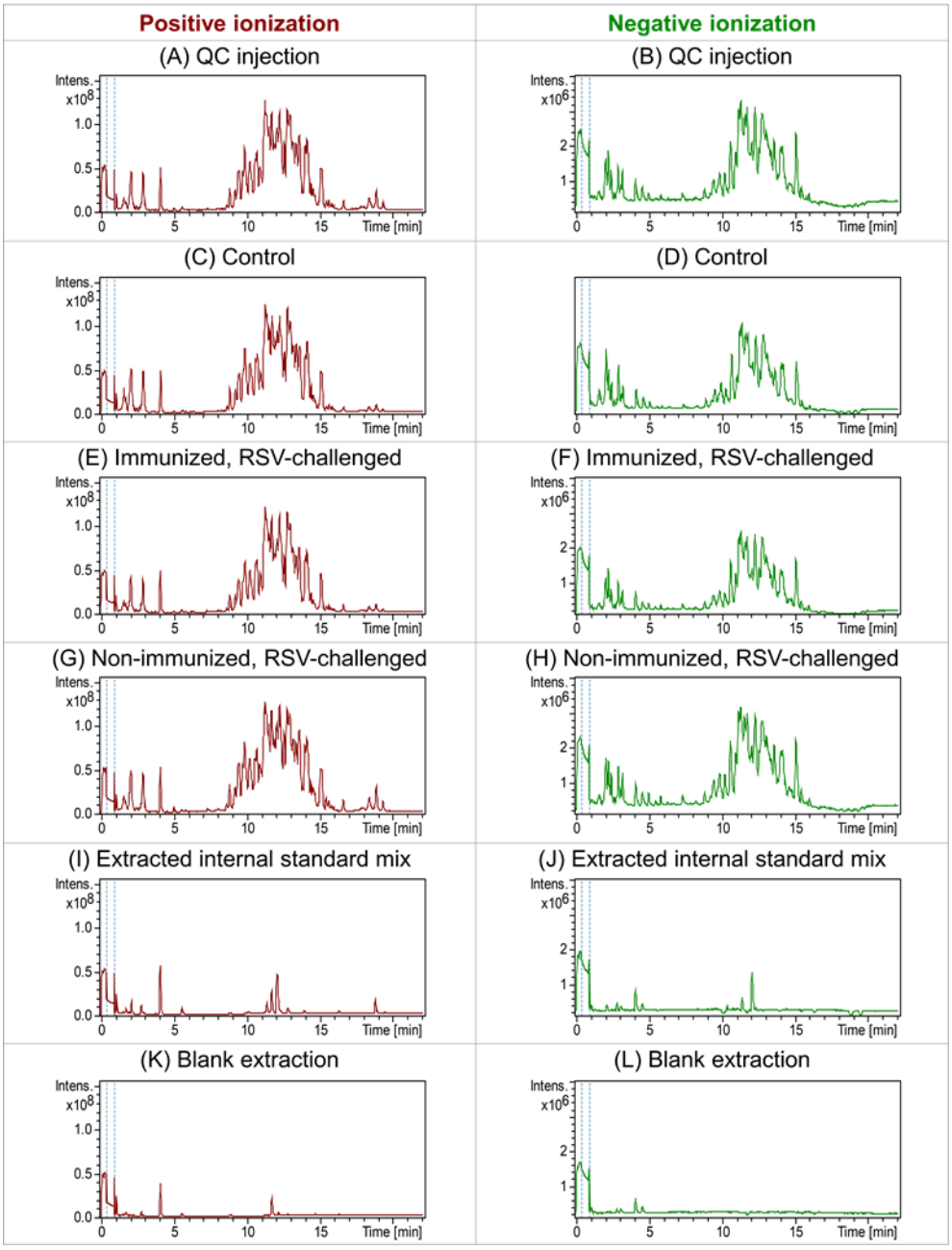

Figure B- 1. Examples of total ion chromatograms obtained under positive and negative ionization for: (A) and (B) QC injections; (C) and (D) the control group; (E) and (F) the immunized, RSV-challenged mice; (G) and (H) the non-immunized, RSV-challenged mice; (I) and (J) extracted internal standard mix composed of 14 deuterated lipids (Table II-1); and (K) and (L) a blank extract (extraction of water instead of sample, following the identical sample preparation procedure).

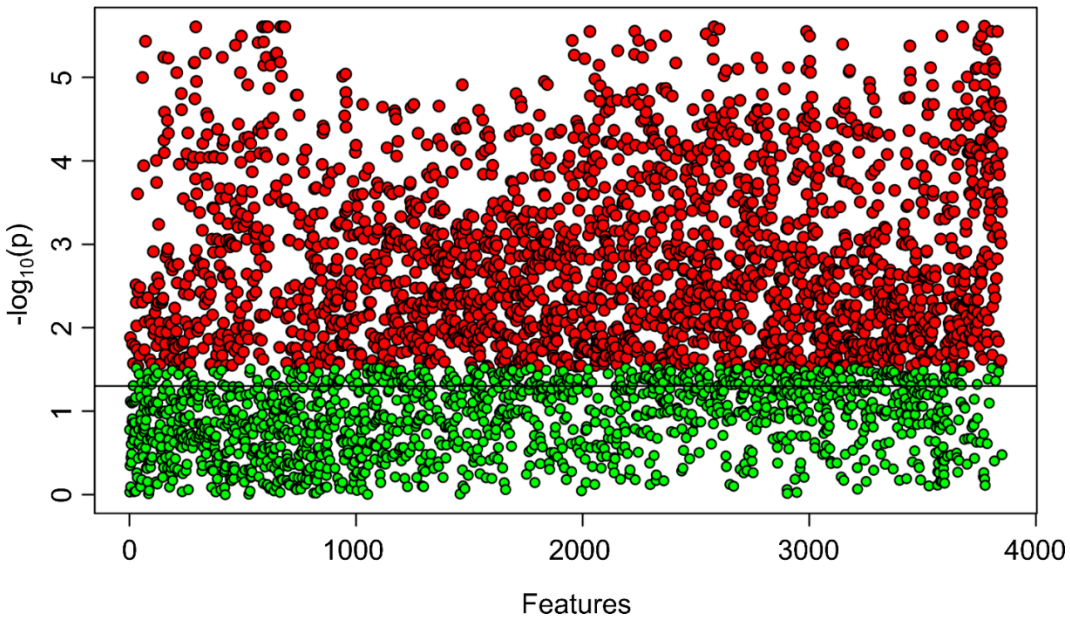

Figure B- 2. Non-parametric ANOVA (Kruskal-Wallis test) for lipidomics of lung tissue of control (group A); immunized, RSV-challenged (group B); and non-immunized, RSV -challenged (group C) mice. The 2316 significant lipids ($p < 0.05$) are shown in red, while non-significant lipids ($p > 0.05$) are in green.

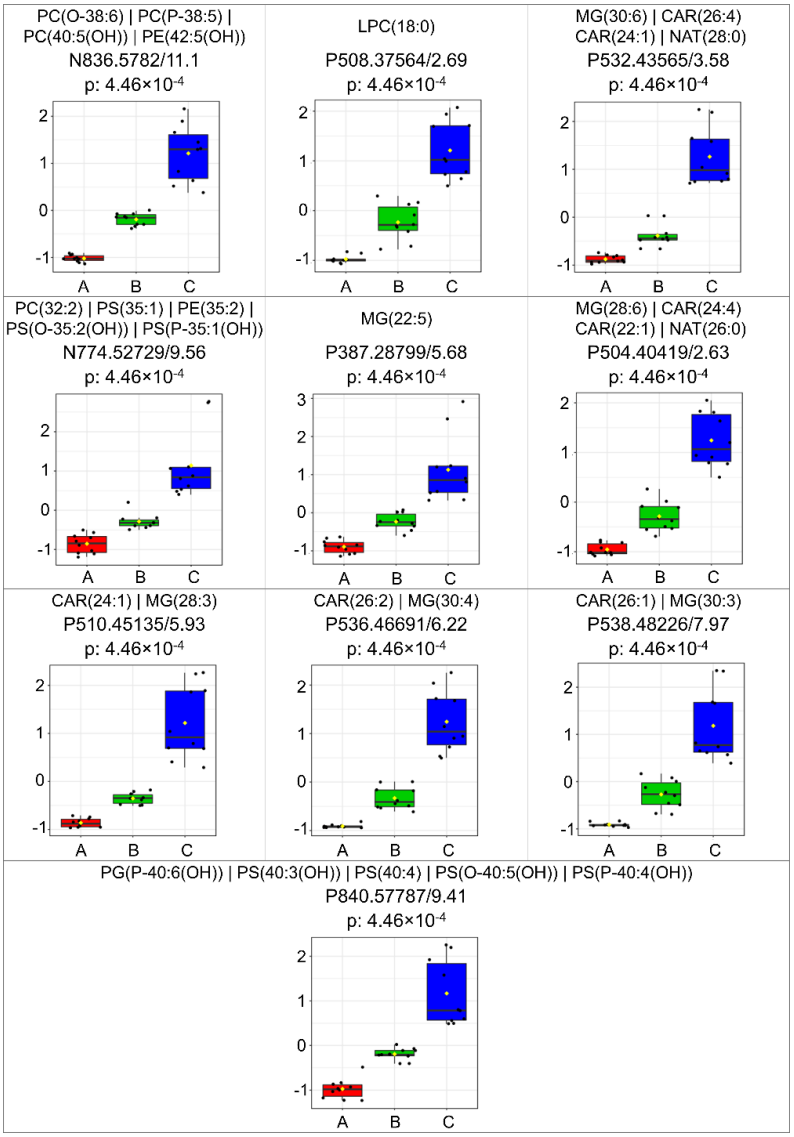

Figure B- 3. Boxplots for ten of the most significant lipids (lowest p values) found by non-parametric ANOVA for lipidomics of lung tissue of control (group A, red); immunized, RSV-challenged (group B, green) and non-immunized, RSV-challenged (group C, blue) mice. Normalized intensities (intensity of each lipid divided by the intensity of the class-matched internal standard) were auto-scaled for statistical analysis. For identification, all isomeric and isobaric possibilities that passed the retention time and adduct filters were ordered according to the filtering layers described in the Experimental section (Chapter V, *5.2.4 Lipidomics*), but a maximum of five are shown.

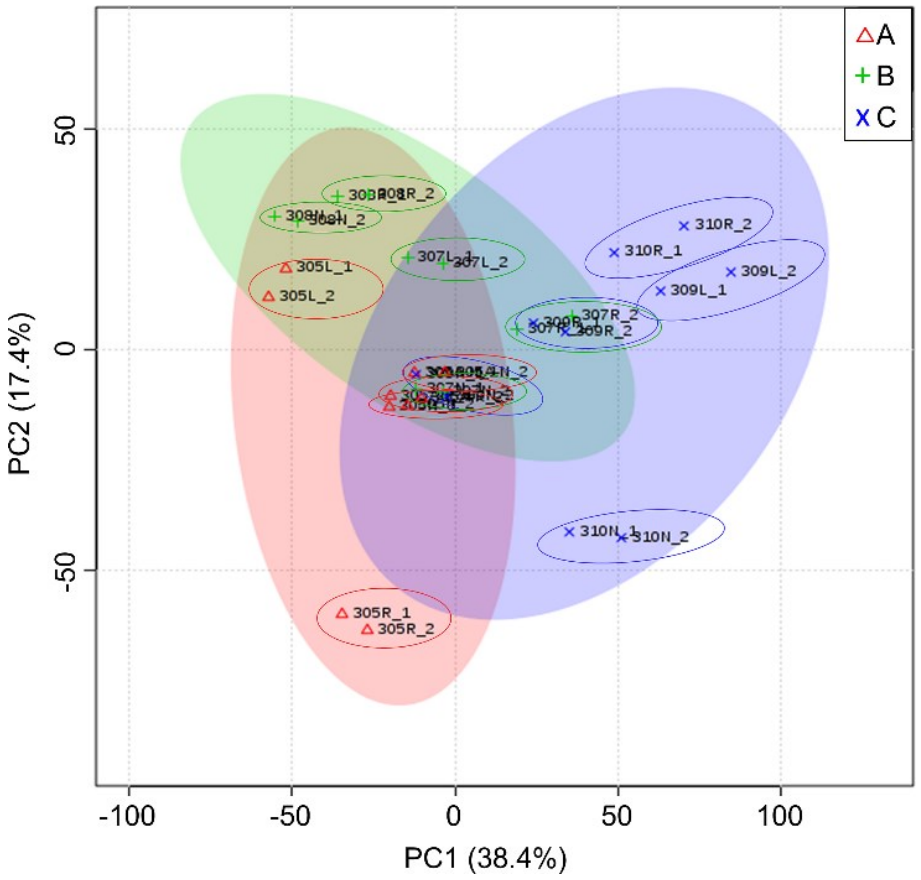

Figure B- 4. PCA score plot obtained for untargeted lipidomics of lung tissue from control (group A, red); immunized, RSV-challenged (group B, green); and non-immunized, RSV-challenged (group C, blue) mice. The sample names of injection duplicates are circled to emphasize their proximity.

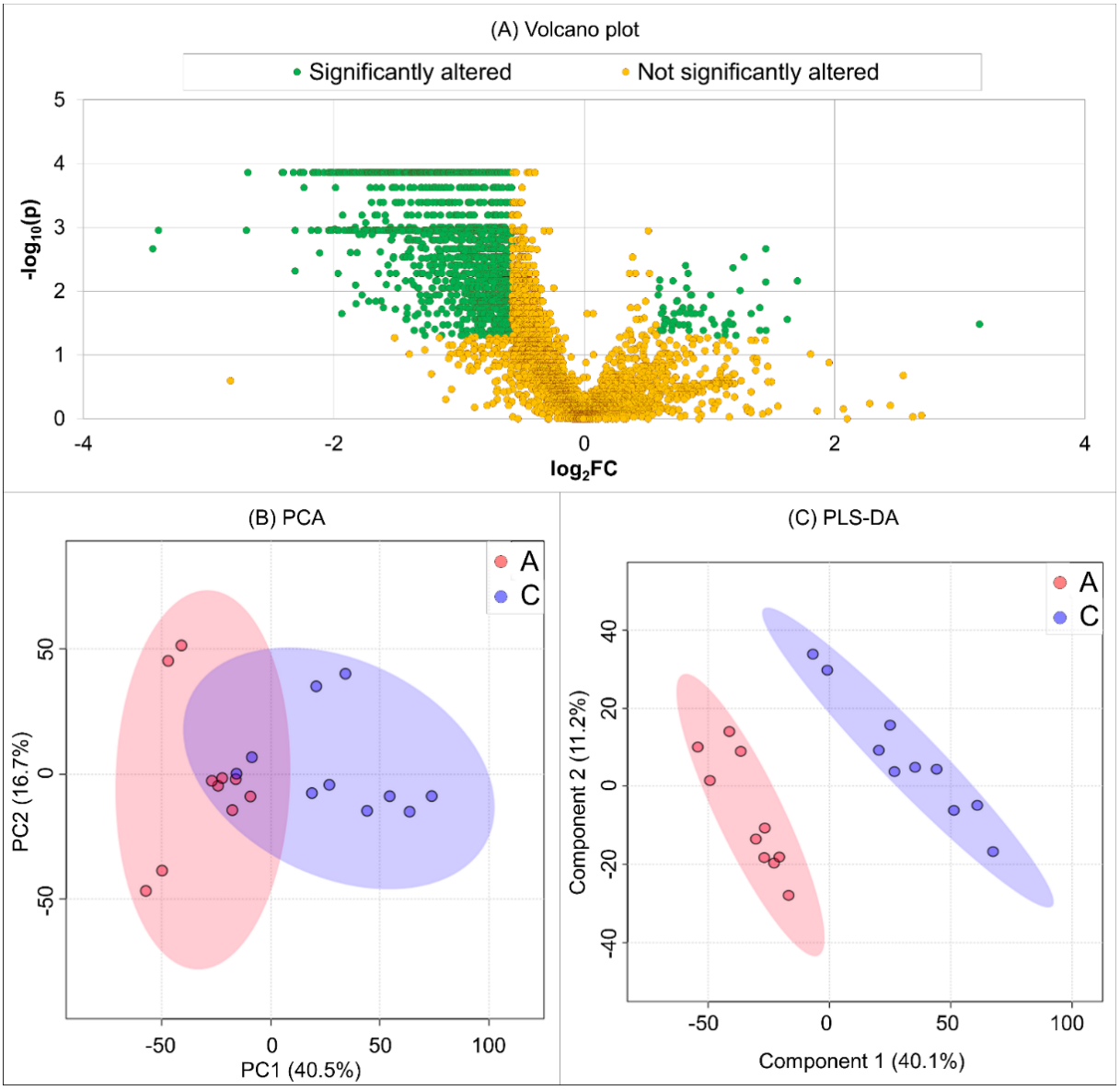

Figure B- 5. (A) Volcano plot, (B) PCA score plot and (C) PLS-DA score plot obtained by lipidomics for healthy controls (group A) *versus* non-immunized, RSV-challenged mice (group C). The PLS-DA model was constructed with 2 components and resulted in R^2 of 0.9616, Q^2 of 0.8735 and p=0.02 for 1000 permutations.

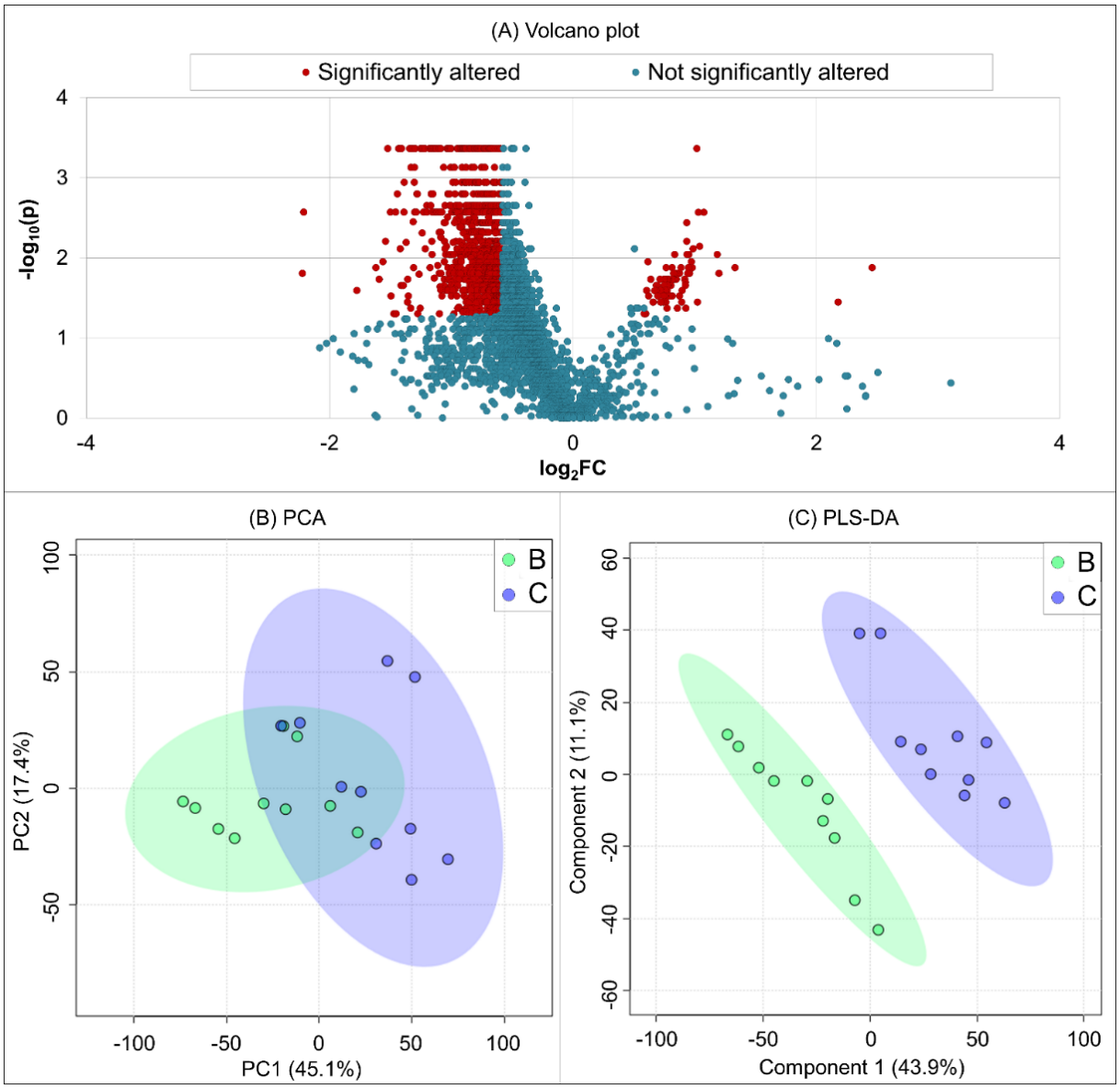

Figure B- 6. (A) Volcano plot, (B) PCA and (C) PLS-DA score plots obtained by lipidomics for immunized (group B) *versus* non-immunized (group C), RSV-challenged mice. The PLS-DA model was constructed with 2 components and resulted in R^2 of 0.9451, Q^2 of 0.8551 and p=0.02 for 1000 permutations.

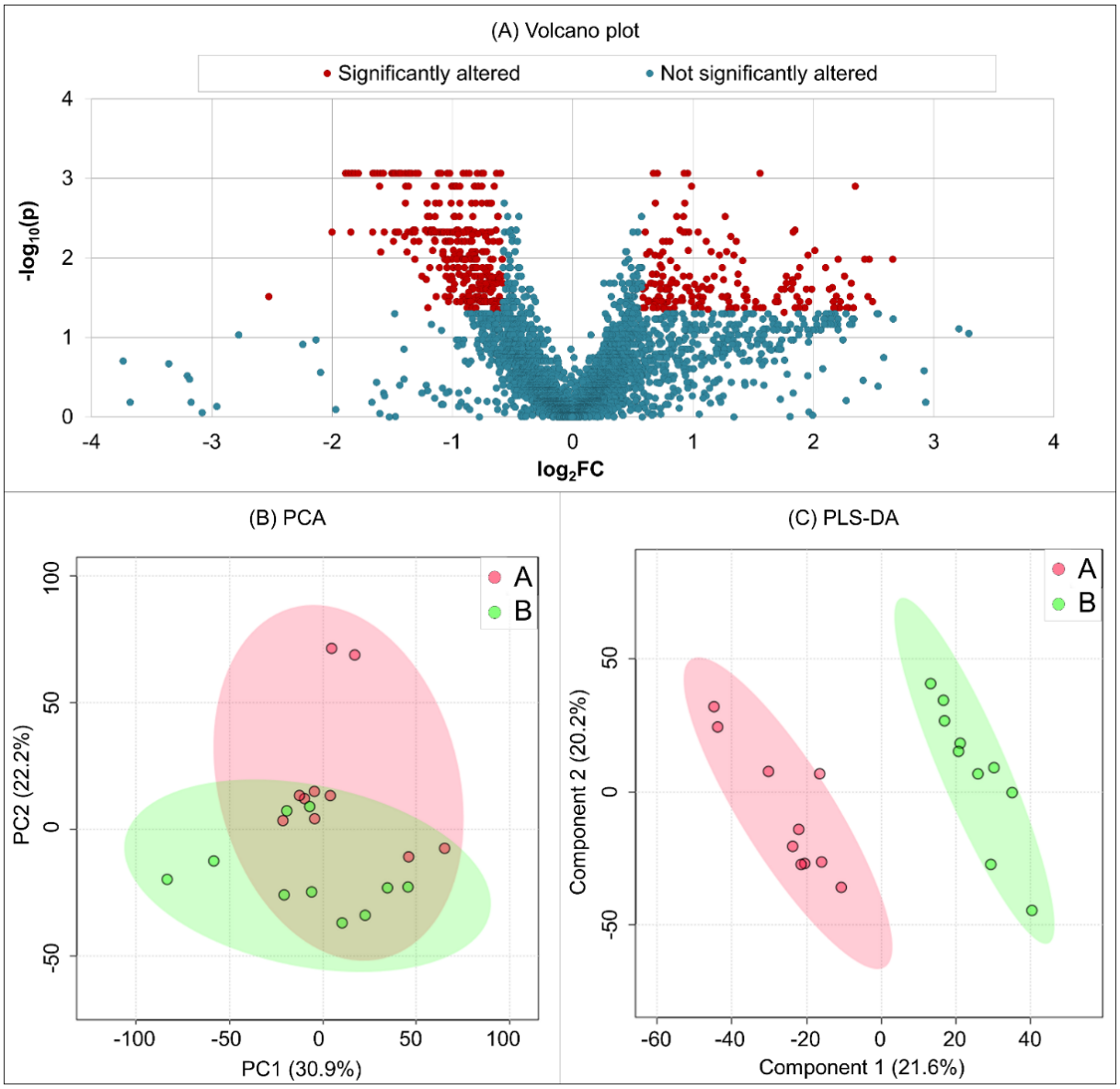

Figure B- 7. (A) Volcano plot, (B) PCA and (C) PLS-DA score plots obtained by lipidomics for healthy controls (non-immunized and unchallenged, group A) *versus* immunized, RSV-challenged mice (group B). The PLS-DA model was constructed with 2 components and resulted in R^2 of 0.9691, Q^2 of 0.9149 and p=0.02 for 1000 permutations.

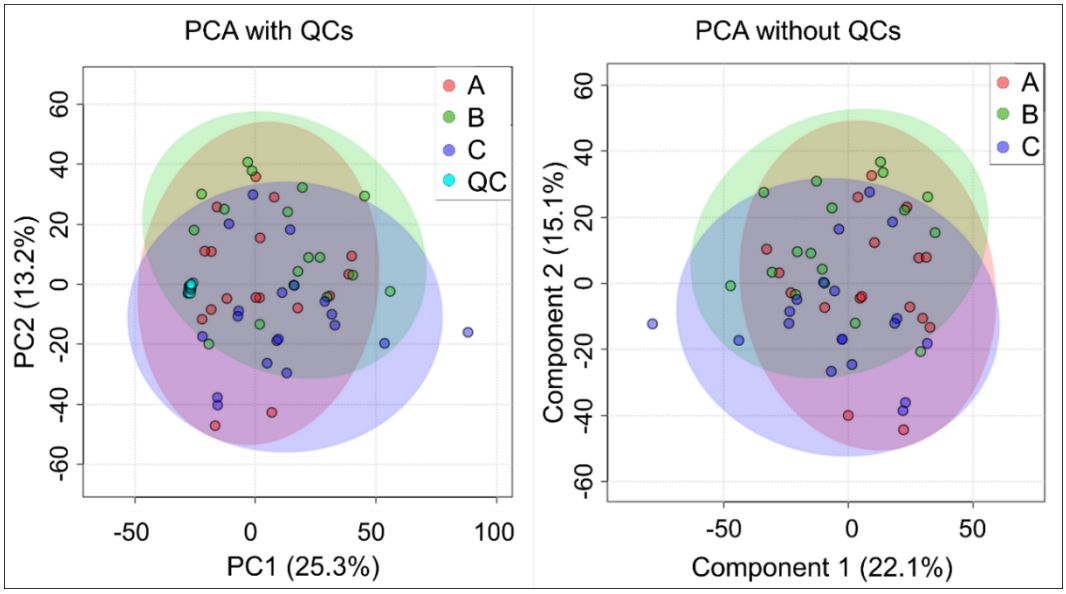

Figure B- 8. PCA score plot obtained for DmPA labeling of lung tissue samples from healthy control (A, red); immunized, RSV-challenged (B, green); and non-immunized, RSV-challenged (C, blue) mice.

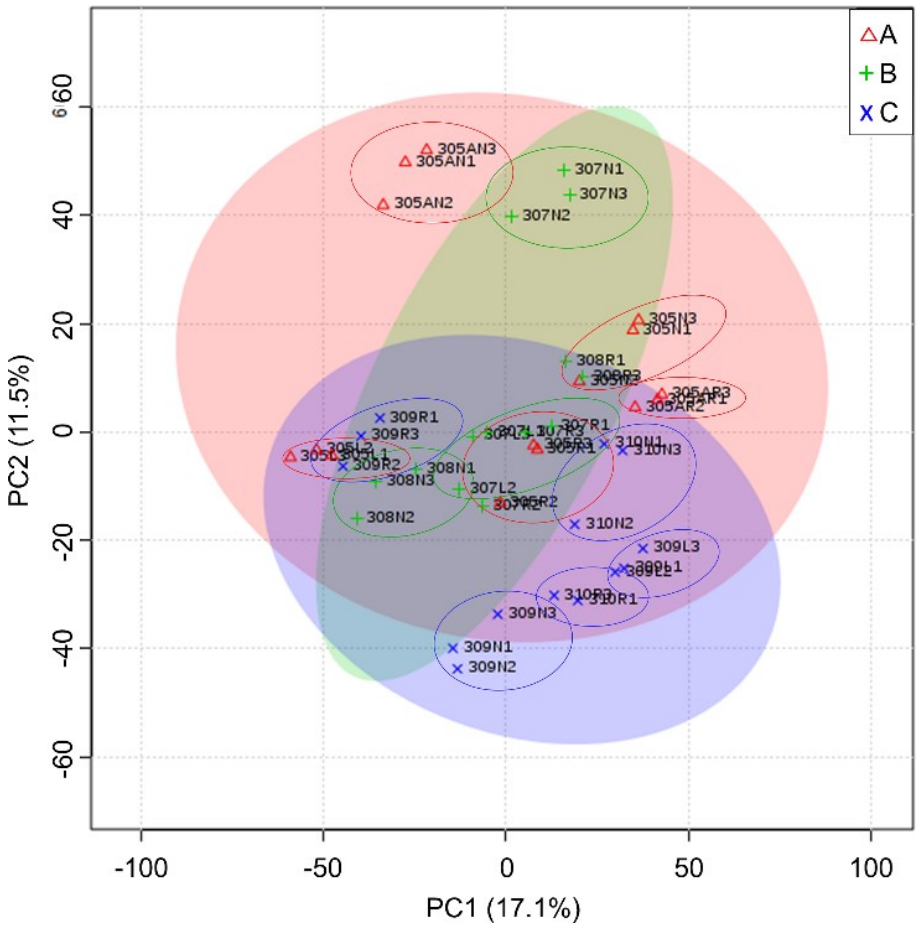

Figure B- 9. PCA score plot obtained for Dns and DmPA labeling of lung tissue samples from healthy control (A, red); immunized, RSV-challenged (B, green); and non-immunized, RSV-challenged (C, blue) mice. The fractions obtained after homogenization were split for labeling triplicates, which are indicated by the colored circles.

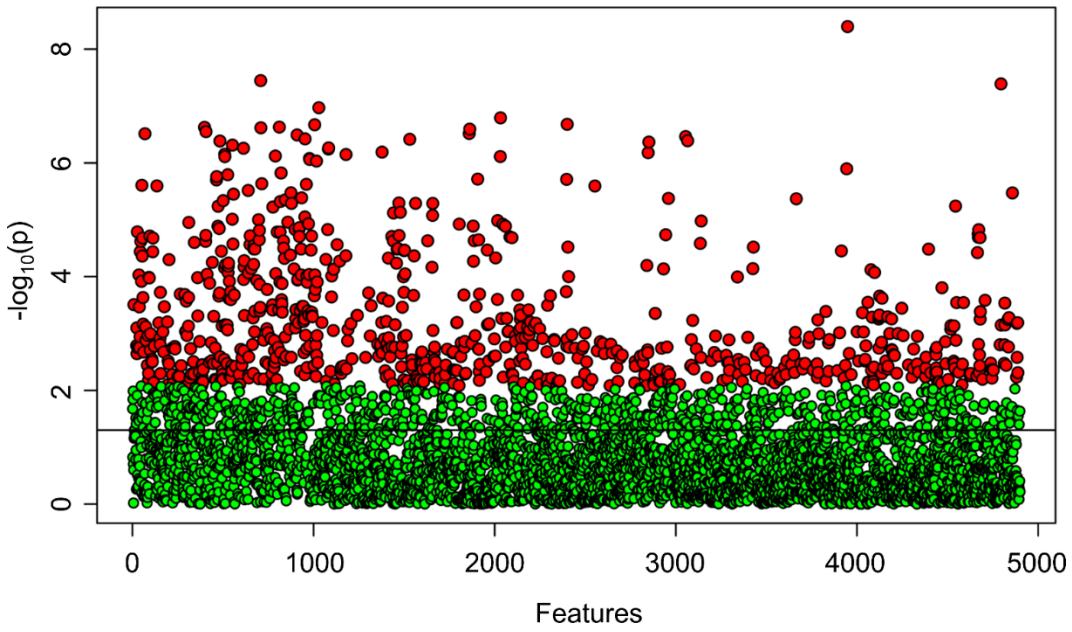

Figure B- 10. Non-parametric ANOVA (Kruskal-Wallis test) for metabolomics of lung tissue via DnsCl and DmPA chemical isotope labeling for control (group A); immunized, RSV-challenged (group B); and non- immunized, RSV-challenged (group C) mice. The 926 significant metabolites ($p < 0.05$) are shown in red, while non-significant lipids ($p > 0.05$) are represented by green circles.

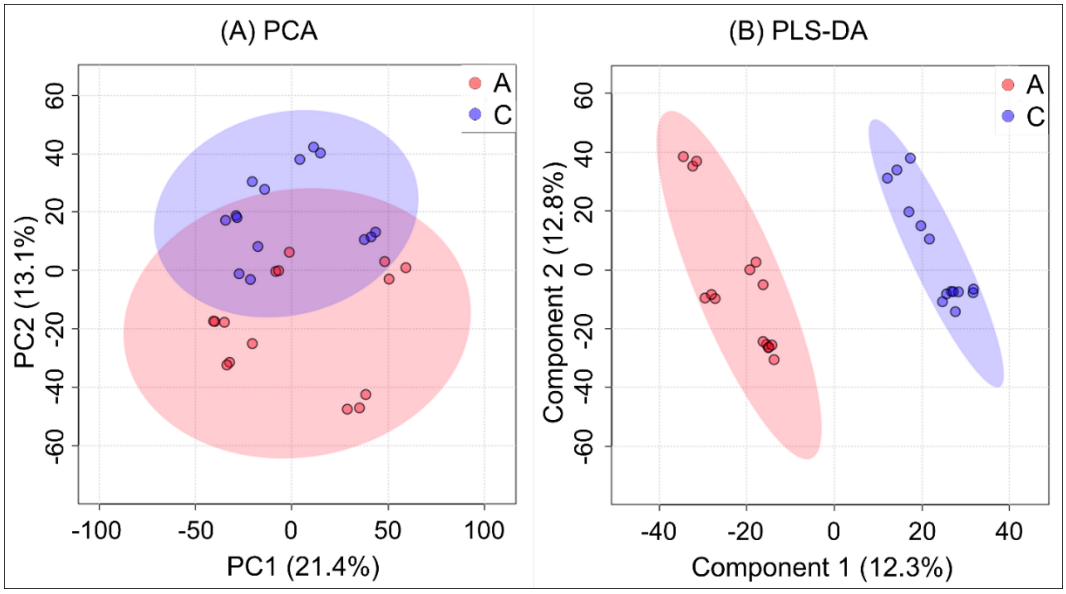

Figure B- 11. (A) PCA and (B) PLS-DA score plots obtained for amine, phenol and carboxylic acid-containing metabolites for the comparison between healthy controls (non-immunized, non-challenged, group A, red) *versus* non-immunized, RSV-challenged mice (group C, blue). The PLS-DA model was constructed with 2 components and resulted in R^2 of 0.9784, Q^2 of 0.9037 and p=0.05 for 1000 permutations.

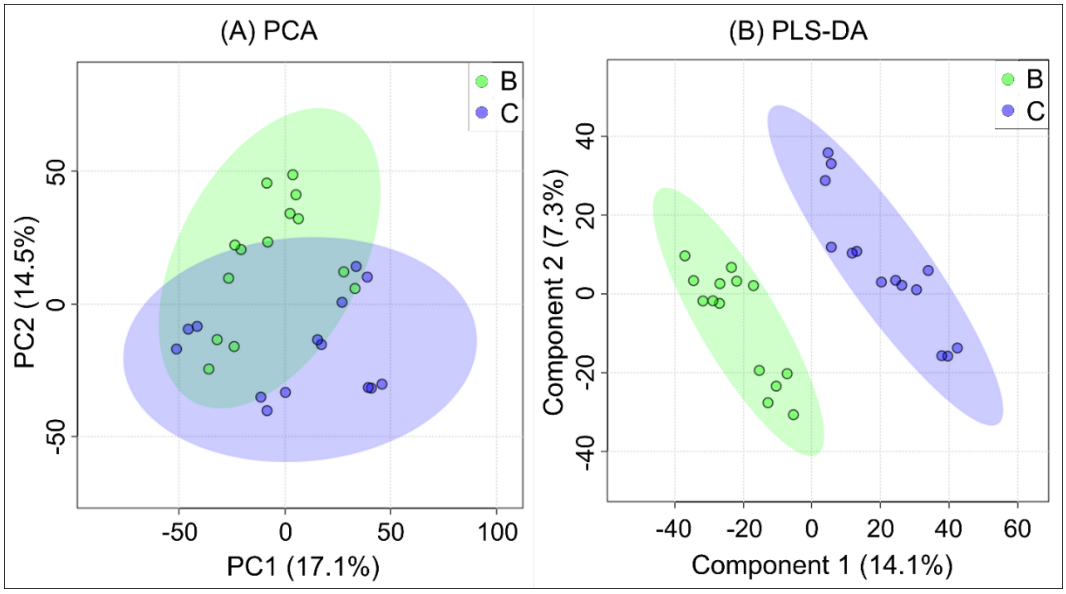

Figure B- 12. (A) PCA and (B) PLS-DA score plots obtained for amine, phenol and carboxylic acid-containing metabolites for the comparison between immunized (group B, green) *versus* non-immunized, RSV-challenged mice (group C, blue). The PLS-DA model was constructed with 2 components and resulted in R^2 of 0.9566, Q^2 of 0.7545 and p=0.05 for 1000 permutations.

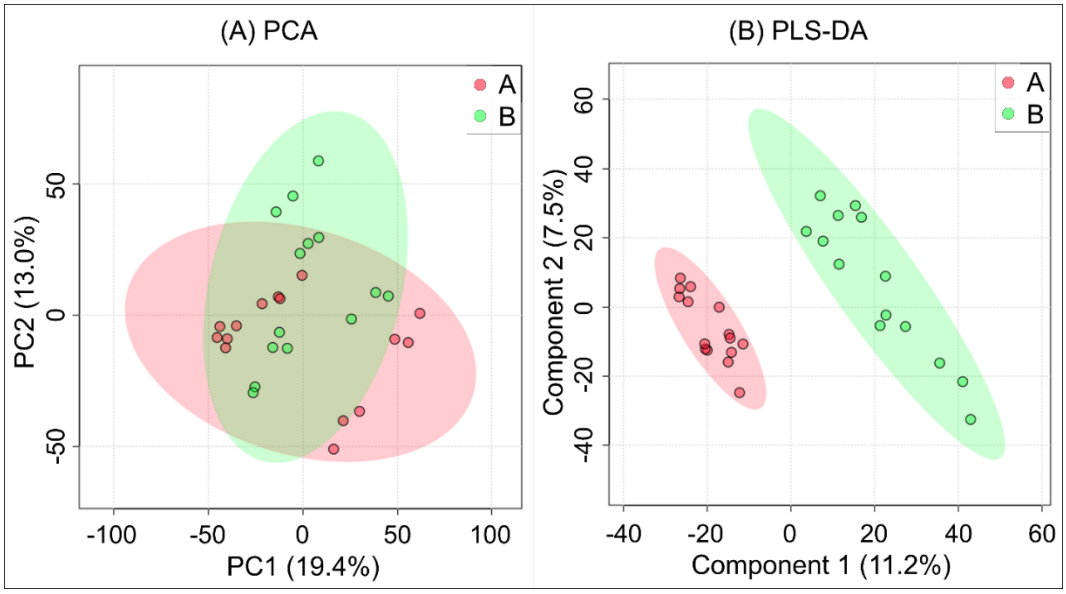

Figure B- 13. PCA and PLS-DA score plots obtained for amine, phenol and carboxylic acid-containing metabolites for the comparison between healthy controls (non-immunized, unchallenged, group A, red) *versus* immunized, RSV-challenged mice (group B, green). The PLS-DA model was constructed with 2 components and resulted in R^2 of 0.9700, Q^2 of 0.8246 and p=0.05 for 1000 permutations.

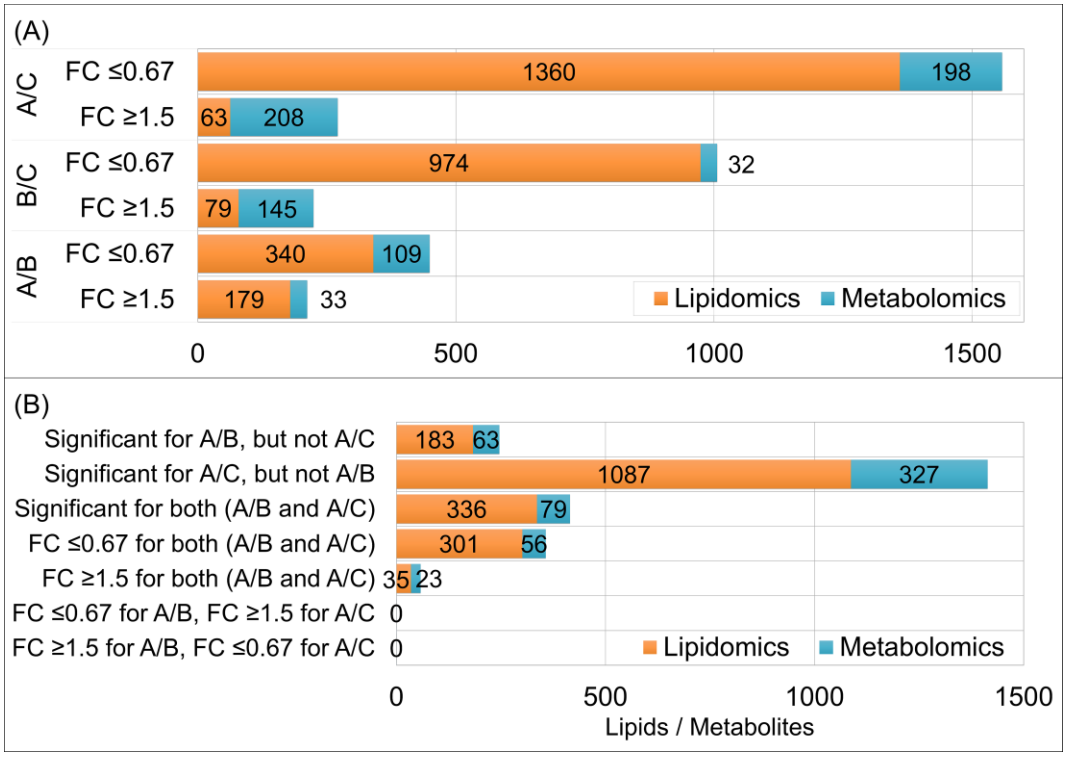

Figure B- 14. Significantly altered lipids and metabolites ($p < 0.05$) for the binary comparison between healthy control mice (group A); immunized, RSV-challenged mice (group B); and non-immunized, RSV-challenged mice (group C). (A) Binary comparisons for the three studied groups; (B) Significantly altered lipids and metabolites for the immunized (group B) and non-immunized (group C), RSV-challenged mice, when compared to healthy controls.

Appendix C

Appendix C: Supplementary Figures for Chapter 6

LIST OF FIGURES

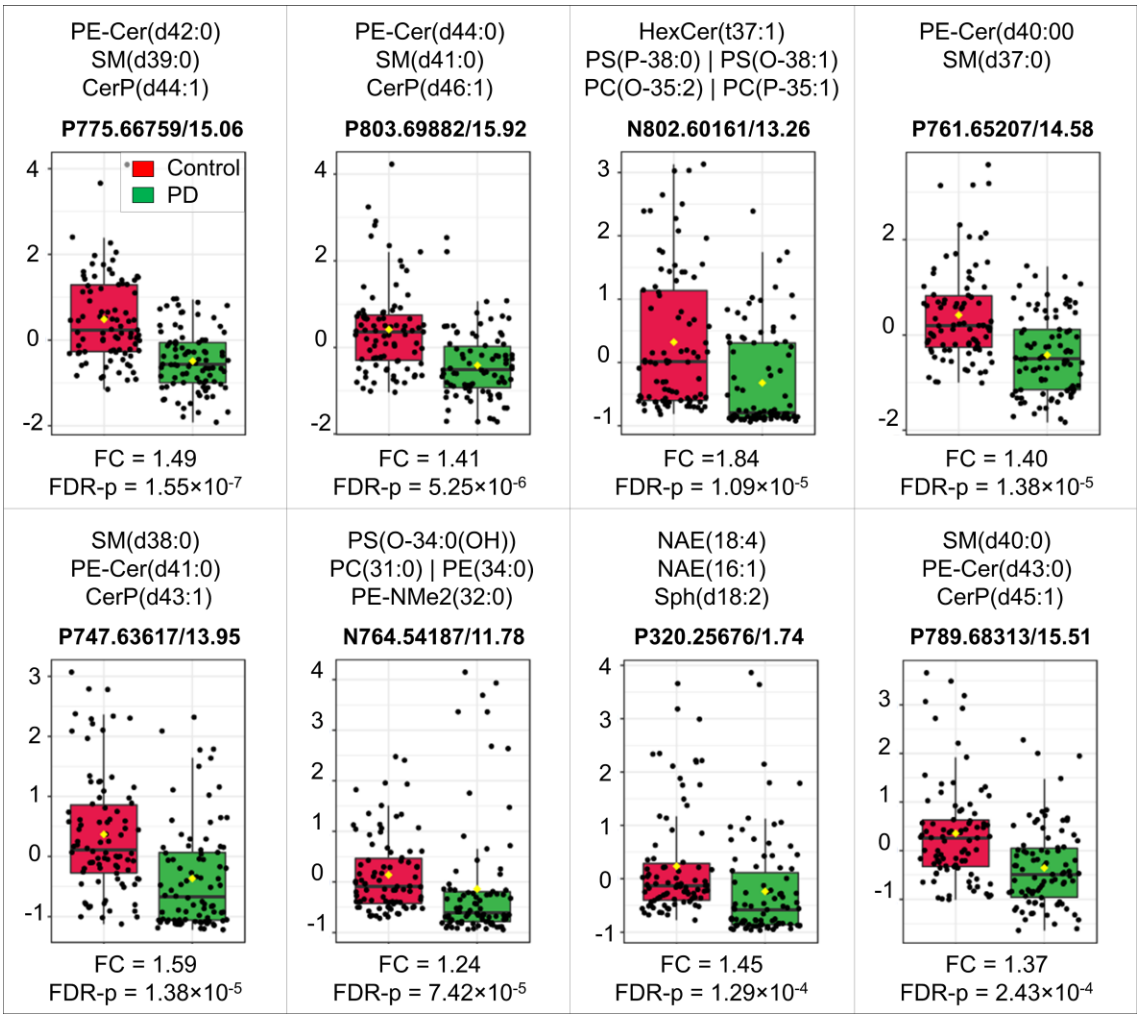

Figure C-1. Boxplots for eight of the most significantly altered lipids for the separation between healthy control subjects (red) and PD patients (green). Features are considered significantly altered if fold change (FC) healthy control / PD ≥1.2 or <0.84 and p <0.05. FDR-p: p value adjusted for false-discovery rate.

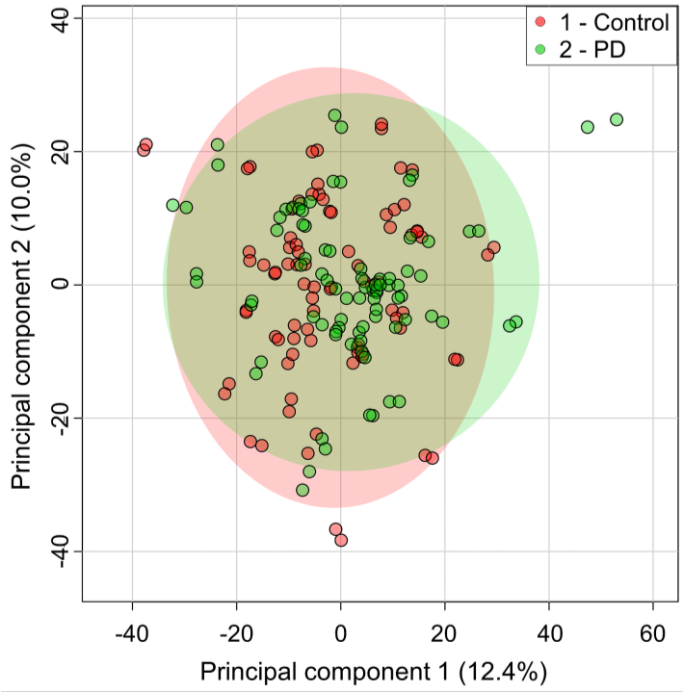

Figure C-2. PCA score plot for the separation between healthy control subjects (red) and PD patients (green). Experimental duplicates for each sample were tightly clustered.

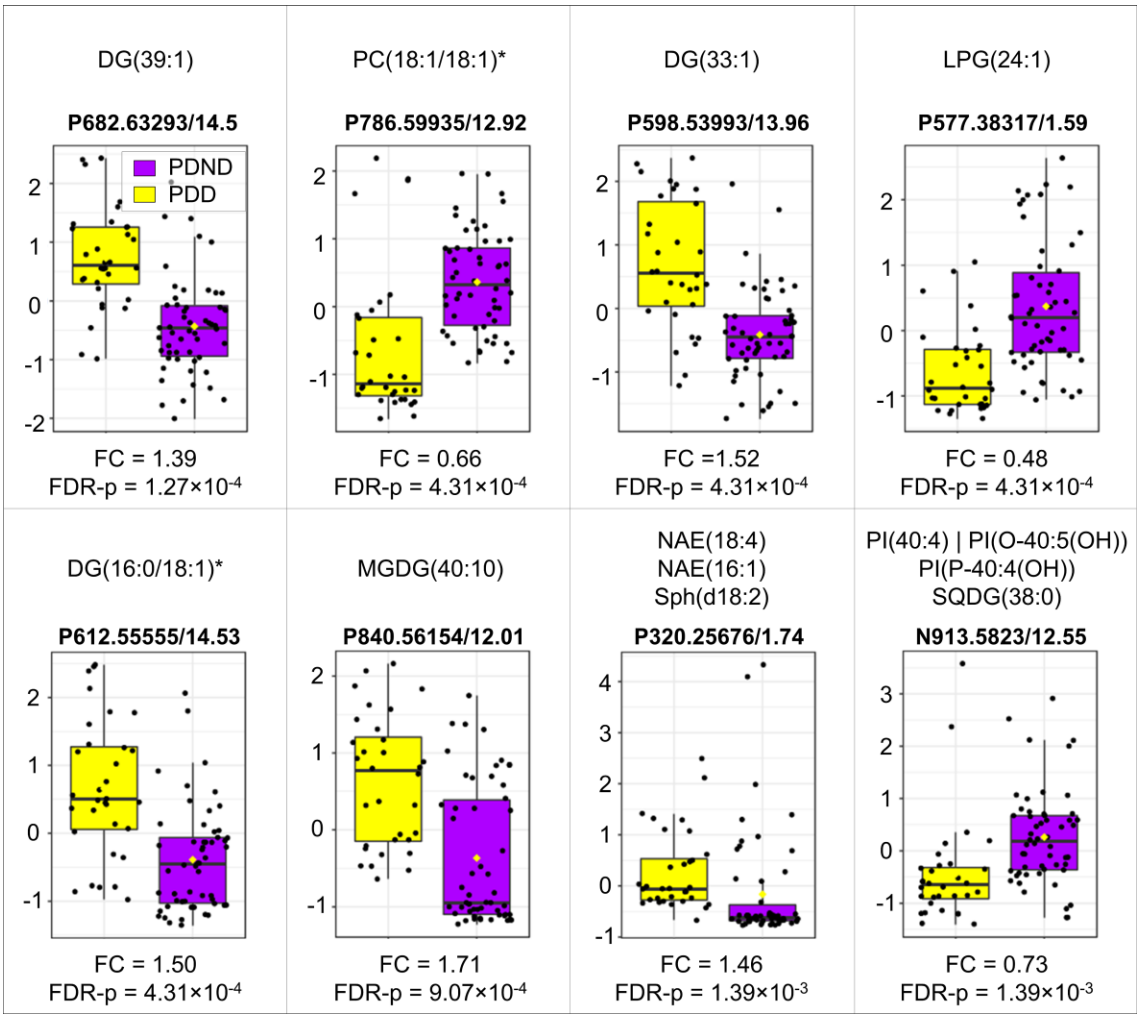

Figure C-3. Boxplots for eight of the most significantly altered lipids for the separation between PD patients with no dementia (PDND, purple) and clinically diagnosed with dementia 3 years after sample collection (incipient dementia, PDD, yellow). Features are considered significantly altered if FC for PDND/PDD ≥1.2 or <0.84 and p <0.05.

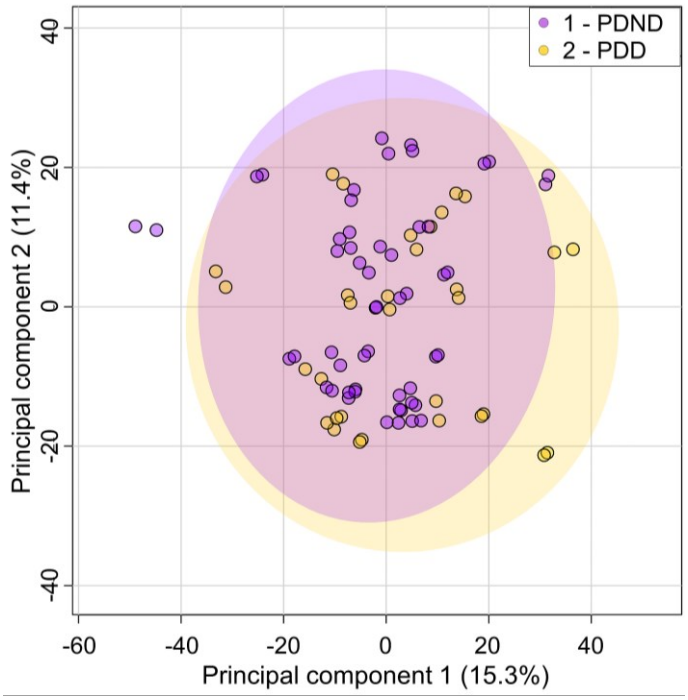

Figure C-4. PCA score plot for the separation between PDND (purple) and PDD patients (yellow). Experimental duplicates for each sample were tightly clustered.

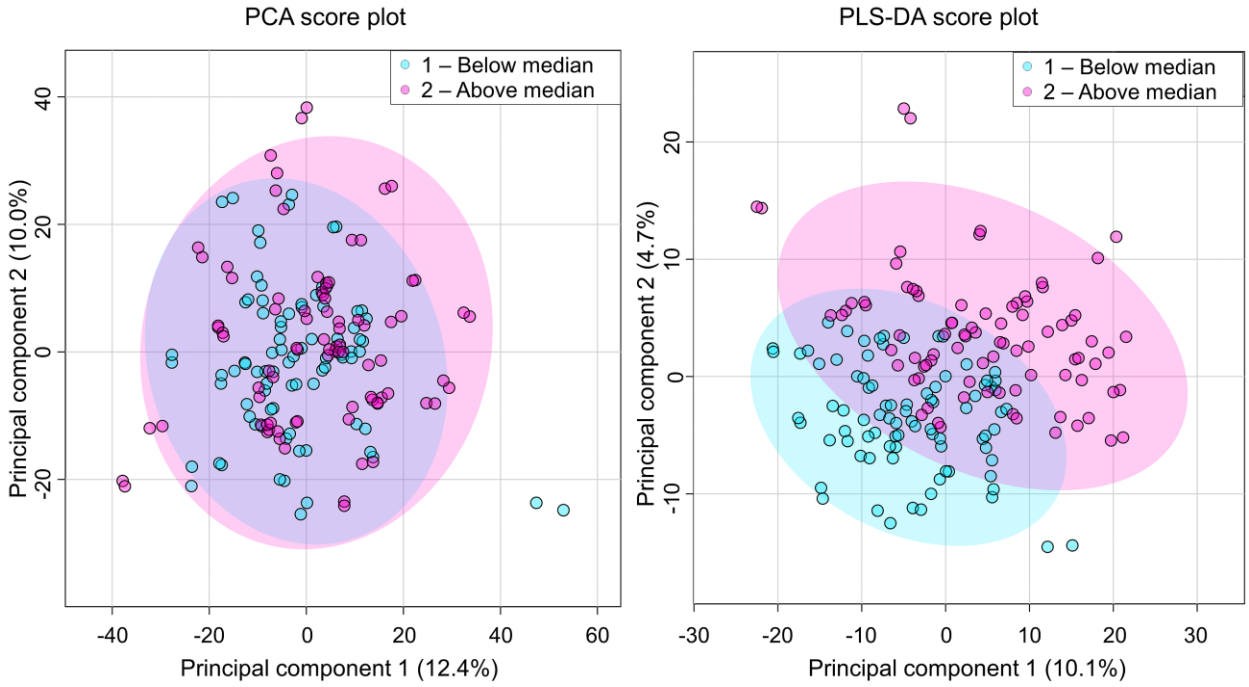

Figure C-5. PCA and PLS-DA score plot for the separation between healthy controls and PD patients aged below (Group 1, cyan) and above (Group 2, magenta) the median of 69.97 years old. The PLS-DA model with 8 components resulted in R^2 of 0.9428, Q^2 of 0.6837 and p of <0.001 for 1000 permutations.

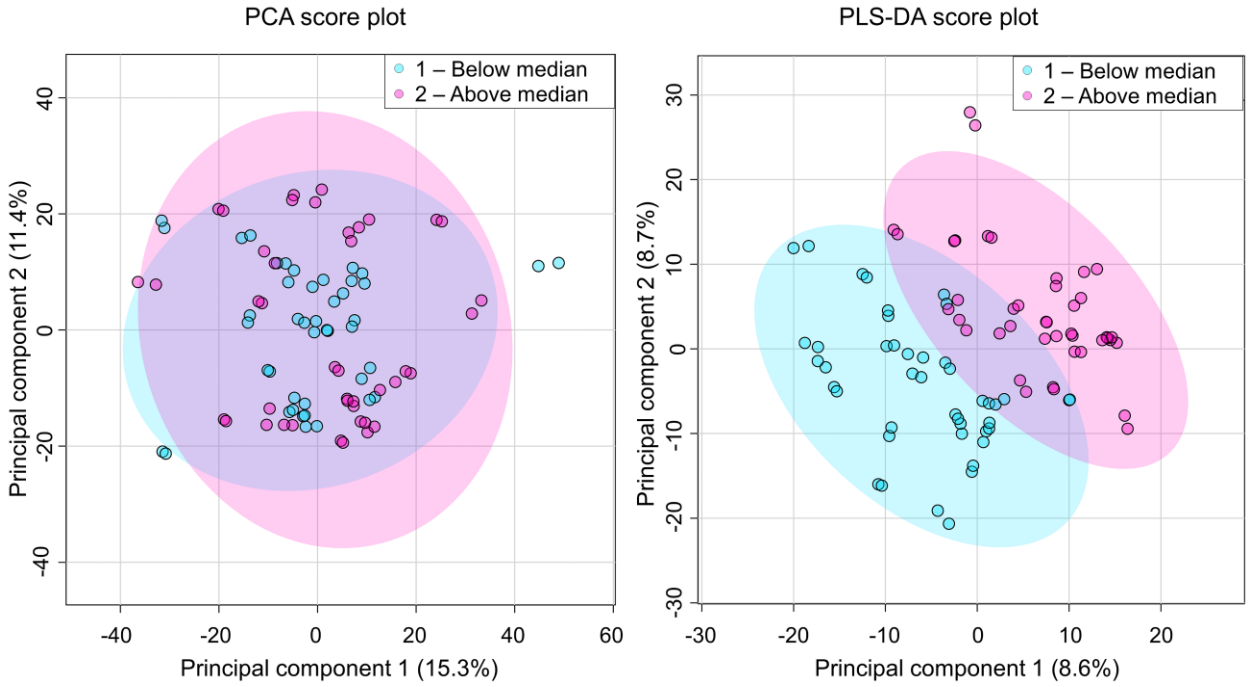

Figure C-6. PCA and PLS-DA score plot for the separation between PD patients aged below (Group 1, cyan) and above (Group 2, magenta) the median of 69.74 years old. The PLS-DA model with 6 components resulted in R^2 of 0.9769, Q^2 of 0.8359 and p of 0.03 for 1000 permutations.

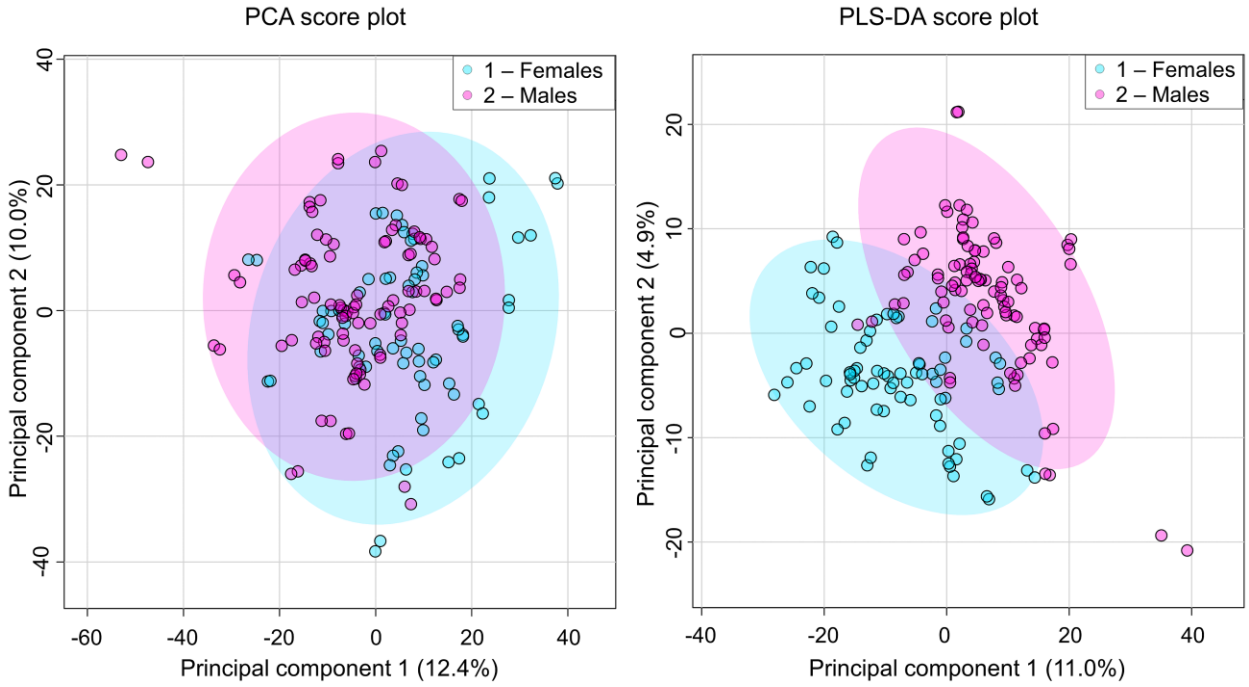

Figure C-7. PCA and PLS-DA score plot for the separation between female (Group 1, cyan) and male (Group 2, magenta) control and PD subjects. The PLS-DA model with 8 components resulted in R^2 of 0.9597, Q^2 of 0.8402 and $p < 0.01$ for 100 permutations.

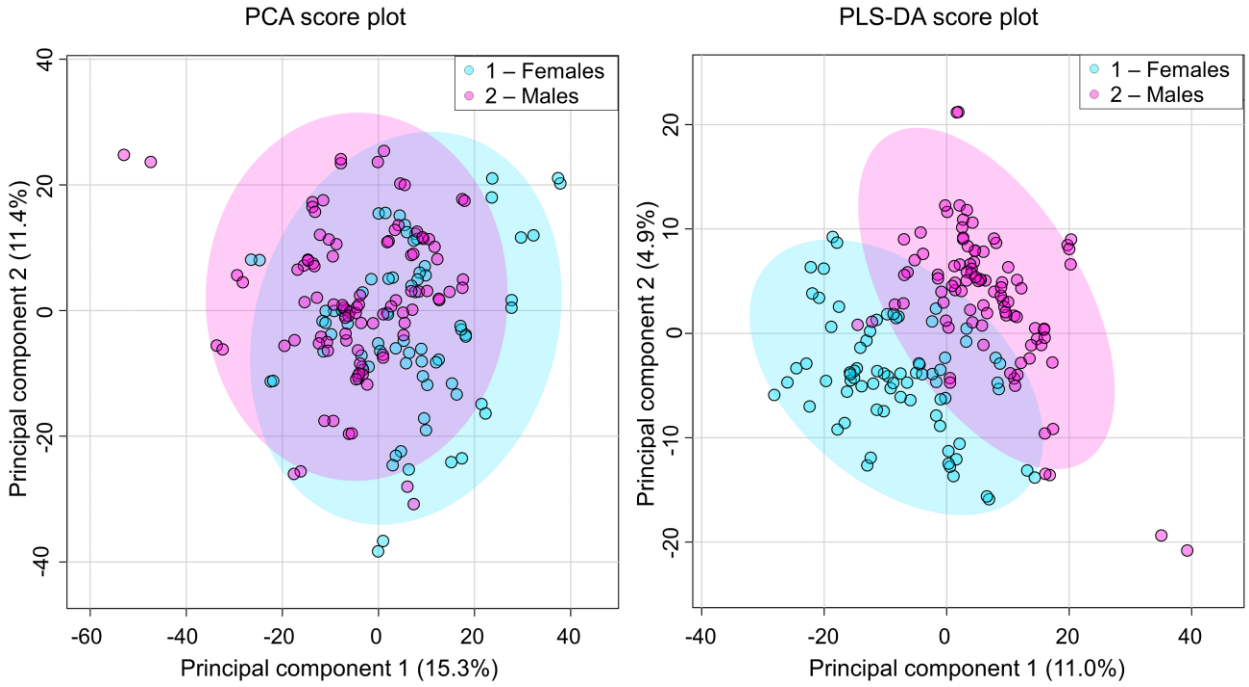

Figure C-8. PCA and PLS-DA score plot for the separation between female (Group 1, cyan) and male (Group 2, magenta) PD patients, including PDND and PDD. The PLS-DA model with 6 components resulted in R^2 of 0.9746, Q^2 of 0.8472 and $p < 0.01$ for 100 permutations.

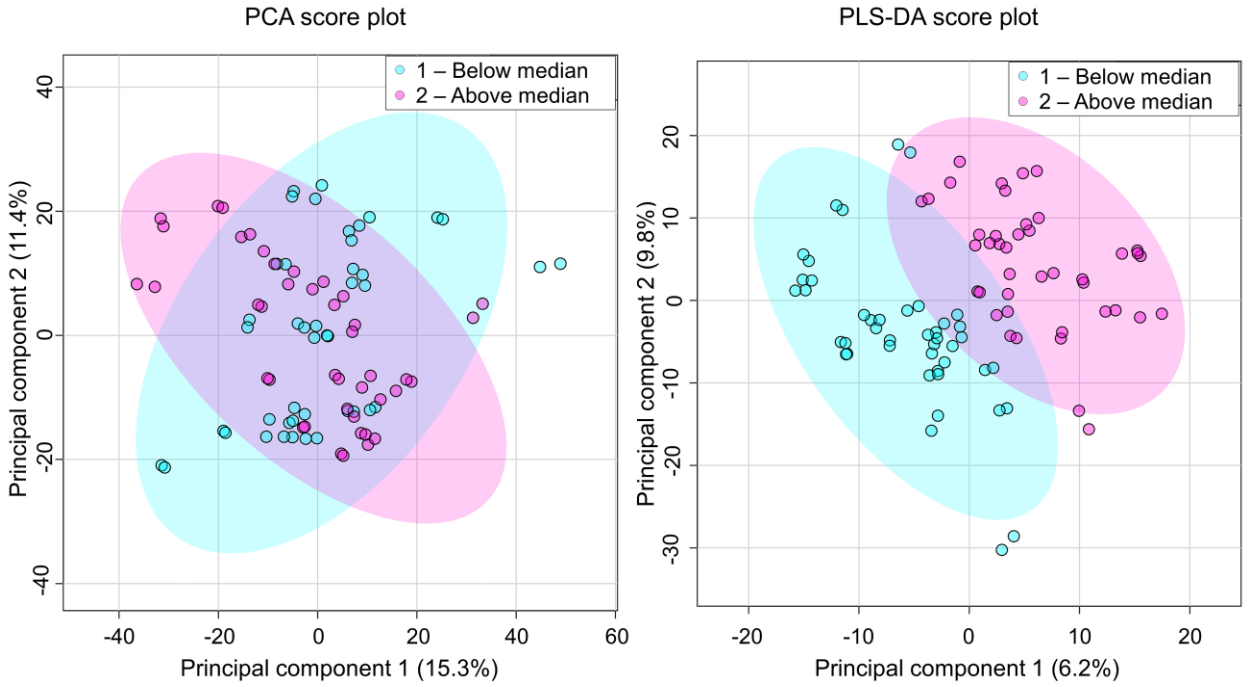

Figure C-9. PCA and PLS-DA score plot for the separation between PD patients, including PDND and PDD, that were diagnosed with PD for less (Group 1, cyan) or more (Group 2, magenta) than the median of 6.5 years before baseline sample collection. The PLS-DA model with 6 components resulted in R^2 of 0.9752, Q^2 of 0.8554 and p of 0.03 for 1000 permutations.

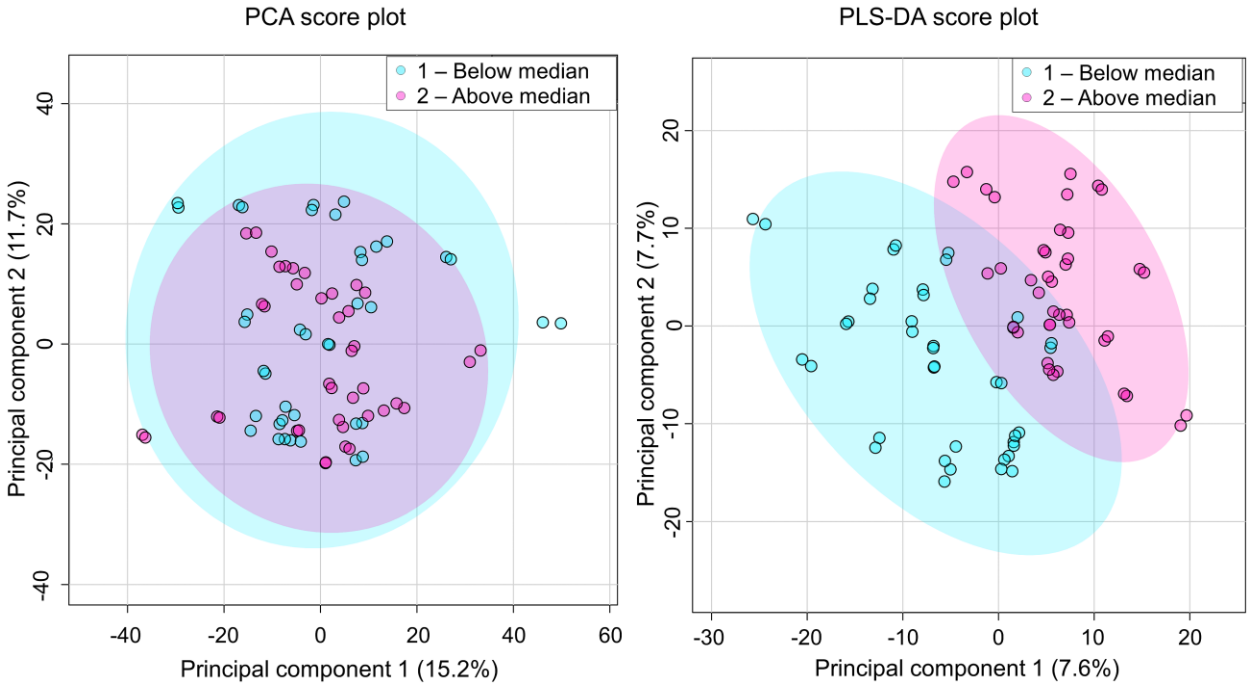

Figure C-10. PCA and PLS-DA score plot for the separation between PD patients, including PDND and PDD, that were treated with levodopa equivalent dosages below (Group 1, cyan) or above (Group 2, magenta) the median of 609.3 mg at baseline sample collection. The PLS-DA model with 5 components resulted in R^2 of 0.9517, Q^2 of 0.7392 and p of 0.003 for 1000 permutations.

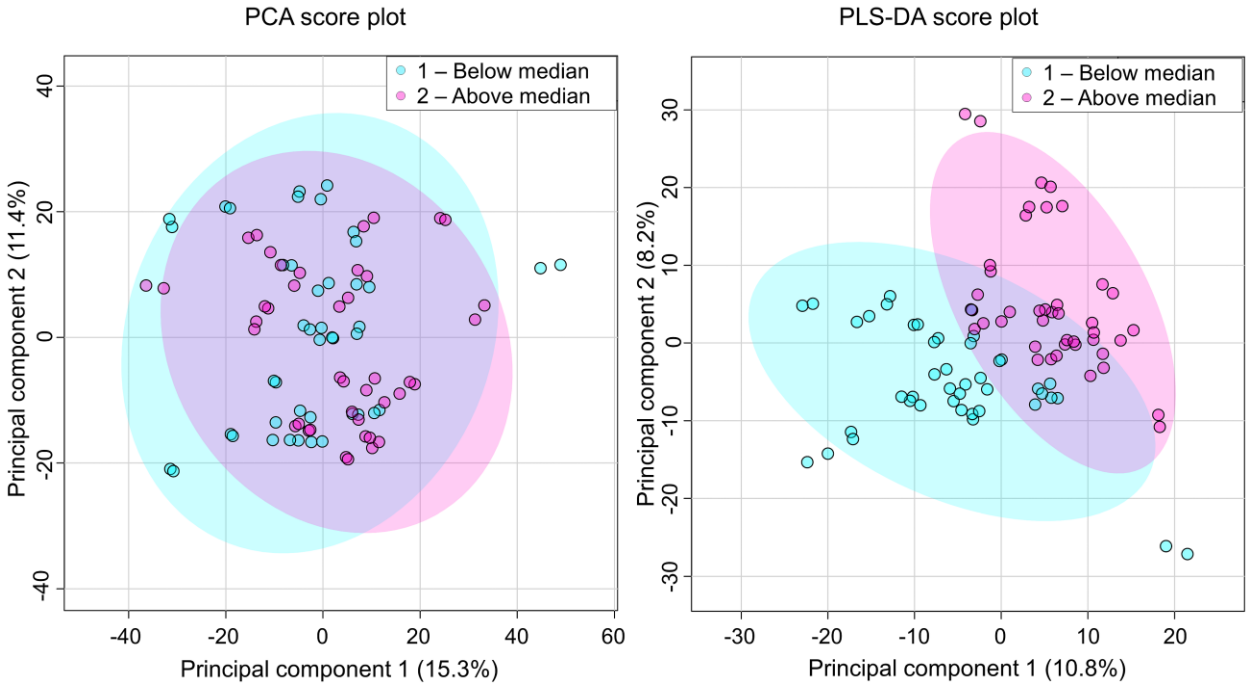

Figure C-11. PCA and PLS-DA score plot for the separation between PD patients that were treated with levodopa for more (Group 2, magenta) or less (Group 1, cyan) than the median of 2.96 years at the moment of baseline sample collection. The PLS-DA model with 6 components resulted in R^2 of 0.9734, Q^2 of 0.7986 and p of 0.02 for 1000 permutations.

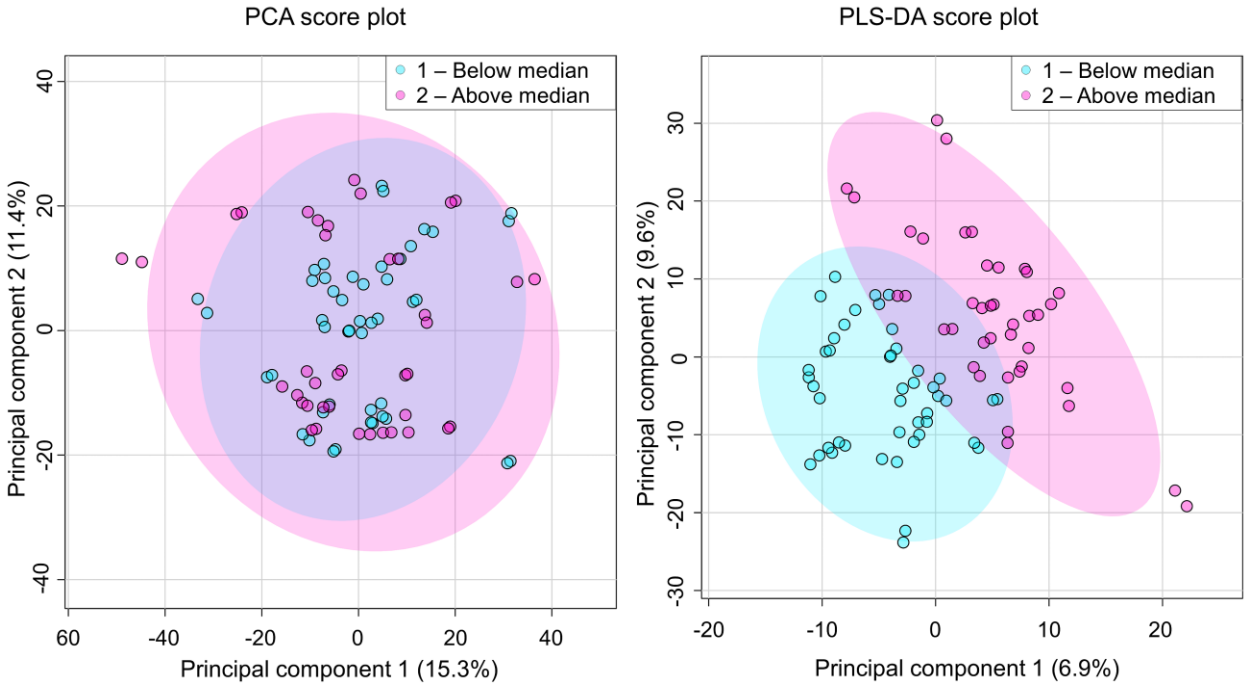

Figure C-12. PCA and PLS-DA score plots for the separation between PD patients with UPDRS – part III (motor examination) scores below (Group 1, cyan) and above (Group 2, magenta) the median of 15. The PLS-DA model with 6 components resulted in R^2 of 0.9714, Q^2 of 0.7838 and p of 0.03 for 1000 permutations.

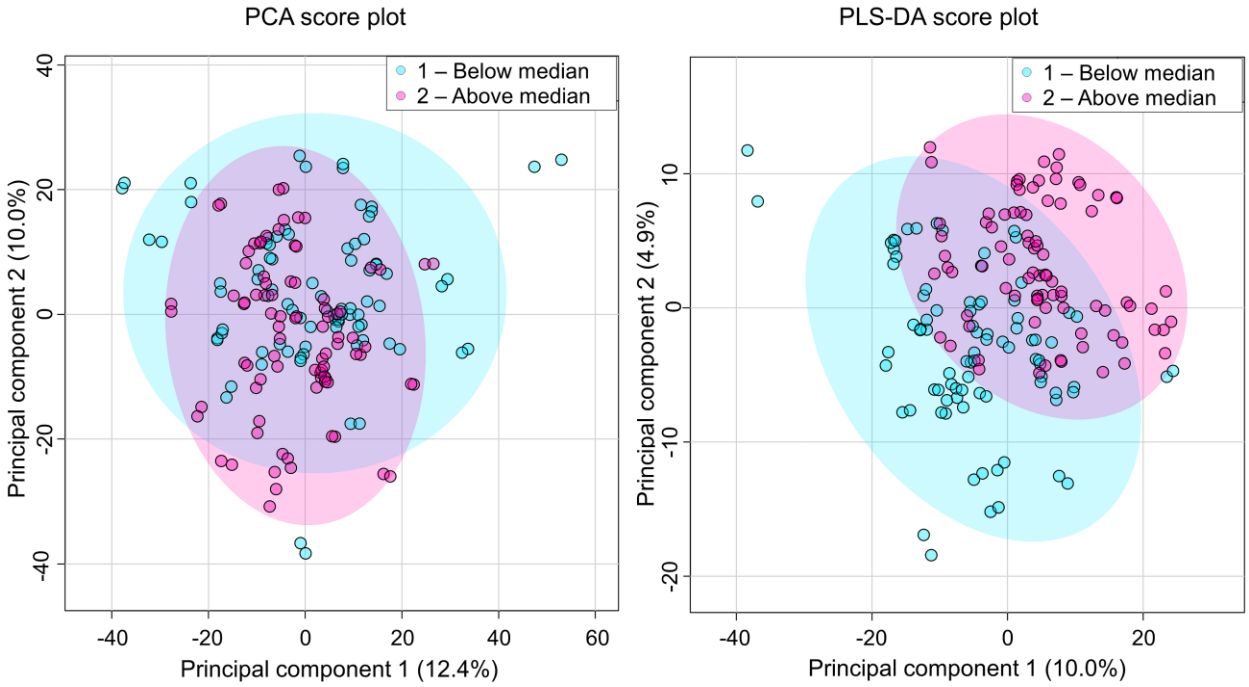

Figure C- 13. PCA and PLS-DA score plots for the separation between healthy controls and PD patients with B12 serum levels below (Group 1, cyan) and above (Group 2, magenta) the median of 304.5 ng/mL. The PLS-DA model with 8 components resulted in R^2 of 0.9534, Q^2 of 0.6422 and p of 0.001 for 1000 permutations.

Appendix D

Appendix D: Supplementary Figures for Chapter 8

LIST OF FIGURES

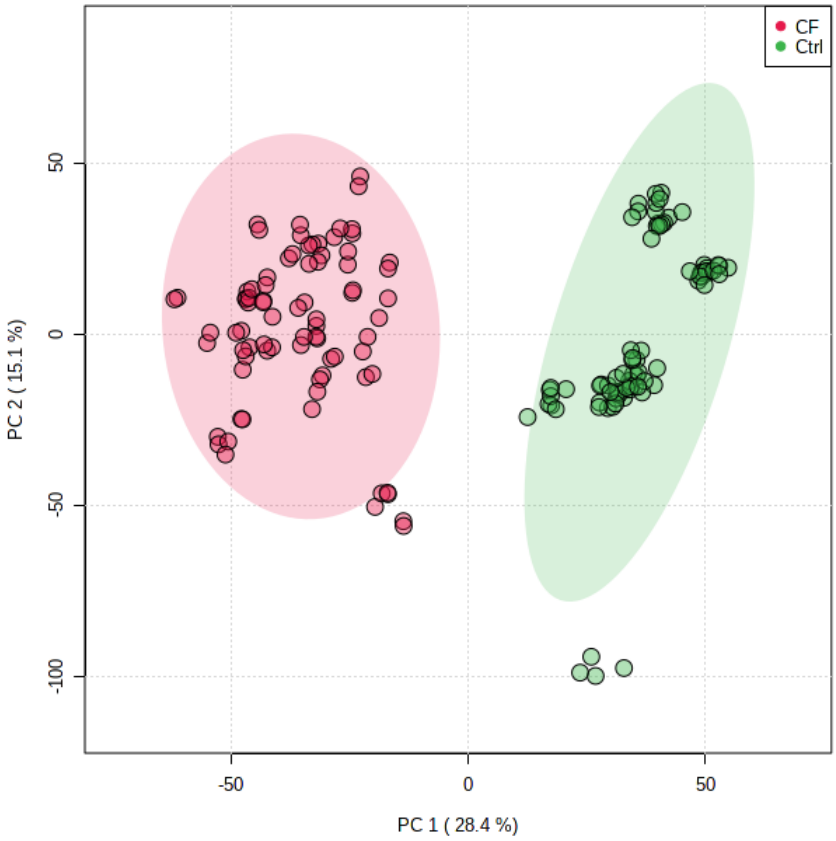

Figure D - 1. PCA score plot for healthy controls (red) compared to cystic fibrosis patients (green). Each sample was analyzed with extraction and injection replicates. The corresponding replicates are tightly clustered, showcasing the reproducibility of the employed methods.

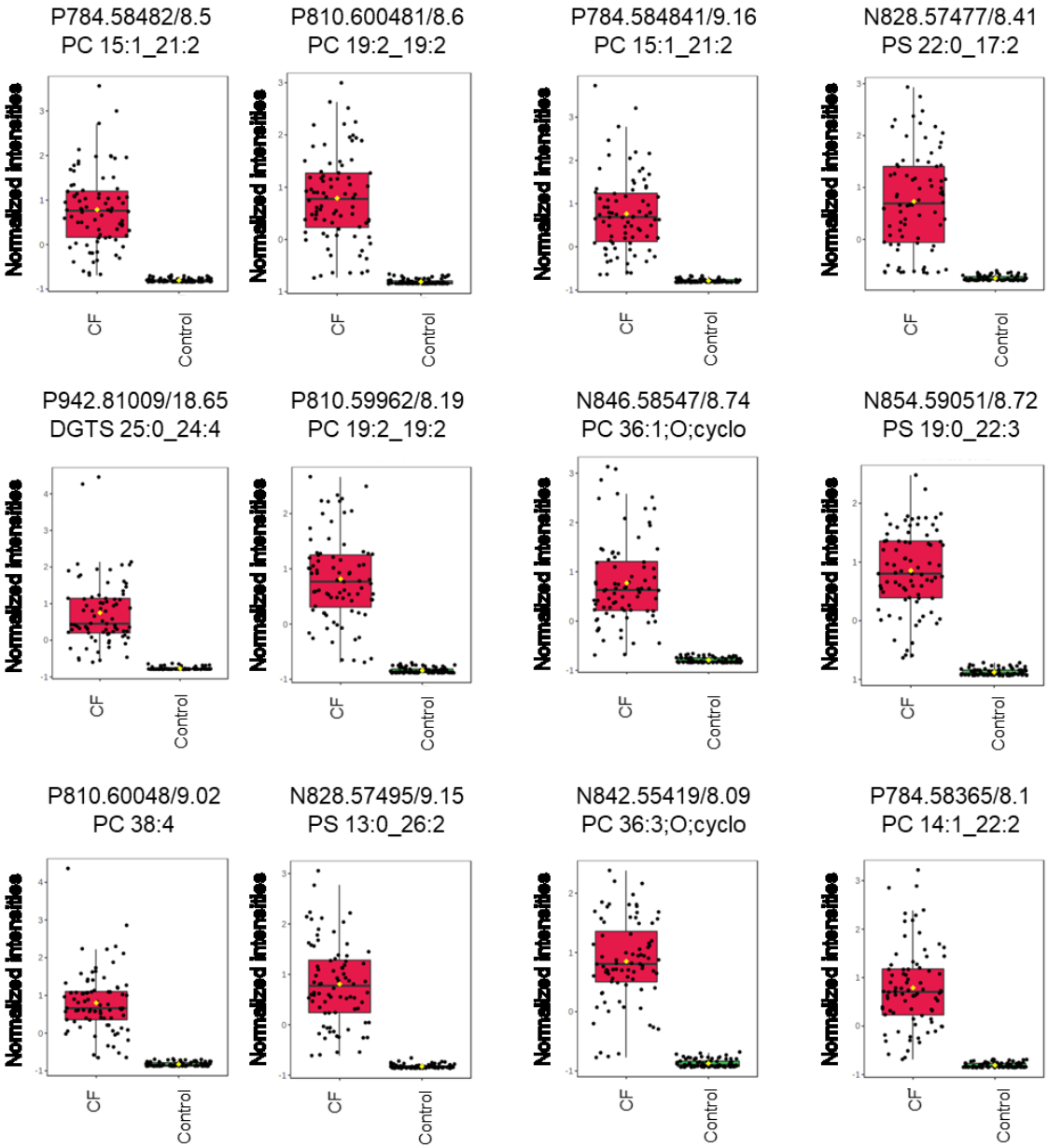

Figure D - 2. Boxplots for significantly altered lipids with the highest fold-changes (FC) for the comparison between cystic fibrosis patients (CF, red) and healthy controls (green), identified in tiers 1 or 2 (MS/MS positive identification). Lipids are identified by polarity (P for positive and N for negative), mass-to-charge ratio (m/z) and retention time (min), as well as the MS/MS identification.

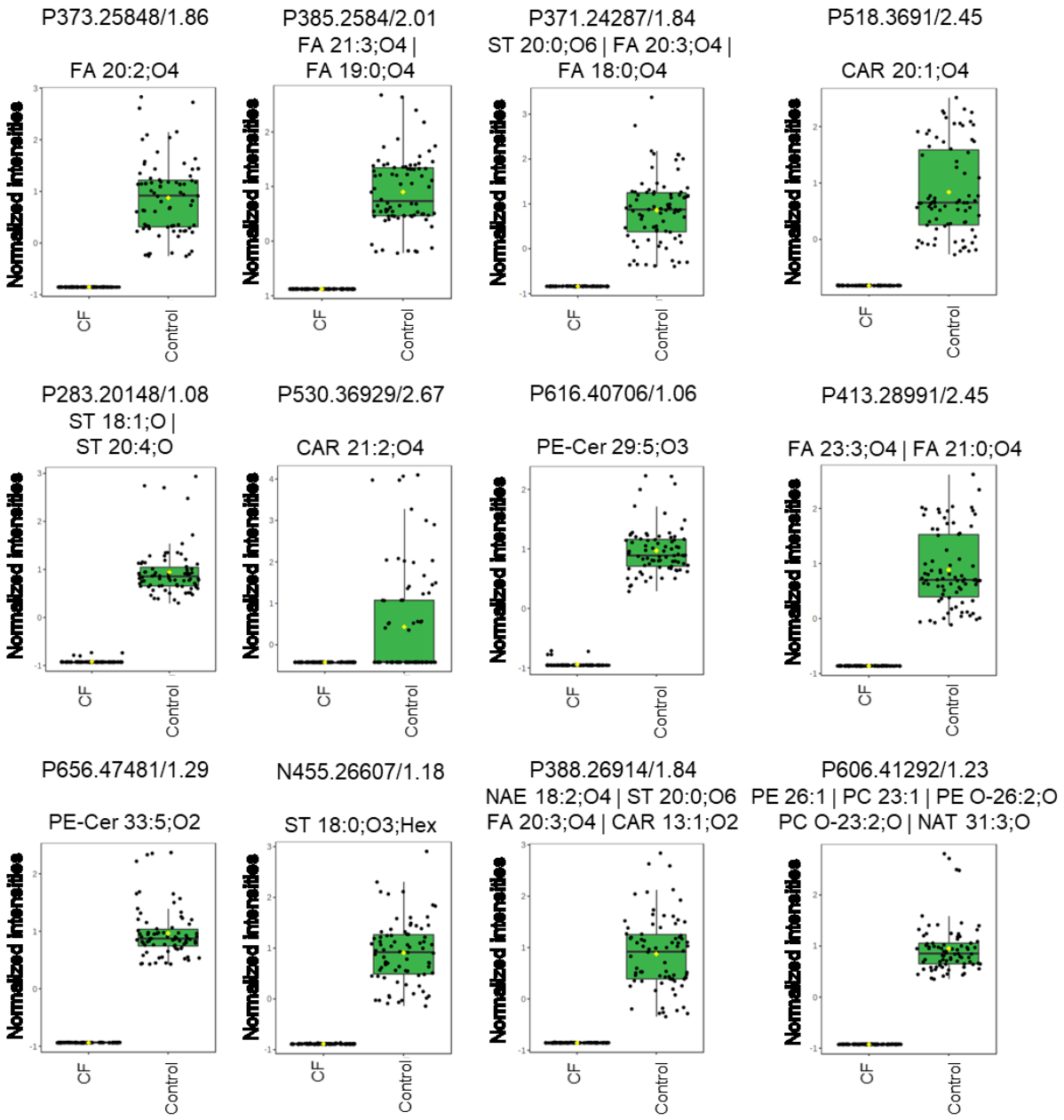

Figure D - 3. Boxplots for significantly altered lipids with the highest fold-changes (FC) for the comparison between cystic fibrosis patients (CF, red) and healthy controls (green), identified in tier 3 (putative mass-match identification). Lipids are named by polarity (P for positive and N for negative), m/z and retention time (min), as well as the putative identification with a maximum of 5 isomers or isobars.

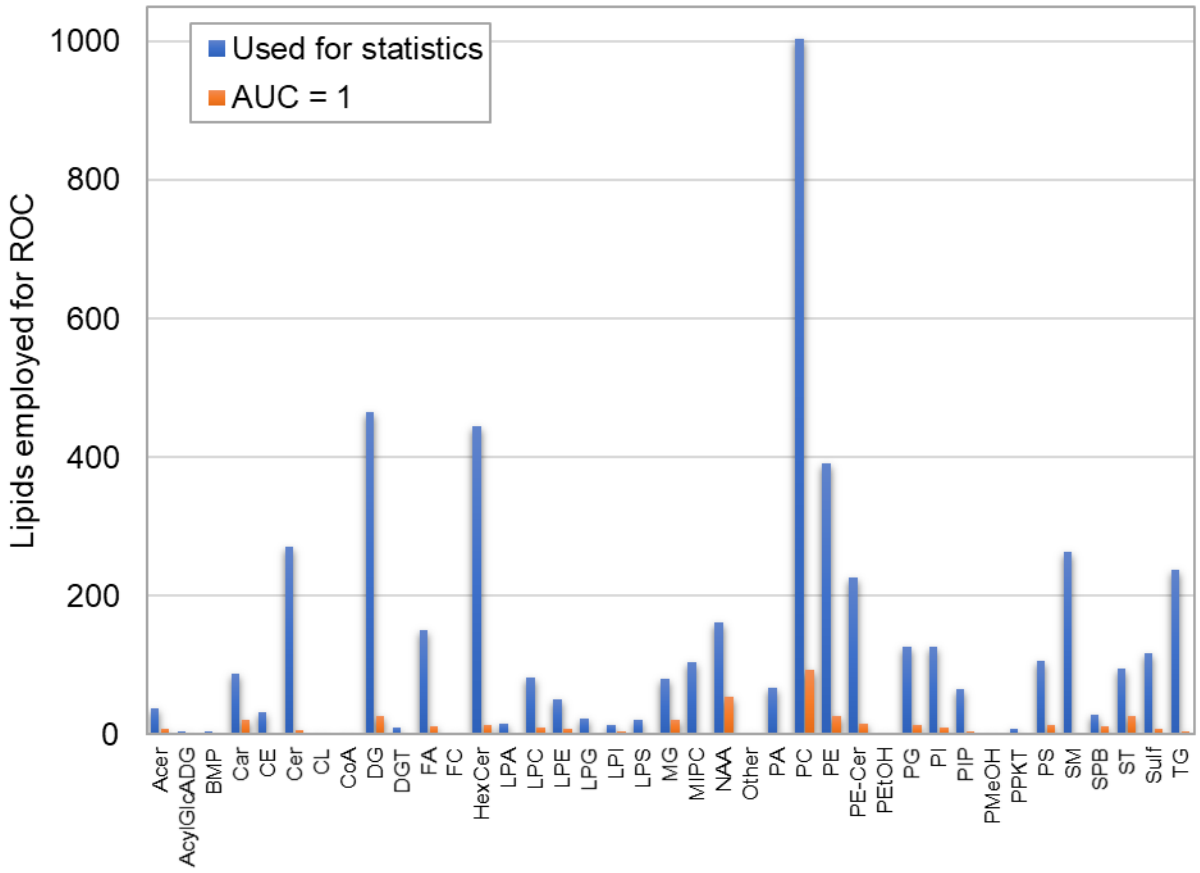

Figure D - 4. All lipids employed for Receiver Operating Characteristic (ROC) analysis (blue) compared to lipids that resulted in area-under-the-curve of 1 (orange), divided by lipid subclasses, for the healthy controls and CF patients. Abbreviations to lipid subclasses are defined in Table II-2.

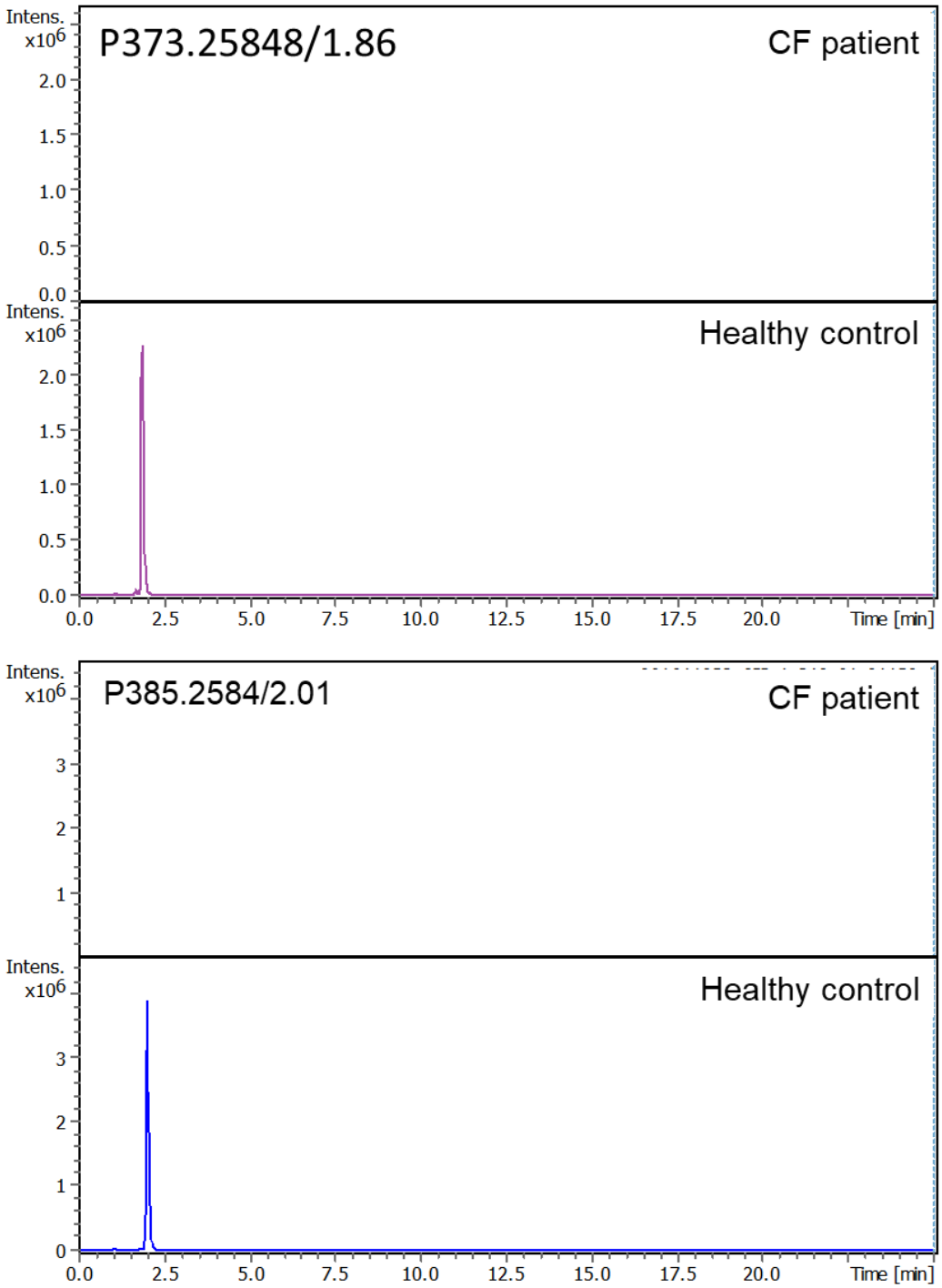

Figure D - 5. Extracted ion chromatograms obtained for two distinct biomarker candidates selected for differentiating cystic fibrosis patients from healthy controls, namely P373.25848/1.86 and P385.2584/2.01. Lipids are named by polarity (P for positive and N for negative), m/z and retention time (min).

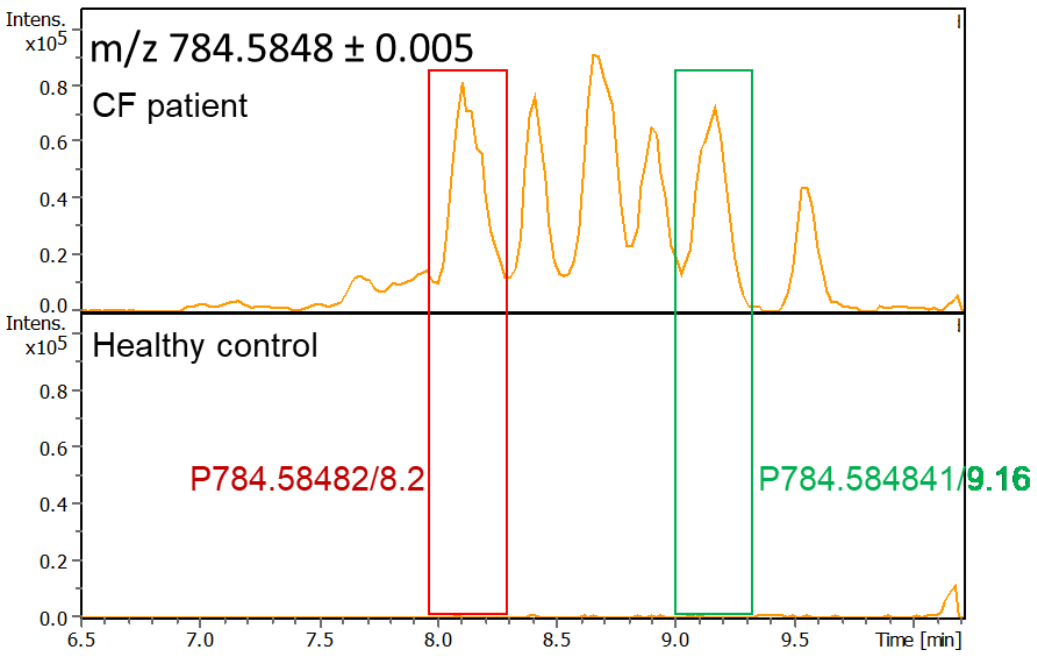

Figure D - 6. Extracted ion chromatograms obtained for two distinct biomarker candidates selected for differentiating cystic fibrosis patients from healthy controls, namely P784.58482/8.2 and 784.584841/9.16, identified as PC 15:1_21:2(tier 1). Lipids are named by polarity (P for positive and N for negative), m/z and retention time (min).

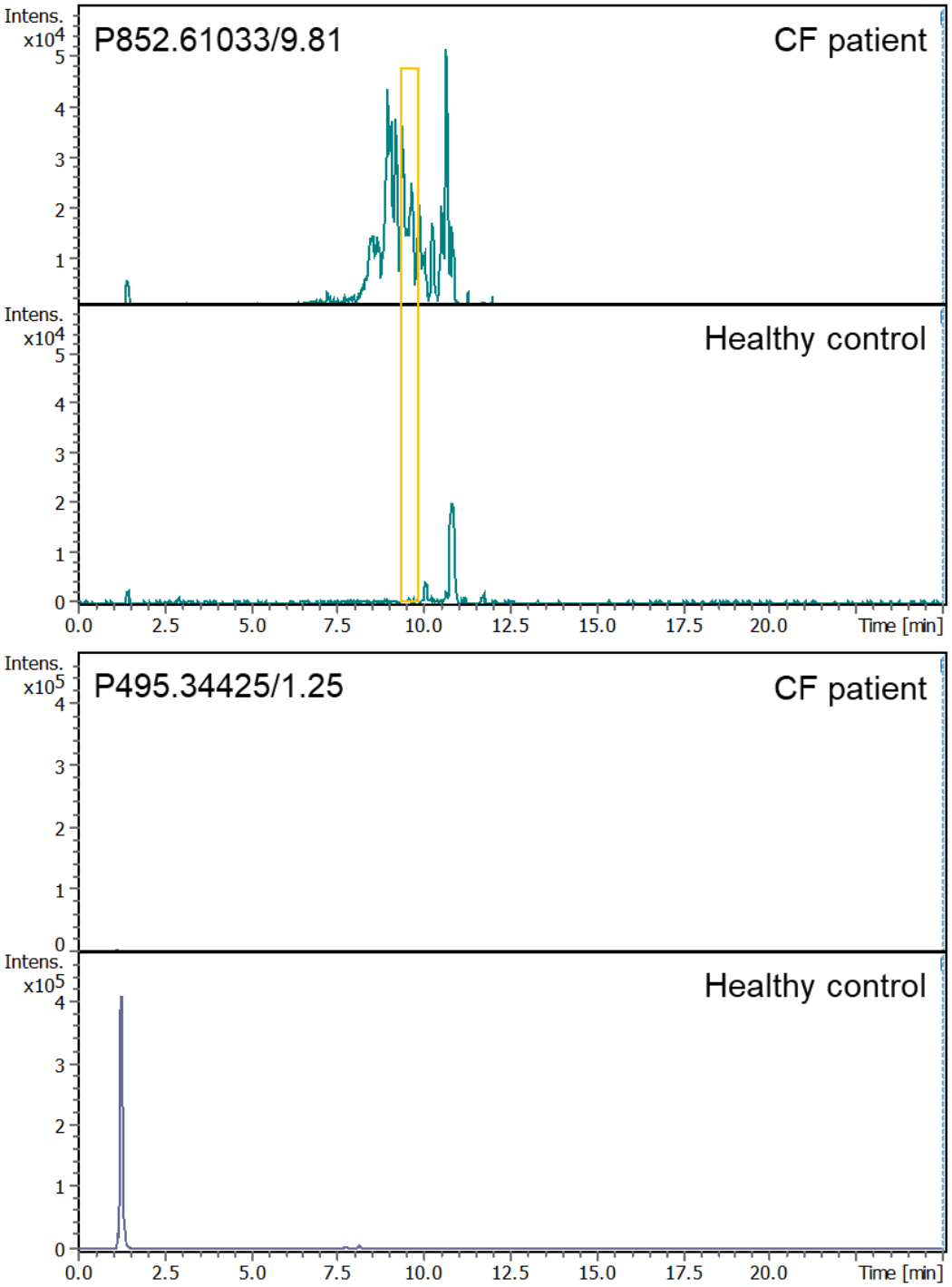

Figure D - 7. Extracted ion chromatograms obtained for two distinct biomarker candidates selected for differentiating cystic fibrosis patients from healthy controls, namely P373.25848/1.86 and P385.2584/2.01. Lipids are named by polarity (P for positive and N for negative), m/z and retention time (min).