Drug Research Program
Division of Pharmaceutical Chemistry and Technology
Faculty of Pharmacy
University of Helsinki
Finland

MASS SPECTROMETRY-BASED APPLICATIONS AND ANALYTICAL METHOD DEVELOPMENT FOR METABOLOMICS

Päivi Pöhö

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Pharmacy of the University of Helsinki, for public examination in lecture room 1041, Biocenter 2, on 31 January 2020, at 12 noon.

Helsinki 2020

©Päivi Pöhö

ISBN 978-951-51-5757-7 (print)
ISBN 978-951-51-5758-4 (online)
ISSN 2342-3161 (print)
ISSN 2342-317X (online)
http://ethesis.helsinki.fi

Unigrafia, Helsinki, Finland, 2020

Published in DSHealth series
'Dissertationes Scholae Doctoralis Ad Sanitatem Investigandam Universitatis Helsinkiensis'

The Faculty of Pharmacy uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

Supervisors

Professor Risto Kostiainen
Drug Research Program
Division of Pharmaceutical Chemistry and Technology
Faculty of Pharmacy
University of Helsinki
Finland

Professor Tapio Kotiaho
Drug Research Program
Division of Pharmaceutical Chemistry and Technology
Faculty of Pharmacy
and
Department of Chemistry
Faculty of Science
University of Helsinki
Finland

Reviewers

Professor Kati Hanhineva
Institute of Public Health and Clinical Nutrition
Faculty of Health Sciences
University of Eastern Finland
Kuopio
Finland

Professor Uwe Karst
Institute of Inorganic and Analytical Chemistry
University of Münster
Münster
Germany

Opponent

Professor Jonas Bergquist
Department of Chemistry - BMC
Uppsala University
Uppsala
Sweden

ABSTRACT

Metabolites are small molecules present in a biological system that have multiple important biological functions. Changes in metabolite levels reflect genetic and environmental alterations and play a role in multiple diseases. Metabolomics is a discipline that aims to analyze all the small molecules in a biological system simultaneously. Since metabolites represent a diverse group of compounds with varying chemical and physical properties with a wide concentration range, metabolomic analysis is technically challenging. Due to its high sensitivity and selectivity, mass spectrometry coupled with chromatographic separation is the most commonly used analytical tool. Currently, there is no comprehensive universal analytical tool to detect all metabolites simultaneously and multiple methods are required. The aim of this study was to develop and apply mass spectrometry-based analytical methods for metabolomics studies.

Neonatal rodents can fully regenerate their hearts after an injury. However, this regenerative capacity is lost within 7 days after birth. The molecular mechanism behind this phenomenon is unknown and understanding the biology behind this loss of regeneration capacity is necessary for the development of regeneration-inducing therapies. To investigate this mechanism, changes in mouse heart metabolite, protein, and transcript levels during the early postnatal period were studied. Non-targeted metabolomics methods utilizing liquid chromatography-mass spectrometry (LC-MS) and two-dimensional gas chromatography-mass spectrometry (GCxGC-MS) were applied to detect the metabolic changes of neonatal mouse hearts. Two complementary techniques increased metabolite coverage. A total of 151 identified metabolites showed differences in the neonatal period, reflecting changes in multiple metabolic pathways. The most significant changes observed in all levels (metabolite, protein, and transcript) were branched chain amino acid (BCAA) catabolism, fatty acid metabolism, and the mevalonate and ketogenesis pathways, thus revealing possible associations with regeneration capacity or regulation of the cardiomyocyte cell cycle.

Insulin resistance (IR), metabolic syndrome, and type 2 diabetes have been shown to induce metabolic changes; the origin of the changes is unknown. In this study, human serum metabolite profiles from non-diabetic individuals were associated with IR. Gut microbiota were identified as a possible origin of the metabolic changes. Serum metabolites were detected with GCxGC-MS and lipids with LC-MS method. In total, 19 serum metabolite clusters were significantly associated with the IR phenotype, including 26 polar metabolites from five separate clusters and 367 lipids from 14 clusters. IR and changed metabolites were further associated with gut microbiota metagenomics and gut microbiota functional modules, showing that gut microbiota impacts the human serum metabolites associated with IR.

Individuals with the IR phenotype had increased BCAA levels, which was influenced by bacterial species with increased BCAA biosynthesis potential and the absence of species with active bacterial inward BCAA transport.

Sample throughput is often limited when chromatographic separation is used in metabolomics applications; a short analysis time is of great importance in large metabolic studies. The feasibility of direct infusion electrospray microchip MS (chip-MS) for global non-targeted metabolomics to detect metabolic differences between two cell types was studied and was compared to the more traditional LC-MS method. We observed that chip-MS was a rapid and simple method that allowed high sample throughput from small sample volumes. The chip-MS method was capable of separating cells based on their metabolic profiles and could detect changes of several metabolites. However, the selectivity of chip-MS was limited compared to LC-MS and chip-MS suffers more from ion suppression.

Many biologically important low-abundance metabolites are not detectable with non-targeted metabolomics methods and separate more sensitive targeted methods are required. An in-house developed capillary photoionization (CPI) source was shown to have high ion transmission efficacy and high sensitivity towards non-polar compounds such as steroids. In this study, the CPI prototype was developed to increase its sensitivity. The feasibility of the ion source for the quantitative analysis of biological samples was studied by analyzing 18 endogenous steroids in urine with gas chromatography capillary photoionization tandem mass spectrometry (GC-CPI-MS/MS). The GC-CPI-MS/MS method showed good chromatographic resolution, acceptable linearity and repeatability, and low limits of detection (2-100 pg mL^{-1}). In total, 15 steroids were quantified either as a free steroid or glucuronide conjugate from the human urine samples.

Additionally, the applicability of the CPI interface for LC applications was explored for the first time using low flow rates. The feasibility of the LC-CPI-MS/MS for the quantitative analysis of four steroids was studied in terms of linearity, repeatability, and limits of detection. The method showed good quantitative performance and high sensitivity at a low femtomole level.

ACKNOWLEDGEMENTS

This work was carried out in the Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki during the years 2014-2019. Part of the work was additionally performed at Steno Diabetes Center, Gentofte, Denmark at 2016 and at VTT Technical Research Centre of Finland during the years 2012-2013. Business Finland 3iRegeneration project, the European Community MetaHIT project, Drug Research Doctoral Program, and The Finnish Cultural Foundation are acknowledged for funding this work.

I am deeply grateful to my supervisors, Prof. Risto Kostiainen and Prof. Tapio Kotiaho, for giving me the opportunity to carry out my thesis under their guidance and support in a highly interesting project. Risto, you have been a great supervisor for me, giving me trust and freedom to perform my work independently, however always willing to advice and comment if needed. Your help especially in the writing process has been invaluable and I have learned a lot from you. Tapio, you have also been a great support and your careful comments and corrections to publications and thesis have been highly valuable. I want to thank also Prof. Tuulia Hyötyläinen and Prof. Matej Orešič for introducing me to the world of research, mass spectrometry, metabolomics, and lipidomics already during my master's thesis work and during the following years at VTT. Without the time spend in your research group I probably wouldn't have started to perform my doctoral studies at all. I also want to thank Prof. Jari Yli-Kauhaluoma and Doc. Tiina Kauppila for acting as my thesis steering group members and for their valuable support during these years. 3iRegeneration project leader Prof. Heikki Ruskoaho I am grateful for great collaboration, positive attitude, and optimism towards our research throughout the project.

Additionally, I want to acknowledge Prof. Kati Hanhineva and Prof. Uwe Karst for their thorough review of the thesis and for their generous and kind comments. I am grateful for Prof. Jonas Bergquist for accepting the invitation to act as my opponent and I am looking forward for discussions during the defense.

I also want to thank all the collaborators and co-authors for their valuable contributions. Special thanks goes to Dr. Anu Vaikkinen, your expertise in analytical chemistry is incredible and your contribution to this work have been essential. I want to thank Jaakko Teppo for sharing this PhD journey with me, for all the help and support during these years and bringing proteomics and data analysis expertise into this work. Furthermore, I want to thank Doc. Virpi Talman and Tuuli Karhu for providing the mouse heart samples, transcriptomics, pharmacological experiments, and valuable output in the data interpretation, Dr. Petri Kylli, Dr. Markus Haapala, Heikki Räikkönen, Niina Kärkkäinen, and Karen Scholz for their help in the laboratory, Prof. Jukka Heikkonen, Assoc. Prof. Tapio Pahikkala, and Paris Movahedi for their

computational and data analysis expertise, Dr. Kajetan Trošt and Dr. Tommi Suvitaival for warm welcome to Steno Diabetes Center and contribution to the GCxGC-MS measurements and data analysis, Doc. Markku Varjosalo for the contribution to proteomics, data-analysis, and data interpretation, Dr. Maxim Bespalov for providing the cell samples, Doc. Tiina Sikanen and Dr. Katriina Lipponen for the microchip measurements and expertise in microfluidics. Additionally, I want to thank all the 3iRegeneration and MetaHIT project members for their contributions.

I am also very grateful to all the past and present colleagues at the Division of Pharmaceutical Chemistry and Technology, especially the MAC group members. Great and supporting atmosphere in the lab, in the office, in the coffee and lunch breaks, during the conference trips and events have been highly valuable and important.

I am most grateful to my family and friends for supporting me throughout this project, for all the precious moments outside the work, and setting life into a balance. Finally, I want to thank my greatest support Lassi, for your understanding, encouragement, calmness, and bringing happiness and joy into my everyday life.

December, 2019

Päivi Pöhö

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Talman V*., Teppo J*, **Pöhö P.***, Movahedi P., Vaikkinen A., Karhu T., Trošt K., Suvitaival T.,Heikkonen J., Pahikkala T., Kotiaho T., Kostiainen R., Varjosalo M., Ruskoaho H. Molecular atlas of postnatal mouse heart development. *Journal of the American Heart Association*, **2018**. 7, e010378. doi:10.1161/JAHA.118.010378.

II Pedersen, H. K., Gudmundsdottir, V., Nielsen, H. B., Hyötyläinen, T., Nielsen, T., Jensen, B. A. H.,Forslund, K., Hildebrand, F., Prifti, E., Falony, G., Le Chatelier, E., Levenez, F., Doré, J., Mattila, I., Plichta, D. R., **Pöhö, P**., Hellgren, L. I., Arumugam, M., Sunagawa, S., Vieira-Silva, S., Jørgensen, T., Holm, J. B., Trošt, K., Kristiansen, K., Brix, S., Raes, J., Wang, J., Hansen, T., Bork, P., Brunak, S., Orešič, M., Ehrlich, S. D., Pedersen, O. Human Gut Microbes Impact Host Serum Metabolome and Insulin Sensitivity. *Nature*, **2016**, 535, 376-381. doi:10.1038/nature18646.

III **Pöhö P**., Lipponen K., Bespalov M., Sikanen T., Kotiaho T., Kostiainen R. Comparison of liquid chromatography-mass spectrometry and direct infusion microchip electrospray ionization mass spectrometry in global metabolomics of cell samples. *European Journal of Pharmaceutical Sciences*, 138, **2019**, 104991. doi:10.1016/j.ejps.2019.104991.

IV **Pöhö P**., Scholz K., Kärkkäinen N., Haapala M., Räikkönen H., Kostiainen R., Vaikkinen A. Analysis of steroids in urine by gas chromatography-capillary photoionization-tandem mass spectrometry. *Journal of Chromatography A*, **2019**, 1598, 175-182. doi: 10.1016/j.chroma.2019.03.061.

V **Pöhö P**., Vaikkinen A., Kylli P., Haapala M., Kostiainen R. Capillary photoionization: Interface for low flow rate liquid chromatography-mass spectrometry. *The Analyst*, **2019**, 144(9), 2867-2871. doi:10.1039/C9AN00258H.

The publications are referred to in the text by their roman numerals.
*all three authors equally contributed to this work.

AUTHOR'S CONTRIBUTION TO THE PUBLICATIONS INCLUDED IN THIS THESIS

I The experimental work and the data analysis related to metabolomics were performed by author with contributions from the other co-authors. Data fusion was performed by author, Jaakko Teppo, Virpi Talman, and Parisa Movahedi. The manuscript was written by the author together with Jaakko Teppo and Virpi Talman and with contributions from the co-authors.

II The experimental part of serum lipidomics was performed by the author with contributions from Tuulia Hyötyläinen and Matej Orešič. A detailed description of other author contributions are available in the manuscript.

III The microchip-MS measurements were performed by Katriina Lipponen and the cell samples were provided by Maxim Bespalov. The LC-MS measurements and all data preprocessing and data analysis were performed by the author. The manuscript was written by the author with contributions from the co-authors.

IV The experimental work was performed by author, Karen Scholz, Niina Kärkkäinen, and Anu Vaikkinen. The data processing and data analysis were performed by the author with the contributions from other co-authors. The manuscript was written by the author with contributions from the co-authors.

V The experimental part of the work was performed by the author with some contributions from other co-authors. The manuscript was written together with the co-authors.

Publication II is also included in the dissertation of Helle Krogh Pedersen from the Technical University of Denmark and Publication I will be included in the dissertation of Jaakko Teppo from the University of Helsinki.

ABBREVIATIONS

2D	two-dimensional
AAS	androgenic anabolic steroid
ACN	acetonitrile
ANOVA	analysis of variance
APCI	atmospheric pressure chemical ionization
API	atmospheric pressure ionization
APPI	atmospheric pressure photoionization
BCAA	branched-chain amino acid
BDH1	3-hydroxybutyrate dehydrogenase 1
BrdU	bromodeoxyuridine
BSTFA	N,O-Bis(trimethylsilyl)trifluoroacetamide
CASMI	critical assessment of small molecule identification
CCS	collision cross section
CE	capillary electrophoresis
CE-MS	capillary electrophoresis-mass spectrometry
Cer	ceramide
ChoE/CholE	cholesterylester
Chol	cholesterol
CI	chemical ionization
CID	collision-induced dissociation
CL	cardiolipin
CPI	capillary photoionization
DART	direct analysis in real time
DDA	data-dependent acquisition
DESI	desorption electrospray
DG	diacylglycerol
DI	direct infusion
DIA	data-independent acquisition
DI-MS	direct infusion mass spectrometry
DTE	dithioerythritol
EI	electron ionization
ESI	electrospray ionization
EtOH	ethanol
eV	electron volt
FA	formic acid, fatty acid
FAHFA	fatty acid esters of hydroxyl fatty acids
FBS	fetal bovine serum
FDR	false discovery rate
FIA	flow injection analysis
FT-ICR	Fourier transform ion cyclotron resonance
FWHM	full width at half maximum
GC	gas chromatography
GC-MS	gas chromatography-mass spectrometry
GCxGC-MS	two-dimensional gas chromatography-mass spectrometry

GMD	Golm Metabolome Database
GNPS	Global Natural Product Social Molecular Networking
GO	gene ontology
HCD	higher energy collision dissociation
HFF	human foreskin fibroblast
HILIC	hydrophilic interaction chromatography
hiPSC	human induced pluripotent stem cell
HMDB	human metabolome database
HMGCL	hydroxymethylglutaryl-coenzyme A lyase
HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase
HMGCS2	hydroxymethylglutaryl-CoA synthase 2
HOMA-IR	homeostatic model assessment for insulin resistance
HPLC	high-performance liquid chromatography
HR	high resolution
HRMS	high-resolution mass spectrometry
i.d.	inner diameter
ICH	The International Council for Harmonisation
IDI1	isopentenyl-diphosphate d-isomerase 1
IMS	ion mobility spectrometry
IMS-MS	ion mobility spectrometry-mass spectrometry
inj. st.	injection standard
IPA	isopropanol
IS	insulin sensitivity
ISTD	internal standard
IT	ion trap
K_2CO_3	potassium carbonate
KEGG	Kyoto encyclopedia of genes and genomes
LC	liquid chromatography
LC-MS	liquid chromatography-mass spectrometry
LCxLC	comprehensive two-dimensional liquid chromatography
LLE	liquid-liquid extraction
LME	linear mixed effect
LOD	limit of detection
LOQ	limit of quantitation
LPL	lysophospholipid
LPA	lysophosphatidic acid
LysoPC/LPC	lysophosphatidylcholine
LysoPE/LPE	lysophosphatidylethanolamine
m/z	mass-to-charge ratio
MALDI	matrix-assisted laser desorption ionization
MeOH	methanol
MG	monoacylglycerol
MgF_2	magnesium fluoride
MoNA	MassBank of North America
MOX	methoxyamine
MQ	milliQ-water
MRM	multiple reaction monitoring
MS	mass spectrometry

MS/MS	tandem mass spectrometry
MSEA	metabolite set enrichment analysis
MSI	mass spectrometry imaging
MSTFA	N-methyl-N-(trimethylsilyl) trifluoroacetamide
MTBE	methyl-*tert*-butyl ether
MVD	mevalonate diphosphate decarboxylase
MVK	mevalonate kinase
MW	molecular weight
$NaHCO_3$	sodium hydrogen carbonate
NH_4Ac	ammonium acetate
NH_4I	ammonium iodide
NIST	National Institute of Standards and Technology
NLS	neutral loss scan
NMR	nuclear magnetic resonance
NP	normal phase
NS	not significantly associated
o.d.	outer diameter
OPLS-DA	orthogonal partial least squares-discriminant analysis
OXCT1	3-oxoacid CoA-transferase 1
PA	phosphatidic acid
PBS	phosphate buffered saline
PC	phosphatidylcholine
PCA	principal component analysis
PCR	polymerase chain reaction
PE	phosphatidylethanolamine
PG	phosphatidylglycerol
PI	phosphatidylinositol
PIS	precursor ion scan
PL	phospholipid
PLS-DA	partial least squares-discriminant analysis
PMVK	phosphomevalonate kinase
PPT	protein precipitation
PS	phosphatidylserine
psig	pounds per square inch gauge
PUR	purine
PYR	pyrimidine
Q	quadrupole
QC	quality control
QQQ	triple quadrupole
Q-TOF	quadrupole-time-of-flight
RI	retention index
ROC	receiver operating characteristic
RP	reversed phase
RSD	relative standard deviation
RT	retention time
S/N	signal-to-noise ratio
SEM	standard error of mean
SFC	supercritical fluid chromatography

SFC-MS	supercritical fluid chromatography-mass spectrometry
SIMS	secondary ion mass spectrometry
SM	sphingomyelin
SMPDB	small molecule pathway database
SPE	solid-phase extraction
SPME	solid-phase micro extraction
SS	stainless steel
SU-8	trademark of an epoxy-based polymer
SWATH	sequential window acquisition of all theoretical fragment-ion spectra
TCA	tricarboxylic acid
TG	triacylglycerols
TIC	total ion chromatogram
TMAO	trimethylamine N-oxide
TMCS	trimethylchlorosilane
TMS	trimethylsilyl
TOF	time-of-flight
UHPLC	ultra-high-performance liquid chromatography
UV	ultraviolet
UVPD	ultraviolet photodissociation
WADA	World Anti-Doping Agency

STEROIDS

11-OH-PROG	11α-hydroxyprogesterone
17-OH-PREG	17α-hydroxypregnenolone
17-OH-PROG	17α-hydroxyprogesterone
21-OH-PROG	21-hydroxyprogesterone
A	aldosterone
CORT	corticosterone
CS	cortisone
DHEA	dehydroepiandrosterone
E1	estrone
E2	β-estradiol
E3	estriol
ETIOL	etiocholanolone
HC	hydrocortisone
PREG	pregnenolone
PROG	progesterone
T	testosterone
ADT	androsterone
Me-T	17α-methyltestosterone
AN	androstenediene

1 REVIEW OF THE LITERATURE

1.1 METABOLITES AND METABOLOME

Metabolites are small molecules (MW <1500 Da) present in a biological system that play important roles in several biological functions, such as energy production and storage, cell signaling and regulation, or as building blocks for multiple biological components.[1–3] Metabolites can be divided into different subclasses based on their origin. Endogenous metabolites are formed during intracellular metabolism in a biological system, whereas exogenous metabolites (i.e. drugs, food nutrients, environmental pollutants) are introduced from outside the system. Metabolites can also originate from interactions between symbiotic biological systems, such as host (e.g. human) and gut microbiota.[4,5] These metabolites represent a diverse group of molecules with varying concentrations from several chemical classes (Figure 1). Metabolites differ in molecular weight and chemical and physical properties, such as hydrophobicity/hydrophilicity, acidity/basicity, volatility, and solubility. The metabolome represents the collection of all small molecule metabolites in a biological system and can be analyzed with metabolomics.[1–3]

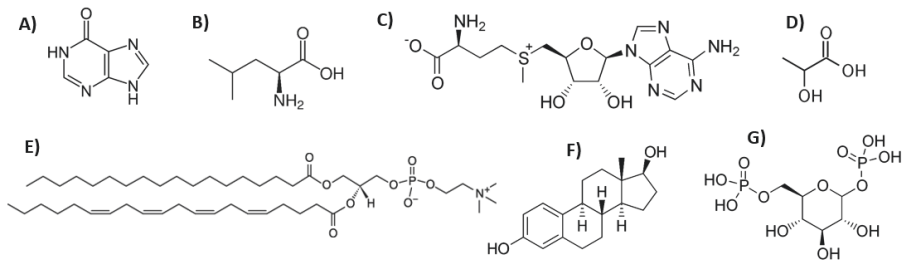

Figure 1 Examples of different metabolite structures; A) Hypoxanthine, B) Leucine, C) S-Adenosyl methionine, D) Lactic acid, E) Phosphatidylcholine(18:0/20:4), F) β-Estradiol, G) Glucose 1,6-bisphosphate

1.2 OMICS CASCADE

The main biological components can be simplified into four categories; genes, transcripts, proteins, and metabolites (Figure 2). The -omics suffix refers to holistic technologies that seek to comprehensively measure all of these biological components, specifically genomics, transcriptomics, proteomics, and metabolomics (Figure 2). Systems biology (multiomics) studies all these biological components and their complex interactions inside the system.[6,7]

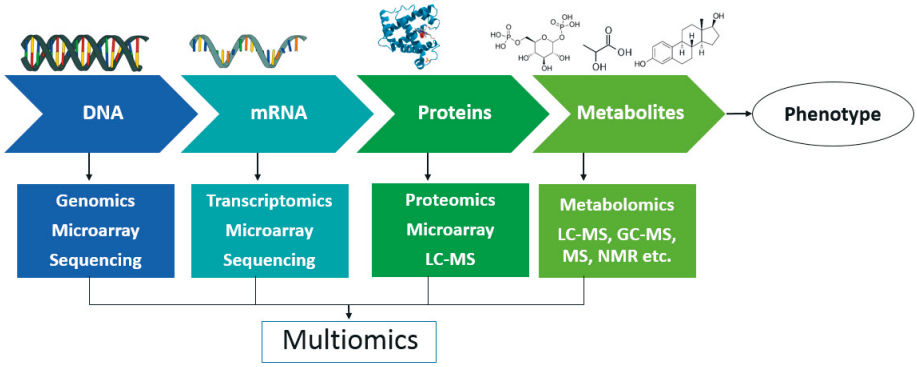

Figure 2 Overview of different -omics platforms and applied analytical techniques. LC-MS; liquid chromatography-mass spectrometry, GC-MS; gas chromatography-mass spectrometry, MS; mass spectrometry, NMR; nuclear magnetic resonance spectroscopy.

1.2.1 METABOLOMICS

Metabolomics is a growing discipline that aims to detect and comprehensively analyze all the small molecules in a biological system simultaneously and compare the levels between different conditions (i.e. disease, diet, treatment, or lifestyle).[1–3] Changes in metabolite levels reflect cell function, as they represent the downstream amplification of changes occurring at the mRNA or protein levels. Whereas genes and genetic risk indicate what might happen, metabolites indicate what is currently occuring and are closest to phenotype, simultaneously representing genetic and environmental alterations (Figure 2).[2,8] It is also known that due to protein modifications, signaling and enzymatic activity does not depend only on the protein levels.[9] Thus metabolomics provides complementary information compared to other omics and a systems-wide understanding of biological function.[8] The metabolome is also highly dynamic in nature and metabolite turnover can be much faster and changes in the metabolite levels can be greater compared to the proteome and transcriptome.[10,11] Metabolomics also offers higher analytical throughput compared to proteomics and transcriptomics, which significantly lowers the costs of analysis. However, the chemical variety of metabolites is large compared to genes or transcripts, which consist only of four nucleotides, or proteins, which are built from 20 amino acid subunits and are commonly detected with a single analytical platform.[10] Due to the high diversity of the chemical and physical properties of metabolites, there is currently no comprehensive universal analytical tool to detect all metabolites simultaneously. Accordingly, multiple strategies are employed for a wide metabolite coverage.[11] An additional challenge is the highly variable metabolite concentrations, which makes analysis technically challenging. The most commonly utilized analytical tools in metabolomics are mass

spectrometry (MS) or nuclear magnetic resonance spectroscopy (NMR).[11] Metabolomics can be further divided into smaller branches with a focus on certain chemical classes, such as lipids (lipidomics),[12,13] steroids (steroidomics),[14] and sugars (glycomics).[15]

1.2.1.1 Non-targeted metabolomics

Non-targeted metabolomics or metabolic profiling aims to detect and compare the levels of all metabolites in a biological system without prior knowledge of the compounds of interest. This is a useful approach in studies with a general hypothesis of expected metabolic differences, but where no specific scientific hypothesis on the differences exists. Thus, non-targeted metabolomics is hypothesis creating and can provide novel insights into metabolic changes related to the biological question.[16]

Non-targeted metabolomics aims to detect metabolites with as wide and universal coverage as possible from several metabolite classes. As typically no standard compounds are applied, the analysis is semi-quantitative with relative abundance; absolute concentrations cannot be determined. The repeatability and reliability of non-targeted methods are not as good as that of targeted methods and validation is more difficult. For reliable results, careful experimental design and quality control is necessary and in an ideal case the findings are later validated with a targeted approach.[17,18]

1.2.1.2 Targeted metabolomics

Targeted metabolomics is focused on a previously determined set of metabolites of interest that are analyzed to test a specific hypothesis.[19] Targeted metabolomics methods are usually also quantitative based on appropriate standard compounds and labeled internal standards. In fact, such methods are classical quantitative methods that have been applied in bioanalysis for a long time. Targeted methods are commonly specific, sensitive, accurate, and are usually applied in the validation and confirmation of the hypothesis. Targeted methods are also widely applied for the analysis of metabolites of which the concentrations are too low to be detected with non-targeted methods. To achieve high specificity, sample preparation and analysis methods can be optimized for certain compound classes, such as bile acids,[20] acylcarnitines,[21] acyl-coenzyme A:s,[22] amino acids,[23] and steroids.[24] The methods can also lie between targeted and non-targeted, referred to as semi-targeted methods.[24,25]

1.2.1.3 Lipidomics

Lipidomics is a branch of metabolomics that aims to analyze all lipid species simultaneously and compare the lipid content between different conditions.[12,13] Lipidomics has become an important research field due to the increased awareness of lipid functions in the cell and their role in many common diseases.[12,13] The polarity of lipids vary substantially from other common metabolite classes, which can be commonly analyzed under classification of "polar metabolites". Thus, the selection of extraction and analysis method in lipidomics is commonly based on the non-polar properties of lipids.[12,13] The challenges in lipidomics, such as structural diversity and the wide concentration range of lipids, are similar to metabolomics in general. However, lipids usually consist of repeating building blocks (fatty acid chains and phospholipid functional groups), which assist in analysis and identification.[12,13]

1.2.1.4 Applications of metabolomics and lipidomics

Metabolomics has become an important and widely applied tool in plant,[26] environmental,[27] food and nutrition,[28] microbial,[29] and mammalian studies.[2,3] Mammalian and human metabolomics have primarily been applied for biomarker discovery in disease diagnostics and prognosis, understanding disease mechanisms, identifying novel drug targets, drug therapeutics, and precision medicine.[2,3] Metabolomics has been used to study several diseases or risk factors of disease progress, such as Parkinson's disease,[30] Alzheimer's disease,[31] diabetes,[32] neuropsychiatric diseases (i.e. schizophrenia, depression, anxiety, psychosis),[33,34] several cancers,[35] multiple sclerosis,[36] cardiovascular diseases,[37] psoriasis,[38] traumatic brain injury,[39] and stroke[40]. The list is endless and multiple diseases, including genetic disorders, have been studied with metabolomics.[41] Screening of metabolic inborn errors is already routinely performed in clinics.[42,43] Yet, the emphasis of metabolomic studies is more in multi-factorial disorders that do not have a single genetic cause but are triggered by multiple factors, interactions, and lifestyle. Such studies aim to identify possible sensitive and specific biomarkers for clinical diagnostics or early prognosis in cases where there are no current markers or the current markers are poor or require expensive analyses.[44,45] Another important application of metabolomics includes studying disease mechanisms or searching for possible drug targets and often in combination with other -omics studies.[2] Cellular metabolic changes and mechanisms related to drug treatment efficacy (pharmacometabolomics) and drug side-effect variation is also interesting and studied field.[46,47] Pharmacometabolomics aims to facilitate personalized medicine and the selection of treatments for different subpopulations of patients to maximize drug efficacy and to

minimize toxicity and side effects.[46] In particular, metabolomics for different cancer therapies and for statins have been investigated. [47,48]

The interaction of human and gut microflora metabolites can have an impact on human health and can offer insights on lifestyle and diet. Accordingly, interest in metabolites produced by intestinal microbes has increased.[4,5,49] The relationships between gut microbiota metabolomic interactions and diabetes,[5] neurodegenerative disorders,[4] and non-alcoholic fatty liver disease have been examined.[50]

Different mechanistic and metabolic regulation studies have also adopted the application of stable heavy isotope labeling (^{13}C, ^{15}N, ^{2}H, ^{18}O) to study the reaction rates and metabolic fluxes inside the system.[51,52] When cells are grown on a heavy isotope-enriched substrate, the heavy isotopes propagate through the metabolic network according to the active metabolic pathways.[51,52] This is referred to as fluxomics, which is a separate branch of metabolomics that uses labeled patterns to identify active metabolic pathways in cells to characterize the metabolic phenotype.[51,52]

Cells even in isogenic culture are heterogeneous populations that encapsulate different cell phenotypes due to genetic, epigenetic, and environmental factors.[53] Single-cell metabolomics aims to study and understand phenotypic heterogenity.[54,55] This can be useful in the study of therapeutic effects for different cell phenotypes or to identify metastatic cancer cells.[54,55] While genes and transcripts can be multiplied with polymerase chain reaction (PCR), amplification of metabolites is impossible, which is a unique challenge in single-cell metabolomics studies.[54,55] In the analysis of tissues, cell heterogeneity is also present. When a tissue sample is homogenized, information on the original distribution of compounds in the tissue is lost. To study the spatial distribution of metabolites and drugs in the sample, mass spectrometry imaging (MSI) is commonly applied.[56]

1.3 ANALYTICAL METHODS IN METABOLOMICS AND LIPIDOMICS

Analytical platforms in metabolomics should be highly sensitive, accurate, reproducible, and able to characterize simultaneously as large portion of the metabolome as possible. These demands can only be partially fulfilled either with MS or NMR.[11,17,57] Although NMR is universal, non-destructive, and suitable for a wide range of chemical structures, it suffers from low sensitivity and detection is limited to mainly high-abundance metabolites.[11,57] On the other hand, MS has high sensitivity and specificity.[11,17] MS-based metabolomics can be performed by infusing sample directly, although MS is commonly coupled with a separation technique such as liquid chromatography (LC),[2,58,59] gas chromatography (GC),[59,60] or capillary electrophoresis (CE).[59,61] All these methods have their own limitations and advantages and none of these methods can detect all metabolites

simultaneously due to varying physicochemical properties and the wide concentration ranges of the metabolites. A typical workflow in metabolomics contains sampling, sample pretreatment, analysis, data processing, data analysis, and data interpretation (Figure 3).

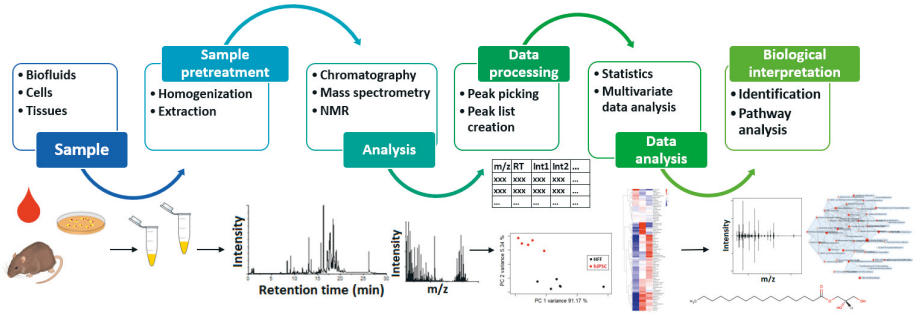

Figure 3 Workflow of non-targeted metabolomics analysis.

1.3.1 STUDY AND EXPERIMENTAL DESIGN

Metabolomics study design is one of the most important parts of a successful experiment.[11,18] Experimental studies compare different treatments or multiple experimental factors at once in controlled manner. However, in metabolomics, the experiments are more often observatorial studies, such as case-control, cross-sectional, cohort, or longitudinal studies.[11,18] A case-control study is where the subjects with a certain condition (i.e diet, disease) (cases) are compared to otherwise similar subjects without the condition (controls). A cross-sectional study compares a population at a certain time point, a cohort compares a group of people with common characteristics (i.e birth, exposure), and a longnitudinal study is a cohort study followed over a long period of time.[11,18] In metabolomics, it is important to discriminate the possible covariables (e.g. age, sex, medication, and clinical variables), which may affect the observed metabolic differences. Thus, cases and controls should be carefully matched considering all possible covariables.[11,18]

A sufficient number of samples is needed for statistical and prediction power in metabolomics. However, compromises are commonly made according to costs, time, and available resources. In biomarker discovery, the study is usually designed to create a training set for biomarker modeling and the test set is used to independently validate the diagnostic performance of the tentative biomarkers.[18,58]

Sampling and sample storage are important parameters that affect the detected metabolite levels. Sampling protocols should be similar even if samples are collected at multiple sites over a long period of time.[62] The impact of blood sample collection conditions (i.e. fasting time, season, and time of day for blood collection, sample collection tubes) on metabolite levels is evident.[62] In experimental design, recommended steps and points to consider include

randomization prior to sample preparation and injection order, avoiding possible errors due to sampling and sample storage, sample preparation, analytical response and correction over time, and application of quality-control (QC) samples.[18]

1.3.2 SAMPLE MATRIXES, COLLECTION, AND STORAGE

The sample matrix in metabolomics can be any biological matrix, although typically biofluids such as serum,[21,63] plasma,[23,38,63] or urine[24,64,65] are practical especially for biomarker search purposes, as such samples are homogenous and easy to acquire. Other matrixes, such as cerebrospinal fluid,[66] saliva,[67] feces,[68] tissues,[59,69] or sweat[70] have also been used. The selected sample type depends on the application, the studied phenomena, and the availability of the sample. Biofluids are easy to collect and provide a snapshot of a mammalian system. For example, cerebrospinal fluid closely reflects concentrations in the brain and has been used in studies of neural disorders.[30,31] On the other hand, tissues are more specific and are frequently used to study the biological mechanisms of organs.[59] However, for clinical diagnostics tissues are not convenient and sampling should preferably be fast and easy.[45]

A proper and objective sample collection, storage, and sample pretreatment are key issues in the success and reliability of metabolic measurements. Biological samples should ideally be stored at low temperatures (e.g. -80°C) immediately after collection and the number of freeze-thaw cycles should be minimal. Sampling should be representative and sample containers should not cause non-specific binding or surface adsorptions.[71] An anticoagulant in plasma sample preparation (e.g. heparin, EDTA, or citrate), sample collection tube selection, and sample collection protocol may influence the detected metabolites and all samples should be treated equally to avoid any bias.[62,72,73]

Quenching is a process that aims to eliminate metabolic fluxes and interconversion to other metabolites by inactivating enzymes in the sample. Quenching is particularly important with cell and tissue samples.[58,71] Quenching can be part of sampling (e.g. cell harvesting and tissue sectioning) or integrated within the sample pretreatment.[74] This is typically performed by adding organic solvent or buffer solution, increasing or decreasing the temperature, or both.[58,71,75,76] The most common quenching methods are addition of ice-cold acetonitrile (ACN), methanol (MeOH), buffer solutions (e.g. ammonium bicarbonate, phosphate buffered saline [PBS] or sodium chloride), or snap-freezing in liquid nitrogen.[58,71,75,76]

1.3.3 SAMPLE PRETREATMENT

Sample pretreatment of biological samples in non-targeted metabolomics should preferably be universal and minimal to prevent potential loss or conversion of metabolites.[17,75,77] Sample pretreatment requirements depend on the matrix, analytes, and analytical method. In metabolomics, sample pretreatment can commonly consist of homogenization, cell lysis, protein precipitation (PPT), liquid-liquid extraction (LLE), solid-phase extraction (SPE), derivatization, evaporation, and reconstitution.[17,69,74,75,77] Sample pretreatment can be quite straightforward, for example removing proteins, salts, urea, or other interfering compounds. Therefore, for biofluids with low protein content (e.g. urine or sweat), the sample is often only diluted prior to analysis.[65] For samples with a high protein content (e.g. serum, plasma, tissue), PPT with organic solvent is commonly used, which at the same time enables extraction of a wide range of various metabolites.[69,75,78] Sample extraction and purification with LLE or SPE is often also necessary to remove matrix interferences and concentrate the analytes. Derivatization is often required in GC-MS applications and sometimes in LC-MS to enhance the ionization efficiency or to increase retention to the LC column.[79–81] Evaporation and reconstitution are often the last steps and are used to concentrate analytes or change the solvent to one compatible with the analysis method, although analyte solubility and potential oxidation should be considered.

1.3.3.1 Extraction and protein precipitation

Extraction and PPT with solvents such as MeOH, ACN, ethanol (EtOH), isopropanol (IPA), acetone, or a mixture of these with water or each other is the most commonly applied sample pretreatment protocol in non-targeted metabolomics.[17,75,78,82–84] Although PPT approaches have been evaluated in terms of protein-removal efficiency, metabolite coverage, precision, repeatability, stability, and extraction recovery in several studies, there is currently no general consensus of the best PPT approach in metabolomics.[78,82,85] Alternative procedures for removing proteins are ultrafiltration and turbulent flow chromatography.[86,87] Thus far, both of these methods have shown poor metabolite recoveries in comparison with solvent-based methods.[86,87]

The selection of sample extraction solvent depends significantly on the polar range of the analytes to be extracted. Figure 4 shows which kinds of metabolites can be extracted with commonly used solvent systems in metabolomics. For the extraction of highly polar metabolites (left side in Figure 4), additional water is essential. In contrast, addition of non-polar solvent (e.g. chloroform) is required for the extraction of non-polar lipids such as triacylglycerols (Figure 4). Addition of chloroform or another non-polar

solvent with a certain solvent ratio forms a two-phasic LLE system, where two solvent layers are immiscible with each other. LLE is widely utilized in non-targeted lipidomics applications as well to simultaneously extract the polar (metabolites) and non-polar (lipid) fractions.[17,75,88] The most popular LLE methods in lipidomics are extractions with chloroform-MeOH mixtures such as Folch extraction,[89] Bligh and Dyer,[90] or extraction with methyl-*tert*-butyl ether (MTBE)/MeOH mixture, referred to as the Matyash method.[91] Furthermore, modifications of these with different solvent ratios are popular.[92,93] Additionally, a two-phase extraction, for example with MeOH/chloroform/MTBE mixture,[94] dichloromethane/MeOH mixture,[95] or two-step extraction with butanol/MeOH followed by heptane/ethyl acetate/acetic acid have been applied in lipidomics.[96] In some studies, a single extraction protocol (chloroform/MeOH or MTBE/MeOH) has been used to collect both layers of biphasic extraction, with the lipophilic solvent containing non-polar compounds and the hydrophilic solvent containing polar compounds.[88] The benefit of two-phase extraction is the wider metabolite coverage from the same sample. However, medium polar compounds are distributed to both phases. LLE can also be performed as a two-step extraction protocol by extracting first polar compounds followed by lipid extraction from the same samples.[97] Subsequent supernatants can be analyzed separately or, alternatively, the polar and non-polar fraction can be pooled into one sample.[98]

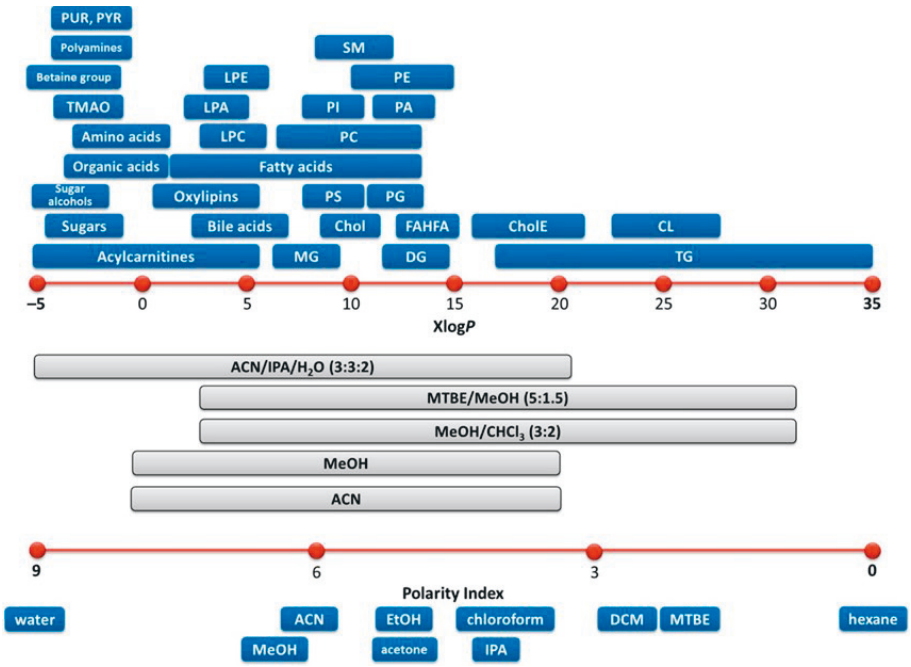

Figure 4 Predicted octanol/water partition coefficient (Xlog*P*) ranges of common metabolite classes detected in blood plasma (top), polarity ranges of isolated metabolites with typical solvents or solvent mixtures used in metabolomics and lipidomics (middle), and polarity indexes of solvents in sample extraction (bottom). Cer, ceramides; Chol, cholesterol; CholE, cholesteryl esters; CL, cardiolipins; DG, diacylglycerols; FAHFA, fatty acid esters of hydroxyl fatty acids; LPA, lysophosphatidic acids; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamines; MG, monoacylglycerols; PA, phosphatidic acids; PC, phosphatidylcholines; PE, phosphatidylethanolamines; PG, phosphatidylglycerols; PI, phosphatidylinositols; PS, phosphatidylserines; PUR, purines; PYR, pyrimidines; SM, sphingomyelins; TG, triacylglycerols; TMAO, trimethylamine N-oxide. Reprinted with permission from [17]. Copyright 2019 American Chemical Society.

1.3.3.2 Solid-phase extraction and solid-phase micro extraction

SPE has high selectivity but often at the cost of metabolite coverage. SPE is thus less frequently used in non-targeted metabolomics than solvent-based extraction methods.[99] SPE is also often more time consuming and the extraction protocol and method optimization can be more complex due to the unique selectivities of SPE sorbent materials.[99,100] SPE has often been applied in non-targeted metabolomics and lipidomics for desalting and concentrating urine samples,[101–103] for removing phospholipids or acylglycerides to enhance sensitivity of other analytes,[104–108] or fractioning of the sample to smaller subsets of analytes.[109–111] Although metabolomics analysis using SPE fractionation is time consuming, the method provides detection of a higher number of metabolites, improved chromatographic separation, and reduced ion suppression in MS analysis.[110] Several different sorbent materials (i.e.

modified silica, graphite, polymer, and zirconium) with different column chemistries such as C18, HILIC, mixed-mode, and ion exchange have been studied in non-targeted metabolomics applications.[101–103,109–111] Mixed-mode cartridges with multiple interaction mechanisms in the cartridge have been shown to be suitable for a wide range of metabolites.[101,102] C18 also provides universal material for a large range of metabolites,[101] but the retention of many highly polar compounds may be poor. SPE is especially useful in targeted methods where the high sensitivity and selectivity is desired. For example, SPE has been applied for the analysis of purines and pyrimides,[112] bioactive lipids,[113] neurotransmitters,[114] and bile acids.[106] Although the SPE protocol can be quite slow, it can also be completely automated (e.g. online-SPE) or multiple samples can be extracted simultaneously using a 96-well format.[100,111]

Solid-phase micro extraction (SPME) has been used in metabolomics primarily to extract volatile and semi-volatile metabolites, such as food flavors and aromas, human breath, or skin biomarkers.[115,116] SPME is less frequently used to extract non-volatile metabolites, such as lipids or amino acids in biofluids.[115,116] For example, SPME has been utilized to study emissions of volatile metabolites during stem cell differentiation,[117] volatiles from pathogens,[118] and cellular metabolites from *E. coli*.[119]

1.3.3.3 Derivatization

Derivatization is required for polar metabolites prior to GC-MS analysis to increase volatility, thermal stability, and chromatographic mobility; to improve chromatographic peak shape; and to reduce peak tailing.[60] Derivatization can also add desired diagnostic fragments to mass spectra for identification purposes.[60] Carboxylic acids, alcohols, amines, and thiols can be derivatized by alkylation, acylation, or silylation.[60,79] The most commonly applied derivatization prior to GC-MS analysis of metabolites is silylation with trimethylsilyl (TMS) reagents. *N*-Trimethylsilyl-*N*-methyl trifluoroacetamide (MSTFA) and N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) are the most popular silylation reagents in metabolomics, although other reagents are available as well.[60,63] In some cases, addition of trimethylchlorosilane (TMCS) as a catalyst to silylation reaction, use of pyridine as a solvent, and heating can be used to accelerate the derivatization with MSTFA or BSTFA.[120] In non-targeted metabolomics, the most popular protocol prior to GC-MS analysis is two-step derivatization with oximation usually with a methoxyamine (MOX) followed by silylation.[63,66] The MOX protects carbonyl moieties and prevents formation of multiple reaction products during silylation.[121] As the silylation reagent and the derivatized extracts are sensitive for hydrolysis, the presence of water in the silylation reaction mixture must be avoided in all phases of sample treatment.[79]

Different techniques, such as offline derivatization,[122] microwave-assisted derivatization,[123] in-time,[66] and injection-port derivatization[124] can be applied

in the derivatization process. Offline derivatization is currently the most applied approach, where derivatization is performed simultaneously on all samples in a batch and analysis is recommenned to be completed within 24-48 h after the derivatization.[122,125] Microwave-assisted derivatization is used to reduce the derivatization time to just a few minutes.[123] In-time derivatization is performed by robotics just before injection into GC. Samples are not exposed to degradation during storage on a sample tray prior to injection as in offline derivatization.[66]

Chemical derivatization can also be applied prior to LC-MS or direct infusion analysis to increase ionization efficacy in electrospray (ESI) ionization of poorly ionizable or non-ionizable compounds, such as neutral alcohols, phenols, and some steroids.[80,126,127] Derivatization can also improve the retention of highly polar compounds to a reversed phase LC column that decreases matrix effects in the LC-MS analysis.[126,127] For example, the derivatization of carboxylic acids in the tricarboxylic acid (TCA) cycle,[128] amino acids,[129] neurotransmitters,[130] phospholipids,[131] and steroids prior to ESI-LC-MS analysis have been used in targeted metabolomics.[80,132] Multiple derivatization reagents can be used depending on the derivatized functional group or the application. For example, o-benzylhydroxyl amine for derivatizing carboxylic acids ,[128] hydroxylamine for ketones,[132] and dansyl chloride for aliphatic alcohols, phenols, and amines have been applied in targeted metabolomics.[126,127,130]

1.3.4 MASS SPECTROMETRY

1.3.4.1 Ionization methods

The most common ionization methods in LC-MS and direct infusion mass spectrometry (DI-MS) are electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and atmospheric pressure photoionization (APPI).[133] ESI is the most popular and widely applied ionization technique in metabolomics.[20,104,134,135] ESI is a soft ionization method suitable for analysis of a wide range of small and large molecules. ESI provides good ionization efficiency for medium polar, polar, and ionic compounds. These compounds are normally ionized via protonation, deprotonation, cation, or anion attachment (i.e. adduct formation). However, the ionization efficacy in ESI towards non-polar neutral compounds may be poor. ESI is also prone to ion suppression that may disturb quantitation of metabolites. Although APCI or APPI are still used quite rarely in metabolomics, APCI and APPI are more applicable for small to medium weight non-polar and polar compounds.[136–139] In particular, APPI has shown high sensitivity towards non-polar lipids, such as steroids.[140–142] Other ionization techniques such as matrix-assisted laser desorption ionization (MALDI), secondary ion mass spectrometry (SIMS),

and various ambient MS methods have also been utilized in metabolomics and MS imaging of metabolites.[143–146]

The two main ionization methods in GC-MS are vacuum techniques, specifically electron ionization (EI)[60,66] and chemical ionization (CI).[60,147] Typically EI with commonly applied 70 eV electrons results in extensive, characteristic fragmentation and the spectra can be searched against widely available EI-MS spectral libraries for identification. However, the fragmentation decreases sensitivity and selectivity and may eliminate the formation of molecular ions, which are important for identification of analytes. CI provides softer ionization and causes less fragmentation than EI. Usually in CI protonated or deprotonated ions are also formed allowing determination of molecular mass. However, CI is not commonly used in metabolomics. GC can also be connected to MS by using API methods, such as APCI,[148,149] APPI,[142,149,150], and ESI.[149,151] In API sources, abundant molecular ions or protonated or deprotonated molecules are formed, which can be further fragmented selectively by MS/MS. When using atmospheric pressure ionization (API) sources, both GC and LC can be connected to the same mass spectrometer equipped with an atmospheric pressure to vacuum ion optics, without the need for separate expensive mass spectrometers. However, the disadvantage of all API techniques is that significant numbers of ions are lost in the ion transfer from atmospheric pressure to the vacuum of MS.[152] Ion transmission in photoionization has been improved with systems in which the photoionization occurs inside a transfer capillary between the atmospheric pressure and vacuum.[153–155] These systems have shown high sensitivity towards non-polar compounds when applied as an interface in GC-MS or DI-MS.[153–155]

1.3.4.2 Mass analyzers

Various mass analyzers have different applicable acquisition speeds, mass resolution, mass accuracy, mass range, dynamic range, and scan modes, which should be considered when selecting a suitable metabolomics method. Mass analyzers can be divided into high-resolution instruments measuring accurate mass such as time-of-flight (TOF), Fourier transform ion cyclotron resonance (FT-ICR), and Orbitrap, or to unit resolution mass analyzers such as quadrupole (Q) or ion trap (IT).[11,17] Some of these analyzers (FT-ICR and IT) are capable for tandem mass spectrometry (MS/MS or MS^n) and selective fragmentation of selected precursor ions. Mass spectrometers can be also hybrid instruments, containing two or more mass analyzers also capable for MS/MS analysis. Hybrid instruments, such as Q-TOF, triple quadrupole (QQQ), and IT-Orbitrap are widely applied in metabolomics.[11,17] The most common collision cells for fragmentation in MS/MS measurements are collision-induced dissociation (CID) and higher energy collision dissociation (HCD), which is a CID-type collision cell available in Orbitrap. Additionally,

ultraviolet photodissociation (UVPD), which reveals the double bond locations of complex lipids, has been applied but with less frequently.[156,157] High-resolution mass spectrometers (HRMS) are often used in non-targeted analysis,[11,17,58] whereas QQQ using multiple reaction monitoring (MRM) scans are more frequently utilized in targeted analysis.[19,20,23]

In non-targeted metabolomics, typically an MS full scan is first recorded and additional targeted MS/MS analysis is performed based on the precursor ions of interest. Different improvements in HRMS analyzers and software have enabled enhanced acquisition rates and made it possible to collect several MS/MS experiments within a single run in addition to the MS scan.[158] These methods are data-dependent acquisition (DDA) and data-independent acquisition (DIA).[159,160] In data-dependent methods, MS/MS acquisition can be triggered with an intensity threshold or by applying inclusion and exclusion lists or other set thresholds.[161] However, the number of produced spectra and the spectral quality is limited. DIA instead aims to fragment and collect MS/MS-spectra from all precursor ions and acquisition is not affected by the acquired data.[160] One DIA method is sequential window acquisition of all theoretical fragment-ion spectra (SWATH), which fragments all precursor ions within defined a RT and precursor *m/z* window.[162] More traditional precursor ion scans (PIS) or neutral loss scans (NLS) with QQQ to detect a certain compound class are commonly applied in non-targeted or semi-targeted lipidomics.[13]

1.3.5 GAS CHROMATOGRAPHY-MASS SPECTROMETRY

Gas chromatography-mass spectrometry (GC-MS) is a widely applied method in metabolomics and is suitable for analysis of volatile and thermally stable low-molecular weight (MW <600 Da) compounds.[60] GC-MS is efficient, robust, sensitive, selective, reproducible, and has good resolution in separation. GC-MS does not suffer as much from common drawbacks such as matrix effects and ion suppression encountered in LC-MS. The separation system is quite simple with only one mobile phase. Separation is based mainly on evaporation in the order of boiling point with influence from analyte interactions with the column. Separation and sample pretreatment in GC-MS may be time consuming and GC-MS is not suitable for non-volatile higher molecular weight biomolecules, such as many lipids and peptides. The most common procedure in GC-MS is extraction of a metabolomics sample followed by chemical derivatization, which is required especially for polar or non-volatile metabolites with poor thermal stability.[79] Thus far, GC-MS methods with unit resolution are most commonly used in non-targeted and targeted analysis.[60] Identification in non-targeted (and targeted) analysis is typically based on characteristic EI-MS spectra that can be searched against broad spectral libraries. In targeted methods, MS/MS with a QQQ is also frequently used, which provides excellent selectivity and sensitivity in monitoring of

metabolites.[142,163] However, HR instruments are becoming more popular in metabolomics, with the ability to measure accurate mass and define the elemental formula.[164,165] GC-MS has often been applied for non-targeted analysis of metabolites such as amino acids, organic acids, and fatty acids after derivatization.[60,66]

GCxGC-MS is also a widely used technique in metabolomics due to its high resolving power in separation.[66,166] The first dimension commonly utilizes a non-polar 5% diphenyl- and 95% dimethyl-polysiloxane column and the second dimension is a short polar column after modulation.[66] The separation in the second dimension lasts only some seconds.[66] Data processing and peak picking are challenges, however, the separation and spectral purity are improved without a significant increase in the analysis time.

1.3.6 LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY

Liquid chromatography-mass spectrometry (LC-MS) is the major analytical technology in global metabolomics due to its high sensitivity, repeatability, and wide coverage of metabolites with diverse chemical properties.[167] Ultra-high pressure liquid chromatography (UHPLC) has replaced traditional high-pressure LC instruments (HPLC) in metabolomics due to improved peak capacity, sensitivity, and faster analysis.[167,168] Sensitivity and peak capacity can be improved even further with nano- or capillary-UHPLC. Nano- and capillary-LC combined with nano-ESI offers improved chromatographic resolution and sensitivity and is especially suitable for low-volume metabolomics samples.[102,169] However, the analysis time increases markedly with lower flow rates and thus nano- and capillary-LC are less frequently used in metabolomics.[102] The most widely applied tool in metabolomics is reversed-phase (RP) chromatography with C18 column and gradient elution, which is highly suitable for a large range of metabolites.[167] Other columns such as C8 or C30 are also applied but less frequently.[170,171] However, highly polar and ionic compounds do not have adequate retention on RP materials and these compounds elute in the void volume. This increases the need for other column chemistries, such as hydrophilic interaction chromatography (HILIC) or normal phase (NP) chromatography.[172–174] Commonly used NP solvents, such as hexane and heptane, are poorly suitable for the ESI ionization technique typically applied in LC-MS, which is why NP is rarely used in metabolomics. This has made HILIC an attractive alternative, as it offers a contrary elution order to RP with common organic solvents (such as ACN and MeOH) and increases elution strength with increasing water content.[172,173] HILIC is nowadays often applied in targeted and in non-targeted metabolomics to detect polar metabolites.[22,170,172,173,175,176] The disadvantages in HILIC are the long re-equilibrium times and poor repeatability of retention times.[167,172] Some highly polar ionic compounds may have better retention

with ion chromatography or ion-pair chromatography using an RP column.[177,178]

Lipids can be separated with NP or HILIC based on their polar head group, which separates lipid classes,[136,179] whereas RPLC separates lipids based on polar head group and number of carbon and double bonds in the fatty acid alkyl chains of lipids.[12] These two main lipidomic approaches can also be combined into a two-dimensional (2D) separation.[136,180] 2D-LC-MS systems have been used to some extent and have used HILIC and C18 colums for the simultaneous analysis of polar metabolites and lipids.[136,181,182] In comprehensive 2D-LC (e.g. LCxLC), the whole effluent from the first dimension is transferred into the second dimension, using a dedicated switching valve and running ultrashort gradients. In heart-cutting 2D-LC,[182] only one or two fractions from the first dimension effluent are transferred into the second dimension with longer gradients than in comprehensive 2D-LC.[181] Although 2D systems increase peak capacity and selectivity, the analysis time and complexity of data processing limits the application of these systems in metabolomics.[181]

1.3.7 DIRECT INFUSION AND MICROCHIP METHODS

Direct infusion (DI) of crude metabolite extracts to MS shows great potential due to its wide metabolite coverage, high throughput, and simple and rapid analysis.[183] However, DI-MS can suffer from extensive ion suppression and the inability to separate isobaric and isomeric substances. DI metabolite analysis is usually performed using a syringe pump or a chip-based nano-ESI source (Nanomate) that constantly delivers the sample.[13,135,183,184] Several other microchip-based methods than commercial Nanomate have also been proposed for the direct analysis of metabolites.[185,186] Alternatively, the sample can be injected with a LC autosampler as a plug to constant solvent flow (i.e. flow injection analysis [FIA]).[135,187] Since no chromatographic separation is used in DI or FIA analysis, HRMS is commonly used to improve peak capacity in non-targeted metabolomics.[135,183] MS/MS methods have been commonly used in targeted DI or FIA metabolomics analysis.[183] Even commercially available kits for quantifying metabolites from different compound classes are available.[188] DI- or FIA-MS/MS is most commonly used in lipidomics (i.e shotgun lipidomics) utilizing several experimental conditions and PIS and NLS for selectively ionizing different lipid subgroups at a time.[13,184]

Ambient MS techniques, in which metabolites are directly analyzed from biological samples without the necessity for sample pretreatment, can be considered as a special kind of DI-MS method.[144,146,189,190] The most commonly applied ambient techniques in metabolomics are desorption electrospray (DESI) and direct analysis in real time (DART), which are utilized especially in tissue imaging applications in addition to MALDI-MS and SIMS.[143–146,189–191]

1.3.8 OTHER METHODS

Capillary electrophoresis-mass spectrometry (CE-MS) is suitable for analysis of ionic compounds, which are separated based to their charge-to-size ratio.[59,61] CE-MS is especially suitable for amino acids, small peptides, nucleosides, nucleotides, sugar phosphates, organic acids, and sulfated compounds, which can be challenging in LC-MS analysis.[59,61,192] CE-MS has high sensitivity, selectivity, has the potential for large-scale metabolic profiling, and is suitable for low-volume samples down to sub-cellular volumes. [59,61,193] CE-MS suffers from variable migration times, technical challenges in coupling CE to MS, and lack of standard methodologies.[59,61] However, a recent cohort study applying CE-MS showed good reproducibility and suitability for large-scale studies.[194]

Ion-mobility spectrometry (IMS) separates ions in the gas phase based on their mobility in a carrier buffer gas when they travel through the ion mobility separation cell under an electric field.[195] The drift time is the time required for an ion to travel through the cell, which again depends on ion properties such as charge, shape, and size.[195–198] The separation is rapid, with drift times at the millisecond scale. Another advantage is that the collision cross section (CCS) values derived from the drift times represent a physicochemical property that can be applied in the identification of metabolites or lipids.[195–198] IMS combined to MS (IMS-MS) provides a very rapid and relatively selective analysis and is increasingly applied in metabolomics and especially in lipidomics.[195] IMS has also been combined to LC-MS systems to improve peak capacity, signal-to-noise, and purity of MS/MS spectra.[195,199,200] CCS structural databases and libraries, commercially available instruments, improvments in separation resolution, and the available preprocessing tools will increase the usage of IMS in the future.[195–198,201]

Supercritical fluid chromatography-mass spectrometry (SFC-MS) is less frequently applied but offers the possibility for complementary separation with properties combining GC and LC.[202] SFC-MS has been applied in lipid profiling and in targeted methods.[203,204] A comparison of SFC-MS with HILIC LC-MS and DI-MS for lipidomics demonstrated that SFC-MS has excellent sensitivity for less polar lipid classes and high sample throughput and the SFC-MS method is applicable for lipidomics of biological samples.[202]

1.3.9 DATA PREPROCESSING AND ANALYSIS

1.3.9.1 Data preprocessing

Data processing and peak integration is straightforward and routine in targeted metabolomics and instrument vendor programs are commonly

applied. Non-targeted metabolic raw data can be preprocessed with commercial instrument vendor software, such as MassProfiler (Agilent), MetaboScape (Bruker), Chromatof (LEGO), MarkerView (SCIEX), Compound Discoverer (Thermo Fisher), and Markerlynx (Waters). These programs are easy to use, well tested, and documented. However, these are commonly fixed for only a certain data format, the applied algorithms are nontransparent, and the software can be expensive. Open-source programs are instrument-independent, freely available, allow modifications, and the functions are usually more transparent. The most commonly applied open-source tools for metabolomics MS data preprocessing are XCMS,[205] MzMine2,[206] OpenMS,[207] MetaboAnalyst,[208] MSDial,[209] Workflow4metabolomics,[210] Galaxy workflow,[211] MetAlign,[212] and AMDIS,[213] among multiple others available.[214] Many of these open-source online platforms include data pre-processing, post-processing, and data analysis steps and are suitable for GC-MS, LC-MS, and DI-MS data. Raw data conversion for common open data formats (mzXML, mzML, netCDF) may be necessary before applying open-source tools. Data can be converted typically with instrument software or the freely available proteowizard.[215] Preprocessing steps aim to reduce the data complexity and noise and to select the metabolic signals to create a data matrix for further data analysis. The data pre- and post-processing pipeline commonly includes noise filtering, baseline correction, peak picking, peak deconvolution, alignment, filtering, normalization, scaling, annotation, and identification. Several algorithms can be applied in these steps and not all are mandatory.[214,216] Two-dimensional separations (i.e. GCxGC and 2D-LC) and IMS-MS applications requires certain chemometrics and data preprocessing algorithms.[201,217,218]

1.3.9.2 Metabolite identification

The non-targeted workflow in metabolomics usually contains HRMS analysis, feature detection, and data analysis, followed by identification of interesting and relevant metabolic features with MS/MS.[158] With modern instruments, MS/MS data can also currently be collected simultaneously with DDA and DIA modes.[158] Metabolite identification is the most time-consuming step and is a bottleneck in metabolomics. Metabolite identification can be classified to different classes (Level 1-4.) based to their identification confidence.[219] Level 1 corresponds identification against authentic chemical standards, level 2 corresponds to putative annotation and identification against MS/MS spectral databases. Level 3 is characterized compound class and Level 4 is an unknown feature with a unique mass-to-charge ratio and possible additional retention time.

For identification purposes of LC-MS or DI-MS data, multiple online databases are available for tentative level-3 identification with molecular mass, including METLIN,[220] Human metabolite database (HMDB),[221] Kyoto encyclopedia of genes and genomes (KEGG),[222] Lipid Maps,[223] and

DrugBank.[224] With MS scan and accurate mass measurements by HRMS, the identification of metabolites is based on the molecular weight of diagnostic ions. In addition to retention time in GC- and LC-MS analysis, isotope abundancies and possible adduct ions may provide additional information for identification. MS/MS spectra provide additional structure information of the detected feature. Measured MS/MS spectra can be searched against commercial spectral databases (NIST 14, MetaboBASE), open-source databases (HMDB,[221] METLIN,[220] MassBank,[225] mzCloud (https://www.mzcloud.org/), the Global Natural Product Social Molecular Networking [GNPS],[226] MassBank of North America [MoNA] (http://mona.fiehnlab.ucdavis.edu/)), and others.[227,228] However, the measured MS/MS spectra are not always comparable against the MS/MS spectra libraries due to different MS/MS istruments and collision conditions used in the measurements. Furthermore, the metabolite coverage in spectral libraries is limited. Therefore, other tools such as MS/MS *in*-silico fragmentation and prediction tools commercially available (Mass Frontier, MyCompound ID) or open-source (MS-FINDER,[229] SIRIUS,[230] MAGMa,[231] CFM-ID,[232] MetFrag,[233] CSI:FingerID,[234] and LipidBlast[235]) have been developed. These tools utilize existing structure and chemical databases and generates mass spectra computationally based on quantum chemistry, machine learning, chemical reactions, or heuristic-based approaches.[236] The performance of different *in*-silico tools have been recently studied in the Critical Assessment of Small Molecule Identification (CASMI) challenge.[237,238]

Identification is usually more straightforward in GC-MS as identification is based on highly repeatable available EI spectral databases and relative retention indexes (RI).[239] The most applied EI libraries are NIST 14, Wiley, Golm Metabolome Database (GMD),[240] and Fiehn library,[241] where the measured EI-MS spectrum and RI are compared to reference data using a similarity score algorithm (e.g. dot product).[242] The GMD provides additional information on the predicted functional groups in a compound, based on supervised machine learning from previously analyzed spectra and RI values.[243] GC-HRMS is currently used with increasing frequency and accurate mass fragments can be used in the identification.[244]

1.3.9.3 Data and pathway analysis

Data analysis in metabolomics commonly includes different statistical and multivariate data analysis methods, which are performed on the data matrix after pre-processing.[245–247] Some post-processing steps, such as scaling (i.e. variance scaling, transferring to logarithmic scale) and imputing missing values, are usually performed prior to data analysis.[245] An unsupervised principal component analysis (PCA) is generally used as an initial data mining and classification method to observe possible differences between sample groups.[247] Supervised classification techniques, such as partial least squares-discriminant analysis (PLS-DA), orthogonal partial least squares-discriminant analysis (OPLS-DA), support vector machines, random forests, and other neural network methods, are also often applied.[247] Clustering is typically applied to study relationships and the similarity of samples or metabolites, or both. The most popular clustering methods in metabolomics are hierarchical and K-means clustering.[246,247] A receiver operating characteristic (ROC) curve is commonly utilized for biomarker modeling and to estimate biomarker prediction power. Different regression models are used to study variable dependencies.[246,247]

Univariate statistics are commonly applied in metabolomics with parametric tests (unpaired or paired t-test and analysis of variance [ANOVA]) or non-parametric tests (Mann-Whitney U-test, Wilcoxon signed-rank test, and Kruskal-Wallis test) depending on the application.[245,246] In metabolomics, the parallel univariate tests exceeds the number of observations, which highly increases the number of false positives. This multiple testing problem is commonly corrected with false discovery rate (FDR) correction or with Bonferroni correction.[245]

In addition to metabolite levels, metabolic reactions and pathways are also important especially in mechanistic studies.[248] Metabolite pathway analysis is commonly enrichment analysis, which computes significantly altered metabolic pathways, or metabolite mapping, which contextualizes the abundances and significances of measured metabolites into network visualizations.[248] Commonly utilized tools and databases for metabolite pathway analysis include KEGG pathways,[222] MetaboAnalyst,[208] small molecule pathway database (SMPDB),[249] Reactome,[250] and MetaCyc[251]. Pathway analysis is limited only to the identified compounds, which prevents a deeper understanding of metabolism in multiple studies. New analysis modes (i.e SWATH, DDA, DIA) and the improvement of identification tools will likely shift identification towards all detected metabolites, not only to possible biomarkers.

2 AIMS OF THE STUDY

The overall aim of this study was to apply and develop mass spectrometry-based analytical methods for metabolomics. The goal was to develop non-targeted methods for screening metabolic changes between different biological conditions (publications I-III). The required analysis times in these non-targeted screening methods with chromatographic separation were quite long and thus prevented high-throughput analysis. To reduce the analysis time, non-targeted chip-MS analysis was studied to detect differences in cell metabolites (publication III). The sensitivity of non-targeted metabolomics is limited and depends on the metabolite chemical properties. Due to low ionization efficacy in ESI and high fragmentation in EI, non-polar low abundant metabolites (such as steroids) are usually not detected with common non-targeted screening methods. The CPI source developed in-house provides high sensitivity towards non-polar metabolites and its feasibility and utilization in metabolomics studies were explored (publications IV and V).

The more detailed aims of the research papers presented in this thesis were as follows:

- To study the metabolic differences in neonatal mouse hearts that reflect changes in the regeneration capacity of cardiomyocytes (I)

- To study the human serum metabolite profile associated with insulin resistance and gut microbiota of non-diabetic individuals (II)

- To evaluate the suitability of the microchip-based direct infusion mass spectrometry for non-targeted analysis to study stem cell metabolites and to compare the performance to the most often applied UHPLC-ESI-QTOF-MS method (III)

- To develop a CPI ion source to increase sensitivity and to develop a GC-CPI-MS/MS method to evaluate CPI performance for the analysis of endogenous steroids from a biological matrix (IV)

- To test the feasibility of the CPI method coupled with LC-MS and low flow rates to analyze selected steroids (V)

3 EXPERIMENTAL

This section briefly describes the chemicals, samples, instruments, and analytical protocols used in this study. More detailed descriptions can be found from the original publications I-V.

3.1 CHEMICALS AND MATERIALS

All chemicals used in this study (I-V) were analytical or chromatographic grade. More detailed information and the material and chemical suppliers are available in the original publications (I-V).

3.2 SAMPLES AND SAMPLE PRETREATMENT

3.2.1 MOUSE HEART SAMPLES

Mouse pups (strain C57BL, n=92, publication I) were acquired from the experimental animal facility of the University of Helsinki with an internal use licence (KEK14-014). Heart tissue samples were collected on postnatal days 1, 4, 9, and 23 (P01, P04, P09, and P23, respectively) at midday without fasting and before weaning. Mice were decapitated, ventricles were excised and cut into three to five pieces, rinsed in PBS, snap-frozen in liquid nitrogen, and stored at -80°C prior to sample pre-treatment.

The tissue samples were homogenized in water using a probe sonicator (sample set 1) or FastPrep-24 5G bead homogenizer with 1-mm zirconia beads (sample set 2). For GCxGC-MS analysis, 25 µL of homogenate and 410 µL MeOH containing labeled internal standards (ISTD) were transferred into an Eppendorf tube. The samples were vortexed, sonicated for 5 min, and incubated on ice for 30 min, after which the extracts were centrifuged for 5 min (14 338 *g*, 4°C) and 180-µL aliquots of supernatants were transferred into vials and evaporated to dryness under a nitrogen flow. The metabolites were then derivatized by automated two-step in-time derivatization. First, 25 µL of MOX was added to the residue and the mixture was incubated for 1 h at 45 °C. Next, 25 µL of MSTFA including the RI standards was added and the mixture was incubated for 1 h at 45 °C. Finally, hexane including injection standard (50 µL) was added to the mixture immediately prior to injection. QC samples were prepared by pooling P4, P9, and P23 tissue sample homogenates and were otherwise prepared similarly as the samples. All the samples were randomized before sample pretreatment and analysis.

For LC-MS analysis, 25 µL of the sample homogenate, 5 µL of ISTD mixture, and 100 µL of MeOH were transferred into an Eppendorf tube followed by vortexing and incubation on ice for 30 min. The samples were then centrifuged for 10 min (7 378 *g*) and the supernatants were transferred into vials and analyzed in randomized order with additional QC, standard, and blank samples.

3.2.2 HUMAN SERUM SAMPLES

Human serum samples (n=397, publication II) were collected as part of the Danish MetaHIT consortium (http://www.metahit.eu) from Danish participants, including normoglycaemic and middle-aged individuals (n=291), type 2 diabetes patients (n=75), and type 1 diabetes patients (n=31). The study was approved by the Ethical Committees of the Capital Region of Denmark (HC-2008-017 and H-15000306) and was in accordance with the principles of the Declaration of Helsinki. All individuals provided written informed consent before participating in the study. Serum samples were collected in the morning after an overnight fast of at least 10 h and without prior morning physical activity using the same personnel and protocol for all the samples. Serum lipids were extracted with modified Folch extraction by mixing 10 µL serum, 20 µL of ISTD mixture, and 100 µL of chloroform:MeOH (2:1, v v^{-1}). The mixture was vortexed for 2 min, incubated at room temperature for 30 min, centrifuged for 3 min (7800 *g*), followed by the collection of the lower phase (60 µL) and addition of second ISTD mixture (20 µL). Extracted samples were stored at -20°C prior to analysis. Sample pretreatment and analysis order were randomized and control serum samples, extracted standard samples, pure standard samples, and blanks were analyzed together with the human serum samples.

3.2.3 CELL SAMPLES

Human foreskin fibroblasts (HFF) and human induced pluripotent stem cells (hiPSC) (publication III) were grown at the Institute of Biotechnology in the University of Helsinki. hiPSC cell line HEL24.3 was differentiated from HFF using non-integrative Sendai virus-mediated transduction with Oct4, Sox2, Klf4, and c-Myc as described earlier.[252] Six replicates of HFF (1 x 10^6 cells), HiPSC (5 x 10^6 cells), and blanks were analyzed. Cells were washed with PBS, snap-frozen in liquid nitrogen, and stored in -80°C prior the analysis.

Cells were lysed by adding 50 µL of MilliQ water (MQ) to the frozen cell pellet and sonicated in an ice-cold ultrasound bath for 15 min two times with mixing between the cycles. Next, 20 µL of the cell lysate, 10 µL of ISTD mixture, and 100 µL of MeOH was mixed. The cell extract was then vortexed and incubated on ice for 20 min. After incubation, the extract was centrifuged

for 10 min (7378 g) and the supernatant was transferred into a vial and stored at -20 °C until analysis. For LC-MS, the samples were injected as such. For chip-MS, 10% formic acid (FA) in MQ water was added to the samples to achieve a final solvent composition of 70:29:1 (MeOH:H_2O:FA).

3.2.4 HUMAN URINE SAMPLES

Urine samples (publication IV) were collected from four male and four female volunteers. Each authentic sample was divided into two aliquots; one aliquot was used to determine free steroids and the other aliquot was used to analyze the sum concentrations of steroid glucuronide conjugates and free forms. The glucuronide-conjugated steroids were hydrolyzed before extraction by adding 0.8 M sodium dihydrogen phosphate solution (1 mL, pH 7) and glucuronidase (50 µL) followed by incubation at 50°C for 1.5 hours. After enzymatic hydrolysis, steroids were extracted and derivatized similarly as the free steroid fraction of urine and artificial urine samples.

17α-methyltestosterone was used as an ISTD with a final concentration of 10 ng mL^{-1}. The sample (2.5 mL) pH was adjusted to approximately pH 8 by dissolving $NaHCO_3$/K2CO_3 (125 mg, 2:1, w w^{-1}) mixture in each sample. After pH adjustment, steroids were extracted with LLE to diethyl ether. Diethyl ether (4 mL) and Na_2SO_4 (1.5 g) were added to samples followed by vortexing and centrifugation (2000 *g*) for 10 min. Part of the organic phase (2.5 mL) was transferred into a new test tube and evaporated to dryness under nitrogen flow. The derivatization reagent mixture MSTFA/NH_4I/DTE (50 µL, 1000:2:4, v w^{-1} w^{-1}) was added to the tube, vortexed, and the sample was transferred to a GC vial. The sample was incubated at 60°C for 15 min to derivatize all hydroxyl and keto groups of steroids with TMS. After offline-derivatization, samples were ready for injection into GC.

3.3 INSTRUMENTATION, ANALYTICAL METHODS, AND DATA PROCESSING

3.3.1 GLOBAL METABOLOMICS OF NEONATAL MOUSE HEART SAMPLES

Global non-targeted metabolomics of neonatal mouse hearts were performed with UHPLC-Orbitrap-MS (Table 1) and GCxGC-MS (Table 2) (publication I).[66] GCxGC-MS raw data were processed with ChromaTOF. Guineu was used for compound alignment based on RI, retention times of both dimensions, and EI spectral matches.[218] Data was further filtered and peaks were identified based on RIs and EI spectral match from in-house library,

NIST 2014 (http://chemdata.nist.gov/), GMB,[240] or Fiehn library.[241] LC-MS raw data was processed with MzMine 2 applying preprocessing steps, mass detection, chromatogram builder, chromatogram deconvolution, deisotoping, alignment, filtering, cap filling, and normalization with ISTD, and tissue weight. [206] Selected features were further identified with MS/MS analysis using mzCloud (https://www.mzcloud.org/), NIST 2014 (http://chemdata.nist.gov/), HMDB,[221] MoNA (http://mona.fiehnlab.ucdavis.edu), LipidBlast,[235] CSI:Finger ID,[234] and known group-specific fragmentation of lipid species for identification based on the MS/MS spectra. Metabolomics datasets were analyzed with R 3.1.2 (https://www.R-project.org) using PCA, one-way ANOVA (LC-MS), or t-tests (GCxGC-MS). Separate datasets were combined and analyzed with linear-mixed effect (LME) model,[253] where the first variable was the postnatal age and the second was the dataset. FDR correction (Benjamini and Hochberg) was used in all statistical tests. Metabolite set enrichment analysis (MSEA) and fuzzy c-means clustering were performed on all significantly changed ($q<0.01$) identified metabolites in the LME model.[254,255]

Table 1 *Experimental parameters for GCxGC-MS (publication I).*

Parameter	**Details**
Instrument	LECO Pegasus 4D and Agilent 7890 gas chromatograph
Retention gap	Deactivated fused silica (1.7 m × 0.53 mm)
Column 1	RTXi-5MS column (10 m × 0.18 mm, 0.2 µm)
Column 2	BPX-50 column (1.5 m × 0.1 mm, 0.1 µm)
Temperature program (1st oven)	50°C (2 min) → 240°C (7 °C/min) → 300°C (25 °C/min) → 300°C (3 min)
Temperature program (2nd oven)	1st oven temperature + 20°C
Modulation time	4 s
Injection volume	1 µL
Injection mode	Pulsed splitless (1.5 min)
Injector temperature	240°C
Transfer line temperature	260°C
Ion source temperature	200°C
Carrier gas	Helium, 40 psig
Ionization	EI, 70 eV
Mass range	*m/z* 45-700
Scan speed	100 spectra/s

Table 2 *Experimental parameters for LC-MS (publication I).*

Parameter	Details
Instrument	Thermo Orbitrap Fusion, Thermo Dionex Ultimate 3000 UHPLC
Column	Acquity BEH C18 (2.1 x 100 mm, 1.7 µm)
Eluents (set 1)	A: 0.1% FA in MQ water, B: 0.1% FA in MeOH
Eluents (set 2)	A: 0.1% FA+ 5% MeOH in MQ water, B: 0.1% FA in MeOH
Gradient of B (set 1)	5% → 100% (15 min) → 100% (10 min) → 5% (0.1 min) → 5% (4.9 min)
Gradient of B (set 2)	0% → 100% (15 min) → 100% (10 min) → 5% (0.1 min) → 5% (4.9 min)
Flow rate	0.3 mL min^{-1}
Injection volume	2 µL or 3 µL
Column temperature	25°C
Sample temperature	10°C
Capillary voltage	3 kV
Resolution	120 000 FWHM
Sheet gas	30 arb
Aux gas	10 arb
Ion transfer tube temperature	333°C
Vaporizer temperature	317°C
Ionization	ESI, positive ion mode
Mass range	*m/z* 100-1000
Scan speed	0.6 s
MS/MS parameters	HCD fragmentation, Normalized collision energies 25% and 40%, resolution 30 000 FWHM

3.3.2 LIPIDOMICS ANALYSIS OF HUMAN SERUM SAMPLES

Lipidomics of human serum samples were performed with UHPLC-QTOF-MS (Table 3) (publication II).[256,257] The data were processed using MZmine 2 preprocessing steps, mass detection, chromatogram builder, chromatogram deconvolution, deisotoping, alignment, filtering, and cap filling.[206] Serum lipids were identified using an internal spectral library or with additional MS/MS or MS^n analysis. Data were normalized using the ISTD representatives of each lipid class. Serum polar metabolites were analyzed with GCxGC-MS.[66] Co-abundant serum metabolites (lipids and polar metabolites separately) were clustered with weighted correlation network analysis across all examined individuals after log2 transformation using the R package WGCNA.[258] All statistical tests were non-parametric, applied FDR correction, and were

performed with R. Serum metabolite clusters and species associations with IR were conducted with a Spearman rank correlation test and all associations for microbial functional modules were identified using Mann-Whitney U-test. Ranks were based on Spearman correlation coefficients for associations of microbial functional modules with homeostatic model assessment for insulin resistance (HOMA-IR), gene richness, and serum metabolite clusters. More detailed description of data analysis can be found from original publication (II).

Table 3 *Experimental parameters for LC-MS (publication II).*

Parameter	Details
Instrument	Waters QTOF Premier, Waters Acquity UPLC
Column	Acquity BEH C18 (2.1 x 100 mm, 1.7 µm)
Eluents	A: 1% 1 M NH_4Ac + 0.1% FA in MQ water, B: 1% 1 M NH_4Ac + 0.1% FA in ACN/IPA (1:1)
Gradient of B	35% → 80% (2 min) → 100% (5 min) → 100% (7 min) → 35% (0.1 min) → 35% (3.9 min)
Flow rate	0.4 mL min^{-1}
Injection volume	1 µL
Column temperature	50°C
Sample temperature	10°C
Capillary voltage	3 kV
Resolution	8 000 FWHM
Source temperature	150°C
Desolvation temperature	450°C
Desolvation gas flow	800 L h^{-1}
Cone gas flow	20 L h^{-1}
Mass range	*m/z* 300-1200
Scan speed	0.2 s
MS/MS and MS^n	Waters Acquity UPLC-TriVersa Nanomate-Thermo LTQ Orbitrap

3.3.3 LC-MS AND CHIP-MS METHODS IN GLOBAL METABOLOMICS ANALYSIS OF CELL SAMPLES

Global non-targeted metabolomics of HFF and hiPSC samples were performed with SU-8-based ESI microchip Orbitrap MS and UHPLC-ESI-QTOF-MS (Table 4) (publication III). The direct infusion ESI SU-8 chips were coupled to a Q-Exactive Hybrid Quadrupole Orbitrap Mass Spectrometer and controlled via Xcalibur software.[259,260] A custom-made high-voltage supply

was used for application of the ESI voltage (2.5 kV) via a platinum electrode inserted into the sample inlet. The spectrum was recorded for 0.3 min and each sample was analyzed in triplicate. The microchip channel was carefully rinsed with MeOH:H_2O (1:1) between runs.

Both datasets (chip-MS and LC-MS) were processed with MZmine 2,[206] applying mass detection, chromatogram builder, chromatogram deconvolution (LC-MS only), deisotoping, alignment, filtering, and cap filling. Both datasets were normalized with ISTD verapamil. For chip-MS, the average abundance of three replicate measurements was calculated and further applied in the data analysis. Identification of detected compounds was based on search with accurate mass (3 ppm) from the HMDB and Lipid Maps databases.[221,223] Data analyses of both datasets were performed with R 3.0.2. Zero-value imputation, log2 transformation, and centering were executed prior the PCA. Unpaired t-test with FDR correction was applied and fold changes were calculated for both datasets.

Table 4 *Experimental parameters for LC-MS and Chip-MS (publication III).*

Parameter	**LC-MS**	**Chip-MS**
Instrument	Waters Xevo QTOF, Waters Acquity UPLC	Thermo Q Exactive Hybrid Quadrupole Orbitrap
Column	Acquity BEH C18 (2.1 x 100 mm, 1.7 µm)	-
Eluents	A: 0.1% FA in MQ water, B: 0.1% FA in MeOH	-
Gradient of B	5% → 100% (15 min) → 100% (10 min) → 5% (0.1 min) → 5% (4.9 min)	-
Column temperature	25°C	-
Flow rate	0.3 mL min^{-1}	-
Sample temperature	10°C	-
Injection volume	5 µL	5 µL
Capillary voltage	2 kV	2.5 kV
Resolution	8 000 FWHM	70 000 FWHM
Source/Ion transfer tube temperature	150°C	200°C
Desolvation temperature	450°C	-
Desolvation gas flow	800 L h^{-1}	-
Cone gas flow	20 L h^{-1}	-
Mass range	*m/z* 100-1000	*m/z* 100-1500
Scan speed	0.15 s	0.33 s

3.3.4 CAPILLARY PHOTOIONIZATION

The CPI ion source (publication IV and V) contained a stainless steel (SS) CPI capillary (i. d. 1.5 mm) connected to a flat SS part with an oval-shaped opening (length 17 mm, width 8 mm at the widest point, height 3 mm) that was covered with a 3-mm thick magnesium fluoride (MgF_2) window. A SS top plate with an 18-mm circular opening and graphite rings used as seals held the MgF_2 window in place. The inlet where the analytes entered the ion source was heated by a resistance wire heater driven by a direct current power supply. The main body of the CPI device was heated (225-350°C) with a 100 or 160 W cylindrical heater (diameter 6.5 mm, length 50 mm) that was controlled by a PID500 temperature controller. The heater was embedded in a cylindrical aluminium block (diameter 35 mm, height 55 mm) that was attached to the bottom of the CPI device. The photoionization was performed with 10.0 and 10.6 eV photons from a radio frequency krypton discharge vacuum UV lamp.

3.3.4.1 GC-CPI-MS/MS for steroid analysis

A gas chromatograph was connected through a capillary photoionization interface to the inlet of a mass spectrometer (Figure 5) to analyze steroids (Table 5) (publication IV). Inside the GC oven, the GC transfer capillary was introduced through a SS T-piece into a SS transfer tube, which was heated and served as a transfer line between the GC and MS. A few millimeters of this transfer tube and of the transfer capillary were introduced inside the CPI inlet capillary. A mixture of auxiliary gas nitrogen and dopant chlorobenzene entered the t-piece through 1/16 in o.d. SS tubing. The dopant was pumped into the t-piece by a syringe pump and vaporized and mixed with auxiliary gas before the t-piece. Inside the SS t-piece, the mixture was directed into the SS transfer tubing and was transferred coaxially with the transfer capillary into the CPI capillary. A nitrogen atmosphere was created around the CPI inlet capillary and the GC transfer tube interface by enclosing them in aluminium housing, which was sealed and insulated with aluminium foil wrapped around the housing. The housing was not completely gas tight to allow the vacuum system of the MS to produce equal gas flow as without the nitrogen housing. However, the excess nitrogen prevented most of the ambient air molecules from entering the ion source.

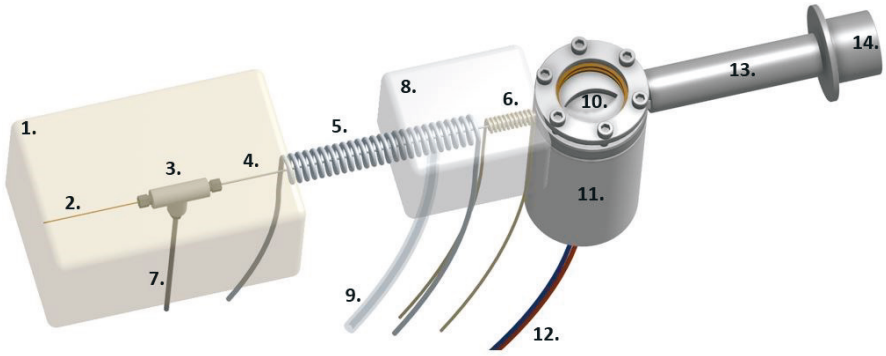

Figure 5 Setup for GC-CPI-MS. Parts marked in the figure are 1.GC oven, 2. GC transfer capillary, 3. SS t-piece, 4. SS transfer tube (i. d. 0.5 mm), 5. Heated transfer line, 6. Heated CPI inlet capillary, 7. Dopant and nitrogen gas line (1/16 in o.d. SS tubing), 8. Nitrogen atmosphere in an aluminium housing, 9. Nitrogen flow through Teflon tubing, 10. Oval-shaped opening under MgF_2 window, 11. Heated aluminium block, 12. Temperature control of cylindrical heater 13. CPI capillary, 14. Inlet into mass spectrometer. Parts 5. and 8. were insulated with aluminium foil (publication IV).

Table 5 *Experimental parameters for GC-MS/MS (publication IV).*

Parameter	**Details**
Instrument	Agilent 6410 triple quadrupole, HP 5892 II gas chromatograph
Retention gap column	Deactivated fused silica (2 m, i.d. 0.25 mm)
Column 1	TR-5MS (15 m, i.d. 0.25 mm, film thickness 0.25 µm)
Column 2	TR-50MS (15 m, i.d. 0.25 mm, film thickness 0.25 µm)
Transfer line	1 m, i.d. 0.15 mm deactivated fused silica
Temperature program	190°C (1 min) → 250°C (10°C min^{-1}) → 260°C (1°C min^{-1}) → 330°C (7°C min^{-1})→ 330°C (3 min)
Injection volume	3 µL min^{-1}
Injection mode	Splitless (1 min)
Injector temperature	250°C
Transfer line temperature	250°C
Carrier gas	Helium, 140 kPa
CPI source temperature	250°C
Auxiliary nitrogen gas	80 mL min^{-1}
Dopant flow rate	3 µL min^{-1}
Nitrogen housing flow rate	4.5 L min^{-1}

3.3.4.2 GC-CPI-MS/MS method validation and analysis of steroids from urine

The GC-CPI-MS/MS method was developed, validated, and applied for human urine samples to analyze 18 endogenous steroids (publication IV). Steroids were quantified with an ISTD method and identification was based on retention time and ion ratios of the selected product ions. The developed method was validated following the validation instructions of The International Council for Harmonisation (ICH) and identification criteria of World Anti-Doping Agency (WADA).[261,262] Limit of detection (LOD), limit of quantitation (LOQ), linear range, linearity, repeatability, LLE recovery, and carryover were studied. In total, 21 different concentration levels in the range 0.001-5000 ng mL^{-1} spiked to artificial urine were tested for determination of linearity, linear range, LOD, and LOQ. LOD and LOQ were determined based on the signal-to-noise (S/N) ratio using criteria S/N >3 and S/N >10 with spiked standard compounds in artificial urine, respectively. A suitable linear range and LOQ fulfilled the following criteria: maximum residual from the curve below 20% or below 30% (LOQ) and quantifier/qualifier ion ratio in the range instructed by WADA.[262] For ADT, ETIOL, and CS, an additional curve from diluted samples was determined and applied to analyze high urine concentrations. 1/x weighting was applied for the calibration curves. Repeatability and LLE recovery were determined at a concentration level of 10 ng mL^{-1} (n=6).

3.3.4.3 LC-CPI-MS

The LCs used in the study (publication V) were Waters Acquity UPLC and Agilent 1100 capillary LC system The column was Waters BEH C18 (2.1 mm x 100 mm, 1.8 µm) in UHPLC-MS experiments. The eluents were (A) MQ water and (B) MeOH with the following gradient: B 10-100%, 0-5 min; B 100%, 5-10 min; B 100-10%, 10-10.1 min; and B 10%, 10.1-15 min. The flow rate was 200 µL min^{-1} and the flow was split after the column with a split ratio of 1:20. The injection volume was 10 µL. The column was Waters Symmetry RP18 (0.3 mm x 100 mm, 3.5 µm) in capillary LC-MS experiments. The eluents were (A) MQ water and (B) MeOH with the following gradient: B 10-100%, 0-10 min; B 100%, 10-15 min; B 100-10%, 15-20 min; and B 10%, 20-60 min. The flow rate was 10 µL min^{-1} and the injection volume was 1 µL.

The eluent from the LC column was directed to the CPI device through fused-silica transfer capillary (0.15 mm i.d.) passing through an SS t-piece and SS tubing (Figure 6). The end of the transfer capillary was positioned inside the CPI capillary in front of the MgF_2 window. Toluene as a dopant was pumped at a flow rate of 5 µL min^{-1} and mixed through a t-piece with nitrogen used as an auxiliary gas (400 mL min^{-1}). The nitrogen-dopant gas mixture was

heated with a resistant wire heater (approximately 110-120°C) that led into the SS t-piece and passed coaxially between the fused silica transfer capillary and the SS tubing inside the CPI device. An Agilent 6410 triple-quadrupole mass spectrometer was used in all experiments applying Mass Hunter for data acquisition and for data processing. DHEA, PROG, E2, and T were used as test compounds. The following optimized (fragmentor voltage and collision energy) mass transitions were used: DHEA, *m/z* 271 ($[M+H-H_2O]^+$) → 213; E2, *m/z* 255 ($[M+H-H_2O]^+$) → 159; PROG, *m/z* 315 ($[M+H]^+$) → 97; and T, *m/z* 289 ($[M+H]^+$) → 97.

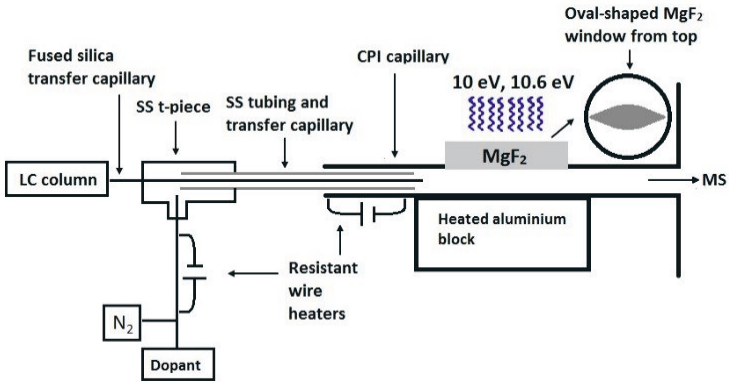

Figure 6 Setup for LC-CPI-MS. Published by The Royal Society of Chemistry (publication V).

4 RESULTS AND DISCUSSION

This chapter summarizes the main results of the work. More detailed descriptions can be found in the original publications (I-V). The overall aim was to apply mass spectrometric methods and to develop new faster and more sensitive methods for metabolomics. Since none of the existing analytical methodologies can detect all metabolites simultaneously, multiple methods with different advantages and disadvantages are required for wide metabolite coverage. Neonatal mouse heart samples (publication I) and human serum samples (publication II) were studied with non-targeted methods applying LC-MS and GCxGC-MS. The aim in both studies was to detect metabolic changes as widely as possible without a previous hypothesis of possible changes. However, sample throughput is limited when chromatographic separation is used and long batches may be prone to retention time shifts or drift in signal sensitivity. To overcome this, the feasibility of a direct infusion ESI microchip MS for global non-targeted metabolomics to detect differences between hiPSC and HFF cells was studied (publication III). The developed method was compared to more traditional LC-MS and the performance of both methods in metabolomics was evaluated.

When quantitative analysis of high sensitivity and specificity is required, non-targeted methods are usually not adequate and targeted methods are required. The ESI ionization technique is applied in nearly all metabolic applications with LC, although ionization efficacy of ESI towards highly non-polar compounds is weak and insufficient to detect non-polar metabolites at low concentrations. On the other hand, in GC applications the ion source is usually EI, where sensitivity is not maximal because of intense fragmentation. To increase sensitivity, a CPI ion source was developed and applied in GC-MS and LC-MS, which enabled high ion transmission efficacy with low fragmentation. A GC-CPI-MS/MS method was developed and validated to detect steroids in human urine to study the feasibility of the interface in metabolic applications (publication IV). The suitability of the CPI interface for LC-MS applications was explored for the first time, using low flow rates (10 μL min^{-1}) that are not optimal with the existing commercial APPI ion source geometries that require higher flow rates (>200 μL min^{-1}) (publication V).

4.1 GLOBAL METABOLOMICS OF NEONATAL MOUSE HEARTS

The aim of this study was to develop a method for global metabolomics and detection of metabolic changes to identify the molecular mechanisms mediating postnatal loss of cardiac regeneration in mammals (publication I). Previous studies have shown that neonatal rodents can fully regenerate their

hearts after an injury. However, this regenerative capacity is lost within 7 days after birth because of cardiomyocyte cell cycle withdrawal.[263] The reported full functional recovery of a newborn baby with a massive myocardial infarction suggests that humans possess a similar intrinsic capacity for heart regeneration after birth.[264] Understanding the biology behind the loss of regeneration capacity plays a key role in the development of regeneration-inducing therapies, which are essential for diseases with massive cell losses (i.e. stroke and myocardial infarction). An integrated multiomics study of mRNA, protein, and metabolite changes was performed from neonatal mouse hearts to identify metabolism-related mechanisms associated with cardiac regeneration. For this study, two individual sample sets of mouse hearts at postnatal days 1, 4, 9, and 23 (P01, P04, P09, P23, respectively) were analyzed (Figure 7).

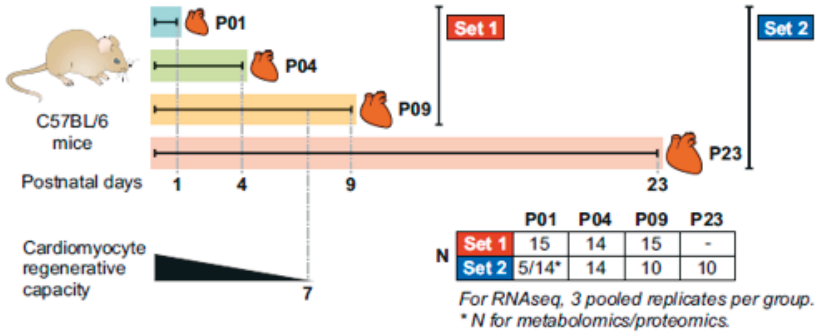

Figure 7 The experimental design of study (publication I).

4.1.1 GLOBAL METABOLOMICS USING LC-MS

Complementary techniques (LC-MS and GCxGC-MS) were used for non-targeted metabolomics to increase the metabolite coverage. Two individual sample sets (set 1 and set 2) were used to increase reliability and decrease the possibility of false findings and artifacts. First, simple PPT with MeOH combined with ESI-RPLC-MS analysis in positive ion mode was selected for global metabolic profiling. In total, 5591 (set 1) and 9043 (set 2) features were detected with LC-MS with 3398 matching features from both sample sets. Metabolic differences were studied with PCA, t-tests, ANOVA, and LMEA model from combined dataset (set 1 and 2). The PCA of all features showed clear separation of the sample groups based on neonatal age (Figure 8A). Statistically significantly changed ($q < 0.01$) metabolic features between the time points (P1 vs. P4, P1 vs. P9, and P1 vs. P23) in the LME model (n=805) were forwarded to identification with additional MS/MS analysis. Metabolites were mainly detected as protonated molecules, although sodium adducts were

also detectable for most of the species. The LC-MS method was able to detect significantly changed polar compounds, including amino acids and purines, lipids such as free fatty acids, lysophospholipids (LPL), monoacylglycerols (MG), acylcarnitines, phosphatidylcholines (PC), phosphatidylethanolamines (PE), ceramides (Cer), and sphingomyelins (SM). Unfortunately, most of the detected features remained unknown. Highly non-polar lipids were not detected (i.e. triacylglycerols [TG] or cholesterylesters [ChoE]) due to selection of PPT solvent and LC eluents and the detected polar compounds were mainly eluted in the void volume. Some phospholipid (PL) classes, such as phosphatidic acids (PA), phosphatidylglycerols (PG), phosphatidylinositols (PI), and phosphatidylserines (PS) and some polar compounds (e.g. organic acids) were not detected and would have required ionization in the negative ion mode. The quality of the method was controlled with pooled QC samples (Figure 8A) and monitoring intensities of ISTDs (Table 6). Repeatability and mass accuracy were adequate and no signal drift was observed when the sample injection order was inspected.

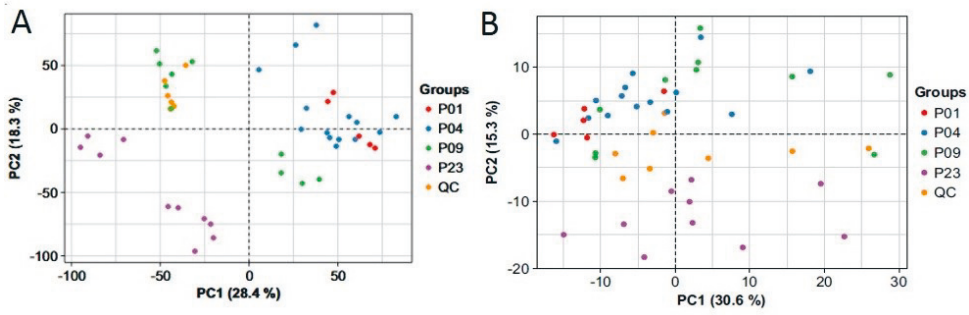

Figure 8 Individual factor maps of PCA of LC-MS (A) and GCxGC-MS (B) data of sample set 2 (publication I). P01, P04, P09, and P23, postnatal days 1, 4, 9, and 23; QC, quality control.

Table 6 *Exact masses, accurate masses, RTs, mass errors, and relative standard deviations (RSD %) of the ISTDs in heart sample sets 1 (n=44) and 2 (n=39), including pooled QC samples in sample sets 1 (n=10) and 2 (n=10) measured with LC-MS using Orbitrap mass spectrometer (publication I).*

Compound	Exact mass	Accurate mass	RT (min)	Mass error (ppm)	RSD% intensity	RSD% normalized intensity
Sample set 1						
LysoPC(17:0)	510.35542	510.35524	15.84	0.35	24.8	0
Propranolol	260.16451	260.16434	9.48	0.67	25	17.7
L-Lysine-d4	151.13790	151.13794	0.78	-0.25	23.4	21.2
Verapamil	455.29043	455.29018	10.15	0.56	16.3	15.4
Sample set 2						
LysoPC(17:0)	510.35542	510.35522	16.01	0.38	11.7	0
Propranolol	260.16451	260.16431	9.24	0.78	12.9	14.5
L-Lysine-d4	151.13790	151.13793	0.78	-0.23	22.8	24.2
Verapamil	455.29043	455.29019	9.9	0.52	12.9	14.5

4.1.2 GLOBAL METABOLOMICS USING GCXGC-MS

Since the polar metabolites were detected with LC-MS mainly in the void volume without a proper separation, an additional GCxGC-MS experiment was designed. GCxGC-MS analysis was performed after PPT with MeOH and automated two-step derivatization. After the first data preprocessing steps, the GCxGC-MS datasets of sample set 1 and 2 contained 760 and 771 metabolic features, respectively. Peak lists were further filtered resulting in a total of 347 and 443 features. Of these, 328 features were found from both datasets and contained in total 162 statistically significantly ($q < 0.01$) changed metabolic features in the LME model. A wide range of polar metabolites, including sugars, organic acids, amino acids, TCA cycle metabolites, and vitamins were detected and identified based on RI and EI spectral match either from an in-house library or spectral databases. However, most of the detected features in this data remained also unknown. The quality of the GCxGC-MS method was monitored with internal and injection standards and pooled QC samples (Figure 8B, Table 7). Repeatability based on ISTDs were acceptable (Table 7). In addition to semi-quantitative non-targeted analysis, some of the metabolites were targeted and quantified against a standard calibration curve (Table 7).

Table 7 *Relative standard deviations (RSD%) of the ISTDs and injection standard (inj. st.) intensities in heart sample sets 1 (n=44) and 2 (n=39) and quantified compounds in pooled QC heart sample sets 1 (n=10) and 2 (n=10) measured with GCxGC-MS (publication I).*

Compounds	**RSD%**	
Internal and injection standards	**Sample set 1**	**Sample set 2**
succinic acid-d4 (ISTD), 2 TMS	11	21
heptadecanoic acid-d33 (ISTD), TMS	9	8
4,4′-dibromooctafluorobiphenyl (inj. st.)	14	14
Quantified compounds		
3-Hydroxybutyric acid, 2TMS	16	14
Alanine, 2TMS	12	18
Arachidonic acid, TMS	50	28
Aspartic acid, 3TMS	62	78
Cholesterol, TMS	12	17
Citric acid, 4TMS	11	26
Fumaric acid, 2TMS	5	21
Glutamic acid, 3TMS	4	5
Glycine, 3TMS	19	37
Isoleucine, 2TMS	23	24
Leucine, 2TMS	31	20
Linoleic acid, TMS	20	13
Malic acid, 3TMS	7	7
Oleic acid, TMS	17	9
Palmitic acid, TMS	19	14
Proline, 2TMS	45	23
Serine, 3TMS	35	30
Stearic acid, TMS	26	19
Succinic acid, 2TMS	14	16
Threonine, 3TMS	24	22
Valine, 2TMS	3	7

4.1.3 CHANGED METABOLITES AND METABOLIC PATHWAYS IN NEONATAL MOUSE HEARTS

Applying the LME model, a total of 151 metabolites were identified with a statistically significant ($q < 0.01$) change related to postnatal age with either LC-MS, GCxGC-MS, or both. Changed features are presented in a heatmap (Figure 9) showing statistical significance, clustering of metabolites, and upregulation or downregulation of metabolites compared to postnatal day 1. The metabolomics data showed a distinct shift from carbohydrates to lipids as the main source of energy and an increase in total energy metabolism over the early postnatal period (Figure 9). The same shift was also detected in

transcriptomics and proteomics. The abundances of glucose and sugar derivatives (glucose, glucose-6-phosphate, fructose-6-phosphate, galactose-6-phosphate) were downregulated at P23 compared with P01 (Figure 9B), whereas the abundances of most fatty acids and components of glycerolipid metabolism (glycerol-3-phosphate, glycerol-2-phosphate, glycerol, glyceric acid) increased after P01 (Figure 9). The abundance of acylcarnitines at P04 and P09 also increased, reflecting increased fatty acid β-oxidation (Figure 9A). Most metabolites of the TCA cycle (citric acid, fumaric acid, pyruvic acid, malic acid, lactic acid), pentose-phosphate pathway (D-sedulоheptulose-7-phosphate), and glycolysis (lactic acid, pyruvic acid) revealed a constant rise with increasing postnatal age (Figure 9B), thus indicating an increase in energy metabolism in general. Various lipid species (PLs, LPLs, MGs, and fatty acids) also showed interesting changes in abundance (Figure 9A). For PLs, there was a general trend toward increasing saturation level along with increasing postnatal age. All PLs that exhibited a constant decrease from P01 to P23 were unsaturated, and many were polyunsaturated. The PL species with saturated medium-to-long-chain fatty acids (e.g. PC(24:0), PC(26:0), PC(28:0)) exhibited an increase at P04 and P09 followed by a decrease at P23. Several LPL species (LysoPC and LysoPE) increased at P04 and remained constant or increased further through P09 to P23, with the exception of LPLs with fatty acids 16:1 and 18:1, which decreased after P01, or 14:0, which decreased after P04. Interestingly, myristic acid (C14:0) revealed a comparable pattern both as a free fatty acid and when incorporated into other lipid species (e.g. Cer, SM, or PC). The abundance of C14:0-containing species increased at P04 and P09 and decreased at P23. A similar pattern was also observed for other medium-chain saturated fatty acid species (e.g. C12:0 and C10:0) (Figure 9A). The levels of most amino acids displayed an initial increase at P04 or P09 (or both) followed by a decrease at P23 (Figure 9B). In contrast to other (proteinogenic) amino acids, the abundance of glutamic acid, alanine, and histidine increased initially but remained significantly higher at P23 compared to P01 (Figure 9B). The increase in amino acid abundance from P01 to P09 likely reflects the active protein synthesis vital for cardiomyocyte growth and maturation. This was also observed in the MSEA, where the most changed metabolic pathway was protein biosynthesis followed by the urea cycle, glycerolipid metabolism, alanine metabolism, and enrichment of multiple individual amino acid metabolic pathways. Significant changes were also observed in purine metabolism, particularly among the metabolites of AMP catabolism. The level of AMP had decreased already at P04, paralleled by an increase in the abundance of its degradation pathway metabolites, such as inosine, hypoxanthine, and xanthine (Figure 9B). Consistent with the postnatal increase in cardiac workload, we also observed an increased abundance of creatine and creatinine at P09 and P23 compared with P01. Furthermore, an increase in ascorbic acid metabolites L-threonic acid and threonic acid 1,4-lactone on P09 and P23 was observed, reflecting increased oxidative stress with increasing postnatal age.

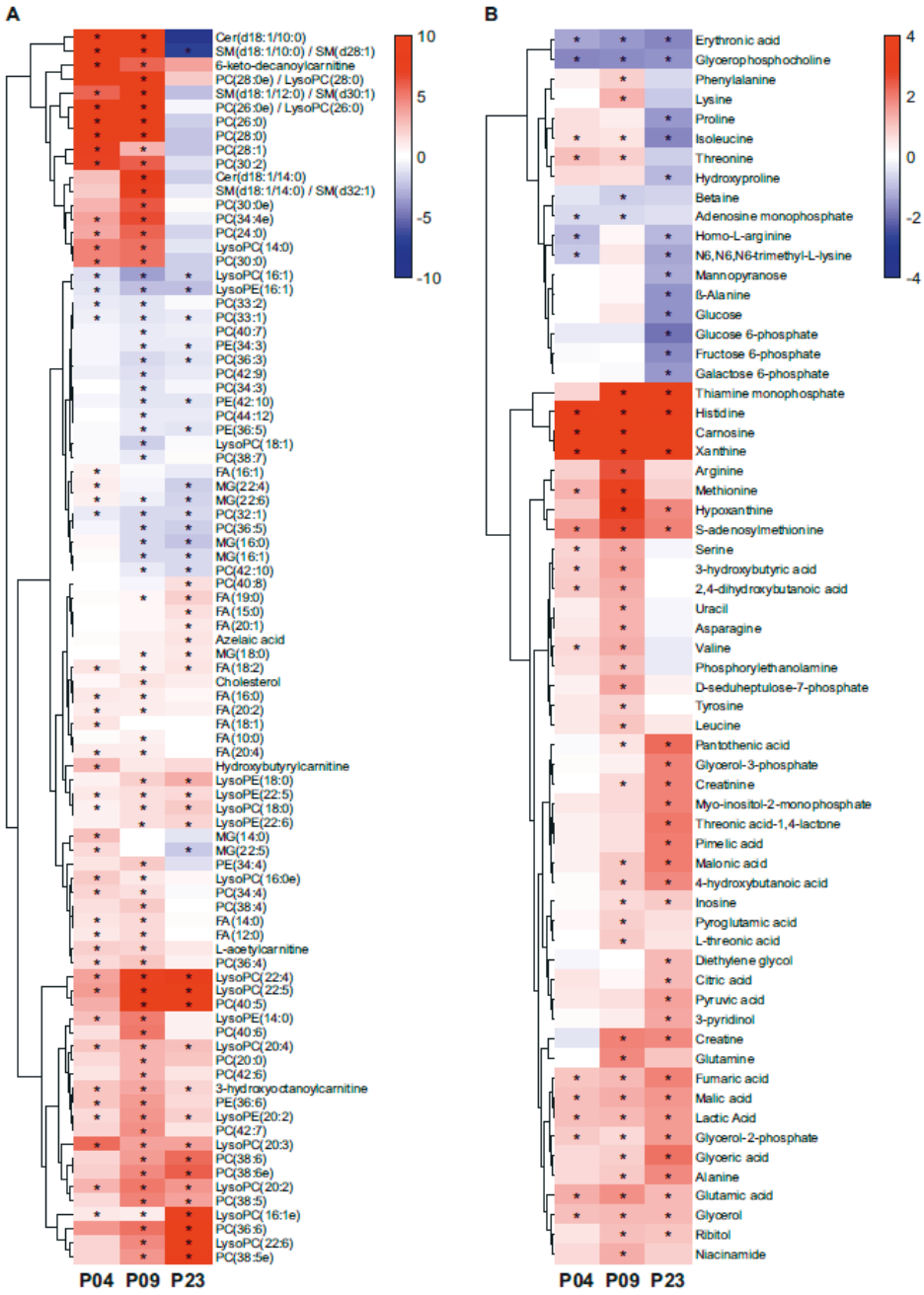

Figure 9 Metabolite changes in the postnatal mouse heart. A, Heatmap of linear mixed effect (LME) model estimates for lipids (A) and polar metabolites (B) compared to postnatal day 1 (P01). *q<0.01 compared with P01. Cer indicates ceramide; FA, fatty acid; LysoPC, lysophosphatidylcholine; LysoPE, lysophosphatidylethanolamine; MG, monoacylglycerol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; SM, sphingomyelin (publication I).

4.1.4 MULTIOMICS OF NEONATAL MOUSEHEARTS

Transcriptomics, proteomics, and metabolomics were also studied in parallel to gain insight into the mechanism of postnatal cardiac maturation. For this, information from fuzzy c-means clustering, enriched biological process gene ontology (GO) terms, RNAseq and proteomics KEGG pathways, ingenuity pathway analysis (to identify potential upstream regulators of gene and protein expression changes), and metabolic pathways enriched in MSEA were utilized. The most significant changes observed at all omics levels were branched-chain amino acid (BCAA) catabolism, fatty acid metabolism, mevalonate and ketogenesis pathways (Figure 10), thus revealing possible associations with regeneration capacity or regulation of the cardiomyocyte cell cycle.

Since hydroxymethylglutaryl-CoA synthase 2 (HMGCS2) was the most significantly upregulated protein from P01 to P04 and cholesterol biosynthesis was one of the enriched metabolism-related biological processes in GO enrichment analysis, the mevalonate pathway and ketogenesis were studied further. Changes were also detected at the metabolite levels of these pathways, including levels of 3-hydroxybutyric acid, the end product of ketogenesis and cholesterol levels, the metabolite of the mevalonate pathway end product. To evaluate the role of the mevalonate pathway and ketogenesis in cardiomyocyte proliferation, the effects of pharmacological HMGCS and 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR) inhibition on the viability and proliferation of neonatal rat ventricular cardiomyocytes was investigated. Simultaneous inhibition of the mevalonate and ketogenesis routes with hymeglusin decreased the percentage of proliferating cardiomyocytes (BrdU levels) with nontoxic concentrations (Figure 10). However, inhibition of the mevalonate pathway alone using the HMGCR inhibitor simvastatin showed no effect on the percentage of proliferating cardiomyocytes at nontoxic concentrations (Figure 10). These results indicate that the HMGCS-mediated ketogenesis may participate in regulating cardiomyocyte cell cycle activity.

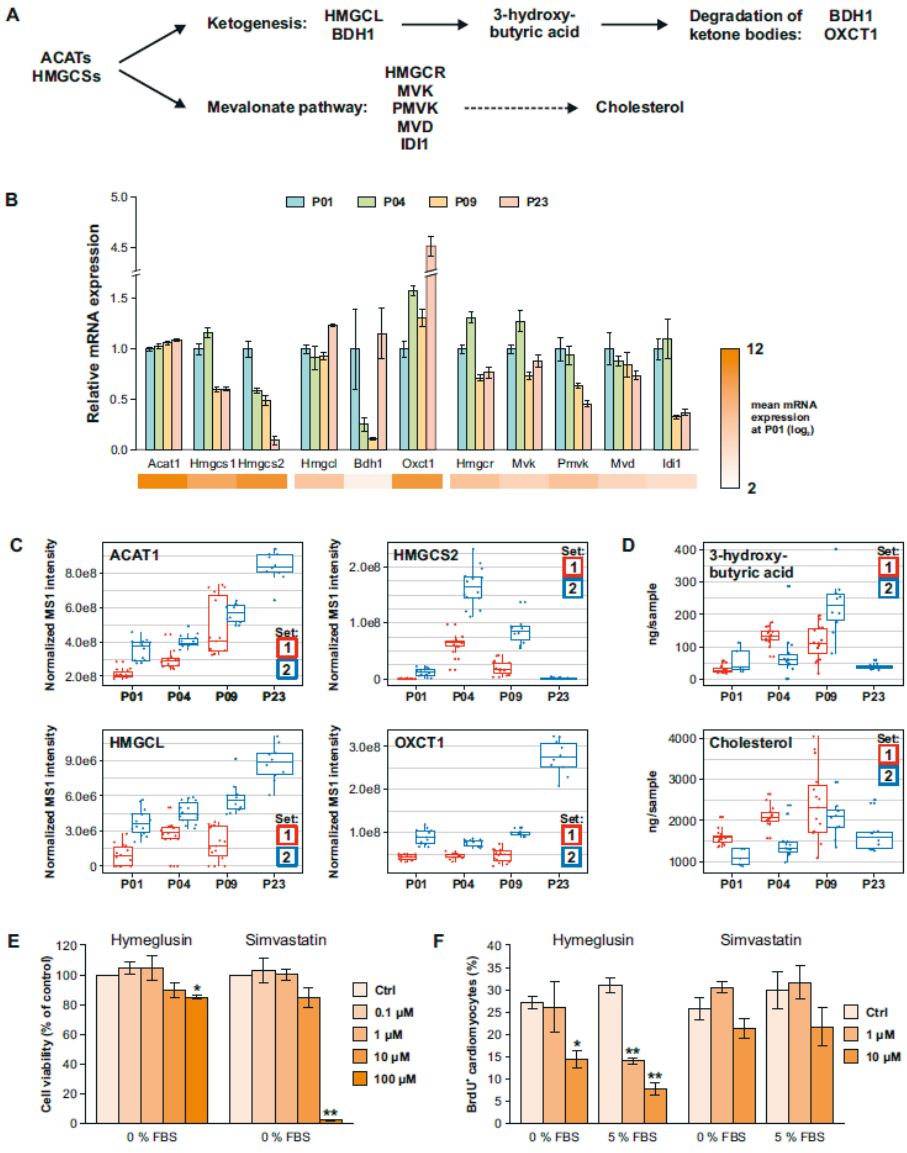

Figure 10 Ketogenesis and mevalonate pathways in the postnatal heart (A). B, Relative mRNA expression of selected ketogenesis and mevalonate pathway components (mean±SEM; n=3 pooled samples, each from 3 hearts). C, Normalized label-free quantification intensities of proteins detected in the ketogenesis and mevalonate pathways. D, Concentrations of 3-hydroxybutyrate and cholesterol over the early postnatal period. E and F, Effect of HMGCS inhibition with hymeglusin and HMGCR inhibition with simvastatin on neonatal rat ventricular cardiomyocyte viability using the MTT assay (E) and proliferation quantified as the percentage of bromodeoxyuridine (BrdU)-positive cells after 24-hour exposure (F). Data are expressed as mean±SEM from 3 independent experiments. *P<0.05, **P<0.01 compared with control (Ctrl); Welch ANOVA followed by Games-Howell. BDH1, 3-hydroxybutyrate dehydrogenase 1; FBS, fetal bovine serum; HMGCL, hydroxymethylglutaryl-coenzyme A lyase; IDI1, isopentenyl-diphosphate d-isomerase 1; MVD, mevalonate diphosphate decarboxylase; MS1, precursor ion mass spectrum; MVK, mevalonate kinase; OXCT1, 3-oxoacid CoA-transferase 1; PMVK, phosphomevalonate kinase (publication I).

4.2 METABOLOMICS OF HUMAN SERUM RELATING GUT MICROBIOTA AND INSULIN SENSITIVITY

In this study, we assessed the role of the gut microbiome as a source for key features of the serum metabolome profile that predicts metabolic and cardiovascular disorders in non-diabetic individuals (publication II). The metabolic phenotype corresponding to an insulin-resistant (IR metabotype) or insulin-sensitive (IS metabotype) phenotype was determined based on HOMA-IR and metabolic syndrome defined according to the International Diabetes Federation. The serum metabolic profile was determined by analyzing polar metabolites with GCxGC-MS and lipidomics with LC-MS. Serum metabolites were associated with the IR and IS metabotypes and results were further validated with type 2 diabetes patients. To evaluate the role of gut microbiota in insulin resistance (IR), gut microbiome metagenomics was analyzed and microbiome functional modules were further associated with IR and IS metabotypes and serum metabolite clusters.

4.2.1 LIPIDOMICS ANALYSIS OF SERUM SAMPLES

Serum lipidomics was performed with UHPLC-QTOF in positive ion mode after modified Folch extraction.[89,256,257] A total of 876 molecular serum lipid features were measured, including 289 identified and 587 unidentified features. The detected and identified lipid species were mainly LPLs, PLs such as PCs and PEs, Cer, SM, diacylglycerols (DG), TGs, and ChoEs. Some polar lipid classes (i.e. PI, PS, PA, PG) were not detected due to poor ionization efficacy in positive ion mode and weak extraction efficacy. Phosphatidic acids and lysophosphatidic acids are inadequately extracted in apolar solvents without pH adjustment with acidification,[265] which would again increase the risk for phospholipid hydrolysis to lysophospholipids.[266] The addition of ammonium buffer to eluent, however, enabled the formation of adducts and detection through adduct formation. The most abundant peaks of LPLs, PLs, Cer, and SM species were protonated ions, whereas non-polar lipids (DG, TG, and ChoE) were detected mainly as ammonium adducts.

Adequate separation of lipid species with an RP column was obtained within a total run time of 18 min (Figure 11), although overlap of some isobaric TGs were observed. Data quality was monitored with ISTD levels spiked to each sample and with additional quality control samples, blanks, control serum samples, and extracted and non-extracted standards. The RSD% of the ISTDs spiked into the samples were between 8-13% and the RSD% of the identified lipids in the control serum samples were on average 12.8%. Altogether, these values showed good analytical performance and neither sample preparation nor analysis order had any significant influence on the results.

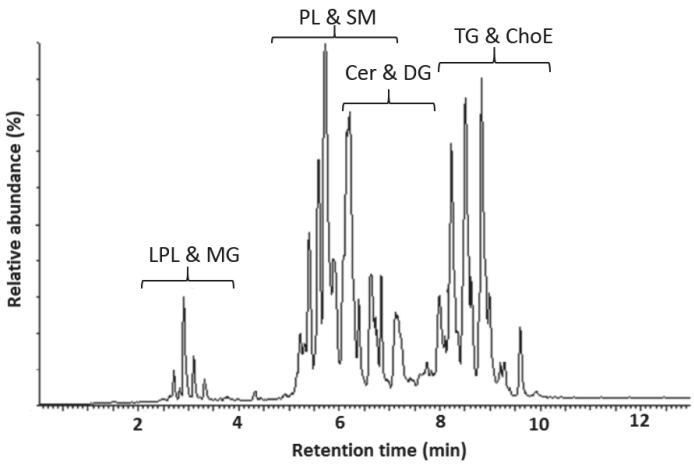

Figure 11 Total ion chromatogram of RP-UHPLC-TOF-MS lipidomics from a serum sample showing separation of major lipid classes. LPL; Lysophosholipids, MG; Monoglycerides, PL; Phospholipids, SM; Sphingomyelins, Cer; Ceramides, DG; Diacylglycerols, TG, Triacylglycerols, ChoE; Cholesterylesters.

4.2.2 METABOLITES CORRELATING WITH INSULIN RESISTANCE

All detected metabolic features were included in the data analyses, including unidentified features. Serum lipids (in total 876, 289 known, 587 unknown) and polar metabolite features detected with GCxGC-MS (in total 325 polar metabolites, 94 known, 231 unknown) were combined and clustered with weighted correlation network analysis to 74 co-abundance clusters across all individuals, containing in total 39 lipid clusters.[66] From these clusters, 19 were significantly and consistently associated with IR and metabolic syndrome from non-diabetic individuals and most of the associations were also confirmed in type 2 diabetes patients. These 19 clusters consisted of a total of 26 polar metabolites from 5 separate clusters and 367 lipids from 14 clusters (Table 8). All clusters were further separated into two groups either correlating positively (IR metabotype) or negatively (IS metabotype) with both IR and metabolic syndrome. The serum metabolites and lipids associated with the IR phenotype were mainly amino acids including BCAA, TCA cycle metabolites, and TGs. Elevated serum BCAA levels, changed TCA cycle metabolites, TGs, and specific membrane phospholipids have also previously been associated with IR and future risk of metabolic and cardiovascular events.[267–271] The IS metabotype contained only serum lipids (mainly PLs and TGs with odd carbon chains and high double bond content) and LysoPCs, LysoPEs, and SMs (Table 8). These results also confirm previous findings that odd-chain fatty acids and polyunsaturated fatty acids are associated with reduced risk of type 2 diabetes.[269,272]

Table 8 *Metabolite clusters correlating with IR or IS phenotype. NS, Not significantly associated with IR, PC, phosphatidylcholine, PE, phosphatidylethanolamine, SM, sphingomyelin, TG, triacylglycerol (publication II).*

Metabotype	Metabolite cluster	N metabolites	Description	Identified metabolites
Insulin resistance (IR)	M02	4	Sugar derivatives	None
	M09	4	Branched-chain keto-acids	4-methyl-2-oxovaleric acid, 2-oxoisovaleric acid
	M10	9	Branched-chain amino acids, other organic acids	Isoleucine, leucine, valine, proline, hydrocinnamic acid, N-acetyl-D-glucosamine, indole-3-lactic acid (NS), alpha-hydroxyvaleric acid (NS), 2-aminobutyric acid (NS)
	M11	4	Amino acids	Aspartic acid, glutamic acid
	M12	5	Tricarboxylic acid cycle metabolites	Pyruvic acid, lactic acid, alanine, alpha-ketoglutaric acid
	L04	5	PEs	PE(40:6)
	L05	78	Polyunsaturated TGs	**Examples:** TG(59:5), TG(56:7), TG(50:5)
	L10	28	Polyunsaturated TGs	**Examples:** TG(53:3), TG(54:4), TG(56:3)
	L11	150	TGs	**Examples:** TG(54:2),TG(54:5), TG(51:1)
	L12	5	Polyunsaturated TGs	TG(56:5), TG(56:5), TG(56:5), TG(58:5)
	L31	13	TGs	TG(46:3), TG(48:4), TG(52:0)
	L32	6	Unknown ChoEs	None
	L36	12	PCs	PC(36:3), PC(38:3), PC(38:4), PC(40:3), PC(40:4), ChoE(20:3)
Insulin sensitivity (IS)	L14	5	Ceramides and SMs	Cer(d18:1/16:0), SM(d18:1/16:0)
	L15	16	Ether lipids and polyunsaturated odd-chain TGs	**Examples:** PC(38:3e), PC(40:4e), PC(40:5e), TG(51:6), TG(53:8), TG(53:7)
	L18	19	PCs and PEs, ether lipids	**Examples:** PE(32:1), PC(34:2e), PC(34:2e), PE(40:7e)
	L19	8		PC(40:3e), SM(d18:1/24:1)
	L20	13		SM(d18:1/26:2)
	L30	9	LysoPCs and lysoPEs	LysoPC(18:1), lysoPC(18:2), lysoPE(20:3), lysoPC(22:5), lysoPC(18:3), lysoPC(20:4)

4.2.3 CORRELATING HOST INSULIN SENSITIVITY AND METABOLIC SYNDROME, GUT MICROBIOME, AND FASTING SERUM METABOLOME

The origin of serum metabolites associated with IR is unknown. In this study, we assessed the role of the gut microbiome as a source of IR-associated serum metabolites. For this, gut microbiome metagenomics of individuals from the study were analyzed and microbiome functional modules were further associated with IR and IS metabotypes. A total of 41 microbiome functional modules were significantly associated with one or more of the IR and metabolic syndrome phenotypes and furthermore associated with the IR and IS metabotypes (Table 8). The functional modules of the microbiome that were positively associated with both HOMA-IR and the IR metabotype contained enzymes for biosynthesis of BCAAs, cofactors, vitamins, lipopolysaccharides, and various transport systems (Figure 12). In contrast, the microbiome negatively associated with HOMA-IR and the IR metabotype contained functional modules important for methanogenesis, pyruvate oxidation, and transport systems, including inward transport of BCAAs (Figure 12). The microbial functional module analysis highlighted an increased BCAA biosynthesis potential, but depletion of genes encoding the transport system for bacterial inward BCAA transport in the gut microbiome of IR individuals. Microbial species contributing to the associations between the microbiome functional modules and HOMA-IR were identified with iterated correlation analysis. The positive associations between HOMA-IR and the biosynthesis modules of BCAA, tryptophan, and lipopolysaccharides were largely driven by *Prevotella copri* (*P. copri)* followed by *Bacteroides vulgatus*. *P. copri* also had a positive correlation with the BCAA-containing metabolite cluster M10 in the 94 individuals with detectable levels of the species. Whereas the functional modules enriched in the gut microbiome of IR individuals were driven by relatively few microbial species each with high impact, all microbial functional modules associated with increased insulin sensitivity were driven by multiple species including *Butyrivibrio crossotus* and *Eubacterium siraeum*, all with minor effect. Accordingly, gene enrichment of the microbiome and presence of multiple species had a positive correlation with insulin sensitivity. These findings also suggest that elevated serum BCAA concentrations, which are known to associate with IR, are a joint effect of increased potential for BCAA biosynthesis and decreased inward BCAA transport of the gut microbiome. Since *P. copri* was the strongest driver species for the positive association between microbial BCAA biosynthesis in the gut and IR, the role of *P. copri* was studied experimentally. *P. copri* (n = 12) and sham-gavaged (n = 12) male mice (C57BL/6J) on a high-fat diet were studied. In 2 weeks, *P. copri* mice aggravated glucose intolerance and showed increased total serum BCAA

levels. Reduced insulin sensitivity of *P.copri* mice was observed after 3 weeks of *P. copri* challenge.

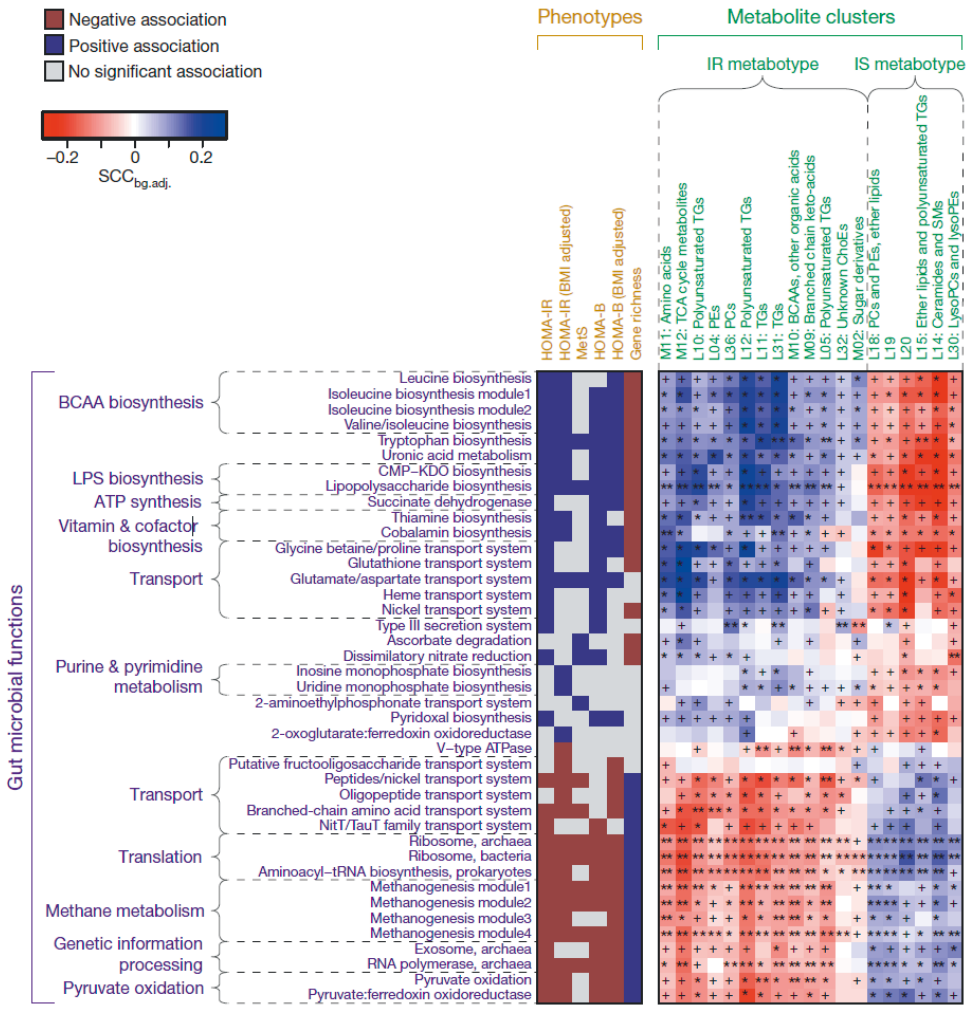

Figure 12 Association map of the three-tiered analyses integrating the phenome, the gut microbiome, and the fasting serum metabolome in 277 non-diabetic individuals with available metagenomic data. The left panel shows significant associations (Mann-Whitney U-test FDR <0.1) between microbial functional modules and the indicated phenotypes; coloring indicates direction of association. The right panel shows associations between the same modules and serum metabolite clusters. Coloring represents the median Spearman correlation coefficient between metabolite clusters and the indicated functional modules. +, FDR < 0.1; *, FDR <0.01; **, FDR <0.001 (publication II).

4.3 COMPARISON OF LC-MS AND CHIP-MS DIRECT INFUSION METHOD IN GLOBAL NON-TARGETED METABOLOMICS

Rapid analysis from low sample volumes would be beneficial in large metabolic studies of clinical samples. Microchip-MS is considered as a potential technique for improving the sensitivity and cost-efficiency of metabolomics research, since it allows a rapid, straightforward analysis from low sample volume, and possibility for completely automated system in the future. In this work, we wanted to study the applicability of microchip-MS for non-targeted metabolomics (publication III). The feasibility of direct infusion electrospray ionization microchip mass spectrometry (chip-MS) was compared to the commonly applied liquid chromatography-mass spectrometry (LC-MS) in non-targeted metabolomics analysis of HFF and hiPSC cells reprogrammed from HFF.

4.3.1 COMPARISON OF ANALYTICAL PERFORMANCE OF LC-MS AND CHIP-MS

Using the custom-fabricated SU-8-based ESI microchip Orbitrap MS method,[259,260] the duration of the global metabolomics measurement was only 0.3 minutes, which was adequate to measure high-quality spectra (Figure 13). The run time was 30 minutes with application of RP-UHPLC-ESI-QTOF-MS (Figure 13). Since high sample throughput is very important in large metabolomics studies (especially in clinical experiments), the short analysis time in chip-MS is a remarkable benefit compared to LC-MS.

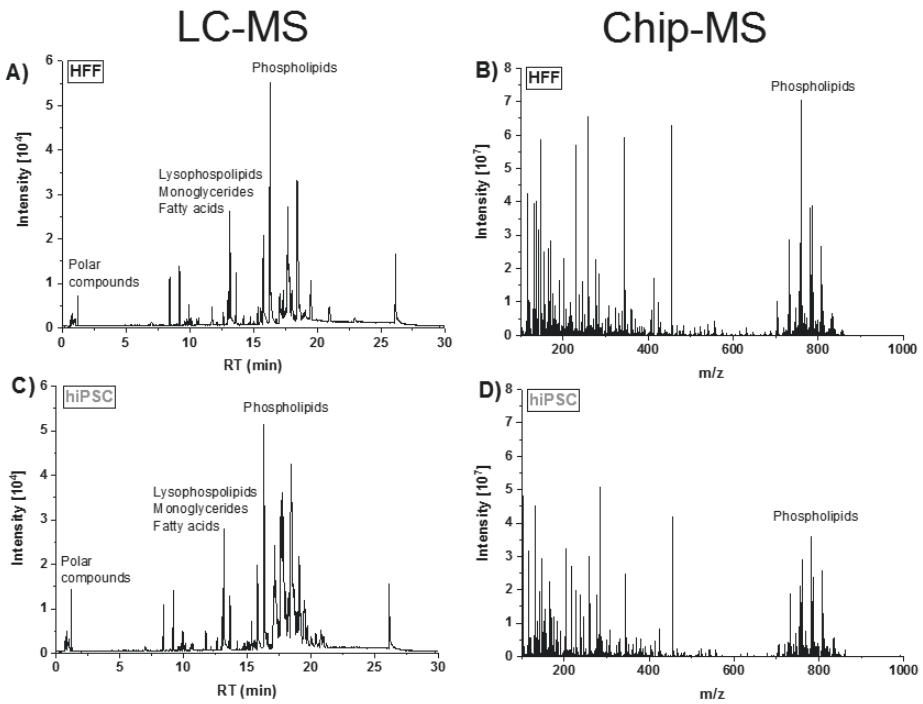

Figure 13 Base peak chromatograms of HFF and hiPSC samples (A and C) with RP-UHPLC-ESI-QTOF-MS with mass range 100-1000 *m/z*. Combined spectra of HFF and hiPSC samples (0.3 min) with SU-8-based ESI microchip Orbitrap MS (B and D) (publication III).

A wide range of metabolites from polar small molecules to lipids were detected with both the chip-MS and the LC-MS methods (Fig 13). The main metabolites identified based on accurate mass were amino acids, amino acid derivatives, acylcarnitines, nucleotides, sugars, sugar phosphates, fatty acids, fatty acid and steroid derivatives, and a high number of lipids from classes such as MGs, PCs, PEs, and SM. In addition, several unidentified metabolites that include oxygen as a heteroatom in aliphatic structures (such as alcohols, esters, and carboxylic acids) were found based on the accurate mass measurements. The metabolites were detected as their protonated molecules or sodium or potassium adduct ions (or a combination of these). Adduct ions were observed as the cell samples contained residual sodium and potassium salts after the extraction. In LC-MS, adduct formation was less abundant due to separation and elution of salts in the void volume and mainly protonated molecules were formed. In chip-MS, significantly more abundant adduct ions were detected, as the residual salts were not chromatographically separated from metabolites and the higher salt concentration in the ESI favored adduct ion formation. Analyte presence as multiple abundant ions may decrease sensitivity. However, salts may also be beneficial and assist ionization of some low proton affinity metabolites, which are poorly ionized via a proton transfer

reaction but are ionized more efficiently via adduct ion formation.[273] Some metabolites (including low proton affinity compounds in compound classes, carboxylic acids, amino acids, carnitines, sugar derivatives, sugar phosphates, phosphates, and nucleosides) were detected only with chip-MS as their adduct ions but were not detected in the LC-MS.

The total number of the detected features with chip-MS and LC-MS were 619 and 1959, respectively. The difference in the number of detected features is largely explained with peak capacity and the sensitivity of the applied analytical method. The lower peak capacity with the chip-MS results in an increased number of overlapping features, particularly with lipids (Figure 14). In the LC-MS data of hiPSC and HFF, the extracted ion chromatograms of m/z 806.57 (Δm/z 0.04) showed several peaks, whereas the chip-MS spectrum showed only one peak between m/z 806.5-806.6; differences of individual lipid species between hiPSC and HFF were not observed (Figure 14). Investigation of the spectra detected in the extracted ion LC-MS chromatogram of m/z 806.57 at retention times of 17.53 and 17.33 min revealed that these most likely represent isomers of $[PC(38:6)+H]^+$ (exact mass m/z 806.5694), and the peaks at 17.88, 18.0, and 18.19 min represent isomers of $[PC(36:3)+Na]^+$ (exact mass m/z 806.5670) (Figure 14). The mass difference between the ions $[PC(38:6)+H]^+$ and $[PC(36:3)+Na]^+$ is only 2.4 mDa, which means that mass resolution over approximately 350 000 would have been required to distinguish these ions in the chip-MS spectra. In addition, the [M+2] isotopes of lipids may overlap with lipids belonging to the same lipid class but with one less double bond (Figure 14). Overlap of metabolic features may especially hinder the detection of differences of less abundant metabolites between samples and the use of LC-MS provides significantly better selectivity and thus more detailed information in non-targeted metabolomics than direct MS.

Another reason for the lower number of detected features with chip-MS is ion suppression, which is more evident in chip-MS than in LC-MS. Salts and some easily ionized phospholipids in ESI (such as phosphatidylcholines) may suppress ionization of metabolites that have lower ionization efficiency.[274,275] For example, the intensities of several LPLs were below the detection threshold in chip-MS but were clearly detected with LC-MS.

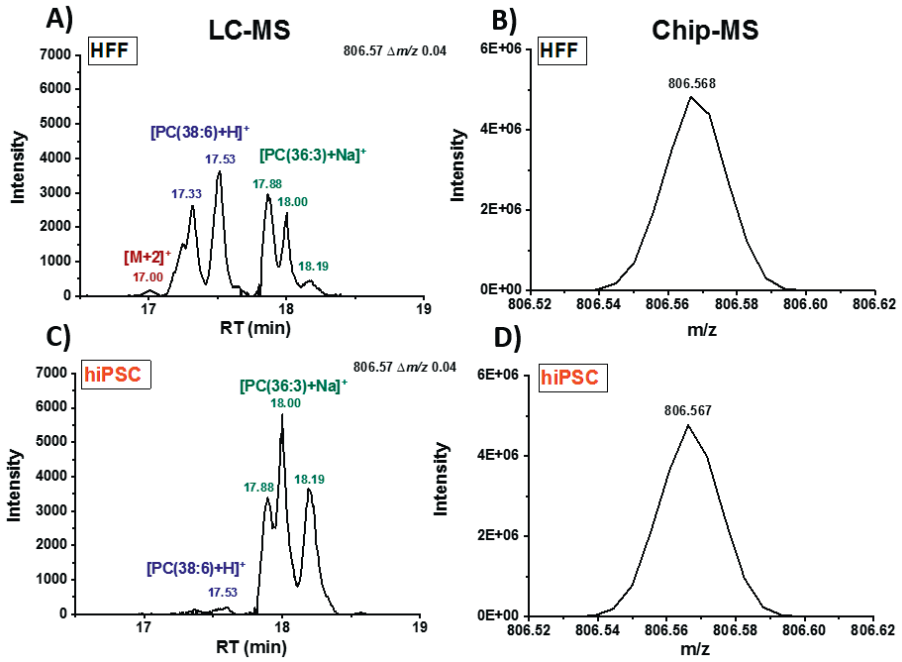

Figure 14 RP-UHPLC-ESI-QTOF-MS ion chromatograms of detected features at *m/z* 806.57 (Δ*m/z* 0.04) from HFF and hiPSC samples (A and C). The peaks in the chromatograms correspond to protonated phosphatidylcholine $[PC(38:6)+H]^+$, phosphatidylcholine sodium adduct $[PC(36:3)+Na]^+$, and (M+2) isotope of m/z 804.55 $[PC(38:7)+H]^+$. SU-8-based ESI microchip Orbitrap MS spectra of HFF and hiPSC samples with mass window of *m/z* 806.57 ± 0.05 (B and D) (publication III).

The repeatability of the methods was evaluated with added ISTDs propranolol and verapamil in the cell samples. The RSD% of ISTD ion intensities measured using LC-MS in HFF samples were 5.1-6.0% and in hiPSC samples 4.2-6.1%. The RSD% measured using chip-MS for ISTDs in HFF samples were 8.6% and in hiPSC samples 6.4-10.6%. These results indicate that the repeatability within each cell sample was good with both methods. However, the signal intensities of the ISTDs were about two times lower in hiPSC samples than in HFF samples in the chip-MS measurements, but no significant differences were observed in the LC-MS measurements (Figure 15). In our experiment, the biomass of the cells was the same in all samples. However, the number of cells was higher in hiPSC samples (5×10^6) than in HFF samples (1×10^6), due to smaller size of hiPSC compared to HFF. Consequently, the total surface area and the amount of phospholipid containing cell membrane is higher in hiPSC than in HFF samples. Certain phospholipids, such as PCs, are known to cause ion suppression in ESI.[274,275] Thus, ion suppression might be stronger with hiPSC than with HFF samples when measured by chip-MS, which is more prone to ion suppression than LC-MS. These results show that normalization using ISTDs is necessary for

comparison of metabolic changes between cell samples when using the chip-MS approach.

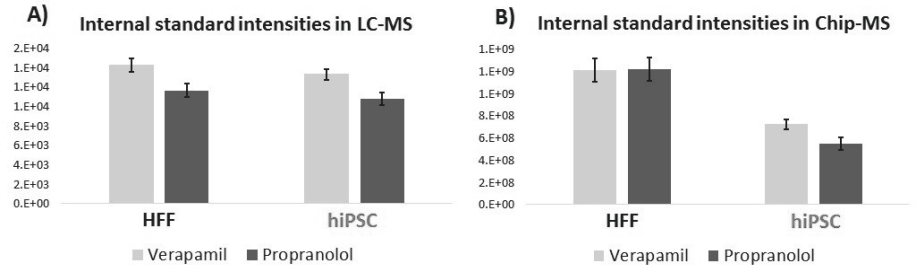

Figure 15 Intensity levels of ISTDs verapamil and propranolol in HFF and hiPSC samples in RP-UHPLC-ESI-QTOF-MS (A) and SU-8-based ESI microchip Orbitrap MS (B) analysis (publication III).

4.3.2 OBSERVED METABOLIC DIFFERENCES BETWEEN CELLS WITH LC-MS AND CHIP-MS

An unsupervised PCA was performed to study the capability of both analytical methods to distinguish metabolic differences of the cell samples. PCA showed clear grouping of the cell types, indicating differences in the metabolic profiles of HFF and hiPSC with both analytical methods (Figure 16). To determine statistically significant changed features between the HFF and hiPSC datasets, we used t-test q-value <0.01 and fold change of at least 2 for the datasets acquired with chip-MS and LC-MS. A total of 435 features detected with the LC-MS fulfilled the criteria (q <0.01, fold >2), of which 304 features were upregulated and 131 were downregulated in hiPSC compared to HFF. The number of changed features (q <0.01 and fold >2) detected with chip-MS was 157, of which 88 features were upregulated and 69 were downregulated in hiPSC. Only about one third of the changed features detected with LC-MS were detected with chip-MS, which is most likely due to the lower selectivity and higher ion suppression of chip-MS. The number of significantly changed features (q <0.01, fold >2) with both analytical methods (with a maximum mass difference of 20 ppm) was 44. These metabolites were tentatively identified based on their accurate masses and these are presented with fold changes (hiPSC/HFF) and q-values in Table 9. The fold changes mainly correlated well between the analytical methods, although three phospholipids (m/z 754.536, 834.599, and 818.603) were changed in the opposite direction. One possible explanation for this is peak overlap and domination of sodium adducts in the chip-MS measurements. The most downregulated compounds in hiPSC compared to HFF were N-methylnicotinamide, creatine, L-carnitine, and PCs and PEs. The most clearly

upregulated metabolites in hiPSC compared with HFF were MGs, aminoadipic acid, choline, niacinamide, and some PCs.

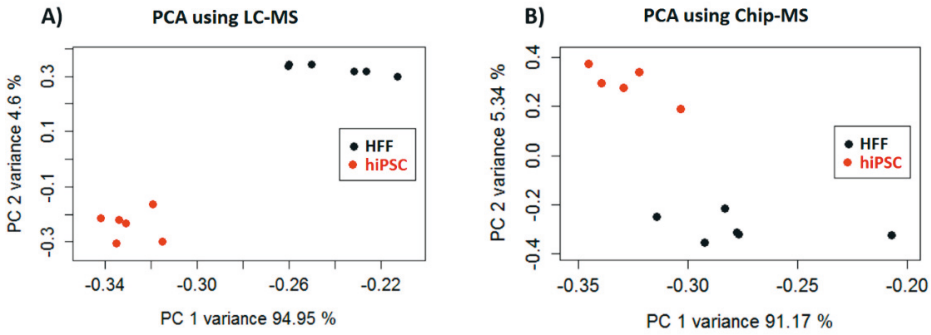

Figure 16 PCA score plots of normalized and scaled intensities in HFF (black) and hiPSC (red) samples using RP-UHPLC-ESI-QTOF-MS (A) and SU-8-based ESI microchip Orbitrap MS (B) (publication III).

Table 9 *Changed metabolites between cell samples detected with both methods (q<0.01, fold >2). Red indicates upregulated and blue indicates downregulated metabolites in hiPSC compared with HFF (publication III).*

LC-MS			Chip-MS			Elemental formula	Tentative identification
q-value	Log 2 hiPSC/ HFF	m/z	q-value	Log 2 hiPSC/ HFF	m/z		
3.5E-05	5	379.2863	1.9E-04	5	379.2816	C21H40O4+Na	MG(18:1)
4.6E-07	4	162.0779	1.6E-04	2	162.076	C6H11NO4+H	Aminoadipic acid
9.5E-06	3	772.6278	1.4E-06	4	772.6197	C44H86NO7P+H	PC(36:2e)
1.3E-04	3	353.2673	1.2E-04	5	353.2659	C19H38O4+Na	MG(16:0)
1.1E-11	3	283.0738	2.7E-04	3	283.0695	unknown	unknown
2.9E-06	3	104.1081	1.2E-04	4	104.1073	C5H14NO+H	Choline
2.4E-06	3	299.0473	1.2E-04	5	299.0437	unknown	unknown
2.0E-07	3	261.0911	4.0E-04	3	261.0877	unknown	unknown
1.3E-05	3	678.5107	8.4E-05	4	678.5065	C36H72NO8P+H	PC(28:0)
3.0E-05	3	351.2481	9.4E-05	5	351.2502	C19H36O4+Na	MG(16:1)
2.8E-05	3	123.0564	9.8E-05	2	123.0554	C6H6N2O+H	Niacinamide
1.6E-07	3	203.1517	3.8E-04	6	203.1501	C8H18N4O2+H	Dimethylarginine
6.2E-06	3	261.0642	9.8E-05	2	261.0579	unknown	unknown
6.7E-03	2	704.5222	3.6E-06	2	704.5226	C38H74NO8P+H	PC(30:1)
8.2E-06	2	258.1129	7.0E-04	7	258.1099	C8H20NO6P+H	Glycerophosphocholine
1.7E-04	2	518.3296	5.7E-04	3	518.3218	C24H50NO7P+Na, C26H48NO7P+H	LysoPC(16:0)Na, LysoPC(18:3)
7.1E-06	2	247.0451	3.1E-05	2	247.0422	unknown	unknown
1.2E-04	2	522.3579	6.5E-04	1	522.3554	C26H52NO7P+H	LysoPC(18:1)
1.2E-04	2	159.0295	8.7E-05	2	159.0276	C5H4N4O+Na	Hypoxanthine
3.7E-06	2	137.0478	5.6E-05	1	137.0458	C5H4N4O+H	Hypoxanthine
3.5E-04	2	544.3386	2.4E-03	2	544.3374	C26H52NO7P+Na, C28H50NO7P+H	LysoPC(18:1)Na + LysoPC(20:4)
8.9E-07	1	206.0575	1.7E-04	3	206.0552	C5H15NO4P+Na	Phosphorylcholine
3.6E-06	1	147.1143	1.1E-04	1	147.1128	C6H14N2O2+H	L-Lysine
1.4E-04	1	169.0605	1.6E-05	1	169.0583	C5H10N2O3+Na	L-Glutamine
6.7E-05	1	175.0024	3.5E-04	1	175.0016	C5H4N4O+K	Hypoxanthine
7.9E-07	1	347.0317	1.1E-03	2	347.0249	C9H13N2O9P+Na	UMP
1.6E-06	1	118.0873	5.6E-05	1	118.0865	C5H11NO2+H	Betaine
6.3E-07	1	780.5554	6.9E-03	2	780.5511	C42H80NO8P+Na, C44H78NO8P+H	PC(32:2)Na, PC(36:5)
2.3E-06	1	147.0787	1.0E-04	1	147.0764	C5H10N2O3+H	L-Glutamine
6.6E-05	1	184.0754	1.2E-04	2	184.0733	C5H15NO4P+H	Phosphorylcholine
7.1E-05	1	305.0047	4.2E-04	2	305.0006	unknown	unknown
1.6E-04	-1	748.5258	1.2E-03	-5	748.5269	C43H74NO7P+H, C41H76NO7P+Na	PE(38:7e), PE(36:4e)Na
2.2E-04	-1	718.5392	7.8E-04	-3	718.5381	C39H76NO8P+H	PE(34:1)
4.2E-05	-2	754.5318	5.6E-03	2	754.5355	C40H78NO8P+Na, C42H76NO8P+H	PC(32:1)Na, PC(34:4)
4.5E-05	-2	834.5949	8.4E-03	1	834.5986	C46H86NO8P+Na, C48H84NO8P+H	PC(38:3)Na, PC(40:6)
8.8E-06	-3	132.0785	1.6E-04	-3	132.0768	C4H9N3O2+H	Creatine
1.2E-04	-4	746.5673	1.2E-03	-3	746.569	C41H80NO8P+H	PE(36:1), PC(33:1)
3.0E-08	-4	830.5710	6.8E-06	-2	830.5663	C46H82NO8P+Na, C48H80NO8P+H	PC(38:5)Na, PC(40:8)
1.4E-06	-4	772.5765	9.7E-04	-3	772.5839	C43H82NO8P+H	PE(38:2)
3.9E-09	-5	856.5875	1.1E-06	-3	856.5813	C48H84NO8P+Na, C50H82NO8P+H	PC(40:6)Na, PC(42:9)
1.7E-05	-5	162.1139	1.7E-04	-4	162.1124	C7H15NO3+H	L-carnitine
5.9E-04	-6	818.5975	2.7E-03	3	818.6029	C46H86NO7P+Na, C48H84NO7P+H	PC(38:4e)Na, PC(40:7e)
9.0E-07	-8	137.0702	1.2E-04	-4	137.071	C7H8N2O+H	N-Methylnicotinamide
1.9E-04	-8	828.5468	3.6E-06	-5	828.5505	C46H80NO8P+Na, C48H78NO8P+H	PC(38:6)Na, PC(40:9)

4.4 CAPILLARY PHOTOIONIZATION

Many biologically important low-abundance metabolites are not detectable with non-targeted metabolomics methods and separate more sensitive methods are required. APPI and CPI have been previously shown to produce high sensitivity towards non-polar compounds such as steroids.[142,150,154,276] However, the suitability of CPI towards quantitative analysis of biological samples of large analyte groups has not been previously established. Accordingly, we sought to develop the previous in-house CPI prototype further

to increase sensitivity and robustness. The new CPI source was designed with a wide and low-depth ionization chamber under the wide MgF_2 window to maximize ionization efficiency using a commercially available photoionization lamp with window diameter 13 mm. The wide surface area of the opening provides high photon flux into the ion source, ensuring high ionization efficiency and thus improved sensitivity. The low depth and oval shape of the ionization chamber were designed to minimize dead volume and dead angles. The volume of the ionization chamber (<0.4 mL) is negligible compared to the total gas flow rate through the chamber (approximately 1 L min^{-1}). The low depth of the ionization chamber also ensures that a large fraction of the molecules entering the ion source are exposed to the ionizing photons. The ionization chamber was heated with a cylindrical heater embedded in a cylindrical aluminium block. PID temperature control of the heater and thermal mass of the aluminium block ensured a stable and sufficiently high temperature to avoid adsorptions of the analytes to the surfaces. Together, these modifications ensured that the design of the ionization chamber did not cause peak tailing, broadening, or carryover. An additional benefit of CPI was that positioning of the end of the transfer capillary between the chromatograph and CPI device is less critical than the positioning of an APPI sprayer in relation to the inlet orifice of a mass spectrometer. CPI was applied in this work as an interface between GC or LC and MS (publications IV and V).

4.4.1 GC-CPI-MS/MS METHOD FOR ANALYSIS OF STEROIDS

A GC-CPI-MS/MS method was developed for the analysis of 18 steroids from urine as their TMS derivatives (Table 10) (publication IV). First, the MS spectra of the TMS-derivatized steroids were measured to select the precursor ions for MS/MS analysis. Steroids formed mainly molecular ions $[M]^{+\bullet}$ as base peaks with CPI, since the low ionization energy of TMS-derivatized steroids favors a charge-exchange reaction with chlorobenzene dopant molecules. For only CS-TMS and HC-TMS was the most intense ion $[M\text{-}TMSOH]^{+\bullet}$, in contrast to molecular ion, which was selected as a precursor ion. Additionally, the ions $[M+H]^+$, $[M\text{-}TMS]^{+\bullet}$, $[M\text{-}TMSOH]^{+\bullet}$, and $[M+O_2]^{+\bullet}$ were formed and changes in ion ratios were observed due to changes in ambient air conditions. As a gas flow towards MS was higher than the gas flow provided by GC and dopant system, ambient air could flow into the CPI, possibly affecting the photoionization process. This was prevented by building up a nitrogen housing around the ion chamber. The nitrogen housing prevented the formation of other ions possibly by preventing reactions with molecules of ambient air such as oxygen and water (Figure 17). The nitrogen housing was further applied in the method development, validation, and application to urine samples. With the nitrogen housing, primarily molecular ion was formed, although the $[M\text{-}TMS]^{+\bullet}$ ion was also seen as a minor peak in the MS spectra of multiple steroids (Figure 17).

Table 10 *Analyzed steroid precursor ions, TMS derivatives, optimized MRM transitions, and retention times.*

compound	RT (min)	precursor ion type	TMS-groups in M	precursor ion (m/z)	quantifier ion (m/z)	qualifier ion (m/z)
ADT	13.38	$M^{+\bullet}$	2	434.3	161.2	195.2
ETIOL	13.56	$M^{+\bullet}$	2	434.3	155.1	221.1
DHEA	15.05	$M^{+\bullet}$	2	432.3	327.3	169.1
AN	16.92	$M^{+\bullet}$	2	430.2	209.1	234.1
T	16.99	$M^{+\bullet}$	2	432.2	209.1	301.2
E1	17.85	$M^{+\bullet}$	2	414.2	155.1	309.2
E2	17.98	$M^{+\bullet}$	2	416.2	285.2	129.1
Me-T (ISTD)	19.04	$M^{+\bullet}$	2	446.3	301.2	314.3
PREG	20.29	$M^{+\bullet}$	2	460.3	157.1	265.2
17-OH-PREG	20.93	$M^{+\bullet}$	3	548.3	231.1	230.1
E3	21.04	$M^{+\bullet}$	3	504.3	311.2	324.2
PROG	21.85	$M^{+\bullet}$	2	458.3	157.1	353.2
17-OH-PROG	22.38	$M^{+\bullet}$	3	546.3	301.2	314.2
11-OH-PROG	23.61	$M^{+\bullet}$	3	546.3	351.3	233
21-OH-PROG	24.77	$M^{+\bullet}$	3	546.3	301.2	230.1
CS	24.82	$[M\text{-}TMSOH]^{+\bullet}$	5	630.3	243.1	283.1
CORT	25.56	$M^{+\bullet}$	4	634.4	230.1	404.3
HC	25.82	$[M\text{-}TMSOH]^{+\bullet}$	5	632.2	234.1	437.2
A	26.02	$M^{+\bullet}$	4	648.3	305.2	233

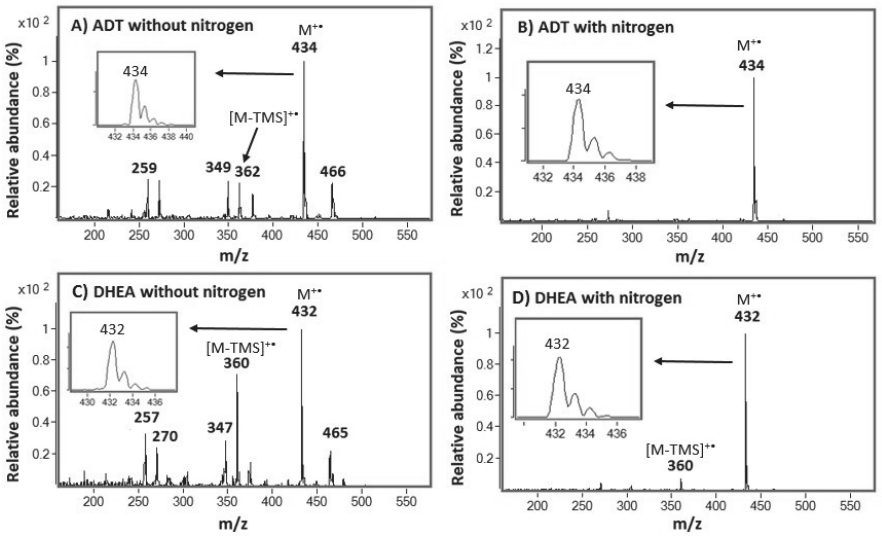

Figure 17 Backgroud-substracted mass spectra of ADT (A,B) and DHEA (C, D), measured from the same samples on the same day with (B,D) and without (A,C) nitrogen housing (publication IV).

The MS/MS spectra of steroid molecular ions showed similar non-specific fragments corresponding to the loss of CH_3 $[M-15]^+$, TMSOH $[M-90]^{+\bullet}$, CH_3 + TMSOH $[M-105]^+$, CH_3 + 2xTMSOH $[M-195]^+$, and $[TMS]^+$ peak with m/z 73 (Figure 18). Specific fragments from the ring structure were also observed for all steroids. For the MRM method, two of the most intense sufficiently specific fragment ions were selected and optimized (Table 10).

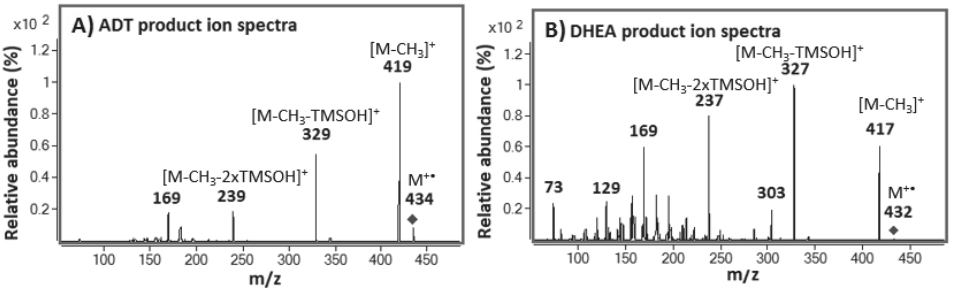

Figure 18 Product ion spectra of ADT (A) and DHEA (B), showing examples of typically formed steroid fractions in MS/MS spectra (publication IV).

Isobaric steroids and isotopes were adequately separated when two GC columns were connected in tandem (Figure 19). All steroids could be unambiguously detected and identified based on chromatographic separation and selective MRM transitions. Retention times of the steroids varied between 13.38 and 26.05 min (Figure 19). Chromatographic separation with the method was good, and retention time RSDs were below 1%. Peaks were symmetric and narrow with peak half widths in the range of 3.1 to 5.3 s. This indicates that the dead volume of the ion source was sufficiently small. The ion source and transfer line materials and temperatures were also adequate to avoid any cold traps, which could cause peak tailing and broadening.

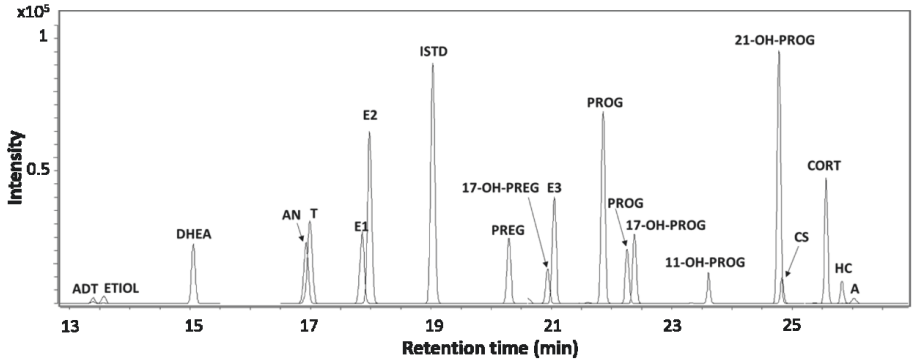

Figure 19 Overlaid MRM chromatograms (quantifier) of steroids analyzed from 10 ng mL^{-1} standard solution spiked in artificial urine (publication IV).

4.4.2 VALIDATION OF GC-CPI-MS/MS METHOD AND APPLICATION TO HUMAN URINE

The method was validated in terms of recovery of LLE, LOD, LOQ, linearity, repeatability, and carryover (Table 11) (publication IV). LODs were in the equal range of 0.002 to 0.02 ng mL^{-1} for most of the steroids, with the exception of A, ADT, ETIOL, and HC, which had slightly higher LODs (0.05-0.1 ng mL^{-1}). LODs reported in the literature for GC-EI-MS analysis of steroids from urine have ranged from 0.5 to 20 ng mL^{-1} for androgenic anabolic steroids (AAS),[277,278] 0.08 to 100 ng mL^{-1} for endogenous androgens,[164,279] 5 to 20 ng mL^{-1} for corticosteroids,[280,281] 0.001 to 0.4 ng mL^{-1} for estrogens,[279,282] and 0.3 to 2 ng mL^{-1} for progestogens.[279] With LC-MS, LODs have ranged 0.1 to 10 ng mL^{-1} for AAS[283] and 0.02 to 113 ng mL^{-1} for corticosteroids.[280,281,283] LOQs for estrogens ranged from 0.02 to 0.05 ng mL^{-1}.[282,284] LODs and LOQs of our GC-CPI-MS method were mainly lower than the values reported previously in the literature,[164,277–284] thus showing the high sensitivity of the method. Repeatability of all steroids was reasonable, with RSD values varying between 5% to 18% for six replicates at the level of 10 ng mL^{-1} and were in the same range as reported previously for GC-EI-MS systems.[164,278,279] Linearity was acceptable with a coefficient of determination (R^2) varying between 0.981 to 0.996. The linear range covered 2 to 4 orders of magnitude in the range 0.05 to 200 ng mL^{-1}. Linear range coverage, linearity, and R^2 values showed similar analytical performance as obtained with GC-EI-MS systems.[164,278–280] LLE recoveries were in the range of 57.3% to 123.7%. Carryover was below 0.01% for all analytes.

Table 11 *GC-CPI-MS/MS validation results; LOD, LOQ, linear range, R^2 of calibration curve, repeatability RSD% (n=6), LLE recovery % (publication IV). *From diluted standard solutions.*

Steroid	LOD (ng mL^{-1})	LOQ (ng mL^{-1})	Linear range (ng mL^{-1})	R^2	Repeatability (n= 6)	Recovery (%)
11-OH-PROG	0.02	0.2	0.2-50	0.991	11.6	82.1
17-OH-PREG	0.01	0.1	0.1-100	0.988	12.9	122.1
17-OH-PROG	0.01	0.5	0.5-100	0.99	10.2	123.7
21-OH-PROG	0.002	0.2	0.2-100	0.995	5	62.9
A	0.1	2	2-200	0.992	12.6	70.3
ADT	0.1	2	2-100	0.988	8.5	70.2
ADT*			100-2000	0.994		
AN	0.02	0.5	0.5-50	0.992	7.5	79.4
CORT	0.01	0.2	0.2-20	0.993	11.1	83.4
CS	0.01	0.5	0.5-100	0.995	17.6	75.4
CS*			100-1000	0.997		
DHEA	0.01	0.05	0.05-100	0.996	5.3	75.7
E1	0.01	0.1	0.1-100	0.994	10.3	72.6
E2	0.01	0.1	0.1-100	0.995	10.3	73.6
E3	0.005	0.1	0.1-100	0.989	5.7	81
ETIOL	0.05	0.1	0.1-100	0.996	9	75.4
ETIOL*			100-2000	0.997		
HC	0.05	1	1-100	0.981	18	57.3
PREG	0.01	0.5	0.5-20	0.99	13.6	64
PROG	0.02	0.1	0.1-100	0.995	5.5	64.3
T	0.02	0.5	0.5-100	0.996	8.3	70.6

The method was applied to analyze the urine of eight volunteers (4 females and 4 males) (Table 12). The determined concentrations of most of the steroid glucuronide conjugates were higher or at the same level compared with the respective free steroids, and only 11-OH-PROG and 17-OH-PROG showed levels below the LOQ in all samples also as glucuronide conjugates (Table 12). The glucuronides of ADT and ETIOL are the main endogenous metabolites of testosterone and were excreted in urine at high concentrations (100-2000 ng mL^{-1}), as reported previously.[285] The corticosteroids CS and HC had the highest observed free-form concentrations (10-100 ng mL^{-1}) and were at the same level as their glucuronide conjugates and consistent with earlier findings.[286] Clear differences in male and female samples regarding sex steroids were observed and steroid concentrations (T, E1, E2, E3, DHEA) were consistent with previous studies.[279,282,284,285] The detected concentrations of DHEA glucuronide were in the range of 10 to 40 ng mL^{-1}; the mean value reported in the literature in a study of over 3000 subjects was approximately 35 ng mL^{-1}.[285,286] The data showed clear differences in male and female

samples regarding sex steroids; testosterone was higher in males (mean 16.9 ng mL^{-1} in males, 1.8 ng mL^{-1} in females) and estrogens were higher in females (>2 fold difference in the mean values), consistent with the literature.[279,284,285] Additionally, small concentrations of AN and CORT were detected in free form and 17-OH-PREG, 21-OH-PREG, A, AN, CORT, and PREG as glucuronide conjugates.

Table 12 *Concentrations of analyzed steroids in the urine of eight volunteers in the free or in glucuronide-conjugated form. All values are given in ng mL^{-1}. <LOQ corresponds to signal above LOD and below LOQ (publication IV).*

	11-OH -PROG	17-OH -PREG	17-OH -PROG	21-OH -PROG	A	ADT	AN	CORT	CS	DHEA	E1	E2	E3	ETIOL	HC	PREG	PROG	T
Free fraction																		
Male 26 y	-	-	<LOQ	-	-	-	0.9	-	18	0.3	<LOQ	-	-	<LOQ	2.7	<LOQ	-	<LOQ
Male 34 y	-	<LOQ	<LOQ	-	-	-	1.3	0.2	47	1.7	<LOQ	-	<LOQ	1.9	37	<LOQ	-	<LOQ
Male 38 y	-	<LOQ	<LOQ	-	-	-	1.1	0.3	66	0.4	<LOQ	-	<LOQ	4.9	38	<LOQ	-	<LOQ
Male 49 y	-	<LOQ	<LOQ	-	-	-	0.9	0.3	87	0.2	<LOQ	-	-	0.8	43	<LOQ	<LOQ	<LOQ
Female 26 y	-	<LOQ	<LOQ	-	-	-	0.7	<LOQ	41	0.7	<LOQ	-	<LOQ	1.3	22	<LOQ	-	<LOQ
Female31 y	-	<LOQ	<LOQ	-	-	-	<LOQ	-	30	0.2	<LOQ	-	<LOQ	0.1	5.2	<LOQ	-	-
Female 34 y	-	<LOQ	<LOQ	-	-	-	0.6	0.2	58	0.2	<LOQ	-	<LOQ	0.8	16	<LOQ	0.1	<LOQ
Female 40 y	-	-	<LOQ	-	-	-	0.7	-	56	0.7	<LOQ	-	<LOQ	1.1	12	<LOQ	-	-
Glucuronide fraction																		
Male 26 y	-	0.3	<LOQ	<LOQ	<LOQ	270[b]	0.2	<LOQ	7.8	11	0.7	0.2	0.8	150[b]	3.0	0.6	-	6.5
Male 34 y	<LOQ	1.2	<LOQ	0.2	2.8	1000[b]	0.5	1.1	48	34	1.7	0.8	3.2	1000[b]	26	1.4[a]	-	23
Male 38 y	-*	0.8	<LOQ	0.3	2.4	1400[b]	0.7	1.6	81	23	1.9	0.6	1.6	1900[b]	31	1.1[a]	-	28
Male 49 y	<LOQ	0.8	<LOQ	0.2	4.2	720[b]	0.3	1.7	90[b]	21	1.1	0.3	1.3	570[b]	33	0.7	-	10
Female 26 y	-	0.7	<LOQ	<LOQ	<LOQ	380[b]	0.2	1.1	27[b]	17	1.4	0.3	0.9	1200[b]	13	0.8[a]	-	1.3
Female 31 y	-	0.6	<LOQ	<LOQ	<LOQ	600[b]	0.6	0.2	10	13	4.2	1.4	3.6	610[b]	4.2	0.9[a]	-	1.9
Female 34 y	<LOQ	0.2	<LOQ	<LOQ	2.1	470[b]	0.2	1.0	42	21	3.0	0.9	3.5	1100[b]	19	0.6[a]	0.1	<LOQ
Female 40 y	<LOQ	0.7	<LOQ	<LOQ	3.6	1100[b]	0.5	0.9	28	37	8.6	3.2	7.6	560[b]	26	<LOQ	<LOQ	2.3

*=interfering background noise above LOQ

a=interfering background in qualifier, quantifier/qualifier ratio above allowed limit

b=from diluted samples with diluted calibration curves

4.4.3 CPI AS AN INTERFACE FOR LOW FLOW RATE LC-MS

The CPI-type ion source has been shown to be sensitive and suitable for volatile and semi-volatile compounds by GC-MS or DI-MS. However, use as an interface in LC-MS has not been shown previously. The feasibility of coupling low flow rate LC to MS with CPI interface was studied by analyzing the steroids DHEA, E2, PROG, and T (publication V). The advantage of CPI compared with common free-space APPI sources is that all eluent and analyte molecules are directed into the CPI device connected directly to the glass inlet capillary of the mass spectrometer. As the vaporization and ionization of analytes take place inside the heated CPI capillary, the ion transfer into the vacuum of the MS is improved.

The main operational parameters (CPI temperature and flow rate) of the CPI device were optimized using a UPLC with a splitter. The effect of CPI temperature to the signal intensity was tested between 225-350°C using an eluent MeOH with flow rate of 10 µL min^{-1} to the CPI. The signal intensity increased when the temperature was raised from 225°C to 250-275°C and decreased at higher temperatures (Figure 20A). At lower temperatures, the vaporization of eluent and desolvation of analyte ions are inefficient, resulting in low signal intensity. The signal decrease at higher CPI temperatures may be partly due to formation of smaller cluster ions with higher temperatures, which increases the mobility of cluster ion and the possibility for charge neutralization through collisions to CPI capillary walls. At higher CPI temperatures, precursor ions of analytes may also thermally dissociate, resulting in decreased sensitivity.

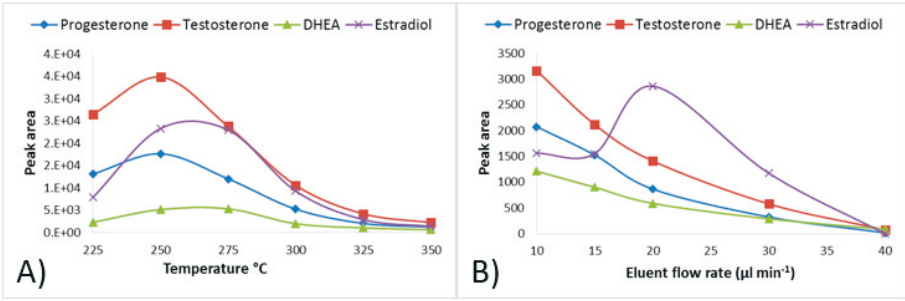

Figure 20 The effect of CPI capillary temperature (A) and flow rate (B) on detected signal from a standard solution of 100 ng mL^{-1}. Published by The Royal Society of Chemistry (publication V).

The effect of LC eluent flow rate on the signal intensity was studied at flow rates between 10-40 µL min^{-1} (CPI temperature 275°C). The optimal flow rate was 10 µL min^{-1} and signal of steroids decreased with higher flow rates (Figure 20B). However, estradiol behaved differentially as its signal increased at a flow

rate of 20 µL min^{-1} and then decreased similarly as the other steroids. Reduced signal intensity with higher flow rates is probably due to a combination of several factors. One possible explanation for the decreased signal at higher flow rates is the inefficient heat transfer that is unable to completely vaporize the eluent before the ionization chamber. As photoionization occurs in the gas phase, incomplete vaporization of eluent can decrease the ionization efficiency of the analytes. The higher flow rate also increases the solvent cluster ion size, which again increases the proton affinity of the cluster ion. It is possible that the protonation of the analytes becomes less favorable when solvent cluster ion size is increased. Additionally, the photoionization efficiency of the dopant might be decreased with higher flow rates, since the photons are absorbed by the eluent in the ion source.

The feasibility of UPLC-CPI-MS/MS for quantitative analysis of selected steroids was studied in terms of linearity, repeatability, and LOD with standard samples (Table 13). The flow rate in UPLC was 200 µL min^{-1}, the injection volume was 10 µL, and the flow was split after the column with a ratio of 1:20. As a result, the flow rate and the sample volume to the CPI were 10 µL min^{-1} and 0.5 µL, respectively. Linearity was studied with a sample concentration range of 5-400 ng mL^{-1} (2.5-200 pg to the CPI). Linearity was good with correlation coefficient (R) levels between 0.990-0.998. The repeatability of injection was determined at the concentration level 10 ng mL^{-1} (5 pg to the CPI). Repeatability was acceptable with RSD% values between 3-20% (n=6). The LODs (S/N >3) were between 0.5-2.5 pg injected to CPI (1-5 ng mL^{-1} in a sample), which indicates good sensitivity of the method.

Table 13 *LODs (amount injected to CPI), linearity, and repeatability of steroids using UPLC-CPI-MS/MS. Published by The Royal Society of Chemistry (publication V).*

Compound	LOD (pg)	Linearity (R)	Repeatability RSD% (10 ng mL^{-1}, n=6)
Estradiol	0.5	0.997	3.2
Testosterone	1.25	0.994	12.7
DHEA	2.5	0.998	20.1
Progesterone	2.5	0.990	13.2

The sensitivity of the UPLC-CPI-MS (flow rate to CPI 10 µL min^{-1}) method was also compared to the standard UPLC-APPI-MS (flow rate 200 µL min^{-1}) method by injecting 10 µL of standard sample including the four test steroids at a concentration of 10 ng mL^{-1}. As the splitter (1:20) was used in UPLC-CPI-MS, but not in UPLC-APPI-MS, the analyte amount eluting to the APPI source was 100 pg and 5 pg to the CPI source. The peak areas were 7.4-15.6 times higher with the standard UPLC-APPI-MS than with UPLC-CPI-APPI, although 20 times more intense signals could be expected with standard

APPI. This shows that the sensitivity with CPI is at least at the same level as with standard APPI.

The feasibility of CPI in coupling of capillary LC to MS was demonstrated by analyzing a standard sample (10 ng mL^{-1}) (Figure 21). The flow rate was 10 µL min^{-1} and the injection volume was 1 µL, corresponding 10 pg injected into the column. All four steroids were detected, showing that CPI is a potential method for coupling of low flow rate LC to MS and is especially suitable for the analysis of low-volume samples.

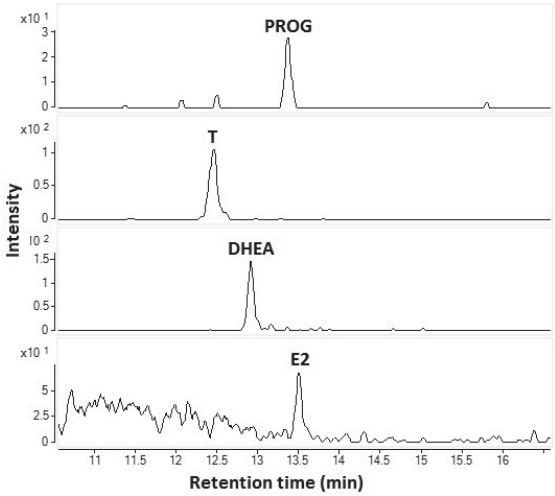

Figure 21 Selected reaction monitoring chromatograms of PROG, T, DHEA, and E2 from a standard mixture of 10 ng mL^{-1} using capillary LC-CPI-MS/MS. Published by The Royal Society of Chemistry (publication V).

5 SUMMARY AND CONCLUSIONS

The overall aim of this study was to apply mass spectrometry-based methods for metabolomics and to develop new analytical methods for metabolite analysis of biological samples.

A multiomics study of neonatal mouse hearts revealed several altered metabolites and metabolic pathways during the early postnatal period. For example, a transition from carbohydrates to lipids as the main energy source, an overall increase in energy metabolism, and increased oxidative stress was observed in the postnatal period based on changes in the metabolite levels. In combination with other omics, BCAA degradation, fatty acid metabolism, mevalonate pathway, and ketogenesis showed a potential role in the regulation of neonatal cardiomyocyte proliferation. Overall, this study provided an extensive resource for molecule abundances for future mechanistic studies. Integration of data from several omics experiments increased the reliability of the findings and provided comprehensive knowledge of neonatal metabolism. However, we noted that all omics had limited coverage of detected species and the datasets did not necessary support each other due to differences in turnover rates, which makes the integration of data a challenging task. The development of more complex mechanistic models and computational tools for multiomics will probably improve the possibilities to gain even further insights.[287]

A metabolomics study of serum from non-diabetic individuals revealed 19 serum metabolite clusters that were significantly associated with the IR phenotype, including in total 26 polar metabolites from five separate clusters and 367 lipids from 14 clusters. The IR phenotype was positively associated with polar metabolites and lipid clusters and negatively (IS phenotype) associated with lipid clusters. These clusters also showed significant associations with gut microbiota functional modules. The serum metabolome of insulin-resistant individuals showed increased levels of BCAAs, which correlated with a gut microbiome that had an enriched biosynthetic potential for BCAAs and was deprived of genes encoding bacterial inward transporters of these BCAAs. *P. copri* and *B. vulgatus* were identified as the main species driving the association between biosynthesis of BCAAs and IR. Mouse studies demonstrated that *P. copri* can induce IR, aggravate glucose intolerance, and increase circulating levels of BCAAs. Our findings showed that the human gut microbiota impacts the serum metabolome and contributes to IR, suggesting that microbial targets may have the potential to reduce IR and the occurrence of metabolic and cardiovascular disorders.

Microchip-MS is considered a potential technique for metabolomics research, as it allows a rapid, straightforward analysis from low sample volume and the possibility for a completely automated analytical system in the future. The feasibility of chip-MS for non-targeted metabolomics was studied and

compared to the commonly applied LC-MS for non-targeted metabolomics analysis of HFF and hiPSC cells. The comparison of chip-MS and LC-MS for global non-targeted metabolomics showed that both methods were capable of detecting metabolic differences between cell types and could distinguish the cells based on their metabolic profiles. The study demonstrated that chip-MS was a rapid and simple method that allowed high sample throughput from small sample volumes. Although selectivity and metabolite coverage of chip-MS was limited compared to LC-MS, chip-MS showed more universal metabolite coverage from polar metabolites to lipids. As a drawback, chip-MS seemed to suffer more from ion suppression.

All the applied non-targeted metabolomics platforms (publications I-III) were capable of detecting metabolic differences and showed acceptable analytical performance. Both LC-MS and GCxGC-MS are suitable for metabolic profiling, although GC separation is limited to low molecular weight compounds and adequate separation of a wide range of metabolites using LC-MS requires more than one column chemistry. HILIC serves as an alternative for GCxGC-MS for detection of polar metabolites without a requirement for time-consuming derivatization.[172] Thus, the simultaneous detection of polar metabolites and lipids with two parallel separations with C18 and HILIC shows promising results for replacing two individual platforms.[88]

Applying two complimentary analytical approaches (i.e. LC-MS and GCxGC-MS) (publication I and II) increased the metabolite coverage and expanded the polarity range of detected metabolites. A separate analysis of lipidomics and polar metabolites also enabled the detection of non-polar lipid classes such as TGs and ChoEs (publication II). Although two complimentary platforms for metabolomics were applied, the overall metabolite coverage was still quite low compared to the number of existing metabolites and some metabolites were not detectable with the methods employed due to low abundance or unsuitable methodologies. Analysis in negative ion mode in LC-MS and chip-MS methods would have increased the metabolite coverage to metabolite and lipid classes such as PA, PI, PG, PS, organic acids, and sugars, which are poorly detected in positive ion mode.[12,288] However, with the applied instruments, this would have doubled the analysis time and limited the sample throughput, which is highly important in metabolomics of large sample series. From the applied methods in this work, only chip-MS (publication III) would allow high throughput even with analysis of both polarities and the metabolite coverage could be feasibly expanded with chip-MS through several measurements without a significant increase in time. Polarity switching in instruments with such capability is becoming more feasible, with the ability to detect both polarities in one run. However, the eluent composition affects the ionization efficacies and maximal sensitivity is not obtained for both polarities simultaneously.[289]

Most of the detected features remained unknown due to limited spectral libraries and identification capabilities. The identification of metabolites is the bottleneck of non-targeted metabolomics,[227,236] as also observed in this study.

In general, EI-MS libraries are considered more applicable due to high repeatability of spectra and RIs available. However, ESI MS/MS libraries are expanding rapidly, quality of comparison spectra are improving, and spectra are obtained with multiple experimental conditions.[227,236] In our studies, EI-MS spectral identification capabilities were also limited. The detected features with GCxGC-MS and LC-MS remained as frequently unidentified. Although non-targeted metabolomics in principal offers the opportunity to discover novel compounds, the findings are quite often limited to already existing databases. Mechanism studies are also limited to already known metabolic pathways and the complexity of the metabolic networks are partly unrevealed. In particular, lipid-containing networks and functions of individual lipid species are widely unknown, which hinders biological interpretation.

In general, non-targeted methods are considered to be less reliable and repeatable than targeted methods.[17] The automated peak picking with different algorithms is more prone to errors when compared with targeted methods where the integrations are usually inspected manually, which is not feasible in non-targeted methods with even thousands of detected features. Several significant findings with LC-MS (publications I-III) were confirmed by also manually inspecting the peaks and alignment, and no false findings were observed. This shows that the applied algorithms work quite well. However, careful selection of peak picking parameters is necessary and caution is required if some drift in the signal intensity or mass accuracy is observed. The signal drift in targeted methods is usually corrected with labeled ISTDs, whereas in non-targeted methods only a few or no ISTDs are used to correct the signal. In lipidomics (publication II), the repeatability was improved with normalization of lipid group-specific ISTDs and good repeatability was obtained for identified lipids in the control serum sample. Even better repeatability could be obtained with an approach using metabolites extracted from ^{13}C-labeled yeast acting as labeled ISTDs for multiple metabolites or lipids simultaneously. Labeled ISTDs can be used to normalize or quantitate the metabolites for better accuracy and wider linear dynamic range.[290–292]

All non-targeted methods, including the methods presented in this thesis, are a compromise between wide metabolite coverage, high throughput, simplicity of operation, good repeatability, and accuracy. Overall, our methods were suitable for non-targeted metabolomics and showed similar and expected analytical performance and throughput as reported.[12,60,63,167,293]

Since non-targeted metabolomics are incapable of detecting biologically important low-abundance metabolites, targeted methods are also required for certain analytes to improve sensitivity. A developed CPI source was applied for the analysis of steroids with GC-CPI-MS/MS from urine samples. The developed ionization chamber showed high ionization efficacy and ion transmission to non-polar steroids. The TMS-derivatized steroids formed mainly abundant molecular ions with minimal fragmentation and formation of other ions. The developed GC-CPI-MS/MS method showed good chromatographic resolution, acceptable linearity and repeatability, and low

LODs. The feasibility of the ion source for the quantitative analysis of biological samples was studied by analyzing 18 endogenous steroids in urine. In total, 15 steroids were quantified either as a free steroid or glucuronide conjugate from human urine samples and the detected steroid levels were consistent with the literature.[279,282,284–286] Overall, the CPI ion source showed potential for quantitative analysis of non-polar low abundance metabolites. The CPI interface would also allow combining GC with hyphenated high-resolution API mass spectrometers, which introduces the possibility to use GC for non-targeted metabolomics analysis to discover elemental compositions of unknowns. Since there is some evidence that the fragmentation of formed molecular ions follows the same odd-electron fragmentation pathways as in EI,[142] it might still be possible to use EI spectral libraries for identification of MS/MS spectra.

The CPI interface was also combined for the first time with LC using low flow rates. The feasibility of LC-CPI-MS/MS for the quantitative analysis of four steroids was studied and the method showed good quantitative performance and high sensitivity. The LC-CPI-MS/MS was applicable with low flow rates, which are not optimal for existing commercial atmospheric pressure ion source geometries. This offers an alternative for nanoESI, which is exclusively used in nano- or capillary-LC applications. In conclusion, CPI ion source showed good potential in combination with both LC and GC and allowed the usage of the same API mass spectrometer with both separation techniques. With CPI, high sensitivity to non-polar compounds can be obtained and the source is applicable for quantitative analysis of biological samples.

6 REFERENCES

(1) Patti, G. J.; Yanes, O.; Siuzdak, G. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263–269.
(2) Wishart, D. S. *Nat. Rev. Drug Discov.* **2016**, *15*, 473–484.
(3) Johnson, C. H.; Ivanisevic, J.; Siuzdak, G. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 451–459.
(4) Luan, H.; Wang, X.; Cai, Z. *Mass Spectrom. Rev.* **2019**, *38*, 22–33.
(5) Kostic, A. D.; Gevers, D.; Siljander, H.; Vatanen, T.; Hyötyläinen, T.; Hämäläinen, A.-M.; Peet, A.; Tillmann, V.; Pöhö, P.; Mattila, I.; et al. *Cell Host Microbe* **2015**, *17*, 260–273.
(6) Bruggeman, F. J.; Westerhoff, H. V. *Trends Microbiol.* **2007**, *15*, 45–50.
(7) Nicholson, J. K.; Lindon, J. C. *Nature* **2008**, *455*, 1054–1056.
(8) Fiehn, O. *Plant Mol. Biol.* **2002**, *48*, 155–171.
(9) Mann, M.; Jensen, O. N. *Nat. Biotechnol.* **2003**, *21*, 255–261.
(10) Greef, J. Van Der; Stroobant, P.; Heijden, R. Van Der. *Curr. Opin. Chem. Biol.* **2004**, *8*, 559–565.
(11) Dunn, W. B.; Broadhurst, D. I.; Atherton, H. J.; Goodacre, R.; Griffin, J. L. *Chem. Soc. Rev.* **2011**, *40*, 387–426.
(12) Cajka, T.; Fiehn, O. *TrAC, Trends Anal. Chem.* **2014**, *61*, 192–206.
(13) Han, X.; Yang, K.; Gross, R. W. *Mass Spectrom. Rev.* **2012**, *31*, 134–178.
(14) Jeanneret, F.; Tonoli, D.; Rossier, M. F.; Saugy, M.; Boccard, J.; Rudaz, S. *J. Chromatogr. A* **2015**, *1430*, 97–112.
(15) Aizpurua-Olaizola, O.; Toraño, J. S.; Falcon-Perez, J. M.; Williams, C.; Reichardt, N.; Boons, G.-J. *TrAC, Trends Anal. Chem.* **2018**, 7–14.
(16) Kell, D. B.; Oliver, S. G. *BioEssays* **2004**, *26*, 99–105.
(17) Cajka, T.; Fiehn, O. *Anal. Chem.* **2016**, *88*, 524–545.
(18) Dunn, W. B.; Wilson, I. D.; Nicholls, A. W.; Broadhurst, D. *Bioanalysis* **2012**, *4*, 2249–2264.
(19) Zhou, J.; Yin, Y. *Analyst* **2016**, *141*, 6362–6373.
(20) Jäntti, S. E.; Kivilompolo, M.; Öhrnberg, L.; Pietiläinen, K. H.; Nygren, H.; Orešič, M.; Hyötyläinen, T. *Anal. Bioanal. Chem.* **2014**, *406*, 7799–7815.
(21) Kivilompolo, M.; Öhrnberg, L.; Orešic, M.; Hyötyläinen, T. *J. Chromatogr. A* **2013**, *1292*, 189–194.
(22) Abrankó, L.; Williamson, G.; Gardner, S.; Kerimi, A. *J. Chromatogr. A* **2018**, *1534*, 111–122.
(23) Casado, M.; Sierra, C.; Batllori, M.; Artuch, R.; Ormazabal, A. *Metabolomics* **2018**, *14:* 76.
(24) Jäntti, S. E.; Hartonen, M.; Hilvo, M.; Nygren, H.; Hyötyläinen, T.; Ketola, R. A.; Kostiainen, R. *Anal. Chim. Acta.* **2013**, *802*, 56–66.
(25) Che, N.; Ma, Y.; Ruan, H.; Xu, L.; Wang, X.; Yang, X.; Liu, X. *Clin. Chim. Acta.* **2018**, *477*, 81–88.
(26) Jorge, T. F.; Rodrigues, J. A.; Caldana, C.; Schmidt, R.; van Dongen, J. T.; Thomas-Oates, J.; António, C. *Mass Spectrom. Rev.* **2016**, *35*, 620–

649.
(27) Lankadurai, B. P.; Nagato, E. G.; Simpson, M. J. *Environ. Rev.* **2013**, *21*, 180–205.
(28) Herrero, M.; Simó, C.; García-Cañas, V.; Ibáñez, E.; Cifuentes, A. *Mass Spectrom. Rev.* **2012**, *31*, 49–69.
(29) Pinu, F. R.; Villas-Boas, S. G. *Metabolites* **2017**, *7,* 43.
(30) Shao, Y.; Le, W. *Mol. Neurodegener.* **2019**, *14*, 3.
(31) Wilkins, J. M.; Trushina, E. *Front. Neurol.* **2018**, *8*.
(32) Pallares-Méndez, R.; Aguilar-Salinas, C. A.; Cruz-Bautista, I.; Bosque-Plata, L. Del. *Ann. Med.* **2016**, *48*, 89–102.
(33) Wood, P. L. *Neuropsychopharmacology* **2014**, *39*, 24–33.
(34) Pedrini, M.; Cao, B.; Nani, J. V. S.; Cerqueira, R. O.; Mansur, R. B.; Tasic, L.; Hayashi, M. A. F.; McIntyre, R. S.; Brietzke, E. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2019**, *93*, 182–188.
(35) Armitage, E. G.; Barbas, C. *J. Pharm. Biomed. Anal.* **2014**, *87*, 1–11.
(36) Boccio, P. Del; Rossi, C.; di Ioia, M.; Cicalini, I.; Sacchetta, P.; Pieragostino, D. *Proteomics - Clin. Appl.* **2016**, *10*, 470–484.
(37) Ussher, J. R.; Elmariah, S.; Gerszten, R. E.; Dyck, J. R. B. *J. Am. Coll. Cardiol.* **2016**, *68*, 2850–2870.
(38) Kamleh, M. A.; Snowden, S. G.; Grapov, D.; Blackburn, G. J.; Watson, D. G.; Xu, N.; Ståhle, M.; Wheelock, C. E. *J. Proteome Res.* **2015**, *14*, 557–566.
(39) Banoei, M. M.; Casault, C.; Metwaly, S. M.; Winston, B. W. *J. Neurotrauma* **2018**, *35*, 1831–1848.
(40) Lee, Y.; Khan, A.; Hong, S.; Jee, S. H.; Park, Y. H. *Mol. Biosyst.* **2017**, *13*, 1109–1120.
(41) Kurko, J.; Tringham, M.; Tanner, L.; Näntö-Salonen, K.; Vähä-Mäkilä, M.; Nygren, H.; Pöhö, P.; Lietzen, N.; Mattila, I.; Olkku, A.; et al. *Metabolism.* **2016**, *65*, 1361–1375.
(42) Miller, M. J.; Kennedy, A. D.; Eckhart, A. D.; Burrage, L. C.; Wulff, J. E.; Miller, L. A. D.; Milburn, M. V; Ryals, J. A.; Beaudet, A. L.; Sun, Q.; et al. *J. Inherit. Metab. Dis.* **2015**, *38*, 1029–1039.
(43) Tebani, A.; Abily-Donval, L.; Afonso, C.; Marret, S.; Bekri, S. *Int. J. Mol. Sci.* **2016**, *17,* 1167.
(44) Kohler, I.; Hankemeier, T.; van der Graaf, P. H.; Knibbe, C. A. J.; van Hasselt, J. G. C. *Eur. J. Pharm. Sci.* **2017**, *109*, 15-21.
(45) Crutchfield, C. A.; Thomas, S. N.; Sokoll, L. J.; Chan, D. W. *Clin. Proteomics* **2016**, *13*, 1–13.
(46) Everett, J. R. *Front. Pharmacol.* **2016**, *7,* 297.
(47) Armitage, E. G.; Southam, A. D. *Metabolomics* **2016**, *12,* 146.
(48) Elbadawi-Sidhu, M.; Baillie, R. A.; Zhu, H.; Chen, Y.-D. I.; Goodarzi, M. O.; Rotter, J. I.; Krauss, R. M.; Fiehn, O.; Kaddurah-Daouk, R. *Metabolomics* **2017**, *13,* 11.
(49) Vernocchi, P.; Del Chierico, F.; Putignani, L. *Front. Microbiol.* **2016**, *7*, 1144.
(50) He, X.; Ji, G.; Jia, W.; Li, H. *Int. J. Mol. Sci.* **2016**, *17*.
(51) Heux, S.; Bergès, C.; Millard, P.; Portais, J.-C. C.; Létisse, F. *Curr. Opin. Biotechnol.* **2017**, *43*, 104–109.
(52) Balcells, C.; Foguet, C.; Tarragó-Celada, J.; de Atauri, P.; Marin, S.; Cascante, M. *TrAC,Trends Anal. Chem.* **2019**, *In press*.

(53) Huang, S. *Development* **2009**, *136*, 3853–3862.
(54) Ali, A.; Abouleila, Y.; Shimizu, Y.; Hiyama, E.; Emara, S.; Mashaghi, A.; Hankemeier, T. *TrAC,Trends Anal. Chem.* **2019**, *In press*.
(55) Duncan, K. D.; Fyrestam, J.; Lanekoff, I. *Analyst* **2019**, *144*, 782–793.
(56) Greer, T.; Sturm, R.; Li, L. *J. Proteomics* **2011**, *74*, 2617–2631.
(57) Larive, C. K.; Barding, G. A.; Dinges, M. M. *Anal. Chem.* **2015**, *87*, 133–146.
(58) Hounoum, B. M.; Blasco, H.; Emond, P.; Mavel, S. *TrAC,Trends Anal. Chem.* **2016**, *75*, 118–128.
(59) Naz, S.; Moreira Dos Santos, D. C.; García, A.; Barbas, C. *Bioanalysis* **2014**, *6*, 1657–1677.
(60) Beale, D. J.; Pinu, F. R.; Kouremenos, K. A.; Poojary, M. M.; Narayana, V. K.; Boughton, B. A.; Kanojia, K.; Dayalan, S.; Jones, O. A. H.; Dias, D. A. *Metabolomics* **2018**, *14*, 152.
(61) Ramautar, R.; Somsen, G. W.; de Jong, G. J. *Electrophoresis* **2019**, *40*, 165–179.
(62) Townsend, M. K.; Bao, Y.; Poole, E. M.; Bertrand, K. A.; Kraft, P.; Wolpin, B. M.; Clish, C. B.; Tworoger, S. S. *Cancer Epidemiol. Biomarkers Prev.* **2016**, *25*, 823–829.
(63) Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-Mcintyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; et al. *Nat. Protoc.* **2011**, *6*, 1060–1083.
(64) Khamis, M. M.; Adamko, D. J.; El-Aneed, A. *Mass Spectrom. Rev.* **2017**, *36*, 115–134.
(65) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2010**, *5*, 1005–1018.
(66) Hartonen, M.; Mattila, I.; Ruskeepää, A.-L.; Orešič, M.; Hyötyläinen, T. *J. Chromatogr. A* **2013**, *1293*, 142–149.
(67) Kaczor-Urbanowicz, K. E.; Carreras-Presas, C. M.; Kaczor, T.; Tu, M.; Wei, F.; Garcia-Godoy, F.; Wong, D. T. W. *J. Cell. Mol. Med.* **2017**, *21*, 640–647.
(68) Karu, N.; Deng, L.; Slae, M.; Guo, A. C.; Sajed, T.; Huynh, H.; Wine, E.; Wishart, D. S. *Anal. Chim. Acta.* **2018**, *1030*, 1–24.
(69) Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2013**, *8*, 17–32.
(70) Mena-Bravo, A.; Luque de Castro, M. D. *J. Pharm. Biomed. Anal.* **2014**, *90*, 139–147.
(71) Nunes De Paiva, M. J.; Menezes, H. C.; Cardeal, Z. D. L. *Analyst* **2014**, *139*, 3683–3694.
(72) Paglia, G.; Greco, F. M. Del; Sigurdsson, B. B.; Rainer, J.; Volani, C.; Hicks, A. A.; Pramstaller, P. P.; Smarason, S. V. *Clin. Chim. Acta* **2018**, *486*, 320–328.
(73) Jørgenrud, B.; Jäntti, S.; Mattila, I.; Pöhö, P.; Rønningen, K. S.; Yki-Järvinen, H.; Orešič, M.; Hyötyläinen, T. *Bioanalysis* **2015**, *7*, 991–1006.
(74) Dettmer, K.; Nürnberger, N.; Kaspar, H.; Gruber, M. A.; Almstetter, M. F.; Oefner, P. J. *Anal. Bioanal. Chem.* **2011**, *399*, 1127–1139.
(75) Raterink, R.-J.; Lindenburg, P. W.; Vreeken, R. J.; Ramautar, R.;

Hankemeier, T. *TrAC,Trends Anal. Chem.* **2014**, *61*, 157–167.
(76) Sellick, C. A.; Hansen, R.; Stephens, G. M.; Goodacre, R.; Dickson, A. J. *Nat. Protoc.* **2011**, *6*, 1241–1249.
(77) Vuckovic, D. *Anal. Bioanal. Chem.* **2012**, *403*, 1523–1548.
(78) Bruce, S. J.; Tavazzi, I.; Parisod, V.; Rezzi, S.; Kochhar, S.; Guy, P. A. *Anal. Chem.* **2009**, *81*, 3285–3296.
(79) Halket, J. M.; Zaikin, V. G. *Eur. J. Mass Spectrom.* **2003**, *9*, 1–21.
(80) Athanasiadou, I.; Angelis, Y. S.; Lyris, E.; Georgakopoulos, C.; Athanasiadou, I.; Georgakopoulos, C. *TrAC,Trends Anal. Chem.* **2013**, *42*, 137–156.
(81) Jiang, R.; Jiao, Y.; Xu, F. *Bioanalysis* **2016**, *8*, 1881–1883.
(82) Lorenz, M. A.; Burant, C. F.; Kennedy, R. T. *Anal. Chem.* **2011**, *83*, 3406–3414.
(83) Michopoulos, F.; Lai, L.; Gika, H.; Theodoridis, G.; Wilson, I. *J. Proteome Res.* **2009**, *8*, 2114–2121.
(84) Sarafian, M. H.; Gaudin, M.; Lewis, M. R.; Martin, F.-P.; Holmes, E.; Nicholson, J. K.; Dumas, M.-E. *Anal. Chem.* **2014**, *86*, 5766–5774.
(85) Chan, W.; Zhao, Y.; Zhang, J. *Rapid Commun. Mass Spectrom.* **2019**, *33*, 561–568.
(86) Snytnikova, O. A.; Khlichkina, A. A.; Sagdeev, R. Z.; Tsentalovich, Y. P. *Metabolomics* **2019**, *15*, 84.
(87) Michopoulos, F.; Edge, A. M.; Theodoridis, G.; Wilson, I. D. *J. Sep. Sci.* **2010**, *33*, 1472–1479.
(88) Schwaiger, M.; Schoeny, H.; Abiead, Y. El; Hermann, G.; Rampler, E.; Koellensperger, G. *Analyst* **2019**, *144*, 220–229.
(89) Folch, J.; Lees, M.; Stanley, G. H. S. *J. Biol. Chem.* **1957**, *226*, 497–509.
(90) Bligh, E. G.; Dyer, W. J. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917.
(91) Matyash, V.; Liebisch, G.; Kurzchalia, T. V; Shevchenko, A.; Schwudke, D. *J. Lipid Res.* **2008**, *49*, 1137–1146.
(92) Ulmer, C. Z.; Jones, C. M.; Yost, R. A.; Garrett, T. J.; Bowden, J. A. *Anal. Chim. Acta.* **2018**, *1037*, 351–357.
(93) Sostare, J.; Guida, R. Di; Kirwan, J.; Chalal, K.; Palmer, E.; Dunn, W. B.; Viant, M. R. *Anal. Chim. Acta.* **2018**, *1037*, 301–315.
(94) Pellegrino, R. M.; Veroli, A. Di; Valeri, A.; Goracci, L.; Cruciani, G. *Anal. Bioanal. Chem.* **2014**, *406*, 7937–7948.
(95) Cequier-Sánchez, E.; Rodríguez, C.; Ravelo, Á. G.; Zárate, R. *J. Agric. Food Chem.* **2008**, *56*, 4297–4303.
(96) Löfgren, L.; Ståhlman, M.; Forsberg, G.-B.; Saarinen, S.; Nilsson, R.; Hansson, G. I. *J. Lipid Res.* **2012**, *53*, 1690–1700.
(97) Rombouts, C.; De Spiegeleer, M.; Van Meulebroek, L.; De Vos, W. H.; Vanhaecke, L. *Anal. Chim. Acta.* **2019**, *1066*, 79–92.
(98) Chen, S.; Hoene, M.; Li, J.; Li, Y.; Zhao, X.; Häring, H.-U.; Schleicher, E. D.; Weigert, C.; Xu, G.; Lehmann, R. *J. Chromatogr. A* **2013**, *1298*, 9–16.
(99) Sitnikov, D. G.; Monnin, C. S.; Vuckovic, D. *Sci. Rep.* **2016**, *6*, 38885.
(100) Zhang, X.; Romm, M.; Zheng, X.; Zink, E. M.; Kim, Y.-M.-M.; Burnum-Johnson, K. E.; Orton, D. J.; Apffel, A.; Ibrahim, Y. M.; Monroe, M. E.; et al. *Clin. Mass Spectrom.* **2016**, *2*, 1–10.
(101) Michopoulos, F.; Gika, H.; Palachanis, D.; Theodoridis, G.; Wilson, I. D. *Electrophoresis* **2015**, *36*, 2170–2178.

(102) Chetwynd, A. J.; Abdul-Sada, A.; Hill, E. M. *Anal. Chem.* **2015**, *87*, 1158–1165.
(103) Fernández-Peralbo, M. A.; de Castro, M. D. *TrAC, Trends Anal. Chem.* **2012**, *41*, 75–85.
(104) Whiley, L.; Nye, L. C.; Grant, I.; Andreas, N.; Chappell, K. E.; Sarafian, M. H.; Misra, R.; Plumb, R. S.; Lewis, M. R.; Nicholson, J. K.; et al. *Anal. Chem.* **2019**, *91*, 5207–5216.
(105) Zhang, Z.; Qin, L.; Guo, M.; Gao, S.; Zhang, Q.; Wang, Q.; Lu, Z.; Zhao, H.; Liu, Y.; Wang, M.; et al. *J. Sep. Sci.* **2016**, *39*, 2616–2625.
(106) Han, J.; Liu, Y.; Wang, R.; Yang, J.; Ling, V.; Borchers, C. H. *Anal. Chem.* **2015**, *87*, 1127–1136.
(107) López-Bascón, M. A.; Calderón-Santiago, M.; Sánchez-Ceinos, J.; Fernández-Vega, A.; Guzmán-Ruiz, R.; López-Miranda, J.; Malagon, M. M.; Priego-Capote, F. *Talanta* **2018**, *177*, 86–93.
(108) Chen, J.-J.; Zhou, C.-J.; Liu, Z.; Fu, Y.-Y.; Zheng, P.; Yang, D.-Y.; Li, Q.; Mu, J.; Wei, Y.-D.; Zhou, J.-J.; et al. *J. Proteome Res.* **2015**, *14*, 3382–3389.
(109) Yang, Y.; Cruickshank, C.; Armstrong, M.; Mahaffey, S.; Reisdorph, R.; Reisdorph, N. *J. Chromatogr. A* **2013**, *1300*, 217–226.
(110) Skov, K.; Hadrup, N.; Smedsgaard, J.; Frandsen, H. *J. Chromatogr. B* **2015**, *978–979*, 83–88.
(111) Li, Y.; Zhang, Z.; Liu, X.; Li, A.; Hou, Z.; Wang, Y.; Zhang, Y. *J. Chromatogr. A* **2015**, *1409*, 277–281.
(112) Liu, C.; Gu, C.; Huang, W.; Sheng, X.; Du, J.; Li, Y. *J. Chromatogr. B* **2019**, *1113*, 98–106.
(113) Yuan, Z.-X.; Majchrzak-Hong, S.; Keyes, G. S.; Iadarola, M. J.; Mannes, A. J.; Ramsden, C. E. *Anal. Bioanal. Chem.* **2018**, *410*, 6009–6029.
(114) Shin, H. J.; Park, N. H.; Lee, W.; Choi, M. H.; Chung, B. C.; Hong, J. *J. Chromatogr. B* **2017**, *1051*, 97–107.
(115) Piri-Moghadam, H.; Alam, M. N.; Pawliszyn, J. *Anal. Chim. Acta*. 2017, *984*, 42–65.
(116) Bojko, B.; Reyes-Garcés, N.; Bessonneau, V.; Goryński, K.; Mousavi, F.; Souza Silva, E. A.; Pawliszyn, J. *TrAC, Trends Anal. Chem.* **2014**, *61*, 168–180.
(117) Capuano, R.; Spitalieri, P.; Talarico, R. V; Catini, A.; Domakoski, A. C.; Martinelli, E.; Scioli, M. G.; Orlandi, A.; Cicconi, R.; Paolesse, R.; et al. *Sci. Rep.* **2018**, *8*, 11056.
(118) Bos, L. D. J.; Sterk, P. J.; Schultz, M. J. *PLoS Pathog.* **2013**, *9*.
(119) Mousavi, F.; Bojko, B.; Pawliszyn, J. *Anal. Chim. Acta.* **2015**, *892*, 95–104.
(120) Pannkuk, E. L.; Laiakis, E. C.; Authier, S.; Wong, K.; Fornace, A. J. *J. Proteome Res.* **2017**, *16*, 2091–2100.
(121) Ruiz-Matute, A. I.; Hernández-Hernández, O.; Rodríguez-Sánchez, S.; Sanz, M. L.; Martínez-Castro, I. *J. Chromatogr. B* **2011**, *879*, 1226–1240.
(122) He, Z.; Wang, M.; Li, H.; Wen, C. *Sci. Rep.* **2019**, *9*, 3872.
(123) Kouremenos, K. A.; Harynuk, J. J.; Winniford, W. L.; Morrison, P. D.; Marriott, P. J. *J. Chromatogr. B* **2010**, *878*, 1761–1770.
(124) Marsol-Vall, A.; Balcells, M.; Eras, J.; Canela-Garayoa, R. *Food Chem.* **2016**, *204*, 210–217.

(125) Shao, W.-H.; Fan, S.-H.; Lei, Y.; Yao, G.-E.; Chen, J.-J.; Zhou, J.; Xu, H.-B.; Liu, H.-P.; Wu, B.; Zheng, P.; et al. *Metabolomics* **2013**, *9*, 433–443.
(126) Qi, B.-L.; Liu, P.; Wang, Q.-Y.; Cai, W.-J.; Yuan, B.-F.; Feng, Y.-Q. *TrAC, Trends Anal. Chem.* **2014**, *59*, 121–132.
(127) Iwasaki, Y.; Nakano, Y.; Mochizuki, K.; Nomoto, M.; Takahashi, Y.; Ito, R.; Saito, K.; Nakazawa, H. *J. Chromatogr. B* **2011**, *879*, 1159–1165.
(128) Gomez-Gomez, A.; Soldevila, A.; Pizarro, N.; Andreu-Fernandez, V.; Pozo, O. J. *J. Pharm. Biomed. Anal.* **2019**, *164*, 382–394.
(129) Sakaguchi, Y.; Kinumi, T.; Yamazaki, T.; Takatsu, A. *Analyst* **2015**, *140*, 1965–1973.
(130) Zheng, X.; Kang, A.; Dai, C.; Liang, Y.; Xie, T.; Xie, L.; Peng, Y.; Wang, G.; Hao, H. *Anal. Chem.* **2012**, *84*, 10044–10051.
(131) Wei, F.; Wang, X.; Ma, H.-F.; Lv, X.; Dong, X.-Y.; Chen, H. *Anal. Chim. Acta.* **2018**, *1024*, 101–111.
(132) Häkkinen, M. R.; Murtola, T.; Voutilainen, R.; Poutanen, M.; Linnanen, T.; Koskivuori, J.; Lakka, T.; Jääskeläinen, J.; Auriola, S. *J. Pharm. Biomed. Anal.* **2019**, *164*, 642–652.
(133) Kauppila, T. J.; Syage, J. A.; Benter, T. *Mass Spectrom. Rev.* **2017**, *36*, 423–449.
(134) Kebarle, P.; Verkcerk, U. H. *Mass Spectrom. Rev.* **2009**, *28*, 898–917.
(135) Draper, J.; Lloyd, A. J.; Goodacre, R.; Beckmann, M. *Metabolomics* **2013**, *9*, 4–29.
(136) Lísa, M.; Cífková, E.; Holcapek, M. *J. Chromatogr. A* **2011**, *1218*, 5146–5156.
(137) Koren, L.; Ng, E. S. M.; Soma, K. K.; Wynne-Edwards, K. E. *PLoS One* **2012**, *7*, e32496.
(138) Cai, S.-S.; Hanold, K. A.; Syage, J. A. *Anal. Chem.* **2007**, *79*, 2491–2498.
(139) González-Domínguez, R.; García-Barrera, T.; Gómez-Ariza, J. L. *Talanta* **2015**, *131*, 480–489.
(140) Imbert, L.; Gaudin, M.; Libong, D.; Touboul, D.; Abreu, S.; Loiseau, P. M.; Laprévote, O.; Chaminade, P. *J. Chromatogr. A* **2012**, *1242*, 75–83.
(141) Keski-Rahkonen, P.; Huhtinen, K.; Desai, R.; Harwood, D. T.; Handelsman, D. J.; Poutanen, M.; Auriola, S. *J. Mass Spectrom.* **2013**, *48*, 1050–1058.
(142) Suominen, T.; Haapala, M.; Takala, A.; Ketola, R. A.; Kostiainen, R. *Anal. Chim. Acta.* **2013**, *794*, 76–81.
(143) Armitage, E. G.; Kotze, H. L.; Lockyer, N. P. *Metabolomics* **2013**, *9*, 102–109.
(144) Cooks, R. G.; Ouyang, Z.; Takats, Z.; Wiseman, J. M. *Science* **2006**, *311*, 1566–1570.
(145) Sturtevant, D.; Lee, Y.-J.; Chapman, K. D. *Curr. Opin. Biotechnol.* **2016**, *37*, 53–60.
(146) Wu, C.; Dill, A. L.; Eberlin, L. S.; Cooks, R. G.; Ifa, D. R. *Mass Spectrom. Rev.* **2013**, *32*, 218–243.
(147) Chu, D. B.; Troyer, C.; Mairinger, T.; Ortmayr, K.; Neubauer, S.; Koellensperger, G.; Hann, S. *Anal. Bioanal. Chem.* **2015**, *407*, 2865–2875.
(148) Ibrahim, W.; Wilde, M.; Cordell, R.; Salman, D.; Ruszkiewicz, D.; Bryant, L.; Richardson, M.; Free, R. C.; Zhao, B.; Yousuf, A.; et al. *BMJ Open* **2019**, *9*, e025486.

(149) Li, D.-X.; Gan, L.; Bronja, A.; Schmitz, O. J. *Anal. Chim. Acta*. **2015**, *891*, 43–61.
(150) Hintikka, L.; Haapala, M.; Kuuranne, T.; Leinonen, A.; Kostiainen, R. *J. Chromatogr. A* **2013**, *1312*, 111–117.
(151) Cha, E.; Jeong, E. S.; Cha, S.; Lee, J. *Anal. Chim. Acta*. **2017**, *964*, 123–133.
(152) Covey, T. R.; Thomson, B. A.; Schneider, B. B. *Mass Spectrom. Rev*. **2009**, *28*, 870–897.
(153) Mirabelli, M. F.; Zenobi, R. *Anal. Chem*. **2018**, *90*, 5015–5022.
(154) Haapala, M.; Suominen, T.; Kostiainen, R. *Anal. Chem*. **2013**, *85*, 5715–5719.
(155) Kersten, H.; Derpmann, V.; Barnes, I.; Brockmann, K. J.; O'Brien, R.; Benter, T. *J. Am. Soc. Mass Spectrom*. **2011**, *22*, 2070–2081.
(156) Klein, D. R.; Brodbelt, J. S. *Anal. Chem*. **2017**, *89*, 1516–1522.
(157) Ryan, E.; Nguyen, C. Q. N.; Shiea, C.; Reid, G. E. *J. Am. Soc. Mass Spectrom*. **2017**, *28*, 1406–1419.
(158) Fenaille, F.; Saint-Hilaire, P. B.; Rousseau, K.; Junot, C. *J. Chromatogr. A* **2017**, *1526*, 1–12.
(159) Zhou, J.; Liu, C.; Si, D.; Jia, B.; Zhong, L.; Yin, Y. *Anal. Chim. Acta* **2017**, *972*, 62–72.
(160) Wang, R.; Yin, Y.; Zhu, Z.-J. *Anal. Bioanal. Chem*. **2019**.
(161) Koelmel, J. P.; Kroeger, N. M.; Gill, E. L.; Ulmer, C. Z.; Bowden, J. A.; Patterson, R. E.; Yost, R. A.; Garrett, T. J. *J. Am. Soc. Mass Spectrom*. **2017**, *28*, 908–917.
(162) Bonner, R.; Hopfgartner, G. *TrAC, Trends Anal. Chem*. **2018**.
(163) Xu, J.; Zhai, Y.; Feng, L.; Xie, T.; Yao, W.; Shan, J.; Zhang, L. *J. Pharm. Biomed. Anal*. **2019**, *171*, 171–179.
(164) Abushareeda, W.; Lyris, E.; Kraiem, S.; Wahaibi, A. A.; Alyazidi, S.; Dbes, N.; Lommen, A.; Nielen, M.; Horvatovich, P. L.; Alsayrafi, M.; et al. *J. Chromatogr. B* **2017**, *1063*, 74–83.
(165) Misra, B. B.; Bassey, E.; Bishop, A. C.; Kusel, D. T.; Cox, L. A.; Olivier, M. *Rapid Commun. Mass Spectrom*. **2018**, *32*, 1497–1506.
(166) Winnike, J. H.; Wei, X.; Knagge, K. J.; Colman, S. D.; Gregory, S. G.; Zhang, X. *J. Proteome Res*. **2015**, *14*, 1810–1817.
(167) Gika, H.; Virgiliou, C.; Theodoridis, G.; Plumb, R. S.; Wilson, I. D. *J. Chromatogr. B* **2019**, *1117*, 136–147.
(168) Kaufmann, A. *TrAC, Trends Anal. Chem*. **2014**, *63*, 113–128.
(169) Karas, M.; Bahr, U.; Dülcks, T. *Fresenius. J. Anal. Chem*. **2000**, *366*, 669–676.
(170) Ciborowski, M.; Kisluk, J.; Pietrowska, K.; Samczuk, P.; Parfieniuk, E.; Kowalczyk, T.; Kozlowski, M.; Kretowski, A.; Niklinski, J. *Electrophoresis* **2017**, *38*, 2304–2312.
(171) Pham, T. H.; Zaeem, M.; Fillier, T. A.; Nadeem, M.; Vidal, N. P.; Manful, C.; Cheema, S.; Cheema, M.; Thomas, R. H. *Sci. Rep*. **2019**, *9*, 5048.
(172) Tang, D.-Q.; Zou, L.; Yin, X.-X.; Ong, C. N. *Mass Spectrom. Rev*. **2016**, *35*, 574–600.
(173) Nováková, L.; Havlíková, L.; Vlcková, H. *TrAC, Trends Anal. Chem*. **2014**, *63*, 55–64.
(174) Josephs, R. D.; Daireaux, A.; Choteau, T.; Westwood, S.; Wielgosz, R. I. *Anal. Bioanal. Chem*. **2015**, *407*, 3147–3157.

(175) Grant, L. K.; Ftouni, S.; Nijagal, B.; Souza, D. P. De; Tull, D.; McConville, M. J.; Rajaratnam, S. M. W.; Lockley, S. W.; Anderson, C. *Sci. Rep.* **2019**, *9*, 4428.
(176) Chen, J.; Hou, W.; Han, B.; Liu, G.; Gong, J.; Li, Y.; Zhong, D.; Liao, Q.; Xie, Z. *Anal. Bioanal. Chem.* **2016**, *408*, 2527–2542.
(177) Wang, J.; Christison, T. T.; Misuno, K.; Lopez, L.; Huhmer, A. F.; Huang, Y.; Hu, S. *Anal. Chem.* **2014**, *86*, 5116–5124.
(178) Michopoulos, F.; Whalley, N.; Theodoridis, G.; Wilson, I. D.; Dunkley, T. P. J.; Critchlow, S. E. *J. Chromatogr. A* **2014**, *1349*, 60–68.
(179) Anesi, A.; Guella, G. *J. Chromatogr. A* **2015**, *1384*, 44–52.
(180) Navarro-Reig, M.; Jaumot, J.; Tauler, R. *J. Chromatogr. A* **2018**, *1568*, 80–90.
(181) Lv, W.; Shi, X.; Wang, S.; Xu, G. *TrAC, Trends Anal. Chem.* **2018**, *In Press*.
(182) Wang, S.; Zhou, L.; Wang, Z.; Shi, X.; Xu, G. *Anal. Chim. Acta.* **2017**, *966*, 34–40.
(183) González-Domínguez, R.; Sayago, A.; Fernández-Recamales, Á.; Fernández-Recamales. *Bioanalysis* **2017**, *9*, 131–148.
(184) Wang, M.; Wang, C.; Han, R. H.; Han, X. *Prog. Lipid Res.* **2016**, *61*, 83–108.
(185) Mao, S.; Li, W.; Zhang, Q.; Zhang, W.; Huang, Q.; Lin, J.-M. *TrAC, Trends Anal. Chem.* **2018**, *107*, 43–59.
(186) Sikanen, T.; Franssila, S.; Kauppila, T. J.; Kostiainen, R.; Kotiaho, T.; Ketola, R. A. *Mass Spectrom. Rev.* **2010**, *29*, 351–391.
(187) Trojanowicz, M.; Kołacińska, K. *Analyst* **2016**, *141*, 2085–2139.
(188) Farshidfar, F.; Kopciuk, K. A.; Hilsden, R.; McGregor, S. E.; Mazurak, V. C.; Buie, W. D.; MacLean, A.; Vogel, H. J.; Bathe, O. F. *BMC Cancer* **2018**, *18*, 26.
(189) Cajka, T.; Riddellova, K.; Tomaniova, M.; Hajslova, J. *Metabolomics* **2011**, *7*, 500–508.
(190) Clendinen, C. S.; Monge, M. E.; Fernández, F. M. *Analyst* **2017**, *142*, 3101–3117.
(191) Cody, R. B.; Laramée, J. A.; Durst, H. D. *Anal. Chem.* **2005**, *77*, 2297–2302.
(192) Gulersonmez, M. C.; Lock, S.; Hankemeier, T.; Ramautar, R. *Electrophoresis* **2016**, *37*, 1007–1014.
(193) Sasaki, K.; Sagawa, H.; Suzuki, M.; Yamamoto, H.; Tomita, M.; Soga, T.; Ohashi, Y. *Anal. Chem.* **2019**, *91*, 1295–1301.
(194) Harada, S.; Hirayama, A.; Chan, Q.; Kurihara, A.; Fukai, K.; Iida, M.; Kato, S.; Sugiyama, D.; Kuwabara, K.; Takeuchi, A.; et al. *PLoS One* **2018**, *13*, e0191230.
(195) Burnum-Johnson, K. E.; Zheng, X.; Dodds, J. N.; Ash, J.; Fourches, D.; Nicora, C. D.; Wendler, J. P.; Metz, T. O.; Waters, K. M.; Jansson, J. K.; et al. *TrAC, Trends Anal. Chem.* **2019**.
(196) Leaptrot, K. L.; May, J. C.; Dodds, J. N.; McLean, J. A. *Nat. Commun.* **2019**, *10*, 985.
(197) Zhou, Z.; Tu, J.; Xiong, X.; Shen, X.; Zhu, Z.-J. *Anal. Chem.* **2017**, *89*, 9559–9566.
(198) Zheng, X.; Aly, N. A.; Zhou, Y.; Dupuis, K. T.; Bilbao, A.; Paurus, V. L.; Orton, D. J.; Wilson, R.; Payne, S. H.; Smith, R. D.; et al. *Chem. Sci.*

2017, *8*, 7724–7736.
(199) Paglia, G.; Angel, P.; Williams, J. P.; Richardson, K.; Olivos, H. J.; Thompson, J. W.; Menikarachchi, L.; Lai, S.; Walsh, C.; Moseley, A.; et al. *Anal. Chem.* **2015**, *87*, 1137–1144.
(200) Paglia, G.; Astarita, G. *Nat. Protoc.* **2017**, *12*, 797–813.
(201) Zhou, Z.; Shen, X.; Chen, X.; Tu, J.; Xiong, X.; Zhu, Z.-J. *Bioinformatics* **2019**, *35*, 698–700.
(202) Lísa, M.; Cífková, E.; Khalikova, M.; Ovčačíková, M.; Holčapek, M. *J. Chromatogr. A* **2017**, *1525*, 96–108.
(203) Bamba, T.; Lee, J. W.; Matsubara, A.; Fukusaki, E. *J. Chromatogr. A* **2012**, *1250*, 212–219.
(204) Taguchi, K.; Fukusaki, E.; Bamba, T. *Bioanalysis* **2014**, *6*, 1679–1689.
(205) Tautenhahn, R.; Patti, G. J.; Rinehart, D.; Siuzdak, G. *Anal. Chem.* **2012**, *84*, 5035–5039.
(206) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *BMC Bioinformatics* **2010**, *11*, 395.
(207) Röst, H. L.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weisser, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H.-C.; Gutenbrunner, P.; Kenar, E.; et al. *Nat. Methods* **2016**, *13*, 741–748.
(208) Xia, J.; Psychogios, N.; Young, N.; Wishart, D. S. *Nucleic Acids Res.* **2009**, *37*, W652-W660.
(209) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; Vandergheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* **2015**, *12*, 523–526.
(210) Giacomoni, F.; Corguillé, G. Le; Monsoor, M.; Landi, M.; Pericard, P.; Pétéra, M.; Duperier, C.; Tremblay-Franco, M.; Martin, J.-F.; Jacob, D.; et al. *Bioinformatics* **2015**, *31*, 1493–1495.
(211) Davidson, R. L.; Weber, R. J. M.; Liu, H.; Sharma-Oates, A.; Viant, M. R. *Gigascience* **2016**, *5*, 10.
(212) Lommen, A. *Anal. Chem.* **2009**, *81*, 3079–3086.
(213) Meyer, M. R.; Peters, F. T.; Maurer, H. H. *Clin. Chem.* **2010**, *56*, 575–584.
(214) Spicer, R.; Salek, R. M.; Moreno, P.; Cañueto, D.; Steinbeck, C. *Metabolomics* **2017**, *13*, 106.
(215) Chambers, M. C.; MacLean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; et al. *Nat. Biotechnol.* **2012**, *30*, 918–920.
(216) Gorrochategui, E.; Jaumot, J.; Lacorte, S.; Tauler, R. *TrAC, Trends Anal. Chem.* **2016**, *82*, 425–442.
(217) Pierce, K. M.; Kehimkar, B.; Marney, L. C.; Hoggard, J. C.; Synovec, R. E. *J. Chromatogr. A* **2012**, *1255*, 3–11.
(218) Castillo, S.; Mattila, I.; Miettinen, J.; Orešič, M.; Hyötyläinen, T. *Anal. Chem.* **2011**, *83*, 3058–3067.
(219) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; et al. *Metabolomics* **2007**, *3*, 211–221.
(220) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. *Ther. Drug Monit.* **2005**, *27*, 747–751.
(221) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.;

Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; et al. *Nucleic Acids Res.* **2007**, *35*, D52i-D526.
(222) Kanehisa, M.; Goto, S. *Nucleic Acids Res.* **2000**, *28*, 27–30.
(223) Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. *Nucleic Acids Res.* **2007**, *35*, W606-W612.
(224) Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. *Nucleic Acids Res.* **2006**, *34*, D668-72.
(225) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; et al. *J. Mass Spectrom.* **2010**, *45*, 703–714.
(226) Wang, M.; Carver, J. J.; Phelan, V. V; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. *Nat. Biotechnol.* **2016**, *34*, 828–837.
(227) Kind, T.; Tsugawa, H.; Cajka, T.; Ma, Y.; Lai, Z.; Mehta, S. S.; Wohlgemuth, G.; Barupal, D. K.; Showalter, M. R.; Arita, M.; et al. *Mass Spectrom. Rev.* **2018**, *37*, 513–532.
(228) Vinaixa, M.; Schymanski, E. L.; Neumann, S.; Navarro, M.; Salek, R. M.; Yanes, O. *TrAC, Trends Anal. Chem.* **2016**, *78*, 23–35.
(229) Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihira, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn, O.; Arita, M. *Anal. Chem.* **2016**, *88*, 7946–7958.
(230) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. *Nat. Methods* **2019**, *16*, 299–302.
(231) Ridder, L.; Van Der Hooft, J. J. J.; Verhoeven, S.; De Vos, R. C. H.; Bino, R. J.; Vervoort, J. *Anal. Chem.* **2013**, *85*, 6033–6040.
(232) Allen, F.; Pon, A.; Wilson, M.; Greiner, R.; Wishart, D. *Nucleic Acids Res.* **2014**, *42*, W94–W99.
(233) Wolf, S.; Schmidt, S.; Müller-Hannemann, M.; Neumann, S. *BMC Bioinformatics* **2010**, *11,* 148.
(234) Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. *Proc. Natl. Acad. Sci.* **2015**, *112*, 12580–12585.
(235) Kind, T.; Liu, K.-H.; Lee, D. Y.; Defelice, B.; Meissen, J. K.; Fiehn, O. *Nat. Methods* **2013**, *10*, 755–758.
(236) Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O. *Metabolites* **2018**, *8,* 31.
(237) Blaženović, I.; Kind, T.; Torbašinović, H.; Obrenović, S.; Mehta, S. S.; Tsugawa, H.; Wermuth, T.; Schauer, N.; Jahn, M.; Biedendieck, R.; et al. *J. Cheminform.* **2017**, *9,* 32.
(238) Schymanski, E. L.; Ruttkies, C.; Krauss, M.; Brouard, C.; Kind, T.; Dührkop, K.; Allen, F.; Vaniya, A.; Verdegem, D.; Böcker, S.; et al. *J. Cheminform.* **2017**, *9,* 22.
(239) Tang, C.; Xiao, H.; Zhang, L.; Cao, D.; Fan, W.; Liang, Y.; Zeng, Y.; Tan, B.; Zeng, M. *TrAC,Trends Anal. Chem.* **2013**, *47*, 37–46.
(240) Kopka, J.; Schauer, N.; Krueger, S.; Birkemeyer, C.; Usadel, B.; Bergmller, E.; Drmann, P.; Weckwerth, W.; Gibon, Y.; Stitt, M.; et al. *Bioinformatics* **2005**, *21*, 1635–1638.
(241) Kind, T.; Wohlgemuth, G.; Lee, D. Y.; Lu, Y.; Palazoglu, M.; Shahbaz, S.; Fiehn, O. *Anal. Chem.* **2009**, *81*, 10038–10048.
(242) Stein, S. E.; Scott, D. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 859–866.
(243) Hummel, J.; Strehmel, N.; Selbig, J.; Walther, D.; Kopka, J.

Metabolomics **2010**, *6*, 322–333.
(244) Kwiecien, N. W.; Bailey, D. J.; Rush, M. J. P.; Cole, J. S.; Ulbrich, A.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. *Anal. Chem.* **2015**, *87*, 8328–8335.
(245) Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. *Metabolites* **2012**, *2*, 775–795.
(246) Hendriks, M. M. W. B.; Eeuwijk, F. A. va.; Jellema, R. H.; Westerhuis, J. A.; Reijmers, T. H.; Hoefsloot, H. C. J.; Smilde, A. K. *TrAC, Trends Anal. Chem.* **2011**, *30*, 1685–1698.
(247) Liland, K. H. *TrAC, Trends Anal. Chem.* **2011**, *30*, 827–841.
(248) Booth, S. C.; Weljie, A. M.; Turner, R. J. *Comput. Struct. Biotechnol. J.* **2013**, *4*, e201301003.
(249) Frolkis, A.; Knox, C.; Lim, E.; Jewison, T.; Law, V.; Hau, D. D.; Liu, P.; Gautam, B.; Ly, S.; Guo, A. C.; et al. *Nucleic Acids Res.* **2009**, *38*, D487.
(250) Croft, D.; Mundo, A. F.; Haw, R.; Milacic, M.; Weiser, J.; Wu, G.; Caudy, M.; Garapati, P.; Gillespie, M.; Kamdar, M. R.; et al. *Nucleic Acids Res.* **2014**, *42*, D477.
(251) Caspi, R.; Billington, R.; Fulcher, C. A.; Keseler, I. M.; Kothari, A.; Krummenacker, M.; Latendresse, M.; Midford, P. E.; Ong, Q.; Ong, W. K.; et al. *Nucleic Acids Res.* **2018**, *46*, D639.
(252) Trokovic, R.; Weltner, J.; Otonkoski, T. *Stem Cell Res.* **2015**, *15*, 266–268.
(253) Bates, D.; Mächler, M.; Bolker, B. M.; Walker, S. C. *J. Stat. Softw.* **2015**, *67*.
(254) Xia, J.; Wishart, D. S. *Nucleic Acids Res.* **2010**, *38*, W77.
(255) Schwämmle, V.; Jensen, O. N. *Bioinformatics* **2010**, *26*, 2841–2848.
(256) Nygren, H.; Pöhö, P.; Seppänen-Laakso, T.; Lahtinen, U.; Orešič, M.; Hyötyläinen, T. *LCGC North Am.* **2014**, *32*, 286–293.
(257) Nygren, H.; Seppänen-Laakso, T.; Castillo, S.; Hyötyläinen, T.; Orešič, M. Humana Press, **2011**; pp 247–257.
(258) Langfelder, P.; Horvath, S. *BMC Bioinformatics* **2008**, *9*, 559.
(259) Sikanen, T.; Tuomikoski, S.; Ketola, R. A.; Kostiainen, R.; Franssila, S.; Kotiaho, T. *J. Mass Spectrom.* **2008**, *43*, 726–735.
(260) Tuomikoski, S.; Sikanen, T.; Ketola, R. A.; Kostiainen, R.; Kotiaho, T.; Franssila, S. *Electrophoresis* **2005**, *26*, 4691–4702.
(261) ICH Harmonised Tripartite Guideline, Validation of Analytical Procedures: Text and Methodology Q2(R1), **2005**, https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1__Guideline.pdf [Last accessed: 12.11.2018].
(262) World Anti-Doping Agency, WADA Technical Document – TD2010IDCR, **2010**, https://www.wada-ama.org/sites/default/files/resources/files/WADA_TD2010IDCRv1.0_Identification%20Criteria%20for%20Qualitative%20Assays_May%2008%202010_EN.doc.pdf [Last accessed: 12.11.2018].
(263) Porrello, E. R.; Mahmoud, A. I.; Simpson, E.; Hill, J. A.; Richardson, J. A.; Olson, E. N.; Sadek, H. A. *Science (80-.).* **2011**, *331*, 1078–1080.
(264) Haubner, B. J.; Schneider, J.; Schweigmann, U.; Schuetz, T.; Dichtl, W.; Velik-Salchner, C.; Stein, J.-I.; Penninger, J. M. *Circ. Res.* **2016**, *118*, 216–221.

(265) Triebl, A.; Trötzmüller, M.; Eberl, A.; Hanel, P.; Hartler, J.; Köfeler, H. C. *J. Chromatogr. A* **2014**, *1347*, 104–110.
(266) Bollinger, J. G.; Ii, H.; Sadilek, M.; Gelb, M. H. *J. Lipid Res.* **2010**, *51*, 440–447.
(267) Wang, T. J.; Larson, M. G.; Vasan, R. S.; Cheng, S.; Rhee, E. P.; McCabe, E.; Lewis, G. D.; Fox, C. S.; Jacques, P. F.; Fernandez, C.; et al. *Nat. Med.* **2011**, *17*, 448–453.
(268) Newgard, C. B.; An, J.; Bain, J. R.; Muehlbauer, M. J.; Stevens, R. D.; Lien, L. F.; Haqq, A. M.; Shah, S. H.; Arlotto, M.; Slentz, C. A.; et al. *Cell Metab.* **2009**, *9*, 311–326.
(269) Rhee, E. P.; Cheng, S.; Larson, M. G.; Walford, G. A.; Lewis, G. D.; McCabe, E.; Yang, E.; Farrell, L.; Fox, C. S.; O'Donnell, C. J.; et al. *J. Clin. Invest.* **2011**, *121*, 1402–1411.
(270) Kotronen, A.; Velagapudi, V. R.; Yetukuri, L.; Westerbacka, J.; Bergholm, R.; Ekroos, K.; Makkonen, J.; Taskinen, M.-R.; Orešič, M.; Yki-Järvinen, H. *Diabetologia* **2009**, *52*, 684–690.
(271) Floegel, A.; Stefan, N.; Yu, Z.; Mühlenbruch, K.; Drogan, D.; Joost, H.-G.; Fritsche, A.; Häring, H.-U.; Hrabě de Angelis, M.; Peters, A.; et al. *Diabetes* **2013**, *62*, 639–648.
(272) 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.; et al. *Lancet Diabetes Endocrinol.* **2014**, *2*, 810–818.
(273) Duffin, K. L.; Henion, J. D.; Shieh, J. J. *Anal. Chem.* **1991**, *63*, 1781–1788.
(274) Ferreiro-Vera, C.; Priego-Capote, F.; Luque de Castro, M. D. *J. Chromatogr. A* **2012**, *1240*, 21–28.
(275) Tulipani, S.; Llorach, R.; Urpi-Sarda, M.; Andres-Lacueva, C. *Anal. Chem.* **2013**, *85*, 341–348.
(276) Hintikka, L.; Haapala, M.; Franssila, S.; Kuuranne, T.; Leinonen, A.; Kostiainen, R. *J. Chromatogr. A* **2010**, *1217*, 8290–8297.
(277) Revelsky, A. I.; Samokhin, A. S.; Virus, E. D.; Rodchenkov, G. M.; Revelsky, I. A. *Drug Test. Anal.* **2011**, *3*, 263–267.
(278) Eenoo, P. Van; Gansbeke, W. Van; Brabanter, N. De; Deventer, K.; Delbeke, F. T. *J. Chromatogr. A* **2011**, *1218*, 3306–3316.
(279) Robles, J.; Marcos, J.; Renau, N.; Garrostas, L.; Segura, J.; Ventura, R.; Barceló, B.; Barceló, A.; Pozo, O. J. *Talanta* **2017**, *169*, 20–29.
(280) Cuzzola, A.; Petri, A.; Mazzini, F.; Salvadori, P. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 2975–2982.
(281) Pujos, E.; Flament-Waton, M. M.; Paisse, O.; Grenier-Loustalot, M. F. *Anal. Bioanal. Chem.* **2005**, *381*, 244–254.
(282) Franke, A. A.; Custer, L. J.; Morimoto, Y.; Nordt, F. J.; Maskarinec, G. *Anal. Bioanal. Chem.* **2011**, *401*, 1319–1330.
(283) Jeon, B. W.; Yoo, H. H.; Jeong, E. S.; Kim, H. J.; Jin, C.; Kim, D. H.; Lee, J. *Anal. Bioanal. Chem.* **2011**, *401*, 1353–1363.
(284) Xu, X.; Veenstra, T. D.; Fox, S. D.; Roman, J. M.; Issaq, H. J.; Falk, R.; Saavedra, J. E.; Keefer, L. K.; Ziegler, R. G. *Anal. Chem.* **2005**, *77*, 6646–6654.
(285) Renterghem, P. Van; Eenoo, P. Van; Geyer, H.; Schänzer, W.; Delbeke, F. T. *Steroids* **2010**, *75*, 154–163.
(286) Palermo, M.; Gomez-Sanchez, C.; Roitman, E.; Shackleton, C. H. L.

Steroids **1996**, *61*, 583–589.
(287) Noor, E.; Cherkaoui, S.; Sauer, U. *Curr. Opin. Syst. Biol.* **2019**, *15*, 39–47.
(288) Khamis, M. M.; Adamko, D. J.; Purves, R. W.; El-Aneed, A. *Anal. Chim. Acta* **2019**, *1047*, 81–92.
(289) Cajka, T.; Fiehn, O. *Metabolomics* **2016**, *12*, 34.
(290) Rampler, E.; Coman, C.; Hermann, G.; Sickmann, A.; Ahrends, R.; Koellensperger, G. *Analyst* **2017**, *142*, 1891–1899.
(291) Wu, L.; Mashego, M. R.; van Dam, J. C.; Proell, A. M.; Vinke, J. L.; Ras, C.; van Winden, W. A.; van Gulik, W. M.; Heijnen, J. J. *Anal. Biochem.* **2005**, *336*, 164–171.
(292) Hermann, G.; Schwaiger, M.; Volejnik, P.; Koellensperger, G. *J. Pharm. Biomed. Anal.* **2018**, *155*, 329–334.
(293) Gika, H. G.; Theodoridis, G. A.; Plumb, R. S.; Wilson, I. D. *J. Pharm. Biomed. Anal.* **2014**, *87*, 12–25.