TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Analytische Lebensmittelchemie

Chromatography and mass spectrometry-based non-targeted metabolomics for Type2 Diabetes studies

Sara Forcisi

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigte Dissertation.

Vorsitzender: Univ.-Prof. Dr. E. Grill

Prüfer der Dissertation:

1. Priv.-Doz. Dr. Ph. Schmitt-Kopplin
2. Univ.-Prof. Dr. M. Rychlik
3. apl Prof. Dr. R. Lehmann
   (Eberhard-Karls-Universität Tübingen)

Die Dissertation wurde am 19.04.2012 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 05.09.2012 angenommen.

ὁ δὲ ἀνεξέταστος βίος οὐ βιωτὸς ἀνθρώπῳ

The unexamined life is not worth living for a human being

Plato, Apology of Socrates

To my Grandmother

Summary

Non-targeted metabolomics is the latest technique in holistic biomarker discovery that gained more interest in modern biology during the last decades. It is a promising tool for the investigation and evaluation of diseases and for understanding their genetic and the environmental influence. Different technologies providing high resolution, accuracy and sensitivity towards a wide range of metabolites were used to discover biochemical patterns distinct for a certain clinical condition and to discover new responsible pathways. In this work the development of a metabolomic platform that integrates the chromatographic resolution of UPLC with the high mass accuracy of FT-ICR-MS for the analysis of human plasma samples was proposed. An application of the techniques optimized for subjects with a pre-diabetes phenotype was developed. Pivotal topics such as sample collection and sample preparation are discussed focusing on the difference and stability of blood anticoagulants and describing diverse sample preparations such as protein precipitation extraction and solid phase extraction. Protocols referring to the latter were investigated in order to define the most appropriate conditions for long term measurements from an untargeted metabolomics perspective. The work flow leads to the experimental application; the study of the untargeted metabolome of fatty liver individuals, whose pre-diabetic state was characterized by their insulin sensitivity according to $ISI_{Matsuda}$ index. The latter is supposed to enable inference with a pre-diabetic state. The integration of UPLC-MS and FT-ICR-MS data enabled identification of discriminative metabolic pattern in insulin sensitive and insulin resistance groups. In order to understand the influence of life style intervention (nine months of integration of diet and exercise), metabolic profiling of the same subjects was as well leading to a discriminative pattern among the two classes of observations. This work provides an overview on the development of the methodologies applied and on their application in the study of insulin sensitivity in fatty liver subjects, suggesting possible discriminative biomarkers in the light of correlation with clinical parameters such as liver fat, $ISI_{Matsuda}$, CRP and fastening glucose. It offers an incipit for future investigations in understanding the biochemical pathways involved in the scenario of Diabetes mellitus.

Zusammenfassung

Ungerichtete Metabolomik ist ein neuer Ansatz für holistische Biomarkeranalyse und, gewann in den letzten Jahrzehnten stark an Bedeutung in der modernen Biologie. Sie ist ein vielversprechendes Werkzeug in der Untersuchung und Evaluierung von Krankheiten und für das Verständins ihrer genetischen und umweltbedingten Verknüpfungen. Verschiedene hochauflösende, hochakkurate und sensitive Technologien wurden verwendet um biochemische Muster zu identifizieren, die spezifisch für definierte klinische Konditionen sind und helfen die verantwortlichen Stoffwechselwege einzugrenzen. In dieser Arbeit wird die Entwicklung einer Metabolomik Plattform präsentiert, welche chromatographische Auflösung mittels UPLC und hohe Massengenauigkeit mittles FT-ICR-MS für die Analyse von humanem Blutplasma verbindet. In diesem Zuge entwickelte Techniken werden später für die Analyse von prä-diabetischen Phänotypen verwendet. Ausschlaggebende Punkte wie die Sammlung und Präparation von Proben werden behandelt. Die unterschiedliche Stabilität und der Einfluss von Antikoagulantien werden im Hinblick auf Probenvorbereitungsverfahren wie Festphasenextraktion und Proteinpräzipitation untersucht. Aus der Perspektive der ungerichteten Metabolomik werden langzeitstabile Konditionen für entsprechende Ansätze erarbeitet. Der Arbeitsablauf wird schließlich auf Plasmata von Fette-Leber-Patienten angewandt, welche der Insulinresistenz nach ISI-Matsuda folgend klassifiziert waren. Die Kombination von UPLC-MS und FT-ICR-MS Analyseermöglichtedie Identifizierung diskriminativer Metabolitmuster in Insulin-sensitiven und Insulin-resistenten Patientengruppen. Nach neun Monaten Lebensstil-Intervetionen (neun Monate spezielle Diät und Sport) konnten in beiden Patientengruppen durch metabolisches Profiling ähnliche diskriminative Muster identifiziert werden. Diese Arbeit erläutert die Entwicklung der schließlich angewandten analytischen Strategien in der Analyse von Insulinsensitivitäts-Markern in Individuen mit nicht-alkoholischer Fett-Leber. Erschlossene Markerkandidaten wurden auf Korrelation mit klinischen Parametern, wie Leber-Fett, $ISI_{Matsuda}$, CRP and Glukose Konzentration nach Fasten überprüft. Diese Arbeit bietet eine Grundlage für das zukünftige Untersuchen und Verstehen der biochemischen Stoffwechselwege in Diabetes mellitus.

Table of Contents

1. Introduction 1
1.1.1 Metabolomics: the concept 1
1.1.2 Metabolomics: the challenge 3
1.1.3 Metabolomics: the tools 5
1.2. Human blood 15
1.3 Human plasma 17
1.4 Diabetes 19
2. Aim of the Thesis 24
3. Theoretical introduction 26
3.1 Mass Spectrometry 26
3.2.1 Fourier Transform Ion Cyclotron Resonance Mass Spectrometers (ICR-FT/MS) 27
3.2.2 Time of Flight mass spectrometers (TOF-MS) 29
3.3 Liquid Chromatography 33
3.4 Sample Preparation 44
3.5 Mass-Mass Difference Networking 46
3.6 Multivariate Statistics 49
4. Methodological aspects of samples analysis 52
4.1 Sample collection 52
4.2 Sample preparation 61
4.3 Long term stability test of the UPLC-MS system 77
5. Non-alcholic fatty liver disease: an overview 80
5.1 Epidemiology 81
5.2 Diagnosis 82
5.3 Pathogenesis 84
5.4 Matbotyping of Non-alcoholic Fatty Liver individuals: the study 91
5.4.1 Introduction 91
5.4.2. The Baseline study – The Results 95
5.4.3 Discussion of the biological results 117
5.5 The puzzle composition: the description of the possible biochemical pathways involved 129
6. Conclusion 131
6.1 Outlook 134
7. Appendix 135

8. Bibliography 218
9. Acknowledgements..... 230

List of figures

Figure 1: Integration of different high resolution analytical technologies in order to establish a metabolite map for each matrix analyzed.. 12

Figure 2: Human blood and its components.. 15

Figure 3: Whole human blood components.. 16

Figure 4: Blood plasma components..... 17

Figure 5: Insulin Secretion and Insulin Action in Type 2 Diabetes pathogenesis. 22

Figure 6: Map of the prevalence (%) of Impaired Glucose Tolerance in adults... 23

Figure 7: Schematic representation of the Competence Network Diabetes interaction..... 24

Figure 8: Schematic representation of the ICR process. 28

Figure 9: Schematic representation of Synapt G1 from Waters.. 31

Figure 10: Evolution of different particle sizes from HPLC until the advent of UPLC. 36

Figure 11: Schematic representation of the electrospray ionization. 42

Figure 12: Scheme of a mass-mass difference network section.. 47

Figure 13: Comparison of Li-Heparinate, NaF, Citrate and EDTA plasma in the LOAD fraction collected..... 54

Figure 14: Comparison of Li-Heparinate, NaF, Citrate and EDTA plasma in the ELUTE fraction collected.. 55

Figure 15: Bland and Altman Plot of the three replications (Plasma EDTA, Load fr.)..... 56

Figure 16: Bland and Altman Plot of the three replications (Plasma Citrate, Load fr.)..... 57

Figure 17: Bland and Altman Plot of the three replications..... 57

Figure 18: Bland and Altman Plot of the three replications (Plasma EDTA)..... 59

Figure 19: Bland and Altman Plot of the three replications (Plasma Citrate) applying PPE.. 59

Figure 20: Bland and Altman Plot of the three replications (Plasma NaF) applying PPE.. 60

Figure 21: Representation of the chemical properties of the sorbent of HLB sorbent material from Waters.. 62

Figure 22: Chromatograms showing the comparison between the load and elute fraction applying HLB cartridge. 63

Figure 23: Pricipal classes of markers present, applying HLB extraction, in the load fraction and in the elute fraction respectively.. 65

Figure 24: Bland and Altman Plot of the three replications applying PPE with Acetonitrile..... 66

Figure 25: Bland and Altman Plot of the three replications applying PPE with EtOH/MeOH(1/1).. 67

Figure 26: Different types of compound classes annotated in positive ionization mode using different anticoagulants. 68

Figure 27: Different types of compound classes annotated in negative ionization mode using different anticoagulants 69

Figure 28: The red PPE mass spectrum shows high salt content based mass defects of 300.7 m/z to 300.9 m/z.. 70

Figure 29: Different SPE cartridges tested in ESI positive mode.. 72

Figure 30: Different SPE cartridges tested in ESI negative mode.. 74

Figure 31: Counts of recovered compound classes as a function of elution volume and evaporation. 76

Figure 32: Long term test of plasma peaks and internal standards. 77

Figure 33: Running RSD showing the day time temperature dependency. 78

Figure 34: HCA of the day time dependency of the mass spectrometric signal variation..... 78

Figure 35: Schematic description of NAFLD. 80

Figure 36: The score scatter plot of the PCA (colored following the classes)..... 96

Figure 37: OPLS/O2PLS score scatter plot in which is visualized the two different group 97

Figure 38: Percentage of annotated masses and unknowns among 173 marker candidates..... 97

Figure 39: Important biomarkers in NAFL via FT-MS-ICR analysis. 98

Figure 40: Dynamic Mass accuracy (ppm) of averaged standard compound measured via TOF-MS during NAFL experiment.. 101

Figure 41: Dynamic Mass accuracy (Da) of averaged standard compound measured via TOF-MS during NAFL experiment.. 101

Figure 42: Variation of internal standards in QCs throughout the study. 102

Figure 43: Partition of peaks with RSD < 30% (blue line) and intensity distribution (red bars). 103

Figure 44: Differentiation of QC and samples. 105

Figure 45: Comparison of OPLS/O2PLS scatter plot describing the separation among the two classes of interest (IR=1 and IS=2). 107

Figure 46: Upper and lower extremities of mass shift in the LC-studies NAFL_Study1 (red) and NAFL_Study2 (blue). 109

Figure 47: Alignment between LC-MS masses and FT-MS mass annotation.... 110

Figure 48: Possible mass annotation candidates in NAFL-LC-MS positive study. 111

Figure 49: Scatter plot of the PCA model.The first component absorbs 32 %... 112

Figure 50: OPLS/O2PLS score scatter plot in which is visualized the..... 113

Figure 51: Classes of important compounds among UP-regulated insuin sensitive group and down reulated.. 114

Figure 52: Characteristic S-lot showing the most significant masses. It is possible to denote the dominace of the phosphocholine class. 114

Figure 53: Typical Lyso-PC pattern yielded at 10 eV collision energy in the LIT partition of the FT-ICR-MS mass spectrometer; standard. 115

Figure 54: Fragment pattern for the same Lyso-PC (16:0) in plasma..... 116

Figure 55: Ammonia loss of Tryptophan as a standard (red) and in Plasma (blue). 116

Figure 56: Chemical structure of glycerol and sphingosine..... 121

Figure 57: General lipid structures of Phospholipids..... 122

Figure 58: Glycerophospholipid synthesis. 122

List of Abbreviations

AA	Amino acid
AcN	Acetonitrile
APCI	Atmospheric pressure chemical ionization
APPI	Atomsohperic pressure photo ionization
BA	Bile acid
BC	Before Christ
CBA	Carboxylic acid
CN	Cyano
CoA	Coenzyme A
CRP	C (complement) reactive protein
CVD	Cardio vascular disease
Da	Dalton
EDTA	Ethylenediamintetraacetate
ER	Endoplasmatic reticulum
ESI	Electrospray ionization
EtOH	Ethanol
FFA	Free fatty acid
ICR-FT/MS	Fourier transform ion cyclotron mass mpectrometry
GC-MS	Gas chromatography – mass spectrometry
GigaHz	Giga Hertz
GLP	Glycerolipid
GLP-1	Glucagon like peptide
HDL	High density lipoprotein
HLB	Hydrophilic-Lipophilic balance
HPLC	High performance liquid chromatography
IL	Interleukin
IR	Insulin resistant
IS	Insulin sensitive
LA	Lipo amide
LDL	Low density lipoprotein
Li-Hep	Lithium-Heparinate
LPL	Lyso-Phospholipid
Lyso-PC/LPC	Lyso-phosphatidylcholine
Lyso-PE	Lyso-phosphatidylethanolamine
m/z	mass over charge
MeOH	Methanol
MHz	Mega Hertz
MNP	Magnetic nano particle
MS	Mass spectrometry
NaF	Sodium fluoride
NF-kB	Nuclear factor ‘kappa-light-chain-enhancer’ of activated B-cells
NMR	Nuclear magnetic resonance spectroscopy
OPLS	PLS with orthogonal signal correction
PEG	Polyethylenglycol
PC	Principal component (only in statistical context)

PC/PE	Phosphatidylcholine/Phosphatidylethanolamine
PCA	Principal component analysis
PheAla	Phenylalanine
PL	Phospholipid
PLA	Phospholipase A
PLS	Projection to latent structures using partial least squares regression
PLS-DA	PLS – Discriminant analysis
PPE	Protein precipitation extraction
ppm	Parts per million
QC	Quality control
RP	Reversed phase
RSD	Relative standard deviation
RT	Retention time
S/N	Signal to noise ratio
SNP	Single nucleotide polymorphism
SPE	Solid phase extraction
TNF-α	Tumor necrosis factor alpha
Trp	Tryptophan
Tyr	Tyrosine
UPLC-TOF-MS	Ultrahigh performance liquid chromatography – time of flight - mass spectrometry
UV-VIS	Ultra violet and visual light spectrum
V	Volt
VLDL	Very low density lipoprotein

1. Introduction

1.1.1 Metabolomics: the concept

During the last two decades many improvements in modern molecular biology were reached and along with it the necessity of a fundamental integration of different disciplines such as transcriptomics, proteomics and metabolomics arose. The integration of different "-omics", with the support of mathematical modeling is defined "top-down" system biology [1, 2]. It represents the key for the integration of different biological levels (subcellular, cellular, tissue, organ) in order to study the connection between genetic variation, environmental factors and possible diseases [3]. In every organism at each complexity level different kinds of biotransformations are occurring, delineating the metabolism. The word metabolism comes from the greek word μεταβολή (metabolē) that means change. The investigation of such as "change" leads to the definition of different approaches applied for studying the *metabolome* and to the quantitative description of the changes of all metabolites (low molecular weight intermediates) present in a cell, tissue or organ [4]. *Metabonomics* is defined as "*the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification*" [5]. This definition was developed after several H-NMR studies on biofluids, cells and tissue [6-11]. *Metabolomics* refers to the comprehensive quantitative detection and identification of the total complement of metabolites of an organism or biofluid [12, 13]. The difference between these two terms is not so distinct and they are often used interchangeably [14]. They both use holistic analytical approaches in order to study biological patterns of low molecular weight metabolites within an organism. Metabonomic studies provide the possibility to examine the end points that directly link genetic variation and environmental interactions [15]. The metabotype is defined as "*the metabolic profile that defines a phenotype which relates to genetic variation of the organism*" [16] and is the results of the interaction with environmental factors such as gender, age, lifestyle, diet, stress level, and gut microbiota; within this interaction a disease risk can be included

[15, 17, 18]. The analysis of the metabotype consists in the analysis of different metabolites (metabolic profiling) in bio fluids such as plasma, serum, urine, exhaled breath condensate in order to achieve information not available at other –omics levels [19,15]. The different metabolites (molecules < 1kDa) can be *endogenous* when they are under the host genome control and cellular function-dependent [20], *exogenous* when they are introduced via environmental exposure or via diet and medicaments or *co-metabolites* when they are originated from the interaction with symbiotic organisms like the gut microbiota [21]. Therefore, perturbations in the gut microbiota modulate the host metabolism [22] with the production of co-metabolites [18]. They can influence the metabolic phenotype and possibly induce different disorders [21] such as intestinal disease, obesity or cancer [19]. Examples of such consequent disorders could be diet-induced insulin resistance and Type 2 Diabetes [23, 24] or Type 1 Diabetes reported in animal models [25].

1.1.2 Metabolomics: the challenge

Based on the various, often redundant, definitions that literature gives for metabolic profiling or metabolomics we can conclude that they aim at the extraction, analysis, identification and quantification of the metabolites present in cells, tissue, organs and bio fluids in order to provide a snapshot of different classes of compounds discriminating various levels of biological conditions. The approach can be *targeted* when it is directed to the detection and quantification of specific classes of compounds or to subsets of known metabolic pathways. One performs selection prior to analysis. This approach finds great application in the pharmaceutical field and the quantification of analytes of interest is often accompanied by the use of isotopically labelled standards. The other side of metabolic profiling is the *non-targeted* approach where from the sample extraction until the data evaluation the aim is to study the widest range of compounds possible, without excluding any. This is the aim of metabolomics research. The number of metabolites can range from 1000 until 200000 [26, 27]. In humans currently 8558 (http://www.hmdb.ca) metabolites of endogenous and exogenous origin are identified. The number of endogenous compounds in humans is smaller than the number of genes (25000) and proteins (1000000) present but the real challenge lays in the chemical complexity that characterizes these different classes of molecules. The chemical complexity of genes is based on the combination of 4 bases and in the case of proteins on the combination of 20 amino acids. Therefore automated sequencing, analyte identification and quantification resulted to be a manageable task [28, 29]. Metabolites range between 50 Da and 1500 Da of mass and from low picomolar concentrations (e.g. hormones) to molar concentrations (e.g. urea). Their chemical and physical properties range from the most polar compounds until the most apolar ones; moreover their stereochemistry leads to different biological functions [30]. Exactly this diversity made the detection of the whole set of metabolites present in a biological sample at different levels (cell, bio fluid, tissue, and organ) an ambitious goal. The achievement of this goal requires the integration of different analytical platforms in order to enable the maximum range of analyte detection

and identification [31-33] through high sensitivity, selectivity and resolution. In biological samples different kinds of chemicals like polar ones, apolar and volatiles can co-exist. The right combination of sample preparation and the integration of appropriate analytical tools define the path to an optimized performance in non-targeted metabolomics. The first important step is the sample collection that needs to preserve the integrity of the metabolites. Temperature effects have to be considered before, during and after sample collection and storage (freezing, freeze drying, thawing). The introduction of contaminations has to be avoided (i.e. PEG present in Li-Heparinate blood collection tubes (see Chapter 4); the technique adopted has to be highly reproducible and robust. Another crucial point is the sample preparation, a source of high variability, for the extraction of the metabolites from a specific biological matrix. The optimal sample preparation for non-targeted analysis requires the least number of passages in order to minimize possible sources of contamination and systematic errors. It has to be robust, sensitive and reproducible. Special treatments are necessary in order to extend as much as possible the coverage of metabolites, considering that in biological system they often are carried from proteins (e.g.albumin in blood [34]. The platforms commonly used in non-targeted metabolomics research involve the use of NMR spectroscopy, Liquid chromatograpy-mass spectrometry (LC-MS) and Gas Chromatography-Mass spectrometry (GC-MS). They are discussed among the metabolomics community in several reviews and books [35-44].

1.1.3 Metabolomics: the tools

As already mentioned in the previous section, due to the wide chemical diversity, global non-targeted metabolomics demands diverse and complementary analytical tools in order to achieve the complete experimental coverage of all the kinds of metabolites, endogenous and exogenous, present in a biological sample. This strategy, in combination with multivariate data analysis will lead to the identification of responsible biomarkers characteristic of a disease status or "alarm bell" of the existence of a pre-disease status (i.e. pre-Diabetes condition due to impaired glucose tolerance). In this section the tools used in non-targeted metabolomics will be introduced, focusing on their difference in terms of resolution, sensitivity, mass accuracy, sample preparation, robustness, in order to delineate their advantages or disadvantages and understand the design used in this work.

Nuclear Magnetic Resonance (NMR)

NMR spectroscopy is one of the most applied techniques for metabolomics studies [6, 45, 37], it is based on the physical phenomenon of absorption and emission of electromagnetic radiation in molecules, under a strong magnetic field. This absorption depends on the nuclei of odd number and /or protons atoms (^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P…) with a magnetic nuclear moment, spin. It provides information on the molecular structure of a molecule, since every molecule is distinct with a characteristic signal pattern. ^{1}H and ^{13}C are the most used, due to their natural abundance in organic molecules. Every metabolite is identified according to their chemical shift measured in parts per million (ppm), the resonance frequency is pH dependent. Generally an internal standard is used in order to allow calibration of the chemical shifts; that then used for the identification of a possible structure against libraries. One of the main advantages in using NMR in non-targeted metabolomic studies is that:

it is a non-destructive analytical technique since the recovery of the sample is possible afterwards. A big advantage is the limited sample preparation that allows the study of tissue and biofluids avoiding metabolites extraction (this is not

possible in the application of MS). It requires limited sample preparation and the quantification results possible in a robust way, allowing for the determination of the chemical structure of known and unknown compounds without derivatization (the high disadvantage of the GC technique). Moreover, chromatographic separation before analysis is not required since every molecule of the biological matrix is independently identified by its own chemical shift; this phenomenon is called “chemical shift chromatography” and it allows the study of crude samples. The disadvantage of this method is its low sensitivity compared to MS (with a limit detection of 1-5µM [1]) and the required volume of sample [41]. Improvement of the sensitivity is possible by the introduction of cryprobes that allow an increase of S/N ratio of four or five times [41] relative to a room temperature probe. Sensitivity can also being improved through the use of small diameter NMR tubes (<3mm diameter), where small volume of sample is necessary (50 µl for 1.7 mm diamter), in combination with high-field magnet (e.g. 800, 900 MHz) equipped with a cryoprobe. The recent introduction of GigaHz NMR instruments [46] aims to improve the sensitivity reached with 800 MHz by 20% [33, 47].

Mass spectrometry techniques

Direct infusion

The different kinds of mass spectrometers are divided according to their characteristic properties such as mass resolution, mass accuracy and sensitivity. Direct infusion of the sample without chromatographic separation *a priori*; represents a possible approach used in non-targeted metabolomics studies. The advantage lays inthe possible high throughput, since when it is compared to hyphenated techniques it is less time consuming. When this approach is applied, ion suppression effects that compromise the quantitative efficacy of ionisation have to be considered.

This phenomenon derives from the interference of the sample matrix competing for a charge with each metabolite [48] (e.g.salts, lysophospholipids, and peptides). Molecules at higher concentration may suppress those in lower concentration or the more polar ones. Multiple chargeable compounds at high mass can suppress the ones at low mass [49], therefore a careful sample preparation is necessary to minimize this effect. Ordinarily, the direct injection approach is performed through electro spray ionization (ESI) [50] or nanospray [51]. The use of high mass resolution and high mass accuracy are prerequisites in order to reach suitable data that allow a good discrimination of the different classes of the studied samples and a high precision in the formula calculation and compound identification of known or unknown bioamarkers. For further mass identification and structure elucidation, additional experiments such as tandem mass spectrometry (MS/MS) at different collision energy are necessary. Fourier transform ion cyclotron resonance mass spectrometers (ICR-FT/MS) [52-54], Orbitraps [55] and Time of Flight mass spectrometer (TOF-MS) [56] have been used in metabolomics applications in direct infusion (**Table 1**).

Table1: Comparison among the highest mass resolution and mass accuracy mass spectrometers.

Mass spectrometer	Mass Resolution	Mass accuracy
ICR-FT/MS	1000000-2000000	0.1-1
Orbitrap	50000-200000	0.5-1
TOF/QTOF-MS	10000-50000	2-5

ICR-FT/MS provides one of the highest mass resolution and mass accuracy, its disadvantage is the high cost and the high scanning time required to improve the sensitivity of the instrument. Therefore LC-hyphenated techniques such as HPLC/UPLC coupling are not possible without excluding loss of resolution. Orbitrap provides a good solution to overcome the high cost of ICR-FT/MS instruments and providing slightly faster scanning times which permits the coupling with separation techniques still preserving high mass accuracy and resolution relative to the maximal values achievable with this instrument [42]. Q-TOF MS was reported in non-targeted metabolomics approaches in combination with chip-based electrospray spray infusion [56] showing highly reproducible results. Direct infusion via ESI-Q-TOF-MS was compared to LC-MS approach using the same mass spectrometer, describing its advantages [57].

Hyphenated mass spectrometry based technologies

The integration of chromatographic techniques represents a staging post in non-targeted metabolomics for extending the investigation on the widest range of compound classes. The major combinations include LC-MS and GC-MS, additionally also electrophoretic separation emerged as a beneficial combination in metabolomics [58]. The application of these strategies surpasses some of the limitations linked to direct infusion mass spectrometry such as the detection of isobars and isomers, the suppression effect caused by competing ions and it gives the possibility to separate and concentrate different classes of compounds according to their physical-chemical properties. All these information allow for further identification of discriminative biomarkers in biological studies.

Liquid-chromatography mass spectrometry (LC-MS)

The advantages of this technique lay in its ability to separate complex matrices via chromatographic separation on analytical packed columns. The variety of sorbent materials is ample but the most used sorbent in holistic approaches are reversed phase (RP), C18 and C8 columns (for definition see Chapter 3). For HPLC-MS analysis, typical column characteristics are between 2.1 and 4.6 mm in inner diameter (i.d.), 5-25 cm in length and 3-5 µm in packing materials size [36, 59]. The advent of the Ultra performance liquid chromatography (UPLC) allowed the use of high pressure level instruments (up to 15000 psi) therefore enabling the use of columns with sub-2 µm particles size. The columns commonly applied range between 5-15 cm in length and 1-2.1 mm in i.d. . This technology is well established in several metabolomics investigations and offers higher chromatographic resolution [60] and peak capacity [61, 37] compared to high pressure liquid chromatography (HPLC). It upgrades the separation performance in terms of velocity [62] and consequently in high throughput analysis. The combination of Hydrophilic interaction chromatography (HILIC) (which has the opposite behaviour than the reversed phase technique) or ion-exchange chromatography, gives the possibility to better separate polar and polar ionic compounds, therefore extending the metabolic information [63, 64]. The design of the integration can be done individually [65, 66] or orthogonally by switching techniques [67]. Attempting to reduce the back pressure of UPLC columns, the application of high temperature (90° until 180°C) aiming in solvent viscosity reduction is reported [68]. Aqueous mobile phases in combination with high temperature and temperature gradients were applied [65] for urine analysis. This analytical set-up may though lead to analytes degradation. Further investigations, via capillary LC applications, show an increase in the chromatographic resolution [69] enabling the reduction of sample volume, particularly important in applications where higher sample volume is not available [70]. Results show further advantages such as a higher number of detected ions and increase in sensitivity compared to ordinary LC-MS analysis [70-72]. Concerning the integration of the mass spectrometer, Q-TOF-MS is the most used due to its peculiar combination of mass resolution, mass accuracy and fast scanning time.

Therefore its integration results more convenient than a possible integration with high resolution instruments (ICR-FT/MS and Orbitrap), where their slower scanning time would compromise chromatographic resolution. The common use of ESI, operating in positive and negative modes, can be integrated with additional combinations of APCI [73] to enlarge the detection window to a higher number of metabolites.

Gas-chromatography-mass spectrometry (GC-MS)

The application of this technology, dated back to the introduction of LC-MS, offers high analytic performances for the analysis of polar volatile compounds, especially non polar volatile compounds. The molecules that do not present this characteristic need to be chemically modified (i.e. derivatization, alkylation and sylation [74, 75] in order to allow the process. In this application technology the sample is separated, after vaporization, onto columns in the gas phase. It is reported in the analysis of fossil fuels/oil, plant metabolites, screening for inborn errors in metabolism [76-78]. Metabolomics investigation showed high resolution power, high sensitivity and reproducibility. Additionally, compound identification is favoured via electron ionization libraries where highly reproducible spectra are stored, taking advantage from the standardized MS electron ionization of 70 eV and from reproducible retention time. The main disadvantage of this technique is the time consuming sample preparation that can lead to high variability, quantification error of a certain metabolite due to incomplete conversion, low experimental reproducibility caused by different conversion rates of the metabolites [79]. Moreover, the formation of additional species after derivatization, can occour. Further developments in the integration of two-dimensional GC (GCXGC) attested good results [80-82] in terms of analysis time reduction, number of features detected. High instrumentation cost and difficult data treatment (the output is a map) do not suggest this technique as possible

strong point in metabolomics research. Different kinds of mass spectrometers are being coupled with GC, such as single quadrupole, triple quadrupole but as well TOF-MS that offer higher resolution, higher mass accuracy and faster scan time [75, 83, 84].

Summary

Non-targeted metabolomics invests new analytical tools towards the study of different biological levels in order to identify a particular state (metabotype) in a biological organism. The strong point lays in its strategy. The aim of metabolomics profiling is to detect and analyze the whole picture of the metabolites present in a biological sample (plant, bio fluids, tissue).

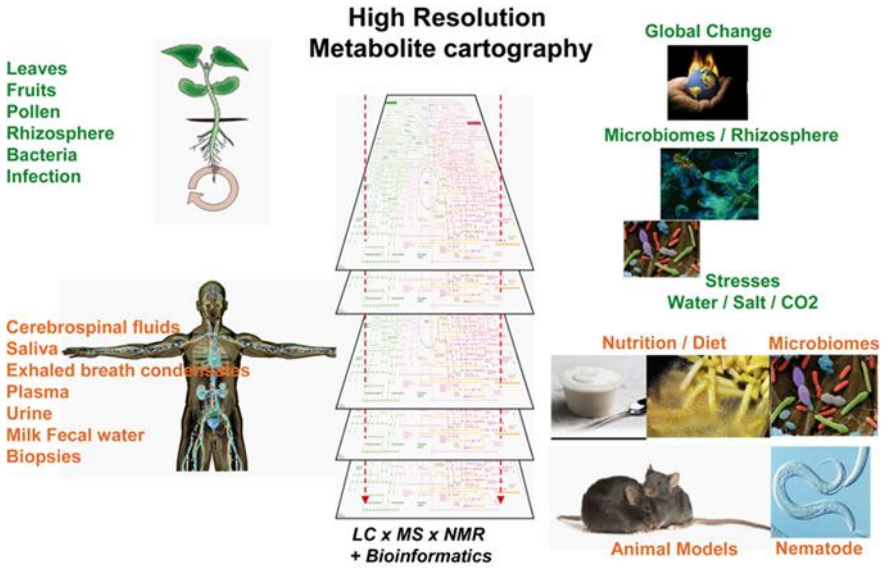


Figure 1: Integration of different high resolution analytical technologies in order to establish a metabolite map for each matrix analyzed. All the maps generated can be superimposed to achieve a space of information that explains the interaction among different samples originated from different biological levels [254]. Reprinted from *Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data*, Wiley-VCH, ISBN: 978-3-527-33089-8, Moritz F., Forcisi S., Harir M., Kanawati B., Lucio M., Tziotis D., Schmitt-Kopplin P., **The Potential of Ultrahigh Resolution MS (FTICR-MS) in Metabolomics**, 117-136, Copyright (2013), with permission from Elsevier.

To achieve this goal, different technologies are available, but none can completely fulfil this task due to instrumental limitations (**Table 2**). Therefore it is necessary to integrate different kinds of techniques in order to build up a high resolution map (high resolution cartography) where different dimensions of resolution, mass accuracy and sensitivity can simultaneously create a space where each point is described under different variables, maximizing the information achieved (**Fig. 1**). UPLC-MS offers high resolution chromatography enabling separations of the main classes of plasma compounds, additionally providing isomer identification.

The integration of ICR-FT/MS analysis gives the advantage of high mass accuracy data extraction that, combined to LC-MS data, provides an additional level of high resolution, the spectrometric resolution. This methodological approach was selected in order to extract matrix information via chromatography, combined with high accurate mass annotations from ICR-FT/MS analyses, in order to validate experiments at different levels and to overcome mass drifts typical for TOF-MS. The ultimate aim is the creation of a common data base encompassing different kinds of metabolites derived from different biological levels (urine, plasma, exhaled breath condensate, faeces, tissue) and different organisms (human, mouse, plants, micro biomes, worms). This would enable a direct comparison of different matrices and the finding of relevant connections within bio-diversity, leading to new hypotheses in terms of new biochemical pathways and their interactions on a global level.

Table 2: Comparison among the principal technologies applied in metabolomics research. Advantages and disadvantages are highlighted

Term of comparison	**NMR**	**GC-MS**	**UPLC-MS**	**Direct infusion (ICR-FT/MS)**
Range of metabolites	All the range of metabolites in high concentration	Volatile compounds, non polar volatile compounds	Polar and apolar non-volatile compounds	Polar and apolar non-volatile compounds
Sample preparation	Minimal	Derivatization of non-volatile compounds	Minimal	Minimal
Sensitivity	Low	Good	High	High
Disadantages	1. pH adjustement after metabolites extraction 2.High sample volumes 3.High cost 4. Peaks overlapping	1.Necessary derivatization for non-volatile compounds,long sample preparation 2.Introduction of artefacts 3. Low mass mass resolution and accuracy	1.Ion suppression 2.RT and m/z drifts 3.bad ionisation at high water contents→necessity of additives 4.Low mass mass resolution and accuracy	1.Ion suppression 2.Missing isomers identification 3.High scan time, limitation for hyphanethion 4.High cost
Advantages	1.Non-distructive 2.Information on molecular structure 3.Robust quatification without standards	1.High reproducibility 2.Ample libraries available, easier identification 3. Application of different chromatography 4. Separation and concentration of different classes of compounds according to physico-chemical properties	1.Detection of isomers and isobars 2.Application of different chromatography 3. Reduced ion suppression relative to direct infusion 4. Separation and concentration of different classes of compounds according to physico-chemical properties	1.High resolution 2.High mass precision formula calculation and compounds identification 3.High throughput 4.Long term ion storage for MS/MS/MS experiments

1.2 Human blood

Since ancient times blood was an object of study and interest for medical purposes. Egyptians practiced bloodletting in order to remove blood from a vein as a remedy for ill individuals; Greeks drank the blood of fallen combatants in order to acquire their vigor. In the following centuries the practice of bloodletting was still conceived as a curative expedient. In 1628 William Harvey had published his results on blood circulation and its properties, so this technique started to be questioned. From that point on it took a long time, until the 20th and the 21st centuries, to discover all blood properties and determine correct ways in handling it from the sampling until storage and analysis [85]. Blood is an important fluid that plays a fundamental role in the transport of oxygen, nutrients and messengers between different organs and tissues; it is also of great importance due to its function of clearing waste products. Blood is considered as tissue, due to its content of specialized cells, and as liquid, due to its liquid portion that is called plasma (**Fig. 2**). It is pumped by the heart in order to reach every organ through the thick net of arteries and veins, while a turbulent flow homogeneously distributes plasma and blood cells and maintains the blood composition passing by the different organs constant. Blood takes up the oxygen in the lungs and releases carbon dioxide originating from the different tissues. Food nutrients are absorbed on the gastro-intestinal level and then they circulate in the blood, hormones, waste products and molecules that are supposed to get recycled are transported through the blood stream as well [85]. Blood shows a typical red colour due to the content of hemoglobin, a protein that contains iron, the actual oxygen transporter (oxyhemoglobin).

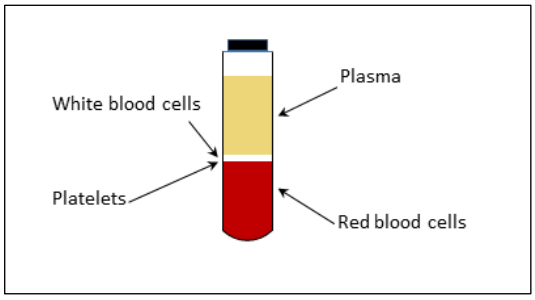


Figure 2: Human blood and its components.

This protein gets darker when is not saturated with oxygen (deoxyhemoglobin), explaining the darker colour of venous blood. In average an adult individual has 5-6 litres of blood; the half of this volume is occupied by three different kinds of cells: erythrocytes constituted by haemoglobin and specialized in the oxygen and carbon dioxide transport; leucocytes from different types which are creating the immune system and the platelets that are important for the coagulation (**Fig 3**). The blood cells (red cells, white cells and platelets) are daily generated from the bone marrow and this process is called hematopoiesis [86]. They are derived from the same hematopoitetic stem cell line, the proliferation and differentiation of different cell lines areregulated from different humeral factors and cell-cell interactions.

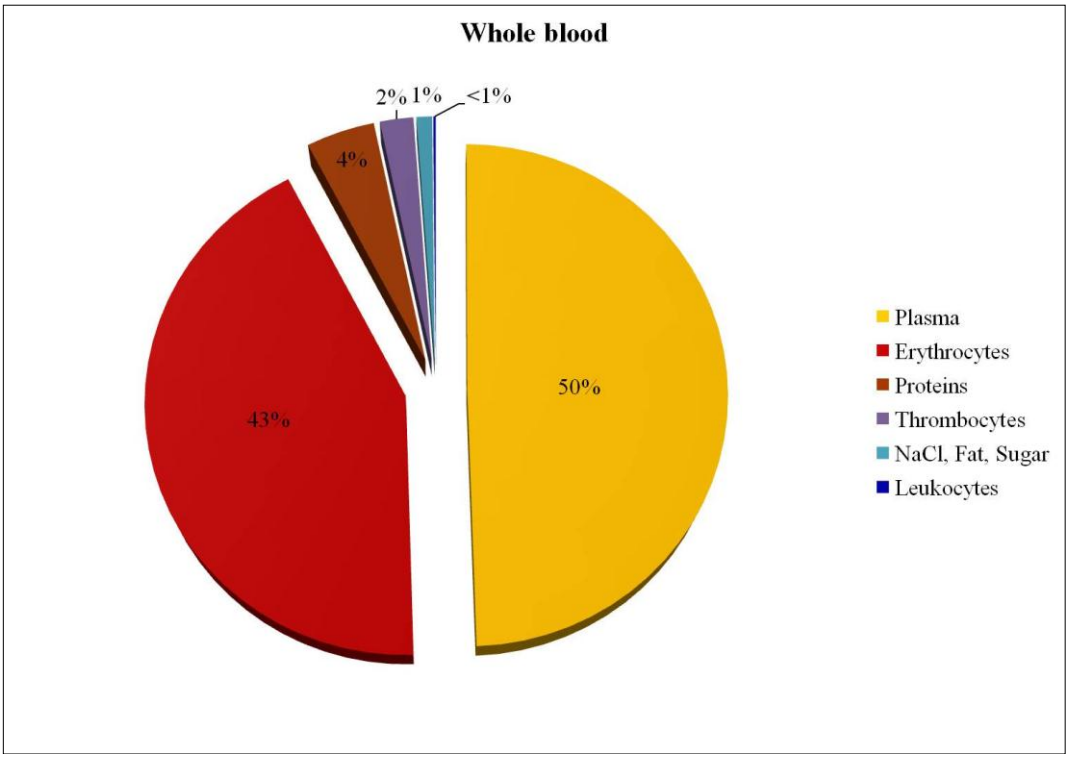


Figure 3: Whole human blood components. According to didactic material from the short course "Sample Preparation for Bioanalytical LC-MS" HPLC 2009 Dresden, Germany. Boos, K.-S., Morello, R., Laboratory of BioSeparation, Institute of Clinical Chemistry, Medical Center of the University of Munich, Germany.

1.3 Human plasma

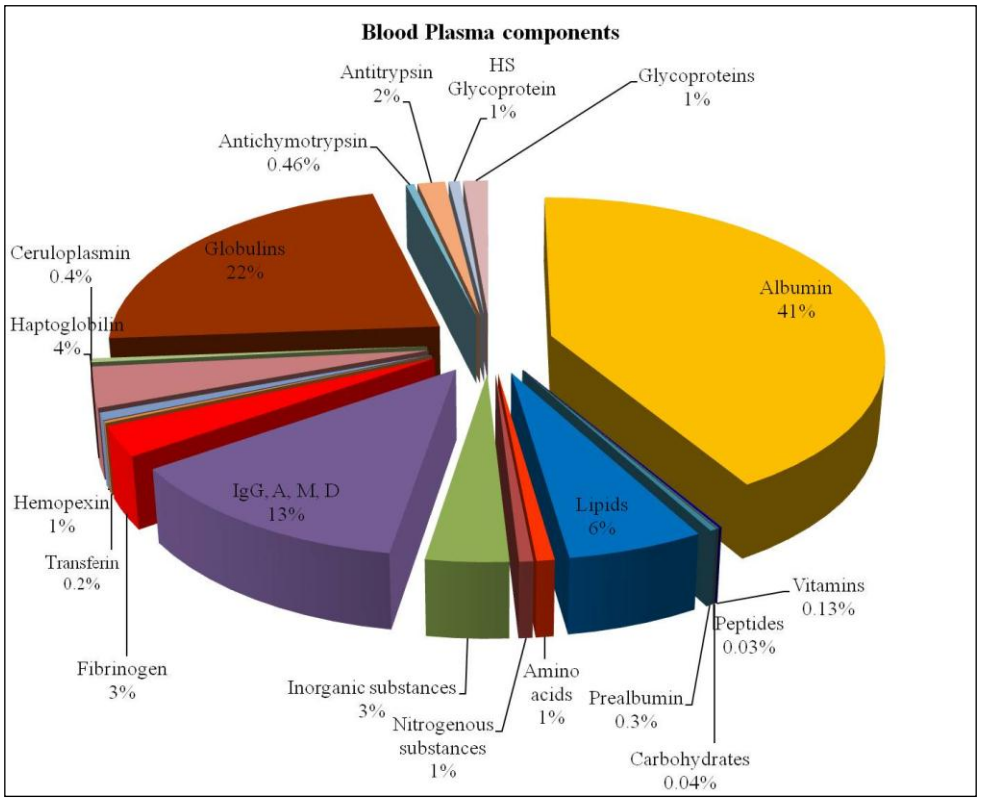


Figure 4: Blood plasma components. According to didactic material from the short course "Sample Preparation for Bioanalytical LC-MS" HPLC 2009 Dresden, Germany. Boos, K.-S., Morello, R., Laboratory of BioSeparation, Institute of Clinical Chemistry, Medical Center of the University of Munich, Germany.

Human plasma is the clear yellow liquid component of blood and it constitutes 50% of the total volume. Plasma plays an important role in the transport of nutrients to the cells of the different organs and directs the cellular waste products to liver, kidney and lungs for excretion. It is controlling the homoeostasis of the body in respect of the heat distribution and pH stability. Plasma consists of water for 90% and for the rest of proteins, electrolytes, amino acids, lipids, vitamins, carbohydrates. The chemical composition of plasma reflects the composition inside the cells (**Fig. 4**); this is the reason of its importance in clinical routine analysis. Over 70 % of the non liquid partition of plasma is made of plasmatic proteins like immunoglobulins, circulating antibodies, lipoproteins at very low density (VLDL), lipoproteins at low density (LDL), lipoprotein at high density

(HDL), albumin, apolipoproteins, proteins involved in the coagulation as fibrinogen and prothrombin and proteic transporter as transferring [87].

Albumin is the most abundant protein (41%) that is synthesized in the liver and has two important functions. It plays a role in the oncotic pressure of the blood, regulating the fluid distribution among the intercellular and extra cellular compartments. The second function, as a non-specific carrier protein, is to transport different kinds of compounds such as bilirubin, free fatty acids, calcium and drugs [85]. Other important families of proteins are globulins with different functions; immunoglobulins have an important role in the immune response to antigen attack or allergen. Other important components are cytokines, originated from cells of different organs and from cells of the immune system and bone marrow. Their role is of messengers responsible for the regulation of the cell formation (hematopoiesis); one cytokine (erythropoietin) is involved in the production of red blood cells via bone marrow stimulation, others are involved in the production of white blood cells and platelets. Cytokines perform as well an important role in the defence mechanism. Additional proteins are involved in transport mechanisms: lipoproteins, important for the transport of lipids; metal-binding protein such as transferring for the transport of iron and ceruloplasmin for the transport of copper. Alpha and Beta globulins are involved in the transport of cholesterol, steroids and sugar inthe transport of the haemoglobin to the reticuloendothelial system where hemolysis takes place. Coagulation proteins and their inhibitors are protein categories that aresynthesized in the liver. The activation of the clotting leads to the conversion of fibrinogen into fibrin for the clot formation, anomalous coagulation is avoided by the intervation of coagulation inhibitors. The widest amount of lipids in plasma is constituted of phospholipids, followed by cholesterol, triglycerides, free fatty acids. The lipids in plasma are associated with different proteins such as lipoproteins; chylomicrons are the biggest lipid aggregates consisting of triglycerides, they are absorbed from the intestine. Other lipids are integrated in the blood circulation via food or from the tissue. The concentration levels of salts such as Na^+, K^+, $Ca^{2+,}$ Mg^{2+} are monitored [86] in order to keep constant levels and to preventsome pathologies. Traces of zinc, iron and copper are necessary for the enzyme synthesis. In order to prevent

clotting, an anticoagulant such as EDTA, NaF, Citrate and Li-Heparinate can be added after blood collection. Blood is then centrifuged to enable the cell separation and obtain plasma.

1.4 Diabetes

The history of Diabetes discovery started already in 1500 BC, when extreme urination frequency along with immoderate thirst and grave loss of weight were observed trying to remedy with fruits, white cereals and sweet beer [88]. The term *Diabete* (διαβήτης) means "*to pass through*", it was introduced in 230 BC in order to depict a condition of high loss of urine. Only in the first century AD, in the "De medicina" from Aulus Cornelius Celsus, the first detailed description of Diabetes status was reported [89, 90]. Similarity with modern Diabetes concepts were introduced in Europe around the 16^{th} and the 18^{th} centuries, when the sweet taste of urine of diabetic patients was discovered and the connection between Diabetes and pancreas contemplated [91]. In the 19^{th} century the presence of glucose in urine came to light using glycosuria as screening toolfor Diabetes detection [92]. Consequently, the introduction of the term "mellitus" (from Latin, *mel*) was introduced to distinguish it from Diabetes insipidus [93]. In this period a customary remedy was a low-calories diet [91, 94], until Frederick Banting and Charles Best discovered insulin in 1921. In 1922 the first insulin treatment started [95], followed by a large scale production of insulin in collaboration with Eli Lilly Company [96], in short time its clinical use was launched. With the advent of genetic engineering, the production of human insulin viarecombinant DNA technology was possible in 1978 [97].

Diabetes is a heterogenous assembly of different disorders associated with hyperglycemia and glucose intolerance, triggered by insulin failure, impaired insulin action or both [98].

It is classified into four different groups:

I. Type 1 Diabetes
II. Type 2 Diabetes
III. Gestational Diabetes
IV. Other specific types

Type 1 Diabetes

This form of Diabetes is a consequence of an autoimmune reaction where β cells of pancreas are recognized as pathogens; therefore they are attacked by the immune system, leading to their destruction and consequent failure in insulin production [91]. This kind of disease is reported in 5-10% of all the cases and can occur at any age [94, 99], but the highest incidence is registered in infants and small children. The treatment with insulin is mandatory for survival.

Type 2 Diabetes

This form of the disease covers around 85-95 % of all diabetic patients in developed countries, showing even higher levels in developing countries. It is delineated by insulin resistance or impaired insulin secretion, or even both. The different abnormalities that are often co-existing and interacting are central and visceral obesity, high triglycerides levels, low high density cholesterol, small particles of low density cholesterol. All these factors are correlating with a risk of cardio vascular disease. The manifestation of the disease can spontaneously arise around 40 years of age, even earlier in groups with high incidence of Diabetes 2. It can manifest without symptoms and it can be subsequently being discovered via routine glucose blood or urine analysis or due to complications development. Differently from Type 1 Diabetes carriers, Type 2 Diabetes patients are not insulin-dependent but they require a strict diet and physical exercise to ameliorate their status; if this is not sufficient, medication is administered with insulin therapy as the last resort. The diagnosis is made by fasting glucose level

measurement, diabetes symptoms observation and OGTT (oral glucose tolerance test) response. Subjects that present impaired glucose tolerance (IGT) and impaired fasting glucose (IFG) are contemplated as "pre-diabetic" and the incidence in diabetes progression is high [100]. ISI Matsuda index (**Eq.1**, [101]) is employed in the evaluation of the whole body insulin sensitivity from the data obtained by OGTT. It represents the whole-body insulin sensitivity including hepatic and peripheral tissues and the ISI (composite) during OGTT, it is calculated bythe following formula:

$$ISI_{Matsuda} = \frac{10000}{\sqrt{\begin{matrix}(FPG \times FPI) \times \\ (Mean\ OGTT[glucose] \times mean\ OGTT\ [insulin])\end{matrix}}}$$

Eq. 1: ISI Matsuda INDEX, according to Matsuda, De Fronzo, *Diabetes care*, 22, 9, 1999

Where FPG is the fasting plasma glucose (mg /dl), FPI (μU/ml) the fasting plasma insulin, 10000 is a constant that consent output numbers between 0 and 12. The square-root conversion was applied for compensation of the non-linear distribution of the data. High plasma insulin concentration in the presence of a normal or increased plasma glucose concentration indicates a state of insulin resistance. Diabetes results from a combination of multiple factors, genetic and environmental; therefore its pathogenesis is still no fullyunderstood. Studies conducted on twins have reported a genetic incidence of 60-90 %, higher in monozygotic twins (70 -90 %) relative to dizygotic twins (15-35%). The rate of incidence increases with age. The development of Type 2 Diabetes and impaired glucose tolerance is inheritable with incidence until 40 % in the case of one first degree relative and 70 % if both the parents are disease carriers [102]. Investigations on ethnic variation reported high Type 2 Diabetes risk in some groups (Asian and Hispanic) and increased risk in populations that rapidly adopted western lifestyle (.i.e. Pima Indians). Another important factor isobesity [103], especially abdominal obesity is a discriminating predictor [104,105]. Additional life style predictors are the sedentariness of a subject [106], the physical exercise which plays a role in the prevention, especially among people under a high prediction risk. Concerning the cause that provokes the incipit of Type 2 Diabetes, the major discussion involves insulin resistance or bad insulin

secretion. Several studies have demonstrated that the main cause was insulin resistance [105, 107]. Insulin resistance describes an abnormal biological state where a quantity of insulin produces a smallerresponse than expected, involving mainly of the insulin goal tissues: skeletal muscle, adipose tissue and liver. When insulin functions start to decline, functions of β cells increase in order to equilibrate the system.

The level of fasting glucose and postprandial glucose will slightly increase during the time leading to glucotoxicity [108] and β cells damage [109]. The following scheme depicts how insulin secretion and insulin action are leading to hyperglycaemia and high blood fatty acids levels (**Fig. 5**).

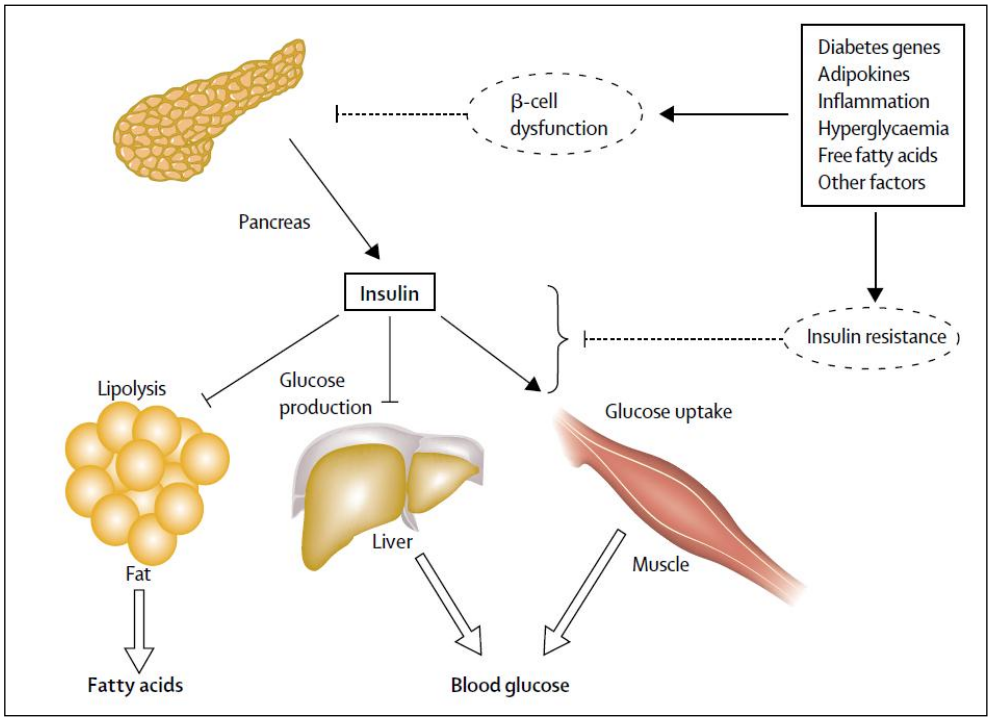


Figure 5: Insulin Secretion and Insulin Action in Type 2 Diabetes pathogenesis.
Reprinted from *The Lancet*, 365, Stumvoll M., Goldstein B. J., van Haeften T. W., **Type 2 diabetes: principles of pathogenesis and therapy**, 1333-1346, Copyright (2012), with permission from Elsevier.

When insulin works correctly it suppresses the release of glucose from the liver, decreases the lipolysis of fat and drives the muscle glucose uptake. If it does not properly work it leads to reduced glucose uptake in the muscle and in fat cells, reduced glycogen synthesis and storage in liver cells. The reduced insulin effect on lipids in fat cells will reduce the uptake of circulating lipids and increase hydrolysis of stored triglycerides and mobilisation of stored lipids. Therefore, the

scenario describes an elevated concentration of fatty acids in blood plasma, a reduced glucose uptake in muscle and an increased glucose release in liver; conducing to a high glucose level in blood. Diabetes mellitus and impaired glucose tolerance are diffused in every country and epidemiological studies suggest that prevention and control are necessary to avoid a global increase [110]. The estimations for 2025 (**Fig.6**) describe 380 milions of diabetic people and 418 milions of people affected byimpaired glucose tolerance, pointing out that the most criticaldiabetes complications are cardiovascular disease related. Additionally, the medical and socio-economic impacts need to be considered. A special importance goes to consortia that enable collaborations among different clinical hospitals and excellence research centres in order to achieve new discoveries and go through the difficult challenge of prevention.

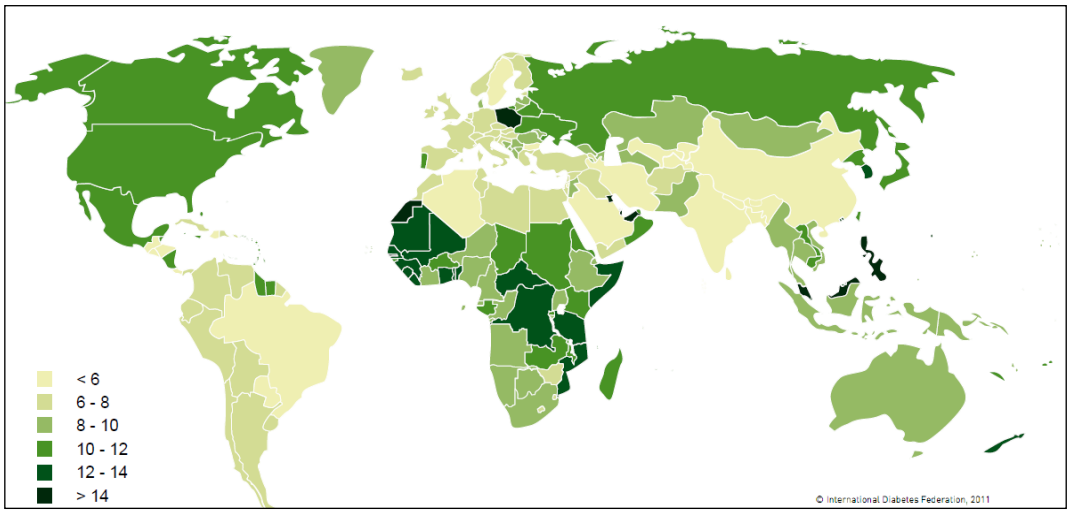


Figure 6: Map of the prevalence (%) of Impaired Glucose Tolerance in adults (20-79 years), 2030. Reprinted with permission from International Diabetes Federation. ***IDF Diabetes Atlas, 5th edn.*** Brussels, Belgium: International Diabetes Federation, 2011. http://www.idf.org/diabetesatlas.

2. Aim of the Thesis

The aim of this work is the development of a robust platform that enables the analysis of human blood plasma through a non-targeted metabolomics approach in order to define biomarkers related to a pre-Type 2 Diabetes condition in different sub-phenotypes of Tulip Lifestyle Intervention Progam (TULIP) cohort subjects, within the competence network Diabetes. In order to accomplish this goal the method development focuses on the integration of high chromatographic resolution by UPLC-MS and high mass spectrometric resolution via ICR-FT/MS. The combination of these techniques enables to reach a higher analytical resolution allowing for higher accuracy of metabolite annotation.

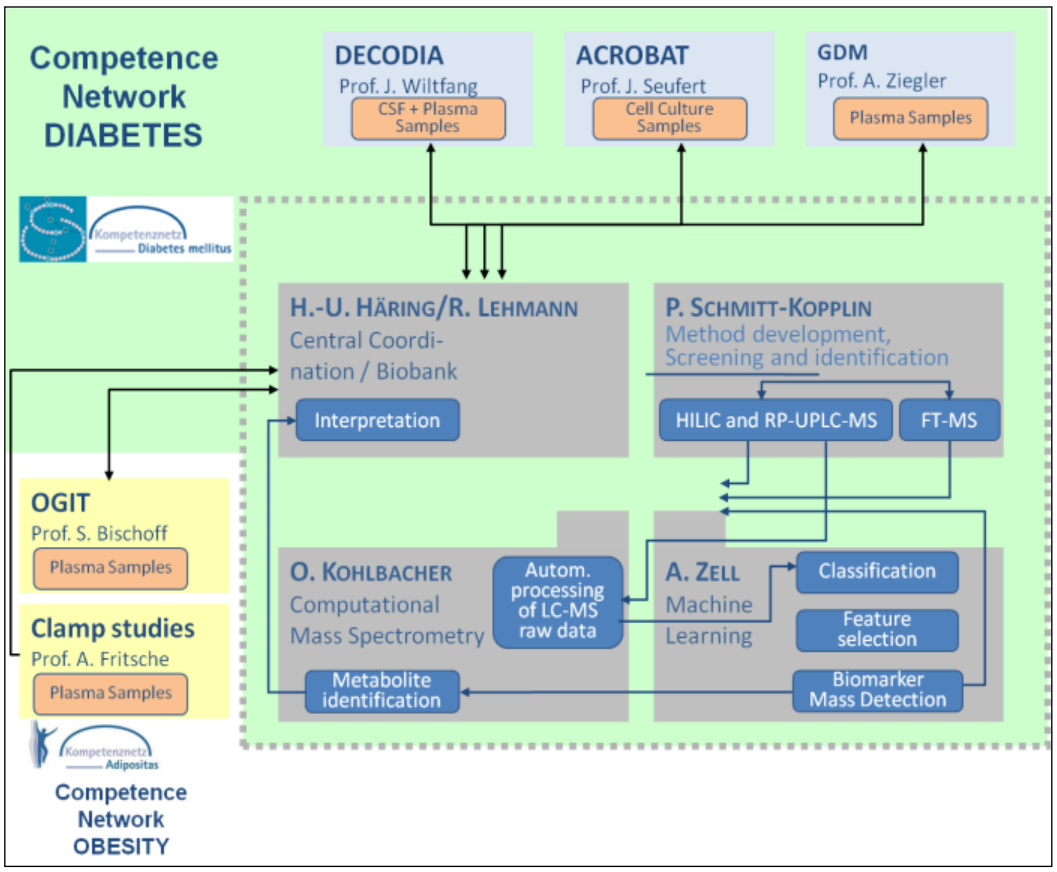
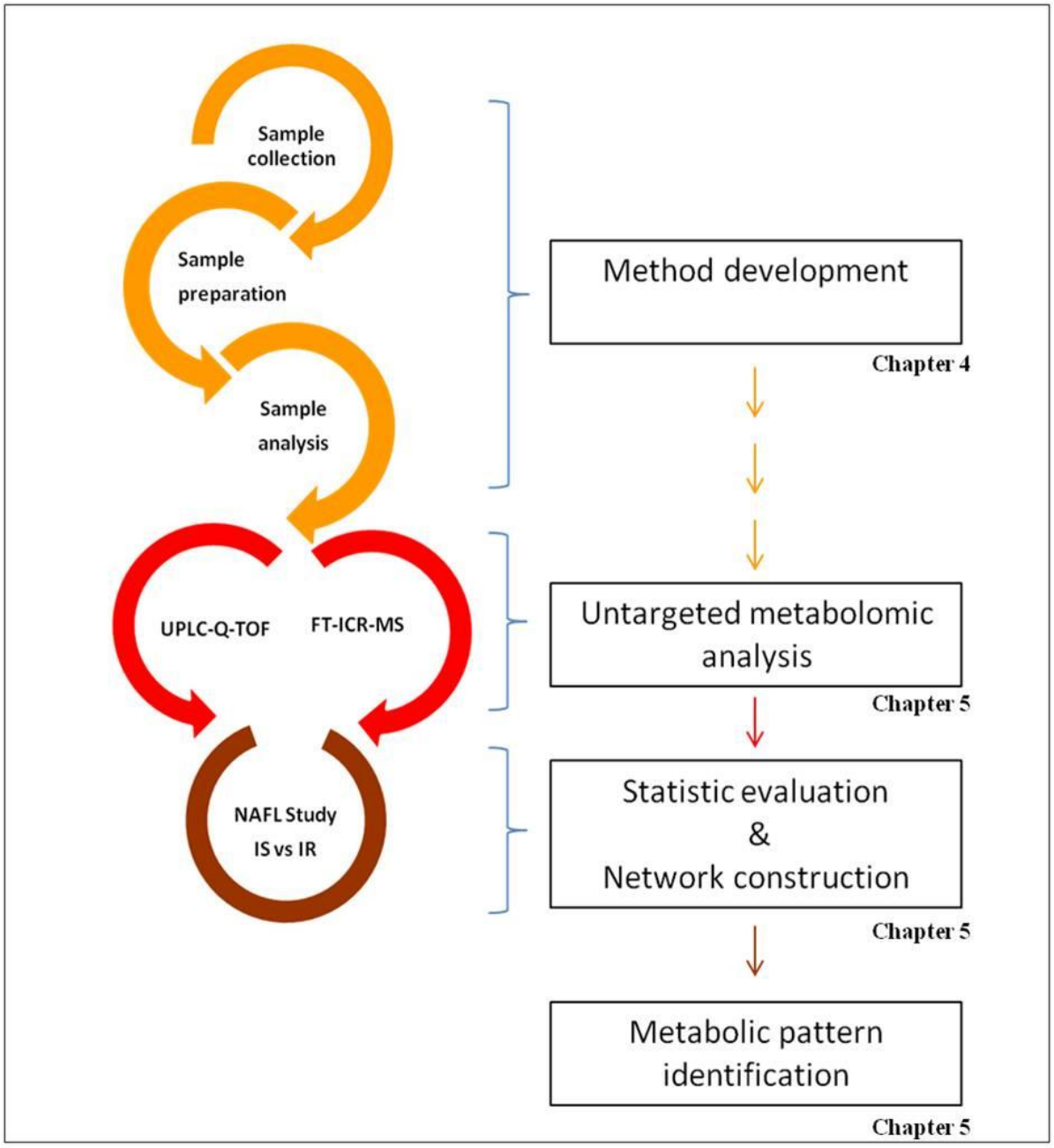


Figure 7: Schematic representation of the Competence Network Diabetes interaction.

The workflow



3. Theoretical introduction

3.1 Mass Spectrometry

Each single metabolite has its specific elemental composition and therefore its specific mass. A qualitative description of the metabolome can consequently be performed by measuring the mass of each metabolite. For this purpose mass spectrometry (MS) is applied. In addition to a qualitative description counting the abboundance of each mass, offers a measure for quantity. In order to gain a deeper understating of the mass spectrometry, it is necessary to define some basic concepts such as mass resolving power and mass accuracy [111]

Mass Resolving power ($m/\Delta m_{50\%}$): the observed mass centroid divided by the mass peak width at 50 % height for a well-isolated single mass spectral peak.

Mass Accuracy: the difference between experimental and theoretical mass of a given sum formula.

It can be calculated either as ($m_{measured}$-$m_{calculated}$) in Da or by

($m_{measured}$-$m_{calculated}$/$m_{calculated}$) x 1,000,000 in ppm.

Signal to noise ratio (S/N): The ratio of the signal intensity of an m/z peak over the standard deviation of the noise amplitude.

Sensitivity: the minimal concentration of a compound leading to a peak intensity bigger than a specified S/N. In this work, as non-targeted application, the yield of mass spectrometric peaks that satisfy a certain S/N was considered; specifically 4 in ICR-FT/MS application and 6 in Q-TOF-MS application.

In the next section a description and comparison of the mass spectrometric techniques applied in the experimental work (ICR-FT/MS and TOF-MS),will be treated.

3.2.1 Fourier Transform Ion Cyclotron Resonance Mass Spectrometers (ICR-FT/MS)

To date, Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (ICR-FT/MS) is the mass spectrometric technology with the highest mass accuracy and resolution. Its principle of mass detection is based on the circular oscillation that charged ions exhibit once they are getting introduced into a homogenous magnetic field according to:

$$F = m\frac{dv}{dt} = zv \times B \qquad \text{Eq. 2}$$

where F is the (Lorentz) force, m is the mass, z is the charge, v is velocity and B is the magnetic field strength in Tesla. A rearrangement of this Eq. yields the relationships

$$w_c = \frac{zB}{m} \qquad \text{Eq. 3}$$

and

$$m = \frac{zB}{w_c} \qquad \text{Eq. 4}$$

where w_c is the cyclotron frequency of the mass m given the charge z and a magnetic field strength B [111]. This relationship states, that a particle of a specific mass has its specific ion cyclotron frequency. In a typical ICR cell all ions of a sample are oscillating with their specific frequenciesw_c around the z-axis of this cell. By superimposing the cyclotron frequency of a specific molecule with an oscillating RF-field of the same frequency this specific molecule can be moved away from the z-axis into an orbit more distant to the z-axis, where detector plates are placed. At this post-excitational orbit the frequency and the presence of this molecule can be detected. The schematic representationof the ICR excitation is depictedin **Fig. 8**.

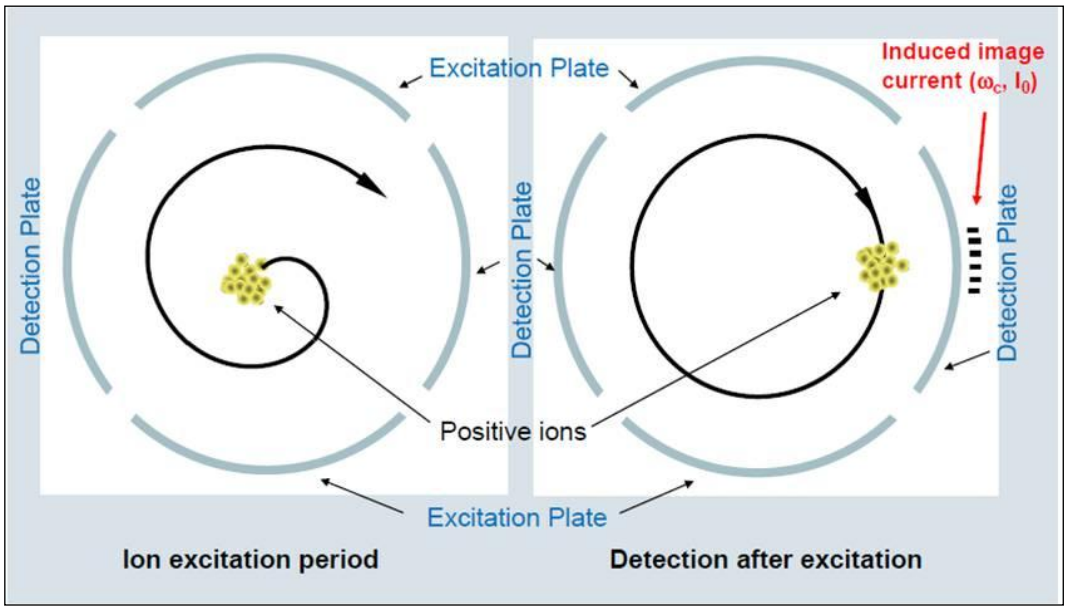


Figure 8: Schematic representation of the ICR process. Ions inserted into the ICR cell start to oscillate at mass specific frequencies at the center of the cell. Superimposition with the resonance RF frequency leads to excitation in proximity of the detector plates. The ion cyclotron frequency is measurable as an image current.is later fourier transformed into mass spectra. Reprinted with permission from Users Manual of Bruker SolariX FTMS.

Eq. 3 shows that the higher the magnetic field strength, the more distinct the frequency difference between two masses. Therefore resolution, mass accuracy and scanning speed increase with increasing magnetic field strength [112].

The frequencies of molecules between 150 and 1000 Da typically ranges between kHz and MHz. In order to acquire each masses frequency for a sufficient time period scanning times of usually between 1s and 2s are necessary.

The per-scan sensitivity of ICR-FT/MS mass spectrometers is generally relatively low [113]. The reason may be that all ions produced from a sample coexist in a spaciously limited orbit around the ICR-cell z-axis. This causes the equally charged ions to interact. Repulsions between ions and consequently different thermal states of the ions lead to off-resonance excitations. Ions of low relative abundance may be absorbed up by noise. For this reason consecutive ICR-FT/MS scans are usually superimposed. This way reproducible occurrence of certain masses enhances peak formation, whereas noise signals cancel each other out.

Consequences:

Due to the long scanning times required to provide high resolution, ICR-FT/MS is not very well suited for LC-MS coupling. This is especially pronounced since developments in LC-MS have focused on increased peak capacity, shorter gradient times and consequently shorter chromatographic peaks. Generally at least 15 data points per peak width are required in order to achieve one mass-spectrometrically well resolved chromatographic peak [114].

For this reason ICR-FT/MS is best performed using direct injection electro spray ionization. This technique delivers constant ion flow enhancing the sensitivity and accuracy of ICR-FT/MS. These assets of ICR-FT/MS as well minimize the lack of chromatographic resolution in terms of differentiation of isobaric masses, a major drawback of other spectrometers.

3.2.2 Time of Flight mass spectrometers (TOF-MS)

The second type of mass spectrometer used in this work was time-of-flight mass spectrometer (TOF-MS). The principle of this mass spectrometer is based on the measurement of the time, that ions take to travel from the beginning to the end of a field-free flight tube, where a detector plate is placed (settled 1 to 2 meters from the source) [115, 116]. The potential energy of a charged particle in an electric filed is expressed in Eq.5, with V the accelerating potential, z the nominal charge of an ion and e the its electronic charge.

$$E_p = zeV \qquad \text{Eq. 5}$$

Once the particles have beenaccelerated into the flight tube, all the ions that are passing through the tube have received the same amount of potential energy. Their final velocities depend on the m/z ratio. The potential energy of each charged ion is equivalent totheir kinetic energy, therefore the Eq.6 is valid. Eq. 6, resolved for the velocity (Eq. 7), displays the inverse quadratic relation between m/z and the velocity. The higher the mass of a molecule is, the lower is its velocity.

$$zeV = \frac{1}{2}mv^2 \qquad \text{Eq. 6}$$

$$v = \left(\frac{2zeV}{m}\right)^{\frac{1}{2}} \qquad \text{Eq. 7}$$

Given a constant distance (L) from the acceleration point to ion detection, molecules of different mass will pass this distance at different, mass specific times. The complete relationship defining the TOF is described in the Eq. 8

$$TOF = \frac{L}{v} = L\left(\frac{m}{2zeV}\right)^{\frac{1}{2}} \qquad \text{Eq. 8}$$

As an example, the time of flight of an ion of m/z = 1000 with a potential difference of 15000V and a distance s of 1.5 meters is calculated:

$$t = \frac{1.5m}{\sqrt{2\times 15000V}}\sqrt{\frac{(1000Da)(1.672621\times 10^{-27}KgDa^{-1}}{1.602\times 10^{-19}C}}$$

$$t = 2.792\times 10^{-5}s$$

Clearly deducible from Eq. 8: the longer the distance of the flight path is, the greater the difference in time of flight. The example calculation for m/z = 1000 shows that 33647 scanning events could be possible in one second of scanning time. Scanning rates of 10 to 20 Hz at simultaneous detection of a mass range of 150 to 1000 m/z are therefore feasible with this type of mass spectrometer which makes it so well fit for LC-MS applications [112]. Since time of flight measurement necessitates a starting time at which one would start to count the flight time, orthogonal TOFs were developed. To enable time discrete ion introduction ions are normally first accumulated into ion packets. These are then forwarded to the beginning of the flight tube, where all ions have a more or less equal starting position. An electric pulse, emitted by a pusher, accelerates the ion packets orthogonally either directly to a detector or to a reflectron. In case of instruments using V- or W-mode the reflectrons reflect the ion front to the detector. These modes increase time of flight and thus accuracy and resolution. However, since ions do not have equal thermal and potential energies before they are pushed to the reflectron or detector due to different instrumental

characteristics (ion optics), each instrument comes with different mass dependent distributions parallel to the pusher. Therefore, polynomials of higher degree need to be used for mass calibration. In order to maintain the consistency of the ion packet, the studies in this work were generally performed in V mode in order to find a good compromise between overall spectrometric performance and detection ranges. The general setup of the SYNAPT-Q-TOF is depictedin **Fig. 9**.

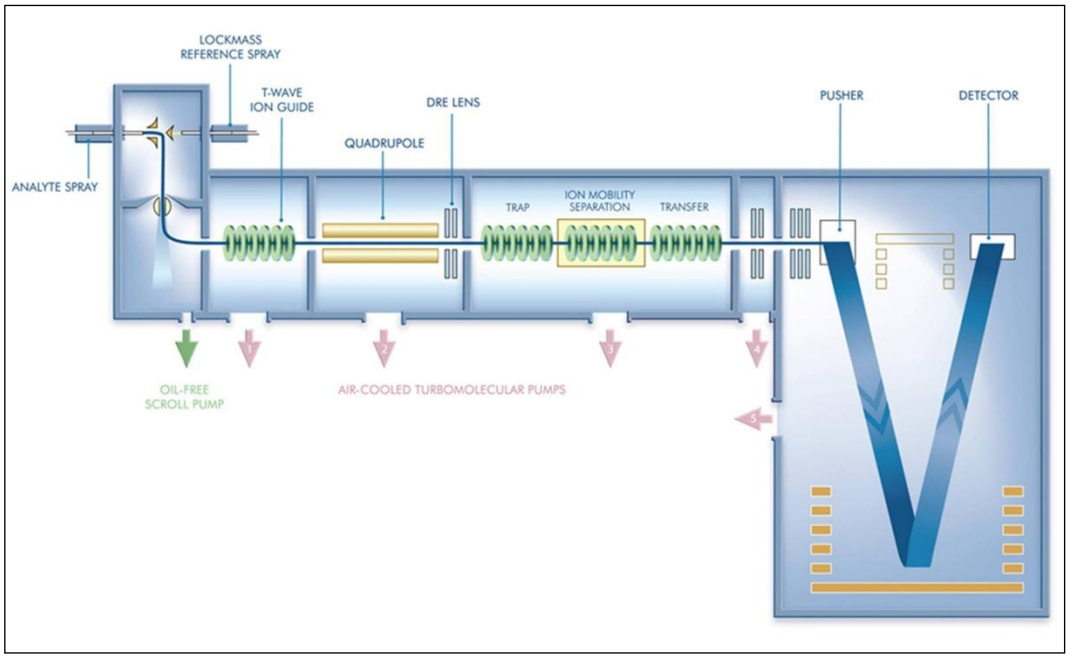


Figure 9: Schematic representation of Synapt G1 from Waters. Reprinted with permission from Synapt users training, Waters Corporation, Milford, USA.

However, TOF mass spectrometers suffer from further limitations such as impaired resolution and mass accuracy, especially when UPLC requirements needto be met (high number of scans in shortest time possible). TOF detectors need to be able to deploarise in very short time scales. Another factor limiting mass accuracy and resolution is the stability of flight trajectory, that can be impaired by temperature and gas pressure change. Slight deviations in the flight direction leads to peak broadening, bad resolution and bad mass accuracy. Additionally, the performance of TOF mass spectrometers strongly depends onthe heat insulation of the instrument and therefore on the room temperature. These limitations and many more, limit the typical resolving power and mass accuracy of TOF-MS to 10000-50000 and 1ppm to 5ppm, respectively. Nonetheless, TOF-MS is currently the type of mass spectrometer best suited for the time requirements of modern high resolution liquid chromatography. The above

mentioned typical specifications of TOF-MS necessitate a chromatographic separation prior to direct infusion electro spray injection, since isobaric masses are often not distinguishable. Independently from chromatographic steps, eventual isotopic masses overlapping each other, such as ^{14}C and ^{14}N isotopes are not differentiable.

Critical comparison of TOF-MS and ICR-FT/MS when coupled to liquid chromatography

Both instruments naturally have their advantages and disadvantages. Where ICR-FT/MS offers unsurpassed resolution and mass accuracy, TOF-MS has high scanning rates and higher perscan sensitivity. Therefore, to maintain the superior mass accuracy of ICR-FT/MS it is necessary to accept long scanning times and inherent quality loss when coupled to liquid chromatography. Another factor limiting resolution and mass accuracy is the ion density in the ICR cell. Interactions and repulsions of equally charged molecules limit the upper limit of its dynamic range. Since LC peaks typically arrive in packets, too strong accumulation of chemical species on the columns (like in UPLC) may cause ICR cell overloading once the center of a chromatographic peak arrives to the spectrometer. To counter limitations at the upper end of dynamic range, selected ion monitoring (SIM)-stitching was developed [117]. This method works by combining multiple narrow-range spectra into one wide range spectrum. Typically quadrupolar ion filtering will only permit ions within a mass window of 30 Da size to populate the ICR cell. Consequently interactions between molecules are supposed to be decreased which increases the systems dynamic range. Unfortunately such workflows are even more time consuming than normal ICR-FT/MS operation modes. Therefore, such developments do not increase the compatibility of ICR-FT/MS to liquid chromatography. Literature references describing LC-ICR-FT/MS coupling reported an average resolution of 50000 [118]. This is comparable to using the newest TOF mass spectrometers but does not justify the difference in costs (around 3Mil € vs. 200000€). Limitations in TOF mass spectrometry that could impaire resolution and mass accuracy are

mostly derived from spatial and velocity spreads across ion packets, stability and linear range of electric devices, mechanical precision and thermal insulation. Consequently, we can state that the most effective compromise, of the presented advantages and disadvantages of both spectrometers, is to scan a sample via direct infusion electrospray ICR-FT/MS and in parallel to perform a UPLC-TOF-MS screening unifying the results yielded from both techniques.

3.3 Liquid Chromatography

The principle of chromatography was discovered by Mikhail Tsvet in 1900. When he extracted plant leaves in chloroform and flushed the resulting fluid over a column filled with sand, he observed that differentially colored fronts would elute from the packed column. These observations gave rise to the term chromatography, which originates from Greek and means "*writing in color*" [119]. The general principle behind chromatography is based on the partition equilibrium of two phase systems. If a compound is introduced into such a system it will be physic-chemically more or less similar to one of the two given phases. In order to minimize the energy of the whole system the compound in question will partition itself between the phases. Consequently, an equilibrium concentration of the compound in both phases will be reached. Such equilibrium exists for every compound within every kind of two phase system, such as liquid-liquid, gas-gas, liquid-gas, gas-solid or liquid-solid. If a two phase system is defined, it should be possible to differentiate compounds based on their compound specific partition constant between the two defined phases. Partition coefficients are formulated as follows:

$$P = \frac{[x]_{in\ phase\ 1}}{[x]_{in\ phase\ 2}} \qquad \text{Eq. 9}$$

P is a concentration ratio between the two phases at equilibrium. If a mixture of two compounds is separeted using a two phase system with compound A having P_A = 10 = 10/1 and compound B having P = 0.1 = 1/10 and both their concentrations in the whole system are 1 mol/L, after reaching equilibrium the following partitions would be reached:

$[A_{P1}]$ = 0.91 mol/L; $[A_{P2}]$ = 0.09 mol/L; $[B_{P1}]$ = 0.09 mol/L; $[B_{P2}]$ = 0.91 mol/L.

Once such extraction processes are run, batch wise or steadily under separation of one phase, the Ps will multiply. After 100 extractions the overall extraction factor for A would be 10^{100}.

Chromatography works with phase A being a fluid (mobile phase), which is loaded with the molecules to be analyzed (analytes), and phase B being a surface (stationary phase) with which the analytes interact. Since A runs over B, leaving B "behind", this process is exactly what was described as a steady process where the overall partition coefficient is finally P^{Nt}, whereN_t isthe number of theoretical plates (i.e. the number of extraction steps). The act of flushing analytes from one theoretical plate to the next is called elution. N_t is generally expressed by:

$$N_t = \frac{H}{Height\ equivalent\ to\ a\ theoretical\ plate} \qquad \text{Eq. 10}$$

Where H is the linear path of an analyte and HETP is the height equivalent to a theoretical plate. Applying thin layer chromatography (TLC), H is basically equal to the length of the used silica or Alox plate or it may be the distance of the furthest band to the spot of analyte application. In chromatography applications wherepacked columns are used, H is the column length. The HETP can generally be explained by the Van Deemter [1] equation (Eq.11), which is formulated as follows:

$$HETP = A + \frac{B}{u} + C \times u \qquad \text{Eq. 11}$$

Where A is a term for Eddy-diffusion, B is a term for longitudinal diffusion, C is a term for mass transfer and u is the linear velocity at which the mobile phase passes by the stationary phase [1]. The separation principle, especially the partition coefficients are hidden in the C term. The other two terms are applied to give an idea of spatial spread. Eddies are vortexes which occur in turbulent regions of a moving fluid. These eddies are used in turboflow chromatography [120] where big beads and high flow rate ensure eddy formation and small molecules are trapped inside them whereas big molecules are getting ejected by the eddies centripetal force. In the chromatographic sense Eddy-diffusion also describes diffusion orthogonal to the z-axis of the chromatographic column. Both

effects cause a deviation of analytes from the linear path of analytes. The bigger these particles are, the stronger the deviation from the linear path. Another factor determined by the bead size is the free path the analyte may take through the solvent without being in contact/interaction with the stationary phase. The bigger the beads are, the bigger are the cavities in the packed column and worse the separation. The longitudinal diffusion term further describes the broadening of the peaks due to back and forward diffusion which is combined with the mobile phase velocity. The bigger the velocity becomes relative to the diffusion, the smaller gets the B-term which in the end decreases the HTEP and increases N_t. The linear velocity of the mobile phase will increase with decreasing column diameter and decreasing bead size, the latter decreasing cavity diameter. Deeper analysis of the van Deemter equation determined the trend in packed column liquid chromatography. As scientists recognized, decreased bead diameter and decreased column diameter would greatly benefit the number of theoretical plates. In order to maintain the linear velocity of the mobile phase, high pressure is necessary to be applied. Consequently, scientists had to switch from packed glass columns with gravity driven elution to metal columns with elution driven by hydraulic pumps. This was the advent of High performance pressure liquid chromatography (HPLC). Consideringthe initial example on partition coefficients, with compound A having $P_A = 10 = 10/1$, we deduced that P^{Nt} would be 10^{100}. If the phase that A had high affinity to was the stationary phase, it is obvious, that A would elute from the column in practically infinite time. This problem persists for many analytes as long as the composition of the mobile phase is kept constant (i.e. whenthe elution is isocratic). To counter this problem the gradient elution was introduced. In this case the partition coefficient P is modulated by gradually mixing the initially pure mobile phase with a second mobile phase, which is closer to the properties of the stationary phase. This second phase has a high affinity to the stationary phase and, as its concentration increases, it would replace the analytes previously adsorbed to the stationary phase washing them off the column. In order to further bypass problems (i.e.density, viscosity, measurement time etc.) chromatography developed from HPLC to Ultra high pressure liquid chromatography (UPLC) [121]. These advancements, especially due to minimized particle sizes, finally led to a minimization of the necessary sample volume, better

chromatographic separation and resolution as well as shorter measurement times (**Fig. 10**).

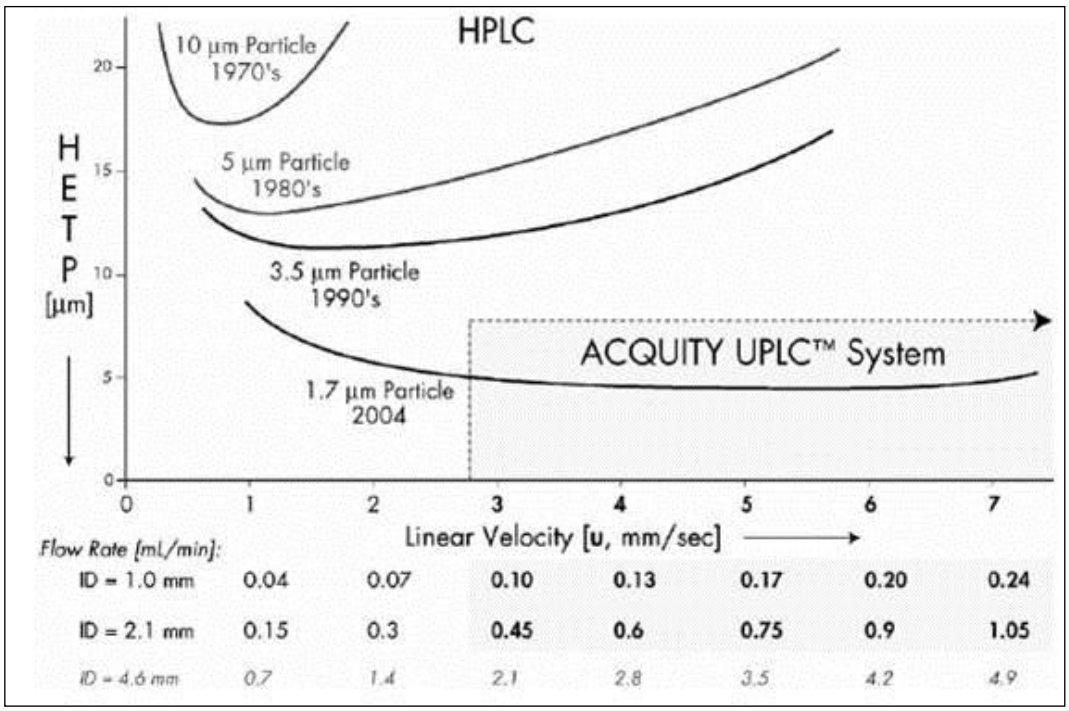


Figure 10: Evolution of different particle sizes from HPLC until the advent of UPLC.
Reprinted from Swartz, M. E., **UPLC ™: An Introduction and Review**, *Journal of Liquid Chromatography & Related Technologies* 2005, 28, 1253–1263.

The basic principle of chromatography has stayed the same over the years but its variety has increased. In the next section the basic types of chromatography, followed by the important issue of metabolite/analyte detection will ba discussed.

Types of Liquid Chromatography

The composition of the stationary and mobile phase, as well as the properties of metabolites/analytes, needto be considered in order to perform an effective separation.

Liquid Chromatography types: Normal Phase Chromatography and Reversed Phase Chromatography

The terms normal phase chromatography and reversed phase chromatography may be confusing since they depend on the target to be analyzed. Both terminologies were implemented targeting polar analytes. For polar analytes it is considered "normal" that they tend to associate with polar phases. Therefore, if the stationary phase (or sorbent) is polar itself the association of likewise polar analytes to it is normal. Consequently, polar analytes get introduced into acolumn, as they are dissolved in an apolar phase, and they get eluted by increasing the hydrophilicity of the eluent. Reversed phase (RP) chromatography, from the perspective of a polar analyte, is based on apolar sorbents instead. Analytes get introduced into a column in an aqueous solution and they interact with the stationary phase. They get eluted by increasing the apolar partition of the eluent. Since normal phase chromatography became more pronounced in the recent years, it was decided to term this chromatographic mode "Hydrophilic Interaction Liquid Chromatography", in short HILIC, in order to avoid the before mentioned terminology. Both techniques RP-LC and HILIC come with their own requirements. As described above apolar stationary phases are used in RP-LC whereas polar sorbents are used in HILC. RP-LC sorbents are typically either silica based, polymeric or a hybrid of both. HILIC sorbents are either naked silica beads, with free or substituted silanol groups on their surface or polymeric or hybrid beads with hydrophilic substitutions (e.g. zwitterionic) [122]. The reason for the late success of HILIC becomes more evident: originally only "naked" silica based materials were available for this purpose. Their production was not as refined as it is nowadays. Due to impurities and the chemical properties of

silicium (SiO)-beads, they were mainlystable over a pH range between 2 and 8. Outside of these limits the beads started to dissolve. RP-LC, in turn, necessitated "end-capping" with the substitution of the free silanol groups. This is the reason for its easier handling. RP-LC is based on sample introduction in aqueous phases which often saved working time and was less prone to errors. RP-LC beads presented some disadvantages as well. Before more refined end-capping and hybridization technologies were developed, RP materials still contained substantial and variable amounts of free silanol groups which made them vulnerable towards extreme pH values as well. Furthermore, the retention mechanism may have not purely been of RP character. Analytes could indeed exhibit hydrophilic interaction with free silanol groups. Consequently, solvents had to be highly buffered and salt needed to be added in order to minimize side interactions (this procedure is still very common in gel permeation chromatography/size exclusion chromatography when proteins get separated using fast protein liquid chromatography). Additionally, an excessive use of salts and additives could damage HPLC and UPLC systems considerably. Due to recent advancements in manufacturing, clean SiO-beads and the production of polymeric materials enable sorbents to withstand the high pressures of UPLC. Therefore, HILIC and RP-LC have become more stable and highly reproducible for LC systems. After a general introduction on basic concepts, an overview of modern sorbents [123] is provided.

Reactivities in RP

Acyl particles. SiOH groups are substituted with C_4, C_8, C_{18} or even C_{30} groups. The shorter the acyl chain is, the lower is the affinity to apolar compounds. If lipids are not elutable from atypical C_{18} material, decreasing the stationary phase hydrophobicity can be a solution. Furthermore, polar compounds may be better resolved. Peak broadening may become a problem. C_{30} groups do not provide significantly higher hydrophobicity than C_8 or C_{18} materials. Due to the flexibility of the C_{30} chain, the separation between acylic cis-trans isomers becomes possible.

Phenyl particles. Mostly an acyl-benzyl group, not a phenyl group, is used. The acyl chain is bound to free SiOH groups in order to provide flexibility. The acyl group is, at the same time, a spacer for the benzyl group which is supposed to exhibit higher selectivity for aromatic compounds. The disadvantage of these particles is that no typical eluotropic row can be defined. According to our experience neither calculated LogD values nor a calculated, sum formula based aromaticity index were able to predict the eluotropic order. Nonetheless, the capacity to separate aromatics is enhanced.
CN particles. They are applicable for both, RP-LC and HILIC. CN-particles exhibit less retention than silica in HILIC mode but they offer an orthogonal alternative to C_8 or C_{18} in RP-LC.

Reactivities in HILIC
Unbonded particles. Undbonded particles used nowadays are rarely purely silica based. Most unbounded HILIC particles are silica-polymer hybrids which are stable against high pressures and corrosion.
Amide particles. Normally a $CONH_2$ group is linked to a spacer, which is connected to silanol groups. This type of functionality is supposed to provide an orthogonal second dimension in HILIC compared to other functionalities.

Implications for Mass spectrometry – Ionization

Various detection modes are possible to be coupled to LC-MS. The most popular systems are UV-VIS detection, fluorescence detection and/or mass spectrometric detection. The first two methods alone are not particularly suited for metabolomics. UV-VIS absorption and fluorescence are specific for compound classes but not as specific for single compounds which we want to detect/identify in metabolomics. The standard detection method is therefore mass spectrometry. While UV-VIS and fluorescence detection do not provide complications in on-line signal detection, mass spectrometry comes with its own draw backs. Before an eluted analyte can be measured using MS it first needs to be ionized and

transferred from liquid phase into gas phase. This is generally done by spraying the sample through a metal capillary circumvented by a turbulent air stream. Based on this principle three major ionization methods are available: Atmospheric pressure photo ionization (APPI), atmospheric pressure chemical ionization (APCI) and pneumatically assisted electro spray ionization (ESI) [124].

Applying APPI, the sprayed sample solvent is immediately vaporized by spraying it through a cylinder which is heated to several hundreds °C. The vaporized sample is exposed to highly energetic UV beams which abstract electrons out of π-bonds. This causes either a dislocation of the charge via conjugated π-systems, and therefore stabilization of radical species, or interaction with these radicals with surrounding molecules such as residual water molecules. These processes may cause chain reactions where positive and negative radical species and de-protonated and protonated ions can manifest. With APPI positive and negative ions can be generated by one source [125]. A big drawback is the complexity of APPI spectra. Moreover, UPLC systems produce such high solvent flows that complete evaporation of the solvent prior to ionization may not always be possible (this problem concerns APCI as well). A solution to this problem may be flow splitting. APCI relies on vaporization of the solvent prior to ionization as well [126]. Downstream to the gas flow a needle shaped electrode is placed at the center of the gas stream. This electrode gets charged either negatively or positively. The strong electric field gradient, around the needle tip, causes naturally occurring electrons in gas phase to accelerate so strongly, that collisions with other gas molecules cause an avalanche of electrons. This process is called corona discharge. Typically vicinal N_2 is converted to N_2^+, which then ionizes sample molecules. Since complex chain reactions are part of this ionization process, this technique is not suitable to metabolomics screening. In targeted studies though, it may indeedbe possible that complicated ionization mechanisms even aid compound identification.

ESI Ionization

The method of choice in most metabolomics MS application is the electro spray ionization (ESI). It is favored by most scientists since it has a well balanced ionization efficacy in terms of chemical compound classes and is a softionization method. This ionization method does not rely on rapid, but on the gentle vaporization of solvent, since ions can be generated in solution. The ESI capillary, which introduces the sample spray into the system, is placed at a close distance to a counter electrode. The potential difference between the ESI capillary and the conter electrodeusually ranges between 3000V and 4500V. This electric potential difference depends on the distance between the ESI capillary and the counter electrode. Usually an electric field of 1000V/cm is maintained [127]. This strong gradient of electric potential, which decreases logarithmically as the distance to the ESI needle tip increases, causes positive ions to accumulate at the tip of the spray capillary. In negative mode negative ions take their place. Sample molecules are therefore embedded in an environment of either positively or negatively charged ions. As droplets are released from the capillary tip, the solvent in which the analytes and the enriched ions are dissolved gradually evaporizes. This enhances the charge density and due to coulombic repulsion the shrinking droplets finally are ripped apart into smaller droplets. At the same time the chemical activity of the respective ions increases, resulting in charge transfer and neutral-ion adduct formation [127]. The whole process is summarized in **Fig. 11**.

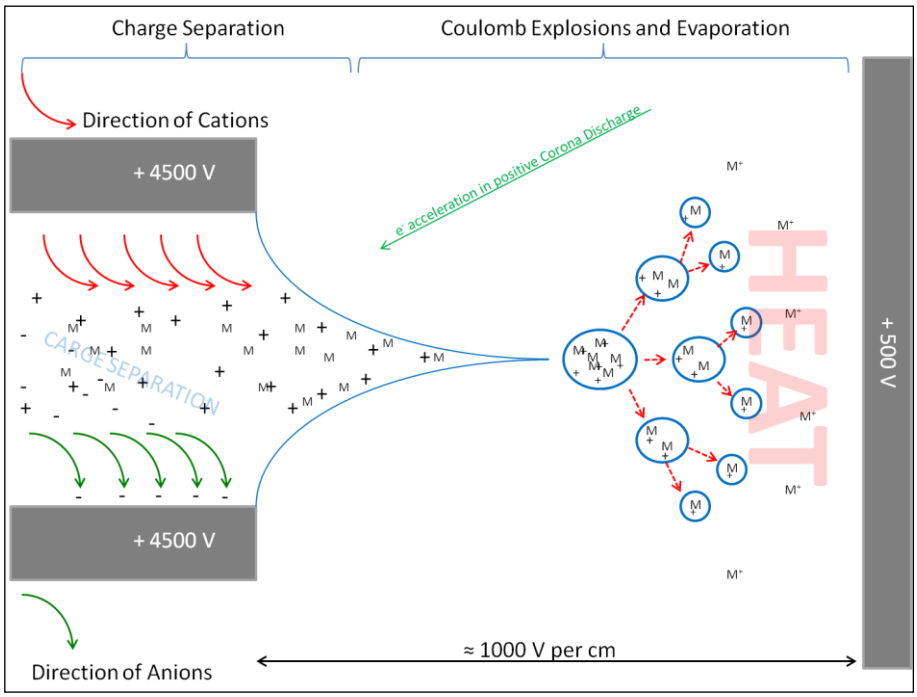


Figure 11: Schematic representation of the electrospray ionization [255].
Reprinted from *Journal of Chromatography A*, 1292, Forcisi S., Moritz F., Kanawati B., Tziotis D., Lehmann R., Schmitt-Kopplin P., **Liquid chromatography-mass spectrometry in metabolomics research: mass analyzers in ultra high pressure liquid chromatography coupling**, 51-65 Copyright (2013), with permission from Elsevier.

Fine tuning of this process even enables to ionize non-covalently bound complexes. This is the major advantage over APPI and APCI and the mass spectra do not get unnecessarily complicated. ESI ionization is not by default more suited for LC-MS coupling, since as well for this ionization method flow rate may easily prove to be too high for effective evaporation. Due to the fact that solvent is not completely evaporised, the ESI process is sensitive to changes in solvent composition as it occurs in gradient elution. The bigger the water proportion, the higher the surface tension the lower the gas pressure of the solvent and the higher the density. This ultimately leads to poor evaporation/ionization and/or instable spray. Countermeasures can be the addition of dopants (ions) in order to increase the ion density and coulombic repulsions. It is necessary to consider that salts in organic phases easily precipitate, given the high temperatures in the ESI chamber, and that this couldlead to clotting of the capillary and instable sprays. Since the major advantage of ESI ionization is its softness and the production of less complicated spectra, it is mandatory to tune the ionization process as to not cause corona discharges. As described above corona discharges are caused by electron

avalanches. The probability for the effective occurance of this event is thus polarity dependent. In positive ionization mode the spraying capillary (which usually is run with pneumatic assistance) would draw anions and electrons to itself. The spray direction and gas pressure counteract electron avalanches in this case. Therefore corona discharge is a minor problem in positive mode ionization. In negative mode ionization, the acceleration of anions ans electronsgoes with the direction of the spray and is thus not hindered. The probability of corona discharge is considerably higher in this mode. Where in positive mode potentials of 1000V/cm do not provide a problem, in negative mode they do. A good countermeasure is to split the flow and therefore decrease to corona discharge driving force. This may enable to still produce a gentle spray at high potentials. Another issue to be aware of, using ESI, is ion suppression [128]. Five factors in ESI exhibit a selective force on analytes: the electrophoresis-like charge separation in the ESI-tip (and therefore the pK_a), the ion size, the surface activity of metabolites (and therefore LogD), the number of ionizable functionalities/molar fraction of an analyte and their gas-phase proton affinity [129, 130]. Accordingly, there is strong competition for charge transfer. For these reasons lipids, which often have high surface activity and are abundant in number, have the tendency to suppress the ionization of other compounds. The same goes for inorganic ions and for proteins. For all these reasons the addition of dopants to the eluent of LC-MS separations needs tobe well evaluated [128].

Implications for building up a routine metabolomics platform

When building up a metabolomics platform the following things have to be considered:

1) As we are applying non-targeted metabolomics studies we need to choose techniques with a broad sensitivity range. Therefore, the chromatographic techniques to be used are RP-LC and/or HILIC; foremost RP-LC due to the solvent exchange necessary for HILIC separations. Other phases are not suitable for metabolomics studies since they are too specific. Cation exchangers retain cations only as much as anion exchangers retainanions only. Affinity chromatography based on antibody-antigen interaction is not suitableas well. Size

exclusion chromatography, which would eventually effectively counter suppression effects in ESI ionization, does not provide orthogonal information.

3.4 Sample Preparation

Independently of any experimental contex, the aim of sample preparation is always to separate a compound of interest (the analyte) from the sample or matrix it is embedded into. Such procedures are supposed to work in a reproducible way and the remaining analytes should, unaltered by the sample preparation step, reflect its abundance in the real sample. Sample preparation procedures in biochemical studies need to be non-destructive in the first place. Performing a metabolomics study, it is necessary to be aware of the composition of the sample matrix used. As shown in Chapter 1.3, human plasma consists for a big partition of proteins/peptides as well as inorganic species. Irrespective of the ionization technique used downstream in the workflow, these compounds either carry several ionizable functionalities or in case of metals have several oxidation states. On one side they are very potent competitors for ionization since they scavenge energy which leaves less energy left for the metabolites of interest. On the other side multiply charged compounds complicate mass spectra and formula annotation and in instruments with lower resolution their peaks may overlay metabolite peaks which interfere with the detection. Consequently, sample preparation in metabolomics must aim at the removal of proteins, polymers and salts/metals without altering the actual relative and, if possible, absolute concentration of metabolites. The techniques most commonly applied for these purposes are solid phase extraction (SPE) and protein precipitation extraction (PPE) [131, 132]. SPE uses the same principles explained above in the liquid chromatography section, namely the partition of compounds between two phases. Typically, SPE is performed using funnel shaped cartridges which are filled with sorbent material. Fluids and aqueous samples or inoculated into the wider, upper end of the cartridge. The respective fluid rests upon the pressed sorbent bed until a force is applied. It can either be performedby application of a pressure through the upper end of the cartridge or by creating an under-pressure at the lower end of it. The

first procedure is faster but eventually more destructive since pressures up to several bars may be applied. The second method is slower and softer since, given a perfect vacuum; the maximal pressure difference is one bar. SPE procedures based on reversed phase silica sorbents commonly follow these steps:

1) Washing of the cartridge by flushing MeOH over the sorbent material
2) Equilibration/activation of the sorbent by flushing with water or water with acidic or basic additives
3) Application/loading of the sample to be prepared
4) Washing off salts and (hopefully) proteins using the water mixture of step 2)
5) Eluting the analytes using MeOH or any other rather apolar solvent

According to our experience these steps alone effectively remove salts and to some extent proteins as well. The reason for contemporary protein removal may be that small metabolites associate with the sorbent more rapidly than proteins do, therefore inhibiting protein-sorbent interactions. One disadvantage of silica is that they must never be allowed to run dry since this interferes with analyte-sorbent interaction. Once they are dry they would have to be reactivated for optimal elution. For these reasons the application of pressure must immediately be stopped as the solvent meniscus touches the sorbent surface. These circumstances lead to substantial dead volumes. Alternative to SPE-cartridges, sorbent filled pipette tips are available. According to our experience these tips efficiently remove proteins. A third alternative would be the application of so called magnetic nano particles or MNPs. Here sorbents are modified with a ferromagnetic layer. These particles are simply to be introduced into sample vials which contain the solvents or the sample. Applying a strong magnet at the lower end of the vial causes these particles to remain there while solvents can be exchanged according to the five steps listed above. An alternative to RP C-18 silica based materials are co-polymeric materials. These showmultiple selectivities and can be used as RP and HILIC phases or even as anion or cation exchangers. Their performance depends on the types of solvents used for conditioning. In order to attain biological activities, in an aqueous environment the amino acid chain of proteins needs to get folded into a three-dimensional structure. It usually happens via aggregation of lipophilic residues in the proteins core and an outward orientation of hydrophilic

residues. In this way proteins dissolve well in water. In order to remove them from biofluids, a possibility is to turn the folding of proteins around. It is possible to exspose the proteins to a lipophilic environment unfolding and denaturating them irreversibly. In this way, proteins can easily be removed from samples via centrifugation [131]. However, PPE also has its adverse effects. While proteins are precipitated, salts do not easily do it as well. There is a risk to loose metabolites, since they can form for example Na^+-salts and precipitate and salts which were before associated with proteins are free in the solution. As a consequence, samples and mass spectra may be contaminated with salts. Additionally, if lipophilic trans-membrane proteins were part of the sample, their removal could not be guaranteed. The suitability of PPE for either LC-MS or ICR-FT/MS metabolomics will be assessed in Chapter 4.

3.5 Mass-Mass Difference Networking

Network analysis which is based on graph theory has become a mainstay in most of the older omics disciplines. By displaying data points as nodes which conditionally get attributed with connections (edges) among each other, networks have been proven ideal for efficient modelling and visualisation of multi-variate real-world scenarios. The properties of such networks were used especially in genomics and interactomics in order to connect e.g. gene expression data to experimentally induced stress on model systems. In order to construct such graphs/networks one needs data points and a hypothesis on possible relations between them. Once these are present, one confronts each possible pair of data points testing for the given experimental hypothesis. Typically such data points will be attributed a connection/an edge once the hypothesis is true and none if it is false. The reason why graph theory has so far not become popular in metabolomics may be the predominance of targeted LC-MS and GC-MS approaches. Such approaches produce size-wise limited (and even biased) data and are mostly not data-driven. Data driven, non-targeted approaches performed via NMR but as well LC-MS are more and more employed. Apparently

Metabolomics is still lacking a framework of hypotheses to perform network analysis. The only testable hypotheses in known approaches can clearly only be developed on the intensity scale.

Testable scales in ICR-FT/MS data are the highly "quantitative" m/z values acquired and their semi-quantitative intensities. The exact mass gives the chemical basis on which we can construct our metabolic pathways in a data driven network based approach. The intensities can be used to perform statistical inference on the data and to therefore place weights in the before constructed mass difference networks.

We exploit the exact mass information of ICR-FT/MS data to its fullest by using such above indicated mass-mass difference networks [133]. They are formed hypothesising that two metabolitic substrates react with each other to give a product and that the difference in mass between the substrates and the product is specific for each reaction type.

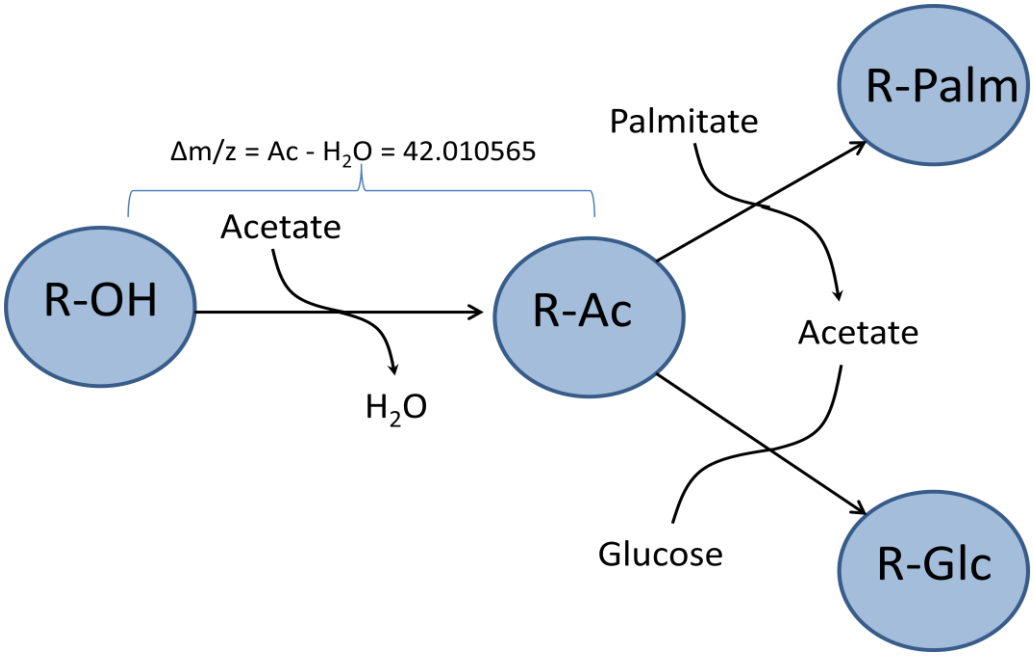


Figure 12: Scheme of a mass-mass difference network section. The identity of each node (circles) is determined by reactions which define the edges (straight arrows). If the identity of R is known, the sum formula of all nodes can be calculated.

As displayed in **Fig. 12** we can for example say that we can express every acetylation as

R-OH + Acetate → R-Ac – H_2O as Ac - H_2O → R-Ac – R-OH.

Therefore "Ac - H_2O" is a universal equivalent for acetylations independent of the identity of R. We can say $C_2H_4O_2 – H_2O = C_2H_2O$ = 42.010565 m/z is universally true for acetylations. Inversely we can say, if we find two exact masses exactly having 42.010565 m/z as mass difference we can deduce that these two masses are connected via an acetylation reaction.

Mass-mass difference networks are generally built up from chains of relationships A + B → C. This has three major consequences:

1) If we know the identity/sum formula of one node in the entire network, we can calculate the identity/sum formula of all other nodes in the network.
2) Since we define one of the partners A or B we can filter our data. We can remove isotopes and therefore unnecessary, often instable, information. We can potentially attribute each connection with further information derived from statistics
3) Only highly exact mass determination in a broad scan scale enables to formulate such kinds of networks.

In fact, as we define a set of reactions to be used for mass-mass difference networking we define the extent to which we can observe the metabolome. Therefore as non-targeted metabolomics is applied, the framework of mass-mass difference networks potentially reflects the metabolic pathways inherent to the acquired data. Combining these pathways which were created in a data driven way with statistical results based on the intensities of ICR-FT/MS spectra therefore potentially gives insight into known pathways and their dynamics. Covering masses which beforehand were not identifiable in terms of sum formulas, new knowledge on before not known interactions can be formed.

3.6 Multivariate Statistics

Multivariate statistics is the multidimensional extension ofunivariate statistics. Univariate statistics is generally used in order to assess the relationship between one variable in the context of different observations. Scenarios in which observations depend on one variable only, are usually experimentally constructed and are far from a biochemical reality. Where univariate analysis with $k = 1$ observations, up to low variate analysis with $k \leq 5$ observations, necessitates variables to be independent. Multivariate techniques can cope with different challenges and are by nature efficient to investigate highly dimensional data with a notable presence of noise and collinearity. Such collinearity may sometimes lead to serious stability problems when statistical analysis is applied [134, 135]. Modern multivariate approaches accept the fact, that an observation (e.g. the expression of an illness), may be a function of correlation or at least inter-dependency of variables (such asobesity, bad diet, few sports and genetic predisposition in Diabetes mellitus).

Many different multivariate techniques have been applied starting from principal component analysis (PCA) and hierarchical clustering analysis (HCA) up to supervised methods as projection to latent structures or partial least squares (PLS), OPLS/O2PLS-DA, PLS-DA (used for discriminant analysis) as well as derivations of them.

Each sample/observation *n* is composed of all the *k* variables that were acquired. As the k variables span up a *k*-dimensional space, each observation *n* may be a point in this space. Once this space is populated with all n sample/observation points one yields a data cloud. It is now possible to calculate a regression line through this cloud. This regression line is a vector (interpretable as point or line) of *k* variables which in each point is the closest to each sample *n*. This regression line is the first component (PC1) and explains the most prominent influence on the data set. The projection of a sample-point n into PC1 is called *score*. With this the position of each variable relative to PC1 is known. Since illnesses such as *Diabetes mellitus* are of multi-factorial nature one may try to elucidate, which other influences independent from PC1 have impact on disease development. For

this purpose in PCA a second PC is calculated. This PC intersects PC1, is orthogonal to it and has the smallest distance to each data point. In general it is possible to create *k*-1 PCs. However the explanatory power of each further PC decreases dramatically. For this reason the first four PC1 are often, in any combination, taken as the axes of a two or three dimensional coordinate system in which the positions of observations are plotted relative to the PCs. The relative position of the samples relative to each other can be observed in the *score*-plot. The importance of variables for the relation between the scores can be examined in the *loadings* plot. These loadings are the angles of each variable axis into the two-dimensional plane spun up by e.g. PC1 and PC2. Each axis in the *k*-dimensional space represents one variable. The angle of each variable into this plane is expressed as function of the cosine. A $\cos\alpha_1 = 1$ relative to PC1 means that k_1 is parallel to PC1 and has the same direction. A $\cos\alpha_1 = -1$ relative to PC1 means that k_1 is parallel to PC1 and has the opposite direction. The original data matrix X is a linear combination of the scores and the loadings. Everything in X which stays unexplained by these values are the residuals, the noise. In order to summarize the quality of such a PCA model the parameters R2X and Q2X can be computed. R2X measures the explained variance of the model. It is defined as R2X = 1 – (residual variation/original variation). It says how good the model fits the data in X. The other parameter Q2X explains the goodness of prediction of the model. It is attained using cross-validation (CV). In SIMCA-P initially the PC1 is computed and its residual variation (R2X) is being calculated. As the second PC is built, several parallel PCA-models are created upon the data set with 5 to 10 predefined data blocks (between $1/5^{th}$ and $1/10^{th}$ of the data per block) being kept out of the model formation. The built models are used to predict these omitted groups. The residuals of the predictions are summed up and are finally compared to the variation of the residuals of the previous PC model. If this sum of predicted residuals is bigger than the sum of residuals of the prior model, the PCA model formation is stopped. Therefore, once the cumulated R2X and Q2X are diverging when a PC is added to the PC model the modeling is stopped. In SIMCA-P the CV is performed in real time.

PCA is a powerful tool in data analysis. Nevertheless it has one big drawback. Prior to analysis of experimental data it cannot be guaranteed that the greatest

variation in the data is coinciding with the experimental hypothesis. It may be that technical variation is much more dominant than the biological variation. In these cases PCA alone is powerless. In order to solve these problems the next step in data analysis would be to build a PLS regression model. This technique does not only take the data matrix X into account, but also a grouping/predictor vector y. The process is similar to the PCA approach, but the PCs are not built on the partial least square regression alone. In PLS the PCs have to fulfill a second criterion. The PCs found have to be of maximal correlation to the predictor vector y. Each newly found PC has to minimize the residual variance in X and as well in y. In addition each PC is again checked using cross validation. If this procedures is still not sufficient in finding any variable or model which describes the data, orthogonal signal correction can be applied prior to PLS (or also PCA). It tries to find a principal component which is orthogonal to the y vector. Once this vector is found in k dimensional space, its respective values are subtracted from X. Afterwards a PCA or PLS model is built. This method effectively improves results of these projection techniques. A tool used to identify loadings/variables of high importance is the S-Plot. It is formed on the results of PLS. It plots the modeled correlation with y over themodel covariance. If variables with high magnitude have the highest correlation to the model, the plot shows an S-shape. In this work PLS results are interpreted based on the S-plot. All statistical evaluations in this work were calculated using SIMCA-P 12. Information in this section was interpreted based on the books “Multi- and Megavariate Data Analysis (Part I and Part II)” by UMETRICS [L. Eriksson, E. Johansson, N.Kettaneh-Wold, J. Trygg, C. Wikström, S. Wold, Multi- and Megavariate Data Analysis (Part I and Part II) 2006, Umetrics Academy, Umeå, Sweden].

4. Methodological aspects of samples analysis

The stability and reproducibility over long time in the development of a non-targeted metabolomic platform relies on important check points, ranging from the sample collection until data analysis. In this section the salient points characterizing the work flow that enables the analysis of plasma samples will be presented and discussed.

4.1 Sample collection

Standardized collection methods are important in order to allow comparison of different samples collected from different laboratories preserving quality, reproducibility and stability in a long term. The latter is defined by the International Standards Organisation (ISO) and indicates the capacity of a sample to preserve its properties from a period of time, based on the observation of defined compounds (a group of compounds that one assumes as reference of one sample and that can assess the variability of the same) when the sample is stored under specified conditions (*ISO Guide 30, 1992).* The principal points of variability involve additives used in the blood collection, the handling temperature, hemolysis, sample storage and freeze-thawing cycles [136]. These check points are getting even more important in the context of non-targeted metabolic technologies where the interest is to detect the maximum extent of the metabolites (50-1200 m/z). Temperature is the greatest factor to be considered in plasma handling; from the collection, trough transport until storage [137-139]. Use of dry ice for transport and iced water in fast handling protocols can help minimizing protein degradation [137-139]. The storage temperature for long periods should also be set at -80°C. In this work every step was performed on iced water. The collected blood was centrifuged at 4 °C, in order to avoid degradation and protein kinase activation that can occur at room temperature. The obtained plasma was then aliquoted in small volumes of 150- 200 µl in order to avoid freeze-thawing cycles. The latter process can have negative impact on the plasma characteristics, as it was shown in non-targeted pre-analytics studies of plasma,

where thawing on ice cycles from one up until four times were compared. It was observed that thawing on ice for two up to four times can undermine the quality of the plasma [256]. The plasma aliquots were then stored at -80° C until sample analysis. For the sample transport the plasma was always delivered on dry ice.

Plasma type selection

In laboratory analysis plasma is obtained from whole blood; in order to prevent clotting an anticoagulant (i.e. EDTA, Heparin, NaF, Citrate) is added into the collection tube or it is alredy present in the tube. Centrifugation leads to plasma. Different components present into the collection tube could interfere with the sensitivity of MS analyses. Silicones, utilized as lubrificant for stoppers or coating, polymeric surfactants (as polyvinylpyrrolidones and polyethylene glycol) can be applied to induce surface wetting in the tube [140, 141]. Therefore, as is also already discussed in proteomics studies [142, 143], the importance of a standardisation of the collection tube is extreme. In order to plan a long term platform for the analysis of plasma, enabling easier cross-comparison of different plasma samples, it is necessary to understand which blood additive conduces to stable non-targeted profiling avoiding interferences. It was necessary to understand which plasma anticoagulant type would be the best choice in terms of matrix effect, robustness and repeatability within sample measurements and the number of features recovered. Four different kinds of anticoagulants in plasma (EDTA, NaF, Citrate and Li-Heparinate) were screened via LC-MC analysis in order to perceive their analytic behavior.

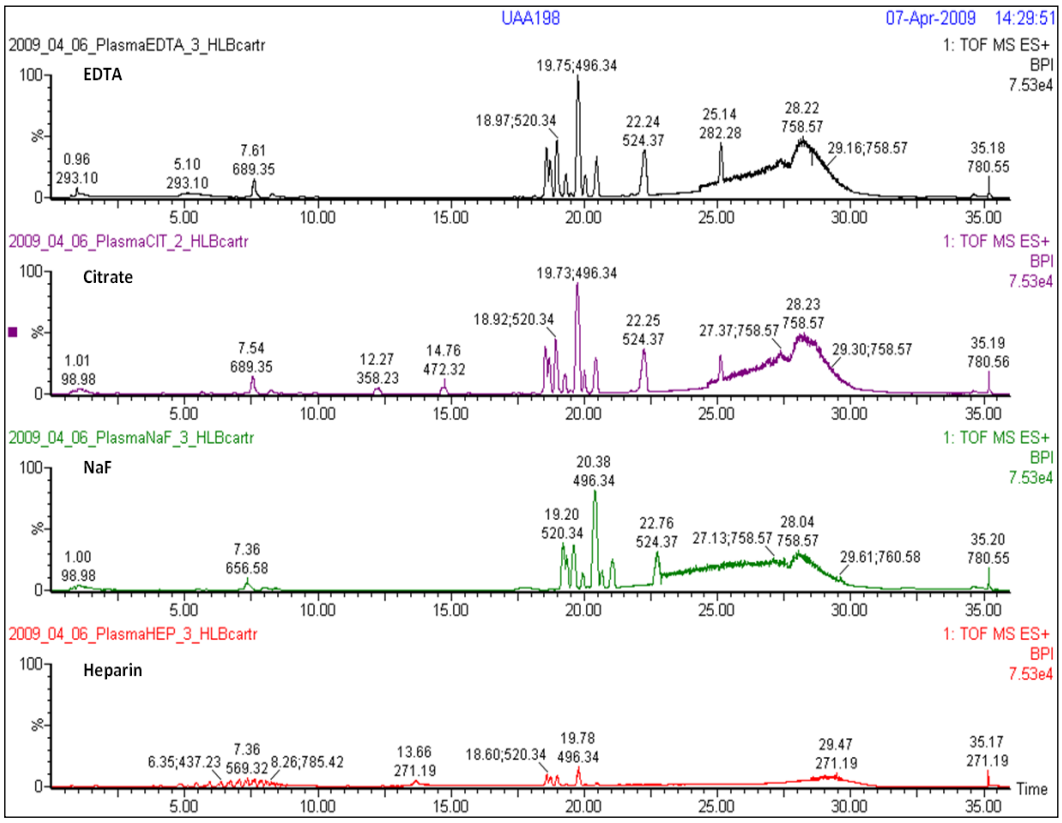


Figure 13: Comparison of Li-Heparinate, NaF, Citrate and EDTA plasma in the LOAD fraction collected. Sample preparation using HLB cartridges. LC analysis performed by C18 HSS UPLC column (1.0 x 150mm, 1.7 µm, Waters, Milford, MA).

The extraction of hydrophilic and hydrophobic components was performed employing the polymer based solid phase extraction cartridges Oasis® HLB (Waters). Every experiment was repeated three times in order to study the repeatability of the different matrices. We have compared every step with the corresponding plasma blank. The latter consists of water poured into a collection plasma tube, in order to visualize interference peaks and to focus on the repeatability and the consistency of interference effects. **Fig.13** and **Fig. 14** show the LC-MS peak base ion chromatograms of different anticoagulants in plasma samples observing the load fraction and the elute fraction of SPE treatment, respectively.

The chromatograms obtained for NaF, Citrate and EDTA plasma samples are very similar in comparison to each other, focusing on the load fraction and the elute fraction.

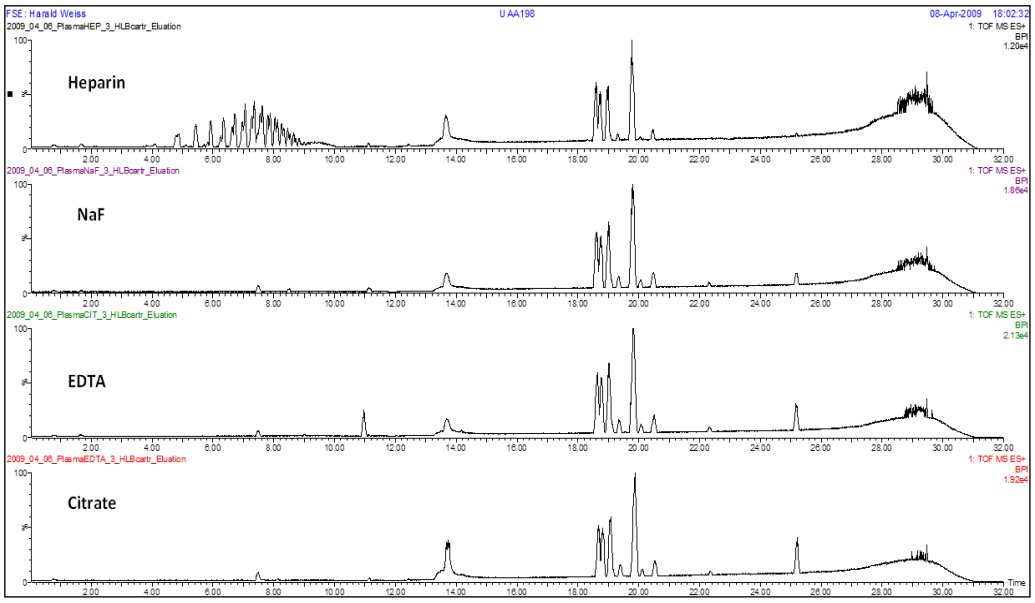


Figure 14: Comparison of Li-Heparinate, NaF, Citrate and EDTA plasma in the ELUTE fraction collected. Sample preparation using HLB cartridges.LC analysis performed by C18 HSS UPLC column (1.0 x 150mm, 1.7 µm, Waters, Milford, MA).

The load fraction presents higher intensity in the base peak ion chromatogram (BPI) and higher number of peaks. It seems like plasma metabolites were not correctly retained on the sorbent material therefore, the information got split into the two partitions. Investigation on the information behind the load fraction will be discussed in the next paragraph. In this preliminary study the interest was focused on the quality of the different anticoagulants much rather than a comparison among sample preparation. Heparin as anticoagulant shows interfering peaks, such as a hump at the beginning of the gradient. This effect is more prominent in the elution fraction but less marked in the load fraction. These peaks are probably related to a signal interference coming from polymeric surfactants, added for favouring surface wetting in the collection tube. Therefore, this kind of anticoagulant is not among the suitable candidates to perform any further non-targeted profiling. The number of features extracted, in each fraction are resumed in **Table 3**. As it is possible to extrapolate from the shown data, the highest number of hits is collected in the load fraction. The numbers of features are calculated as the averaged number of three experimental replications, analyzed

in triplicates. The information between EDTA, Citrate and NaF is quite similar in respect to the total sum of detected features.

Table 3: Features counting in plasma anticoagulants screening experimental triplicates of load fractions and elute fractions collected. Sample preparation performed used HLB cartridges. LC analysis using C18 HSS UPLC column (1.0 x 150mm, 1.7 µm, Waters, Milford, MA)

	Features collected in the Load fraction		**Features collected in the Elute fraction**		**SUM of the features collected**
Plasma Type	**Average**	**RSD %**	**Average**	**RSD %**	
EDTA	1257	18	485	7	1742
Citrate	1280	1	487	7	1767
NaF	1249	8	504	2	1753

In order to investigate the reproducibility of each anticoagulant and to conceive their stability, the experimental data were analyzed [144] with a Bland and Altman Plot [145, 146] to achieve a direct comparison. The analyses were performed on both fractions collected (load and elute) and the load fraction showed better reproducibility (**Fig. 15**- **Fig. 17**).

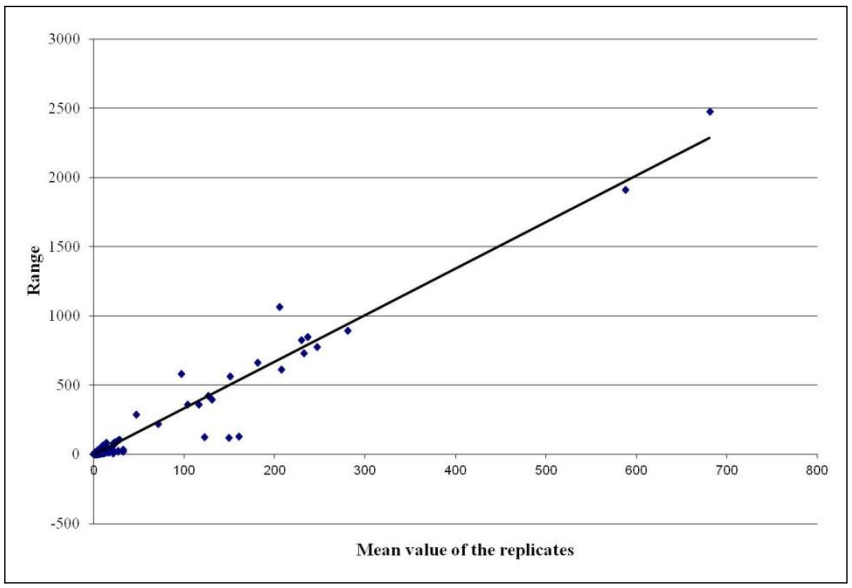


Figure 15: Bland and Altman Plot of the three replications (Plasma **EDTA,** Load fr.). Correlation R^2=0.95 ($p<0.01$). Slope=3.359. Intercept =-1.70 ($p<0.05$). The calculation was done at the confidence level of 95%.

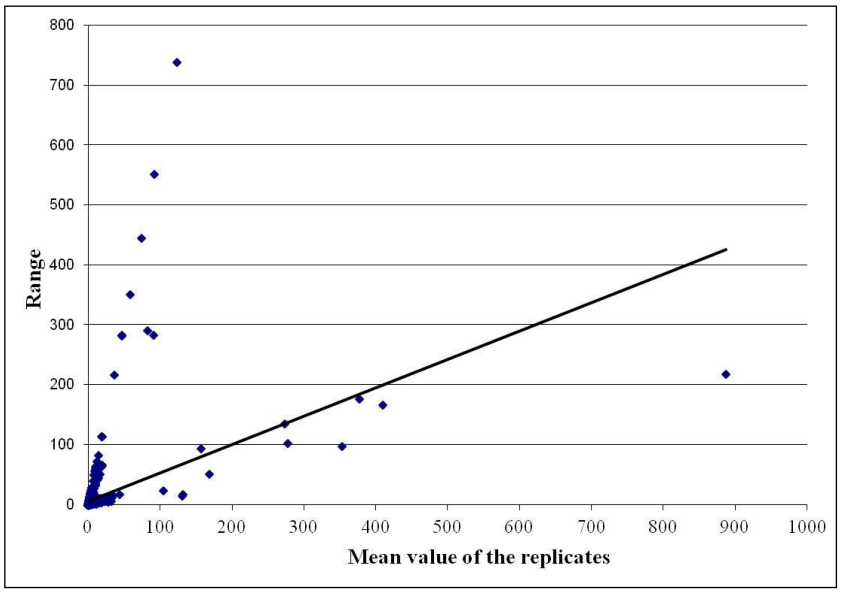


Figure 16: Bland and Altman Plot of the three replications (Plasma **Citrate**, Load fr.). Correlation R^2=0.19 ($p<0.01$). Slope=0.47. Intercept =4.91 ($p<0.05$). The calculation was done at the confidence level of 95%.

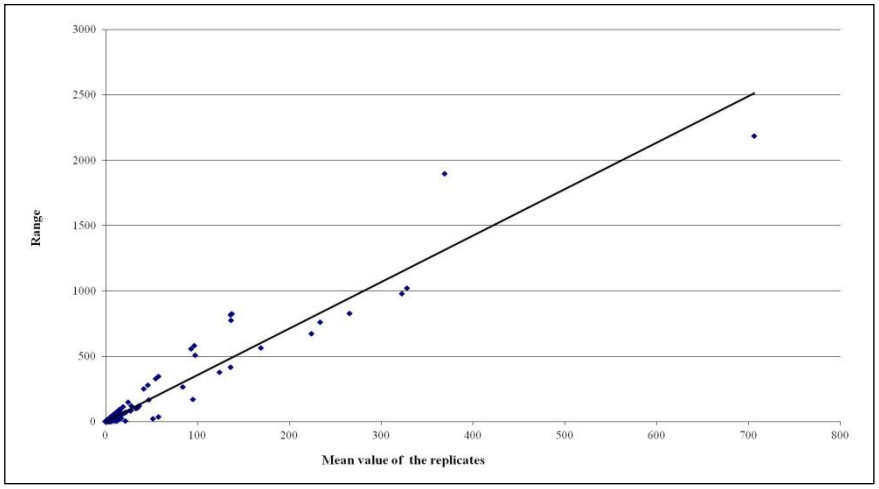


Figure 17: Bland and Altman Plot of the three replications (Plasma **NaF**, Load fr.). Correlation R^2=0.92($p<0.01$). Slope=3.53. Intercept =1.43($p<0.05$). The calculation was done at the confidence level of 95%.

As we can gather from the correlation coefficient R^2 in the linear regression model, EDTA plasma (**Fig. 15**) is the most solid matrix in terms of reproducibility. Citrate and NaF exhibit two linear relationships, which may point to interferences with the chromatographic procedure. Furthermore, a flat slope of the regression line indicates small increase of variability over the dynamic range. Strong slopes indicate the opposite. Citrate has low slopes but is multimodal. EDTA and NaF are more similar, but EDTA has a better R^2 and looks definitely unimodal. This was the reason why EDTA was chosen for the following experiments. In order to confirm the results found, indipendently from the sample preparation applied, EDTA, Citrate and NaF were compared after protein precipitation extraction. Further investigation on organic solvent precipitation as sample preparation, will be handled in the next section. In **Table 4** it is possible to visualize the number of features collected according to different plasma anticoagulant profiles.

Table 4: Features counting in plasma anticoagulants screening experimental triplicates.
Sample preparation performed via PPE (AcN/MeOH: 3 /1).
LC analysis using C18 HSS UPLC column (1.0 x 150mm, 1.7 µm, Waters, Milford, MA)

Plasma type	Features collected
EDTA	464
Cit	432
NaF	468

The analysis was conducted on the experimental triplicates applying Bland and Altman Plot as well (**Fig.18- Fig. 19**).

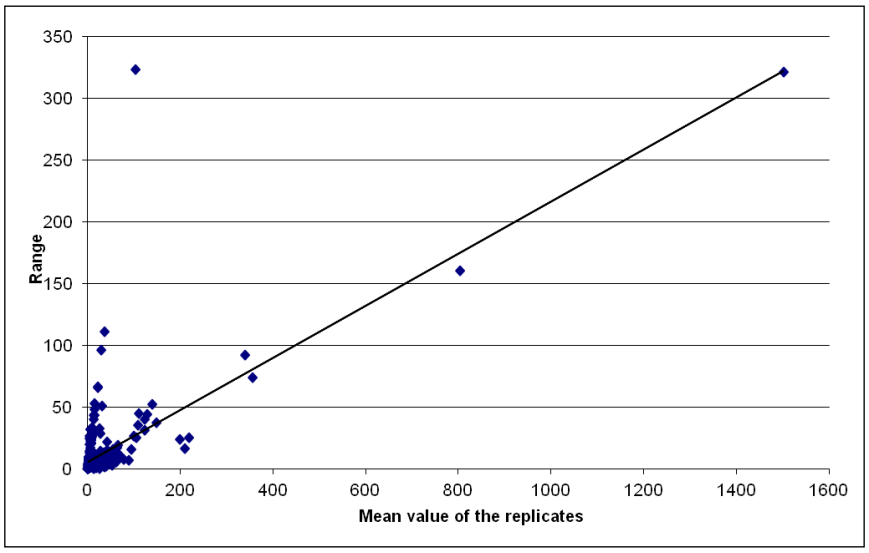


Figure 18: Bland and Altman Plot of the three replications (Plasma **EDTA**) applying PPE. Correlation R^2=0.51 ($p<0.01$). Slope=0.2. Intercept =5.21 ($p<0.05$). The calculation was done at the confidence level of 95%.

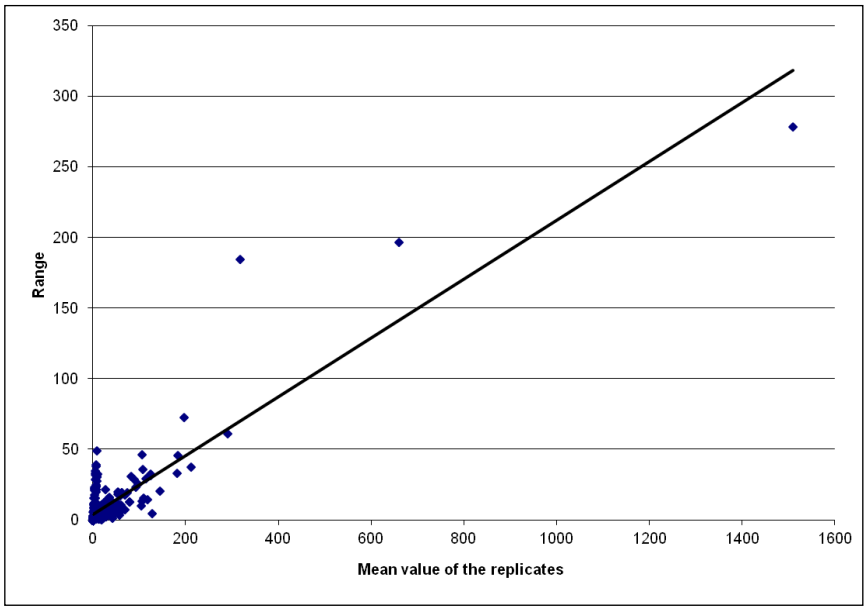


Figure 19: Bland and Altman Plot of the three replications (Plasma **Citrate**) applying PPE. Correlation R^2=0.77 ($p<0.01$). Slope=0.20. Intercept =3.67 ($p<0.05$). The calculation was done at the confidence level of 95%.

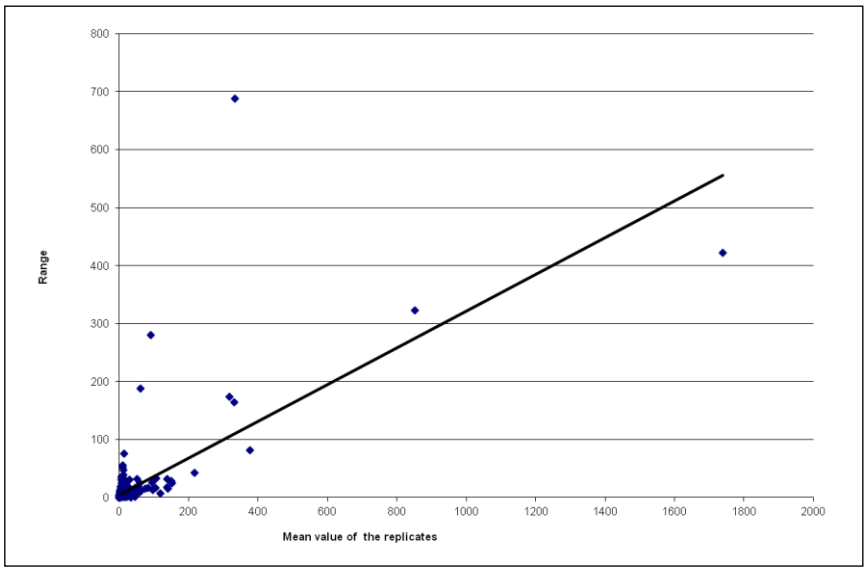


Figure 20: Bland and Altman Plot of the three replications (Plasma **NaF**) applying PPE.Correlation R^2=046 ($p<0.01$).. Slope=0.31. Intercept =4.54($p<0.05$).The calculation was done at the confidence level of 95%.

From the analysis of the different plots, Citrate and EDTA are showing the best regression coefficient value, therefore the best performance. NaF produced stronger outliers. Comparing the results of the previous experimental sets, EDTA is confirmed to be the most suitable anticoagulant.

4.2 Sample preparation

Sample preparation is a crucial step especially in non-targeted metabolomics studies. Plasma is a very complex matrix "*per se*". Its manipulation requires the avoidance of direct injection onto the UPLC-column, due to the high protein content. Therefore, an incisive sample pre-treatment is fundamental. The factors to be considered in choosing the most reliable method are the number of the extracted features, the effectiveness of protein removal and the repeatability. Different techniques such as PPE using organic solvent [131] or acid conditions [132] or SPE [147] are already discussed in literature concerning the handling of this matrix. The sample preparation of plasma is an analytical challenge, due to its high content of small polar metabolites and apolar compounds, such as lipids. In non-targeted metabolomics the principal aim is to extract the widest number of metabolites, and therefore to remove compounds that could interfere with mass spectrometer (e.g. salts) or with the chromatographic system. A suitable sample preparation must be characterized due to minimal number of steps in order to limitate the degradation of thermo labile compounds and to minimize contaminations and inter-laboratory errors as well. Additionally, the ideal sample preparation requires the lowest volume of plasma possible and needs to consider a possible integration with other analytic platforms, enabling a high throughput of sample analyses. The aim was to find a robust and fast method using the already selected EDTA plasma. In literature a wide variety of examples explains the advantages of PPE with organic solvent such as acetonitrile [148] methanol [132] and methanol-ethanol and methanol-acetonitrile-acetone mixtures [131].

Sample preparation in UPLC-MS

In this section the passages that lead to the optimization of an adapted sample preparation will be explained. The investigated technologies involve solid phase extraction and protein precipitation extraction. Solid phase extraction is largely used in targeted metabolomics studies in order to quantitatively isolate one class of compounds of interest. It is an appealing technology due to its benefits, such as the minimization of the ion suppressionand possibility of automatation. The

integration of this technique into a non-targeted metabolomics approach requires, as the first important stage, the selection of the sorbent material. In order to maximize the classes of chemicals extracted, a macroporous copolymer made from a balanced ration of two monomers (lypophilic divinylbenzene and hydrophilic N-vinylpirrolidone) was selected (**Fig. 21**). This combination enables reversed phase capability joined to a polar analytes capture capacity.

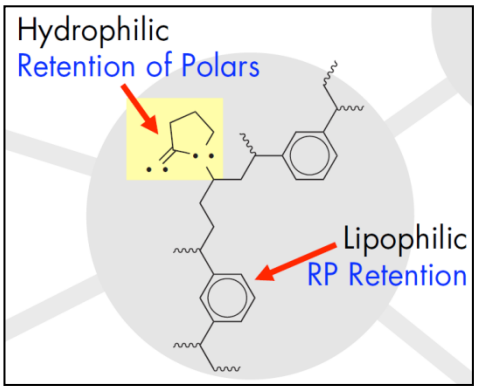


Figure 21: Representation of the chemical properties of the sorbent of HLB sorbent material from Waters. Figure reprinted with permission from sample preparation seminars, ©Waters, Waters Corporation, Milford, USA.

Prior to the loading of the sample onto the cartridge, the sample was pre-treated with 2% acid buffer (2% H_3PO_4) at the dilution ratio of 1/1. This step is necessary in order to disrupt protein binding. Load, wash and elute fractions were collected in order to asses the information relative to each fraction. The wash fraction did not show important loss of markers, therefore will not be considered into the discussion. The load fraction showed a consistent number of features collected compared to the elute fraction as it is shown in **Fig. 22**.

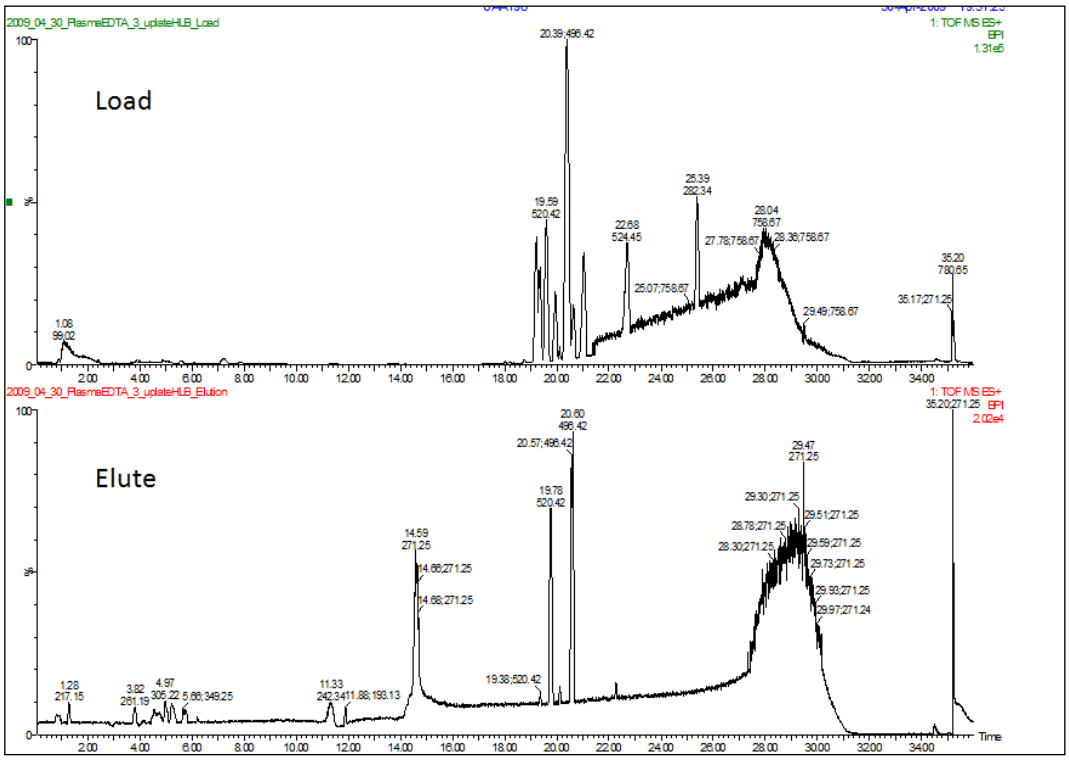


Figure 22: Chromatograms showing the comparison between the load and elute fraction applying HLB cartridge.

Relative to this preliminary study, it appears that a partition of important information needs to be observed in the load fraction as well. In order to understand whether the cause of this breakthrough is due to the acid pre-treatment, the same amount of plasma was treated as well in basic condition using NH_4OH 2% to demonstrate that the cause of the breakthrough was not due to overload of the sorbent material (30 mg), the latter one was increased, loading the same amount of diluted plasma onto a cartridge with 1 g of sorbent material. Additionally, the elute step was optimized in order to increment the features collected according to volume efficiency where methanol, as eluting solvent, was tested at different volumes (200µl, 500µl and 1 ml). A second strategy counted on the increase of solvent strength as eluent (MeoH, MeoH-AcN: 2/8 (v:v), AcN/IPA: 4/6 (v:v), IPA). All the results are summarized in the **Tables 5- Table 7**.

Table 5: Number of features collected in the load fraction after HLB extraction (pre-treatment with acid) and in the elute fraction of different experiments.

Plasma Load Acid	Plasma Elute 200 µl MeoH	Plasma Elute 500 µl MeoH	Plasma Elute 1ml MeoH	Plasma Elute 500 µl MeOH/AcN	Plasma Elute 500 µl AcN/IPA	Plasma Elute 500 µl IPA
7918	4235	4042	4716	3667	4033	4021

Table 6: Number of features collected in the load fraction after HLB extraction (pre-treatment with base) and in the elute fraction.

Plasma Load Basic	Plasma Elute 500 µl MeoH
4855	3017

Table 7: Number of features collected in the load fraction after HLB extraction (pre-treatment with acid) and in the elute fraction.

Plasma Load Acid (1G)	Plasma Elute 500 µl MeoH
7771	2547

From the tables, it is possible to observe that the higher number of features lies in the load fraction, independentely from the kind of pre-treatment applied (acid or basic) and independentely from the quantity of sorbent used (30 mg or 1G). The number of features gained in the elute fractions remains inferior, independentely of the optimization protocol applied to improve the elution process.

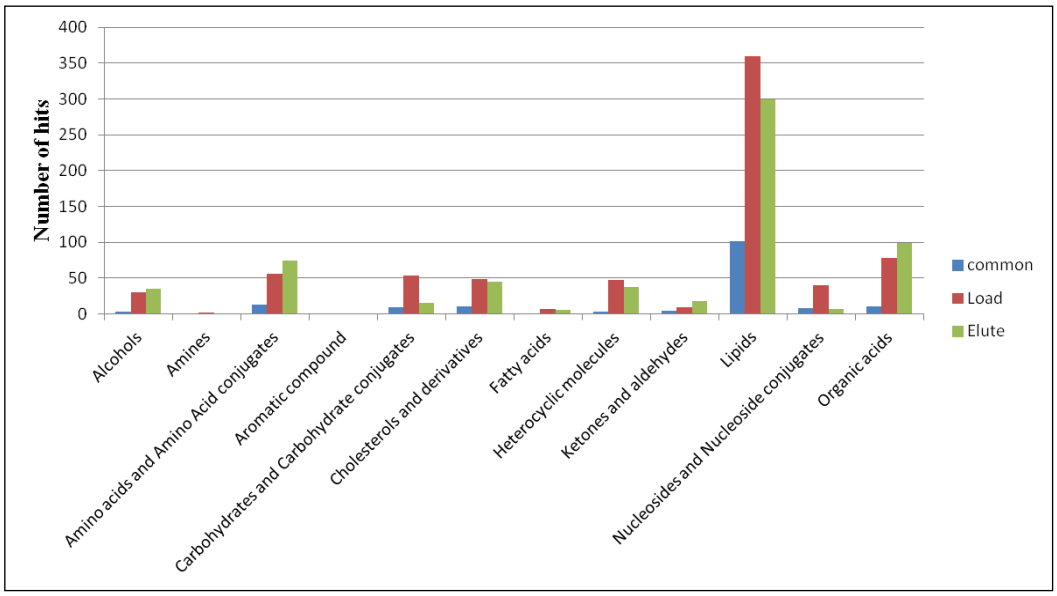


Figure 23: Pricipal classes of markers present, applying HLB extraction, in the load fraction and in the elute fraction respectively. The common compounds are as well shown.

In order to assess whether behind the number of feature into the load, there are biomarker information, annotation via HMDB data bases was provided with an error tolerance of 12 ppm.

From **Fig. 23** it is possible to understand that the collection of the load fraction should be an unavoidable step in a non-targeted metabolomic experiment. The load fraction contains higher content of lipids and nucleosides and carbohydrates.

The integration of the information from the two different fractions can lead to a good point for the investigation in non-targeted metabolomic studies, but the passages required could not avoid a high variability among each sample preparation. Therefore, a suitable technique should be found. For this reason, protein precipitation using organic solvent could represent a possible alternative. Different kinds of organic solvents and mixtures were tested: AcN, AcN /MeOH/Acetone:1/1/1 (v:v), EtOH/MeOH: 1/1 (v:v) [131] and AcN/MeOH:3/1 (v:v). The ratio between the plasma and the organic solvent was 1/4 and all the solvent were added cold. Every sample preparation was prepared in triplicate in order to asses the robustness of the experimental method.

Table 8: Number of features collected, comparing different kinds of organic solvent in protein precipitation experiments.

Solvent	AcN	AcN/MeOH/Acetone	EtOH/MeOH	AcN/MeOH(3/1)
Number of features	4260	3696	4019	3421

Table 8 shows that, according to the number of features possible to achieve with each different protocols, acetonitrile and EtOH/MeOH (1/1) mixture are the most effective. In order to asses their stability, Bland and Altman Plot was proposed.

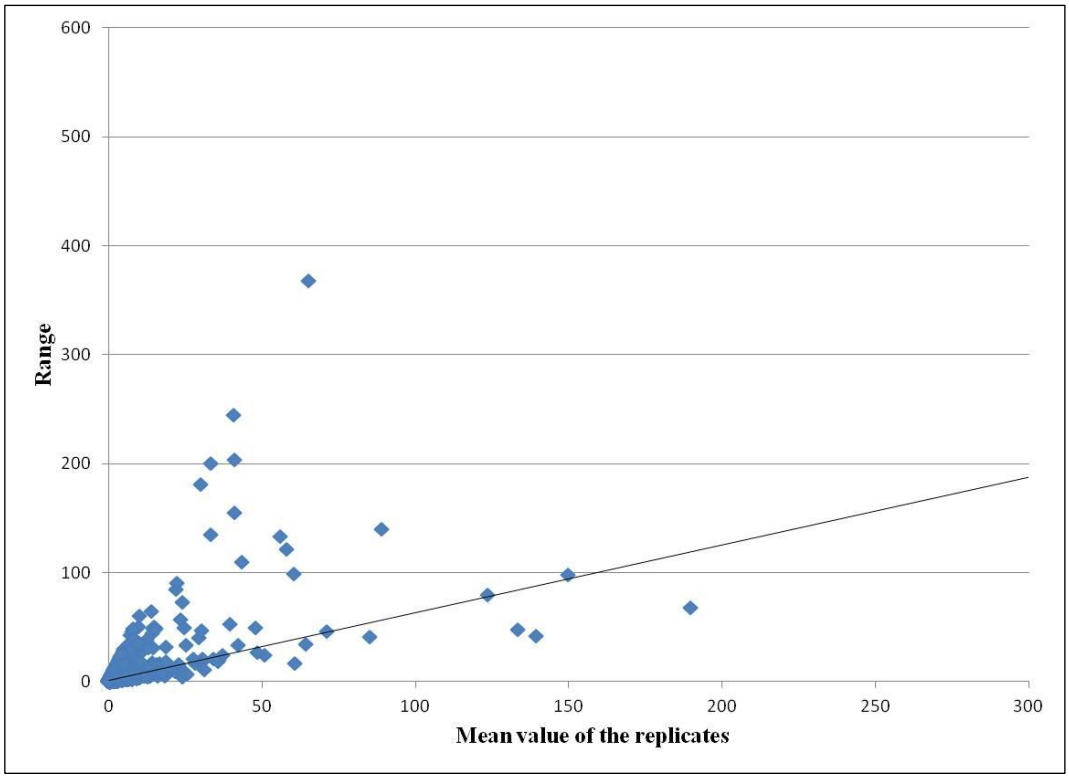


Figure 24: Bland and Altman Plot of the three replications applying PPE with Acetonitrile.Correlation R^2=0.74 ($p<0.01$). Slope=0.62. Intercept =0.57($p<0.05$).The calculation was done at the confidence level of 95%.

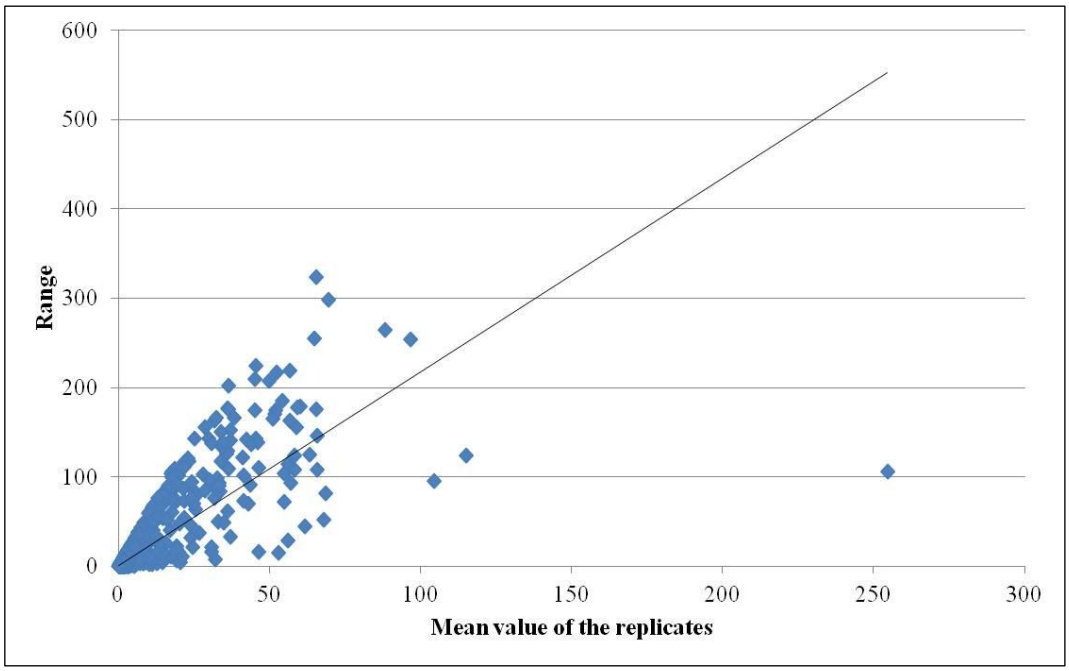


Figure 25: Bland and Altman Plot of the three replications applying PPE with EtOH/MeOH(1/1).Correlation R^2=0.6 ($p<0.01$). Slope=2.1. Intercept =0.69 ($p<0.05$).The calculation was done at the confidence level of 95%.

According to the regression cofficient value, acetonitrile shows better consistency. The slope in the latter one is less steep; therefore this would indicate a lower degree of variability over the magnitude. These results indicate a slightly better performance of acetonitrile.

Sample Preparation in ICR-FT/MS

The first question necessary to be answered was the suitable kind of anticoagulant in ICR-FT/MS applications in terms of information, ion suppression due to the specific matrix effects and artefacts possibly derivedfrom the different collection tubes.

Since the optimal sample preparation for LC-MS applications is based on the use of EDTA plasma treated viaPPE using acetonitrile, the first attempt was to verify whether the same aniticoagulant and the same sample preparation would have been feasible in order to have the same sample preparation protocol. Different kinds of anticoagulants (EDTA, NaF, Li-Heparinate, Citrate) were compared (**Fig. 26**) after PPEof 80 µl of each anticoagulated plasma with acetonitrile. Ionization mode was investigated in positive and negative modes.

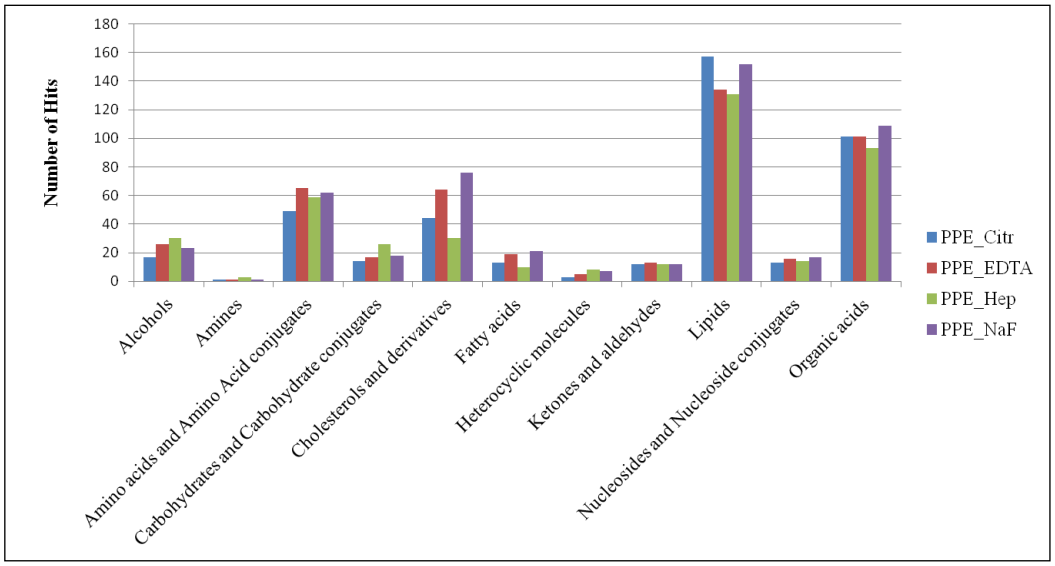


Figure 26: Different types of compound classes annotated in positive ionization mode using different anticoagulants.

Especially Citrate and NaF plasma seem to show a positive effect in metabolite recovery. The overall numbers of annotations are none the less very similar. Only Li-Heparinate has a clearly worse performance in respect to the other kinds of plasma. This result was expected, due to the above documented observations in terms of polymers and stability in this anticoagulant type.

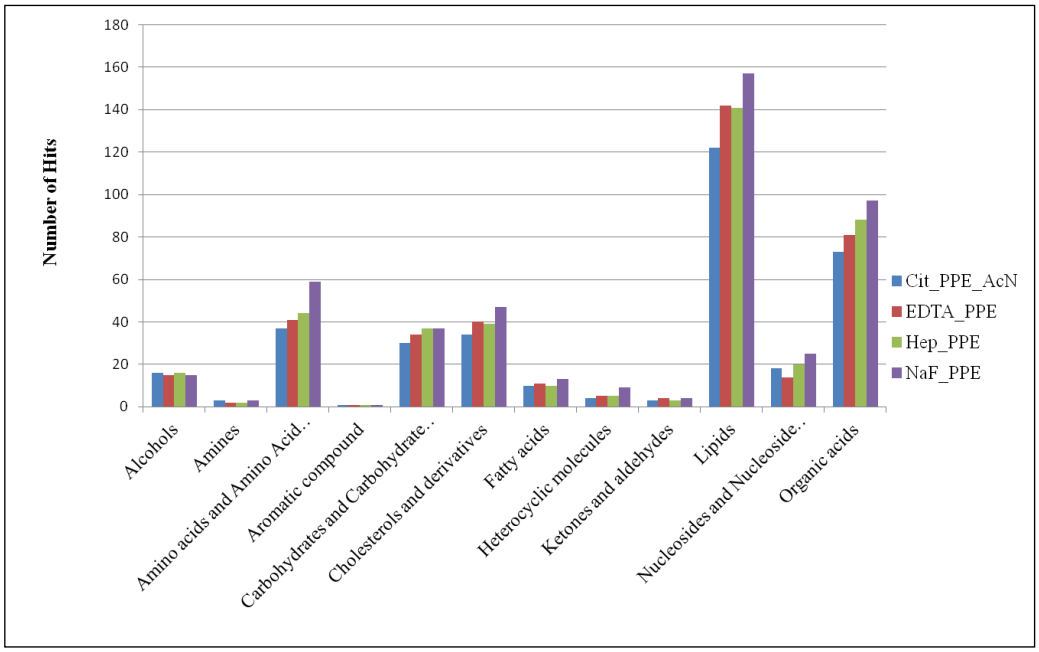


Figure 27: Different types of compound classes annotated in negative ionization mode using different anticoagulants

In negative ionization mode Citrate plasma was the worst in terms of metabolites information (**Fig. 27**). We suspected a suppression effect due to the multifunctionality of citrate but the mono-deprotonated ion corresponding to citrate was not found in the spectra. In fact none of the applied anticoagulants was ever detected in negative mode. It appears that their precipitation and their way of action aremost effective. These preliminary studies revealed EDTA as one of the best anticoagulant of these application techniques. This point allows a homogenization with the established LC-MS platform in order to collect the same kind of plasma. Further investigation already started (data not shown) will be developed in the future in order to verify the use of this kind of anticoagulant in the integration withthe NMR platform.

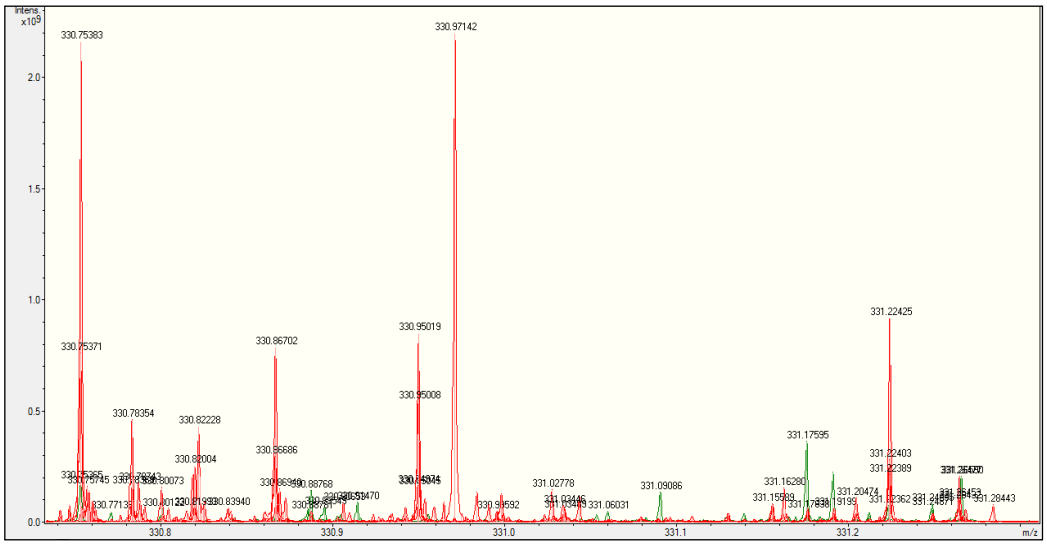


Figure 28: The red PPE mass spectrum shows high salt content based mass defects of 300.7 m/z to 300.9 m/z. The green SPE mass spectrum shows lowe peak heights but clearly fewer salts than the PPE.

Any further investigation on PPE in the ICR-FT/MS context was halted because of a dominant salt content in PPE. **Fig. 28** shows particular peaks acquired in negative ionization mode with an absolute mass defects between 0.7 and 0.9. Such peaks typically occur with salty samples, since typical salt constituents such as Cl^- or Na^+ themselves have a mass defect of 0.969 in the first and 0.997 in the latter case. Red peaks represent plasma prepared via PPE and green peaks represent plasma prepared via SPE using cartridges or ZipTips. During the PPE process not only proteins, but salts as well precipitate. They are not effectively separated from the liquid phase via centrifugation and get retained in the supernatant. As a consequence SPE is the method to be preferred when performing ICR-FT/MS metabolomics. Conclusively, the next step in sample preparation optimization for ICR-FT/MS consists of the examination of SPE. In reference to the results found during the development of the sample preparation protocol for LC-MS application, the breakthrough observed in the load fraction could not be anlyzed via ICR-FT/MS due to the content of orthophosphoric acid that could be harmful for the mass spectrometer in direct infusion without a prior chromatography. Furthermore, the content of peptides and small proteins could result in ion suppression for the same reason of lack of chromatographic separation. The point of discussion during the development of a suitable sample preparation protocol for ICR-FT/MS, wasthat the number of features in the elute fractions were not

reproducing the totality of the sample. This problem can be bypassed since the resolution and the overall sensitivity of ICR-FT/MS is far higher than in UPLC-TOF-MS. Therefore, sufficient and additional information detection was expected. Different kinds of sorbents were tested in order to study the chemical selectivity of different phases as a function of the polarity (C2, C8, C18, CN, and Phenyl silica based material were compared; HLB copolymer was as well studied). An additional application of "in-house" synthesized magnetic nano particles prepared with C18 and HLB sorbent were tested and compared as well. The number of hits throughout all metabolite classes in the normal phase comparison points to the absolute superiority of the HILIC sorbent. CN and especially C2 perform poorly (**Fig. 29**). The reversed phase comparison indicates a superiority of HLB with C18 being the second best sorbent. In the final comparison of groups C18-MNP, HLB-MNPs, C18- OMIX tips, C18 and HILIC SPE cartridges were compared. C18 was chosen rather than HLB in order to get a better comparison to the OMIX tips. OMIX tips have the advantage of far smaller dead volume and it is possible to use them with only 25 µL of Plasma. Furthermore no vacuum pump needs to be applied, which makes them per se fit for a generalized SOP. As we can see in this last comparison, C18 SPE cartridges and C18-MNPs perform worse than HILIC SPE cartridges (being on rank two). HLB-MNPs and C18-OMIX tips result almost equally efficient. As expected, HILIC SPE cartridges performes the best with hydrophilic compounds. A disadvantage in using HILIC/normal phase techniques, which makes them unfit for a platform method, is that the sample first has to be evaporated in order to transfer it into a hydrophobic solvent. This is necessary for the normal phase to work optimally.

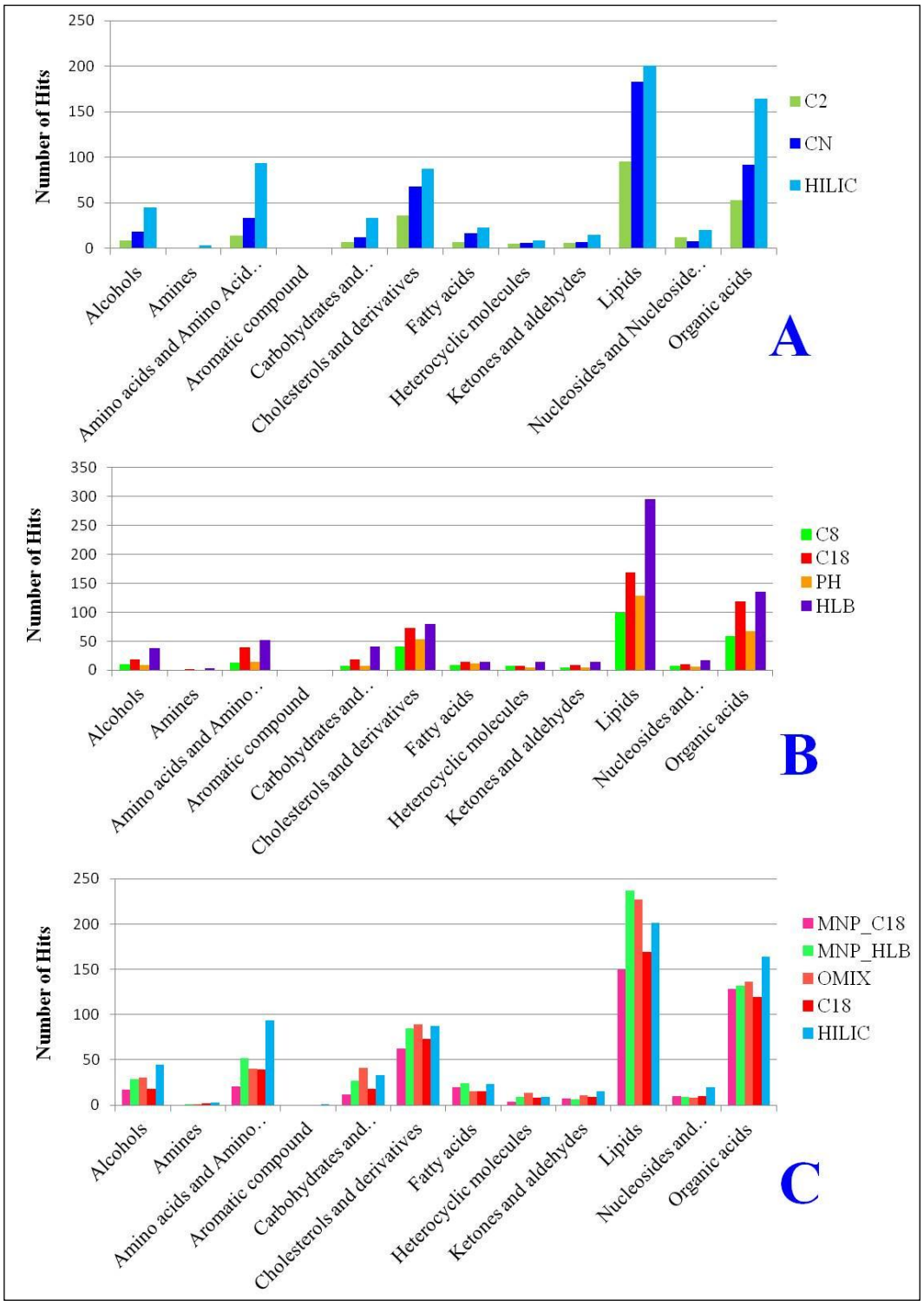


Figure 29: Different SPE cartridges tested in ESI positive mode. A) is a comparison of hits in normal phase sorbents. B) Compares reversed phase sorbents and HLB C) compares the final candidates as well as MNPs and the OMIX tips.

The performance of these sample preparation methods, was screened via negative ionization mode as well.

Interestingly the behaviours of the different techniques are remarkably different in negative ioinization mode. C2 and CN appear to have a much greater capability to retain organic acids than HILIC (**Fig. 30**). Since the normal phase is based on pure silica and it would therefore rather exhibit negative surface charges, it appears to be more suited for retention of bases. In the lipophilic comparison the changes are surprising as well. HLB is less effective in anionretention. Phenyl sorbent performs much better for negatively ionisable compounds. It can be hypothesized, that aromatic compounds are more stable anions because they can dislocate the charges. In order to maintain a parallelism with the positive mode observations, the same MNPs, OMIX tips and SPE cartridges were evaluated. Strickingly, in negative mode, MNPs perform the worst. It may be hypothesized that carboxylic acids strongly associate with the iron layer on the surface of MNPs and therefore do not get eluted. OMIX tips and C18 SPE cartridge worked as the best in this comparison.

Based on the presented results, C-18 OMIX tips were selected for plasma sample preparation prior to ICR-FT/MS analyses. They show the best average performance and have the above described advantages in sample expenditure and independency of vacuum pumps.

The results concerning MNP application are promising in order to overcome the issue of missing chromatographic separation prior to direct infusion. One disadvantage is the necessity of several centrifugation steps prior to dilution and direct injection of the sample, in order to remove the magnetic particles which remained suspended in the elute fraction. These steps are necessary in order to avoid problems with the high magnetic field generated from the ICR-FT/MS. Since the development of this protocol is intended to optimize the application of high throughput analyses of plasma samples, where the minimization of the sample preparation steps is desirable, this protocol does not present the imaginable optimum.

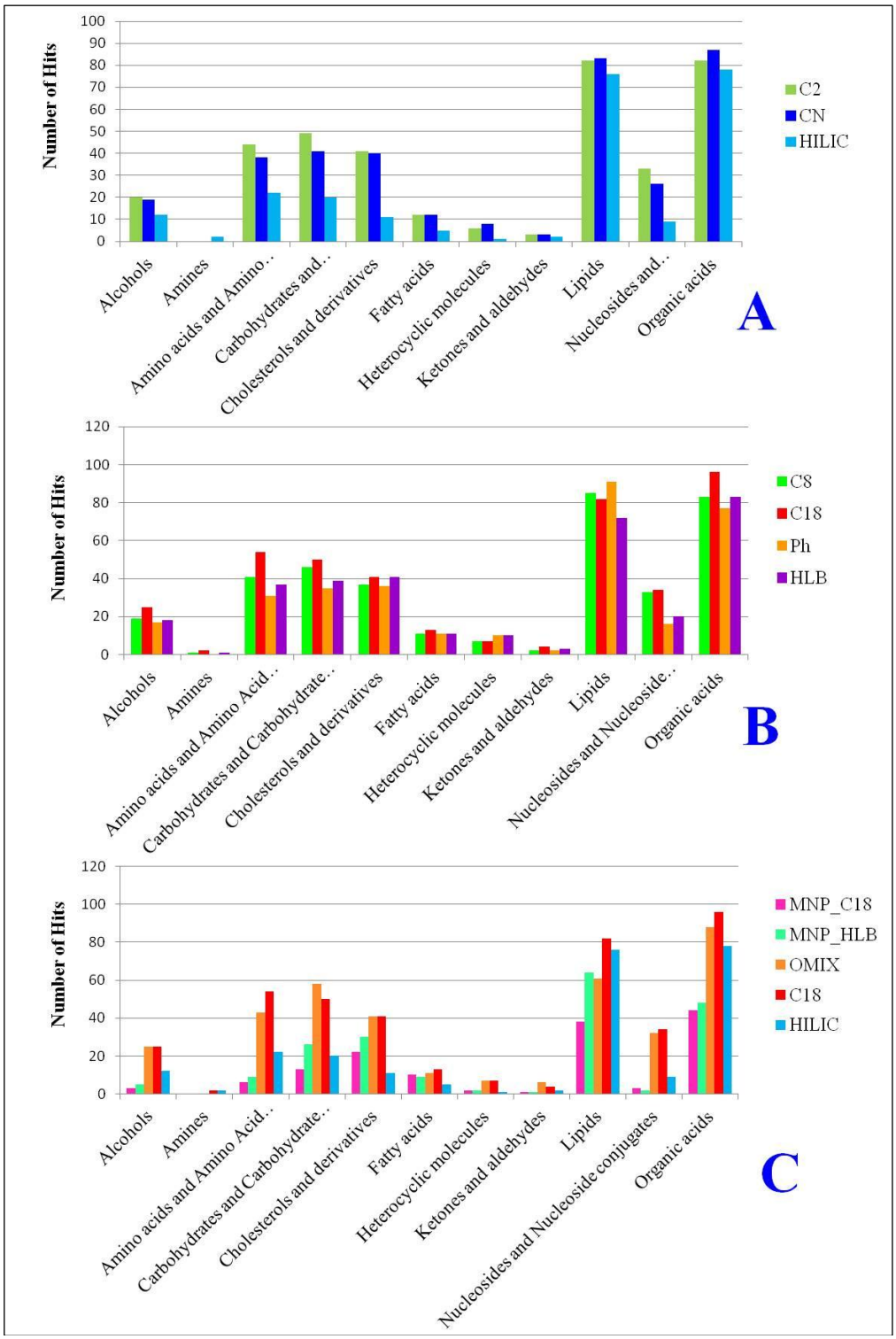


Figure 30: Different SPE cartridges tested in ESI negative mode. A) is a comparison of hits in normal phase sorbents. B) Compares reversed phase sorbents and HLB C) compares the final candidates as well as MNPs and the OMIX tips.

For the optimization of the protocol using C18 SPE cartridges or C18 OMIX tips in the sample preparation, two questions are still necessary to be answered.

1. Is an increase of elution volume beneficial?
2. Which is the effect of lyophilisation/evaporation?

C18 SPE cartridges of 1g sorbent amount were loaded with 300 µL of plasma. Different MeOH volumes such as 500 µL, 1000 µL, 1500 µL and 3000 µL were tested in the elute step (**Fig.31**). All collected eluents were consequently evaporated in a SpeedVac system and reconstituted to the desired dilution in MeOH/H_2O: 8/2 (v:v). Plasma was reconstituted to a dilution factor of 50 in respect to the original concentration (according to experience the best compromise between suppression effect and dilution). One additional experiment, using a MeOH elution volume of 500 µL, was performed once again without evaporation and therefore immediate dilution to the desired concentration in MeOH/H_2O:8/2 (v:v). The results are depicted in **Fig.31**.

Increased elution volumes and therefore longer evaporation, led to substantial losses in compound content relative to the non-evaporated sample (> 50% in the case of lipids). The loss is the most pronounced in the case of amino acids, cholesterols, lipids and organic acids. All metabolites containing carboxyl groups, except for fatty acids, were largelyeliminated from the sample. A reason could lay in MeOH esterification and subsequent increase in volatility. On the other sides the time factor may have played the major role. Conclusively, increased volumes to evaporate need to be avoided. This concerns especially HILIC applications and PPE.

The experiments in this section have revealed that MNPs are a promising alternative to SPE and PPE since they can be applied in small sample volumes, without evaporation and with free choice of sorbent material.

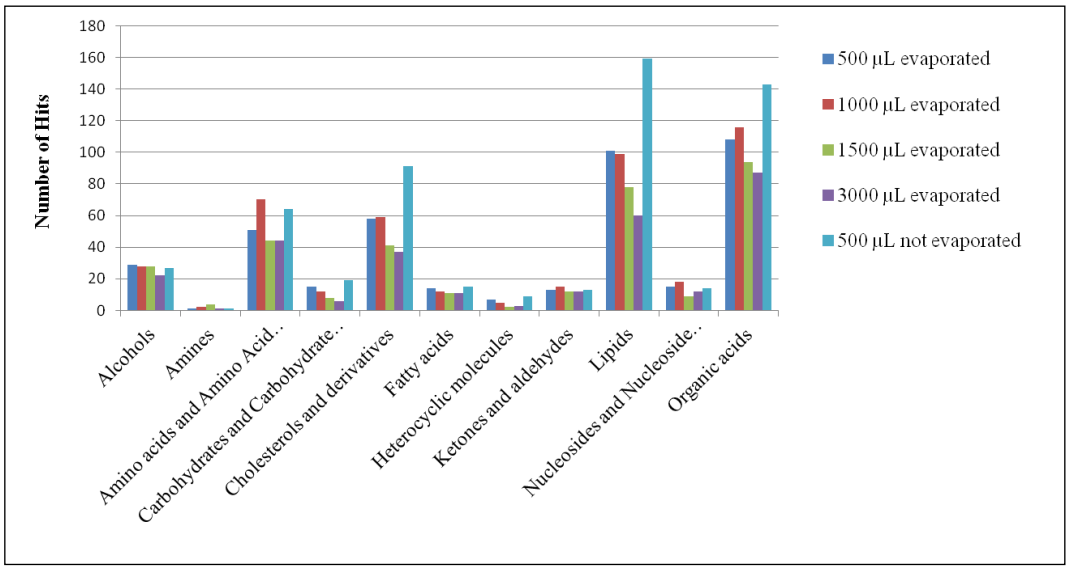


Figure 31: Counts of recovered compound classes as a function of elution volume and evaporation.

Mixing of different sorbents should be possible as well, in order to use the specific capabilities of different sorbents at once. SPE cartridges tested had substantial metabolite loss into the load fraction. As ICR-FT/MS appeared to provide enough sensitivity to capture the remaining information in the elute fraction, UPLC-TOF-MS could not cope with the substantial loss of concentration. For this reason PPE needed to be applied in UPLC-TOF-MS screening, whereas this same procedure does not seem to be beneficial for ICR-FT/MS. As mentioned above, a solution at different sample preparation requirements may be possible applying MNPs if the loss of metabolites detected in negative mode can be dealt with. Given the material available on the market we tested the most common variancies and found the at the time best solutions. These are the use of AcN-PPE for UPLC-TOF-MS applications and the use of C18 OMIX tips for ICR-FT/MS applications. The latter reduces the necessity for evaporation as well. Alternative procedures, such as on-line SPE, further developments in the field of MNPs and the application of 2D-LC will be tested in future.

4.3 Long term stability test of the UPLC-MS system

In order to test the reproducibility of the chromatographic system as well as of the mass spectrometer, a long term study over 100 quality control injections (i.e. 60 hours) was conducted.

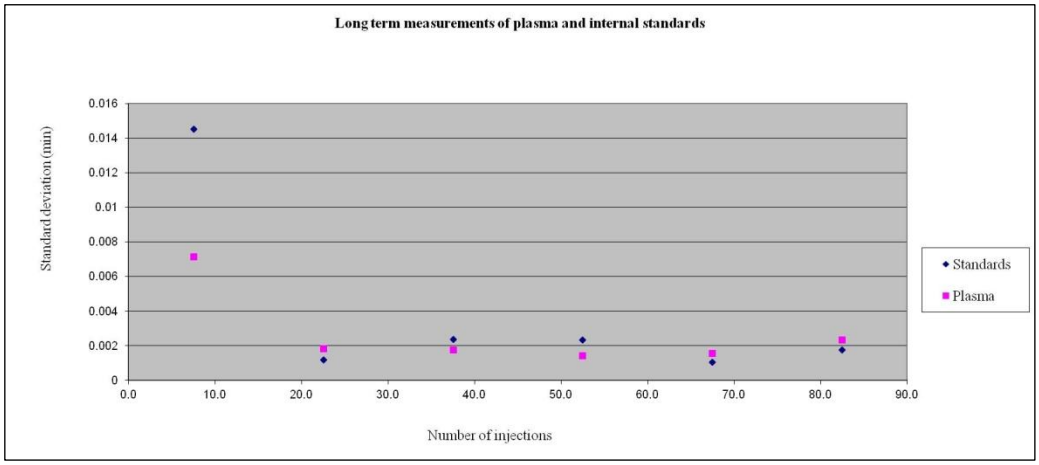


Figure 32: Long term test of plasma peaks and internal standards.

Retention time reproducibility of all internal standards and selected plasma peaks was tested in order to evaluate the equilibration time that the column needs to run with minimal variability. Retention time profiles, as function of injections, are congruent for both internal standard and plasma peaks. The column needed 20 injections in order to be equilibrated withthe matrix. Starting from injection 20 the reproducibility of the chromatographic system was excellent (**Fig. 32**).

Peak area and mass shift of standard compounds and of selected plasma peaks were as well calculated, the same sapproach will be used in the evaluation of the QC plasma in the fatty liver study (see Appendix).

In order to investigate the stability of the mass spectrometer we first calculated the RSD of all peak intensities.7.5% of all peaks showed RSD values < 30% which, according to literature references, is a weak reproducibility [209]. Similar studies in literature barely reported run times longer than 24 hours, nevertheless the reproducibility was evaluated in time intervals of 10 samples, in form of a "running RSD". The first RSD-calculation was conducted on the first 10 samples. The second calculation was based on QC 2 to 11. The third calculation was based

on QC 3 to 12 and so on. At the end a set of RSD values of the size n = 100 – 10 was yielded. For each n the averaged RSD, of the before mentioned 7.5% of peaks which had an overall RSD < 30%, was calculated. Plotting these values over the measurement time revealed a strong diurnal variation of the reproducibility of the signal intensities (**Fig.33**). As it is depicted in the following figure, a continuous oscillation about 12h peak to peak could be observed.

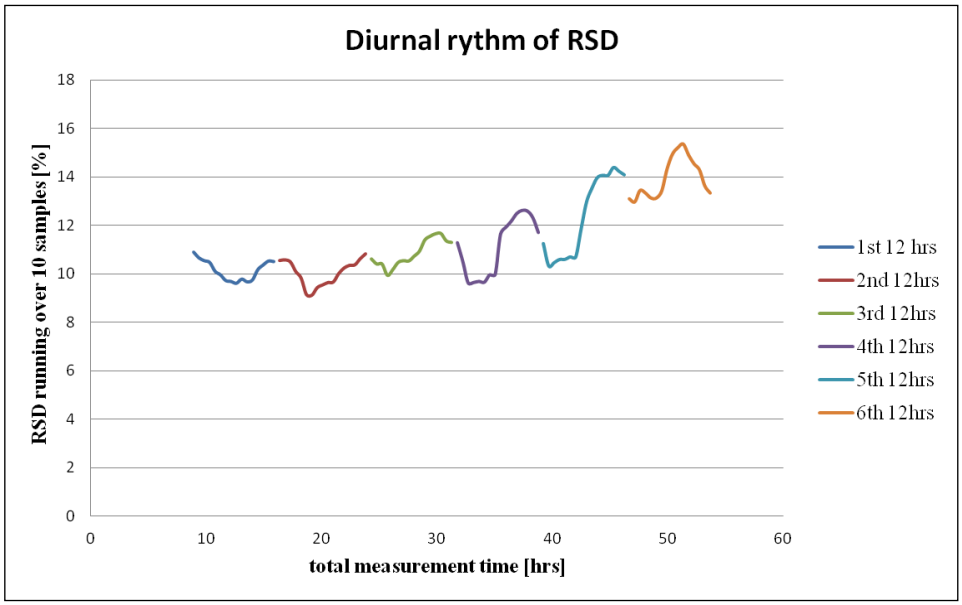


Figure 33: Running RSD showing the day time temperature dependency.

In order to exclude mistakes in the procedure, a hierarchical clustering analysis on the raw intensities of the same features responsible for the previous figure (using HCE software) was performed.

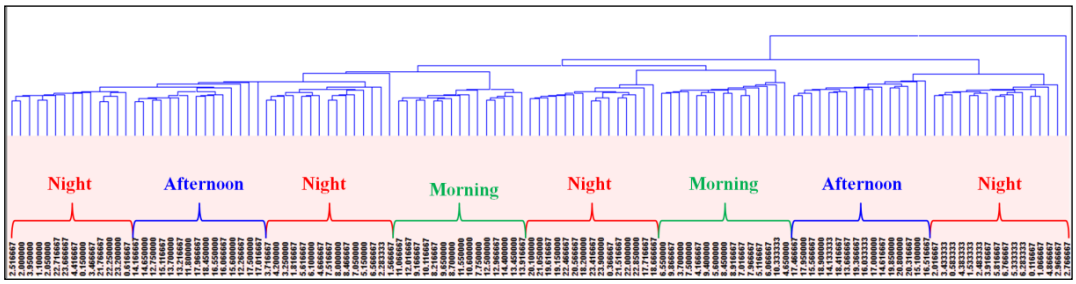


Figure 34: HCA of the day time dependency of the mass spectrometric signal variation.

Based on the HCA (HCE software) analysis using Euclidean distances and row-wise normalization on the mean the diurnal rhythm could be reproduced. The findings of the long term measurement revealed that the reproducibility of the chromatographic system was excellent. The mass spectromter response was

clearly influenced by temperature changes as a function of day time (**Fig.34**). This indicates, that the climatization of the UPLC-MS laboratory was insufficient and that the temperature insulation of the mass spectrometer was not satisfactory.
Conclusively, experiments longer than half a day might suffer from non-linear variation which, if possible, has to be considered in the data analysis despite the non-linearity of the effect.

5. Non-alcholic fatty liver disease: an overview

Non-alcholic fatty liver disease (NAFLD) is a disease of our generation that belongs to the metabolic syndrome and it is related to obesity and diabetes mellitus. This disease includes different disorders that range from the early status of *steatosis*, where patients present fat accumulation in the liver without presence of inflammations and in most of the cases asymptomatic; to the status of *non-alcoholic steatohepatitis* (NASH),where steatosis comes with lobular inflammation and ballooning degeneration with or without fibrosis; to the last phase consisted of *cirrhosis* [149, 150]. Before 1980 the status of hepatic *steatosis* was reported in liver of obese patients [151] during biopsy, then in 1980 Ludwig described histologically some changes in the liver of some patients not having reference to alcoholic history but comparable with the ones of the alcoholic liver disease [152]; calling this status *non-alcoholic steatohepatitis* (NASH).

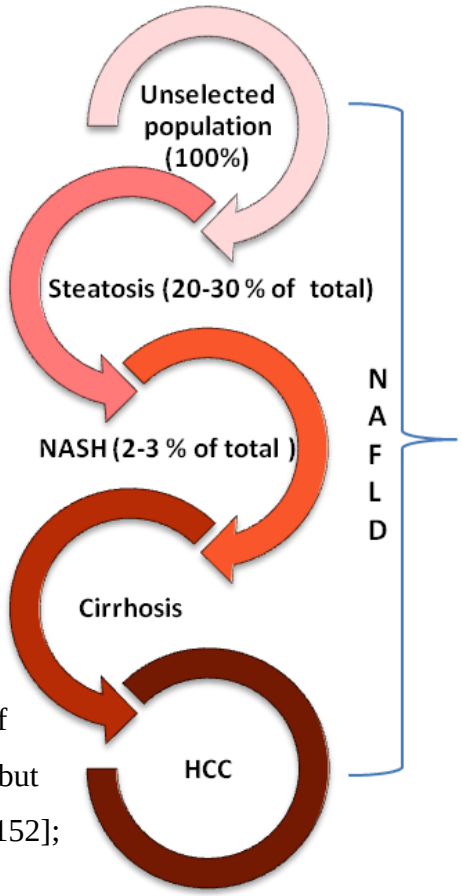


Figure 35: Schematic description of NAFLD. Adapted with permission, from Preiss D., Sattar N., 2008, *Clinical Science*, **115**, 141-150. © *the Biochemical Society*.

5.1 Epidemiology

The rate of NAFLD is very high in industrialized countries, it was shown to involve 10-20% of the population [153] with incidence of 20-30 % [154, 155] of steatosis among whom 1-2% may progress to cirrhosis and its complications in 15-20 years, while 12% of the chunk affected by NASH (2-3 %) may evolve to cirrhosis over 8 years [156]. Among the ones that suffer from cirrhosis 30-40 % can be affected by liver-related mortality. [157]

The incidences of NAFLD and NASH can increase with BMI (body mass index) reaching percentages of 65-75 % [153, 158] with BMI > of 30kg/m^2 (obese individual) and 85-90% [159] with BMI > of 35 kg/m^2 (morbidly obese individual) for NAFLD and 15-20 % [160] for NASH in obese individuals. Studies of NASH patients have shown that an incidence of 40-95% will be affected by obesity, more than half by Type 2 Diabetes mellitus and up to 80 % by dyslipidaemia [161]. Female alike males can be affected from this disease, but some studies show a higher incidence in men than in female morbidly obese patients. In the Asian population [162, 163] as well as the Hispanic population [164] it was shown that they present higher levels of visceral adiposity at lower BMI compared to the Caucasian population, so it was necessary to set different standard levels for the BMI evaluation [165, 166]. Data has also shown an increase in children [167]. The link between the adiposity position and NAFLD development was investigated among histologically confirmed patients and it was found that the peripheral adiposity (arm and hip) is more predictive of fibrosis than the central one (measured at the waist); it was also found that the connection of the latter can be independent of the overall obesity (defined by BMI) [168]. The visceral adipocytes are more resistant to insulin than the subcutaneous ones and they are combined with high levels of inflammation. Insulin resistance and hyperlipidemia (especially concerning triglyceride concentration) are independently associated with the development of the disease .

5.2 Diagnosis

NAFLD mainly is present in asymptomatic subjects, in some cases some of them can refer fatigue and right upper quadrant pain and possible hepatomegaly during manual visit. Cases of Acanthosis nigricans are being reported in children [169].

Serum Parameters

Clinical biochemistry parameters (**Table 9)** like raised fastening glucose, low HDL level and high fastening triglycerides can be an indicator of NAFLD presence, of course before alcohol excess related pathogenesis needs to be excluded. There are no biochemical markers that can validate the development of NAFLD and to identify the disease stage; some serum abnormalities are common in NAFLD individuals and they can be an alarm bell that induces appropriate imagine investigations. Fatty liver related abnormal biochemical patterns can be characterized form elevated serum levels of alanine aminotransferase (ALT), aspartate aminotransferase (AST) commonly no greater than four times the normal upper limit and elevated γ-glutamyltransferase (GGT); the ratio of ALT to AST (aspartate aminotransferase) is variable and should usually be less than one [170]; in some cases the ratio can be reverse when the level of AST is getting higher, and this is a sign of cirrhosis development. Some individuals present elevated levels of alkaline phosphatase as well [171]. ALT is strongly associated with waist circumference and hyperinsulinaemia, indicators of insulin resistance, and GGT is not a specific marker but usually present at high levels in NAFLD; both are correlating with the amount of liver fat measured via MRI (magnetic resonance imaging) in adults and in children. So the liver enzyme level can be used as a rough estimate of the presence of fatty liver disease. Some anomalous serum biochemical markers can be resumed in the following table. Additional abnormalities like hypoalbuminemia, prolonged prothrombin time and hyperbilirubinemia can be present in individuals already in cirrhosis status [172].

Table 9: Prevalent abnormal biochemical serum makers in NAFLD individuals. Adapted with permission, from Preiss D., Sattar N., 2008, *Clinical Science*, **115**, 141-150, © *the Biochemical Society*.

Feature	NAFLD
ALT	↑
AST	→
ALT/AST	> 1
GGT	→↑
Weight	↑↑
Fasting plasma glucose	→↑
HDL-cholesterol	↓
Triacylglycerols	↑
Alchol intake	↓

Imaging

Beside the clinical serum routine analysis, the diagnosis of fatty liver demands a specific and accurate imaging analysis. The possible tools are ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI). Ultrasound is the cheapest non-invasive tool for the diagnosis of fatty liver disease, but is less accurate when the detection of liver fat below 30% is necessary and it is unable to detect inflammation or fibrosis. Another kind of ultrasonography resulted in using a contrast agent (Levovist) in order to detect NASH status. This contrast agent is based on galactose and palmitic acid that are being taken up by the hepatocytes; NASH individuals show reduced uptake of Levovist [153]. CT is also another possibility to evaluate the steatosis stage with a degree major to 30%. Some studies have shown that non-contrast CT results are better than the one using contrast-enhanced scans.

Magnetic resonance imaging (MRI) is a reliable tool for the detection of steatosis even with a degree of 3%. Fatty changes are evaluated via differential chemical shifts between fat and water. A variant of MRI is proton magnetic resonance spectroscopy (MRS). Liver elasticity is another alternative to the other ultrasound techniques; an ultrasound probe emits a vibration that creates a shear wave in the liver, this one corresponds to the liver rigidity; liver rigidity can also be measured by MRI. In MRE (magnetic resonance elastography) a mechanical wave is created and MRI is used to measure the liver displacement that is getting converted as measure of elasticity. Histology is still the only methodology sufficient for the evaluation of the NAFLD stage. Diagnosis of NAFLD can be done though integration of serum analysis, history and abdominal imaging, but to verify the severity of the disease a biopsy can be necessary, especially if it is been considered that it would allow the monitoring of cirrhosis complications and screening for hepatocellular carcinoma (HCC) [154].

5.3 Pathogenesis

The development and the progression of NAFL expose the affected individuals to possible advancement of cardiovascular disease (CVD) and type 2 Diabetes, if they are not already affected from the latter one. Therefore this disorder can be seen as a possibility to watch into the window of diabetes prediction in order to understand the correct way leading to prevention and avoidance of the diabetic status. Important clinical fatty liver related markers as ALT and GGT, are independently linked with the Diabetes development.

Fatty Liver and Insulin Resistance

The increase in body fat is an indicator of insulin resistance but the accumulation of ectopic fat (in the liver and skeletal muscle) is even more related with insulin resistance. Liver fat is made up from products of *de novo* lipid synthesis and

circulating non-esterified fatty acids (NEFA). Approximately 60 % of the liver fat in NAFLD patients, under a normal fat intake diet, is coming from NEFA [174]. NEFA concentrations are higher in obese individuals, systemic values are lower when the parameters are referring to the total body fat. Furthermore rising flux of NEFA is linked with hepatic steatosis.

In skeletal muscle the presence of NEFA induces insulin resistance development: in normal subjects the insulin stimulation of IRS (insulin receptor substrate)-1 activates intracellular P13K (phosphoinositide kinase) that stimulates GLUT (glucose-transporter)-4, which finally allows glucose to enter the cell. Increase of NEFA levels and with this of diacylglycerol leads to a diminished activity of PI3K and consequently impairs glucose entrance [175]. Analogy to the skeletal muscle is being presented in the hepatic cells where under insulin resistance the increased flow of NEFAs conduces to intracellular accumulation of diacylglycerol and therefore decreased activity of IRS2 via serine kinase cascade activation. As in muscles the decreased action of PKI3K will drive into abnormal glucose production/ gluconeogenesis.

Liver Fat and Inflammation

The pathogenesis and progression of NAFLD is described as an integration of insulin resistance, oxidative stress and inflammation. The liver is involved as well in the immune response [176, 177]. 2/3 of the total cells in the liver are hepatocytes. Other kinds of cells are biliary epithelial cells, sinusoidal endothelial cells, Kuppfer cells, stellate cells, dendritic cells and lymphocytes.

Insulin resistance evolves as consequence of increased flow of free fatty acids (FFA) that are being absorbed by the liver which entails the steatosis status (the first hit). This condition extends to a scenario where the interaction (the second hit) of hepatocytes, stellate cells, adipose cells, Kupffer cells, inflammatory mediators and reactive oxygen species (ROS) finally yields into the development of the NASH status or cirrhosis. Fatty acid metabolism is linked with the beta-oxidation in mitochondria and peroxisomes. Mitochondria are the main source of ROS production that can lead to steatohepatitis and fibrosis via lipid peroxidation

leading to cytokine induction and Fas ligand induction. Steatohepatitis subjects present abnormalities in the mitochondria structures, a state not present in the case of steatotic individuals and healthy ones. The link between fat accumulation and inflammation is also being described by the role of the transcriptor factor NF-kB (nuclear factor kB). In animal models it was shown to be up-regulated in the liver during high fat diet. This factor is activating the production of TNF (tumor necorsisi factor)-α, IL (interleukin)-6 and IL-β that possibly can have a role in the activation of Kupffer cells, macrophages in the liver tissue and progression of NAFLD. It was additionally found that the increased expression of NF-kB induces insulin resistance. This indication offers an alternative to the "double-hit" hypothesis [178] where the first hit leads to the fat accumulation and the second one to the progression.

Inflammated adipose tissue in obesity as well secretes high amounts of pro-inflammatory cytokines as TNF-α and ILs (IL-6), that compromises the production of the insulin-sensitizing adipokine adiponectin. The impaired secretion pattern of the adipocytokines is considered a link between obesity and fatty liver [179].

The action of adiponectines is related as well with an increase in lipid oxidation in liver and skeletal muscle and decrease of the activity of enzymes involved in fatty acid synthesis such as acetyl-CoA carboxylase and FA-syntase. Another important regulator of the liver fat is considered leptin, even though the mechanism of action is still unclear. It may have antisteatotic effects by enhancing lipid oxidation and inhibiting lipogenesis in tissue. Additional cause of the hepatic inflammation is indicated as endoplasmatic reticulum (ER) stress. In murine liver and adipose tissue of models, with genetic or diet induced obesity, activation of multiple stress response that induces the activation of CREBP (an hepatocyte-specific transcriptor factor that maybe is involved in the hepatic acute phase response) was observed [180]. ER is involved as well in the generation of ROS species and therefore oxidative stress. Systemic subclinical inflammation can be estimated by circulating CRP levels, the plasma levels of this protein are low in healthy people under normal conditions but increase under inflammatory processes; this protein represents a useful early non-specific marker of inflammation (half-life of 18 h) and is produced mainly by hepatocytes regulated by IL-6 and other

proinflammatory cytokines. Circulating CRP is positively correlated with liver fat [181]. Its levels are higher in NASH patients compared with steatosis ones .

Causes of Fatty Liver and Role of Nutrition

The principal origin of FFAs in the systemic plasma pool, during the fasting state, is subcutaneous fat [182]. Independently from the origin of FFAs, increased hepatic lipid supply is probably contributing to hepatic fat accumulation [180, 183] ; supported by experiments showing that exogenous lipid infusion and high fat diet increase liver fat content and hepatic insulin resistance, where low fat diets show the opposite effect [180, 183]. When subcutaneous adipose tissue is absent or deficient, the excess of calories cannot be stored in the insulin sensitive tissue. The expansion of the visceral fat mass, as well as ectopic fat accumulation in the liver and skeletal muscle is a consequence of the inability of the body to store energy. This is a state that is driven by insulin resistance of subcutaneous adipose tissue [184]. De novo lipogenesis (DNL) may also induce insulin resistance. Dietary glucose and fat are important regulators of DNL via activation of ChREBP and SREBP-1c, two transcription factors [185]. The first one stimulates pyruvate kinase increasing the glycolysis of glucose into pyruvate that is involved in the formation of AcetylCoA and MalonylCoA, the latter one is required for the synthesis of FFAs. The activities of these two transcription factors are increased in animal models of fatty liver [186]. Human studies show that fatty liver affected patients have higher saturated fat and cholesterol intake compared with the healthy controls, as well as lower intake of polyunsaturated fatty acids, fiber and antioxidant vitamins like Vitamin C and E [187]. Low-calory diet was demonstrated to have positive results and reduce liver fat in obese patients with NASH [188]. The restriction of saturated fat intake was shown to be particularly effective. Life intervention strategies aim at weight reduction by restriction of total and saturated fats intake, in combination with an increase in physical activity.

The Treatment of NAFLD

The care of NAFLD lies in reducing the risk factors and improvement of clinical and histological parameters. In the worst case of decompensated cirrhosis liver transplantation is inevitable. The pathophysiology of fatty liver is rooted in behavioral factors. Diet composition was found to be relevant. Moreover some studies showed that a sedimentary lifestyle with reduced physical activity, independent of the diet, is another determinant. Life style intervention, diet and physical exercise have been demonstrated to bring improvements in clinical and histological parameters. An amelioration of the steatosis status [188] was observed but not significant reduction in fibrosis. Modification in the diet composition shows as well good improvements. The benefit of lower carbohydrates intake was indicated [189]. A moderate weight loss is recommended. A too rapid weight loss was found to deteriorate liver histology, possibly due to increased lipolysis [190]. Additionally, the risk of developing NAFLD with high fructose intake was presented [191] seeing that lipogenesis, hypertriglyceridemia and insulin resistance raised.

Exercise intensity is not an independent determinant of liver fat. Mitochondria, whose function can be estimated by measurement of aerobic fitness, maybe are involved in the pathophysiology of hepatic steatosis. Recent studies have shown a close relation between aerobic fitness and liver fat [192]. The mechanisms elucidating the link between fitness and liver fat involve factors regulating hepatic lipid oxidation. Mitochondria are playing an important role in hepatocyte metabolism, since it is the primary site for the oxidation of fatty acids and oxidative phosphorylation. Hepatocytes are rich in mitochondria which occupy around 18% of the liver cell volume; therefore mitochondria are the major regulators of the liver fat [193]. When mitochondrial functions are impaired or when an excess of FFAs is available, as it occurs in fatty liver, reactive oxygen species (ROS) can increase leading to oxidative stress. This plays an important role in the development of NASH and fibrosis [194]. The pharmacological approach is another alternative in subjects where the life style intervention was

not effective. The application of pancreatic lipase inhibitors and appetite-suppressive agents with GLP-1 agonists was reported. Larger human trials are necessary in order to evaluate their efficiency. Baratric surgery is considered necessary in patients with BMI greater than 40 kg/m^2, or for the ones with BMI greater than 35 kg/m^2 and affected from obesity–related conditions (hypertension, insulin resistance, hyperlipidemia and obstructive sleep apnea). This kind of treatment shows benefit in clinical and hitological NAFLD markers, in subjects with higher BMI and severe insulin resistance the status of steatosis prevailed after the operation. This treatment is being considered an alternative when the patient does not answer positively to the life style intervention via diet, exercise and pharmacological therapy. Additional kinds of pharmacological treatment involve the use of insulin-sensitizing agents like the thiazolidinediones, which are agonist of the peroxisome proliferator-activated receptor (PPAR-γ). These kinds of medications are recommended in patients without heart failure history [195]. Another pharmacological possibility is treatment with Biguanide-Metformin; it was shown to have positive effects in reduction of clinical parameters, inflammation and fibrosis. It is suggested in the cure of NAFLD patients who are reporting hyperglycemia but they do not present any chronic renal insufficiency, congestive heart failure or sepsis [196].

Antioxidants like vitamin E and betaine were also investigated as possible treatment for NAFLD showing the benefit in terms of diminished inflammation response involved in the development of NAFLD [197] but larger trials are necessary. In malignant status where NASH degenerates to liver failure liver transplantation is necessary; some cases show that NASH can revert [198].

Correlation with Insulin Resistance

Fatty liver and obesity strongly correlate with insulin resistance, which leads to Type 2 Diabetes manifestation and cardio vascular disease. As fat accumulates in hepatic tissue it interferes with insulin signaling in liver cells of animal models. Impaired stimulation of the insulin receptor substrate (IRS)-s and IRS-2 may lead to hepatic insulin resistance and increased gluconeogenesis [199]. A strong

correlation between fat accumulation in the liver and whole body insulin resistance exists in humans. Nonetheless, NAFL may develop independent of insulin resistance in skeletal muscle and adipose tissue [200]. Fetuin-A, a human protein (α_2-Hermans-Schmid glycoprotein (AHSG)) is one factor regulating insulin sensitivity. This protein is predominantly expressed in the liver. Placenta and tongue show low expression levels as well. The liver is the only organ which regulates the circulating levels of of Fetuin-A. This protein naturally inhibits the insulin receptor tyrosine kinase in liver and skeletal muscle [201]. Experiments with mice which were deficient for the gene encoding fetuin-A, showed improved insulin signaling indicating that this protein has a key role in the insulin sensitivity regulation in animals [202]. SNPs in the fetuin-A gene (AHSG) are associated with Type 2 Diabetes in humans. Circulating fetuin-A correlates positively with liver fat and as well negatively with insulin sensitivity [203]. High levels of this protein are predictive for insulin sensitivity changes and point towards incident type 2 diabetes. Fetuin-A is increased in NAFL and insulin resistant individuals. It may as well be associated with the metabolic syndrome and correlates positively with inflammation markers like CRP levels. Fetuin-A is as well the promoter of cytokine expression in monocytes and adipocytes and inhibitor of the insulin-sensitizing adipokine adiponectin.

5.4 Matbotyping of Non-alcoholic Fatty Liver individuals: the study

5.4.1 Introduction

As already described in the introductive part, non-alcoholic fatty liver (NAFL) is characterized by many phases ranging from insulin resistance to inflammation. Interesting is the condition of those subjects affected by NAFL but not showing an insulin sensitive pattern. Accumulation of fat interferes with correct insulin signaling in hepatocytes, the origin of this condition can be located in the fatty acid and diacylglycerol (DAG) metabolism and in endoplasmic reticulum stress. The latter may cause ineffective intracellular trafficking and signalling.

Triglycerides (TAGs) represent the least toxic form for the storage of the fat in excess in ectopic tissues. Therefore the role of the triglycerides in conjunction with oxidative fatty acid degradation and stearoyl-CoaA desaturase SCD 1 that converts saturated fatty acids in monounsaturated fatty acids can be seen as a short time shield to overcome lipotoxicity. When these temporarily safeguarding mechanisms are beaten, abnormal metabolism is caused by cell injury from fatty acid action. Genetic studies have indicated the clarification concerning TAG's detoxification role in the single-nucleotide polymorphism SNP rs738409 in the patatin-like phospholipase 3 gene PNPLA3. It was found to be highly correlated with liver fat but not with insulin resistance, hyperlipidemia and inflammation. Additionally, in fatty liver biopsy the G allele was found associated with steatosis, steatohepatitis, fibrosis and cirrhosis but with a decrease of Type 2 Diabetes risk, lower serum lipid concentration and lower blood pressure. It was appointed as the orchestra director of the above mentioned shield mechanism [204]. Controversially in alcoholic fatty liver (AFL) saturated fatty acids protect against the progression and degeneration of the disease, probably due to the alterations in fatty liver metabolism due to alcohol presence [205].

The aim of this study is to investigate the metabolic profile of 40 different non-alcoholic fatty liver disease (NAFLD) patients divided into two groups according to their whole-body insulin sensitivity and to understand the pattern responsible

for the discrimination of those patients that show normal levels of insulin sensitivity, even though they are affected by NAFLD. Of additional interest is the measurement of plasmata sampled from the same fatty liver patienrs in a follow up study after nine months of life style intervention.

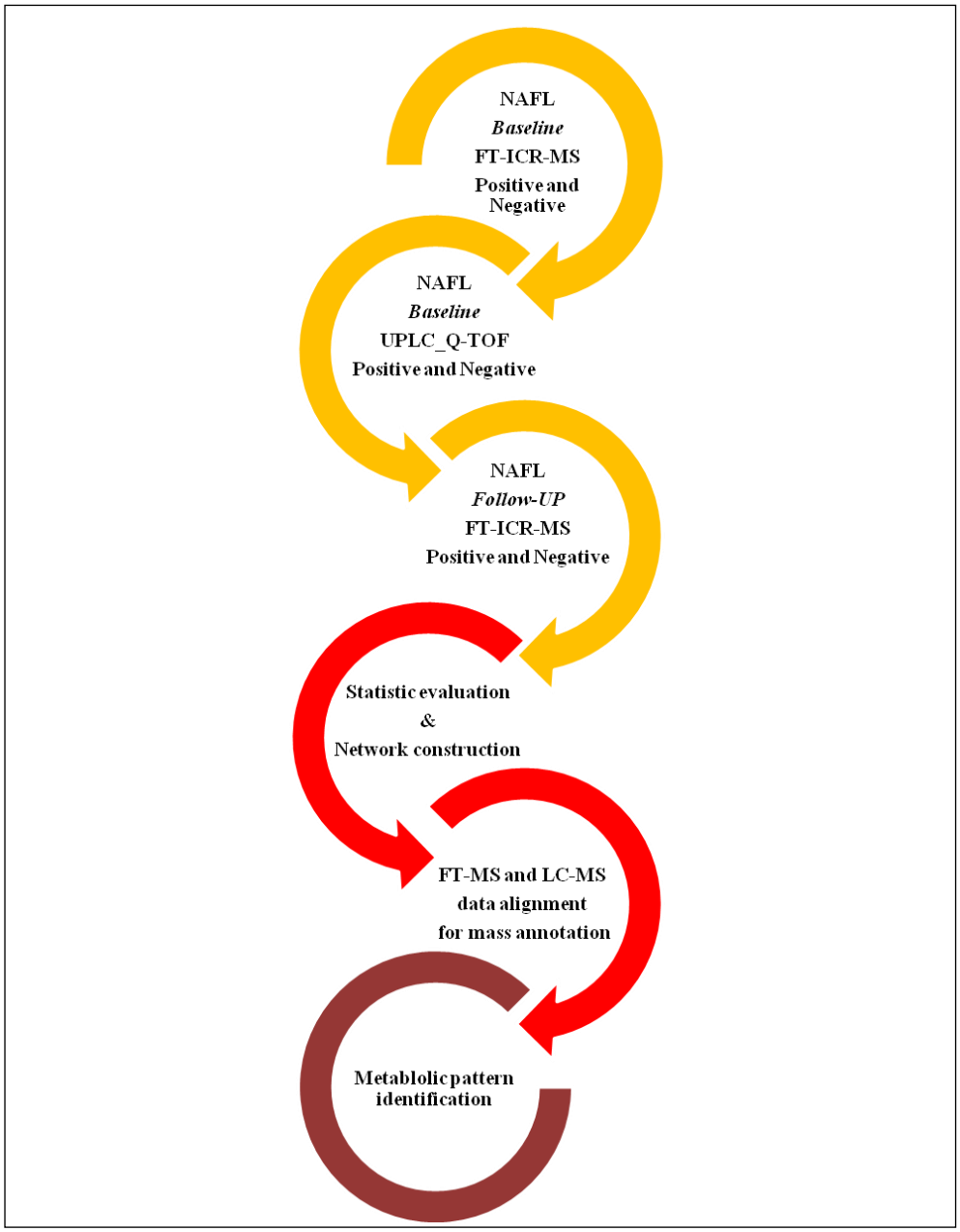
In **Table 12** (Appendix) the clinical parameters of the patients involved in the study are shown. The population study group consisted in 75% males in the insulin sensitive (IS) group and 65% of females in insulin resistance (IR) group. Insulin sensitivity was calculated via $ISI_{Matsuda}$ Index [101] that derives from oral glucose tolerance tests (oGTT) and mirrors insulin sensitivity in the muscle and in the liver showing lower values of $ISI_{Matsuda}$ in pre-diabetic status. $ISI_{Matsuda}$ values differentiated in respect to the IS and IR groups and they ameliorated (increased) in both groups after life-style intervention. Liver fat measured via magnetic resonance tomography (MRT) decreased in both groups from baseline to follow up. Insulin resistant individuals showed higher CRP (C reactive protein) values at baseline than insulin sensitive individuals. CRP values notably diminished in IR after the life style intervention. No remarkable difference was noted in the fasting glucose in both groups with slight amelioration after intervention. Elevated values of plasminogen activator inhibitor 1 (PAI-1) and Fetuin A in insulin resistance indicate hepatic inflammation status [206].

Experimental Strategy

Global plasma metabolic profiling of NAFL subjects was performed through a non-targeted metabolomic platform based on the analytic combination of ion cyclotron resonance Fourier transform MS (ICR-FT/MS) with high magnetic field (12 Tesla) and UPLC coupled to Q-TOF-MS. The strategic power of this approach lies in the integration of ultra high chromatographic resolution with the high resolving power of ICR-FT/MS that allows exact mass annotation of the detected ions. UPLC additionally provides the possible quantification of the interesting compounds and information concerning isomer patterns. The coupling of UPLC with the Q-TOF mass spectrometer offers a good chromatographic separation of the sample matrix combined with a mass spectrometric resolution of 10000. Therefore, the integration of UPLC-MS data with the annotation of ICR-FT/MS data is an appealing resort to ambit high precision mass accuracy annotations.

The issue that suggests to point at this direction lays in the limitation of the mass accuracy performance of the Q-TOF mass spectrometer after long time measurement. In order to maintain high mass accuracy, and to compensate for mass drift, mass calibration is required supplying lock mass as reference (Leucine Enkephalin). This reference compound is injected at regular intervals of 15 seconds from an orthogonal spray (the lock mass spray). The lock mass correction is applied on the experimental masses following the calibration curve generated during the instrument calibration. Due to the fact that only one mass is being applied for mass correction, the surveillance of the dynamic range of mass shifts is limited. Additionally, as shown in Chapter 4, diurnal variation of temperature in the laboratory and high sensitivity of the system to changes of temperature, increase the intensity variability of the detected features and influence the mass accuracy. The integration with ICR-FT/MS data results in an elegant alternative to lock mass post calibration because it allows the alignment of UPLC-TOF/MS data with masses that were annotated within a 1 ppm error window on the basis of the NetCalc algorithm.The best NetCalc-reaction assignment is generally within an error window of 0.1-0.2 ppm which is an infeasible condition in the case of assignment with Q-TOF detected masses.

The following scheme describes the experimental work flow followed in the study of fatty liver samples. ICR-FT/MS measurements of the baseline samples in both ionization modes were performed. The same samples were analyzed via UPLC-Q-TOF-MS in order to complement the isomeric information. The follow-up samples were acquired via ICR-FT/MS in both modes. Here the results of the positive mode study are reported, since they were the most rich in information. Nevertheless the important makers which were as well confirmed in the negative ionization study will be mentioned as evidence of their discriminative behavior among the two observation classes (IR and IS).



5.4.2. The Baseline study – The Results

Statistical strategy

In order to find a possible metabolic patterns that could differentiate the two respective groups of observations (IR versus IS), multivariate analysis is the most appropriate approach for screening this complex human data set [50]. For each data set outliers as well as features with less than 10% abundance among all samples were excluded. Wilcoxon signed rank sum test was then applied to the data in order to get examine univariate discrimination among the two groups of interest (IR vs IS) according to the two-sided p-value within a significance level of 0.05. Before proceeding with the combination of different multivariate techniques all the data were scaled using unit variance scaling. An orthogonal signal correction (OSC) has been applied [207] to the data in order to remove the information that is orthogonal to Y (in this case represented by a vector with dicotomic values identifying the sample classes) from the matrix X (m/z - sample matrix). Additionally, principal component analysis (PCA) and orthogonal-partial least squares (OPLS/O2PLS) models were generated obtaining valid models (see Chapter 3 for theoretical introduction). The multivariate approach led to the selection of a list of masses discriminative for the different classes. The selection has been estabilished by the absolute value of the regression coefficients in parallel to the reported values of the Wilcoxon test.

ICR-FT/MS results

Through ICR-FT/MS in positive mode 17934 ions were detected. After a 10% cut off in frequency of masses the assignment and the statistical analysis was conducted on 8432 masses of which 30% were assigned through MassTRIX. The rest of the unknowns were annotated via NetCalc. The statistical procedure was applied as described above. PCA analysis (**Fig.36**) did not provide good separation among the two classes; the reason lies in the fact that the trend of

biofluids data within metabolomic studies needs more sophisticated statistical analysis for optimization and adaptation of the models.

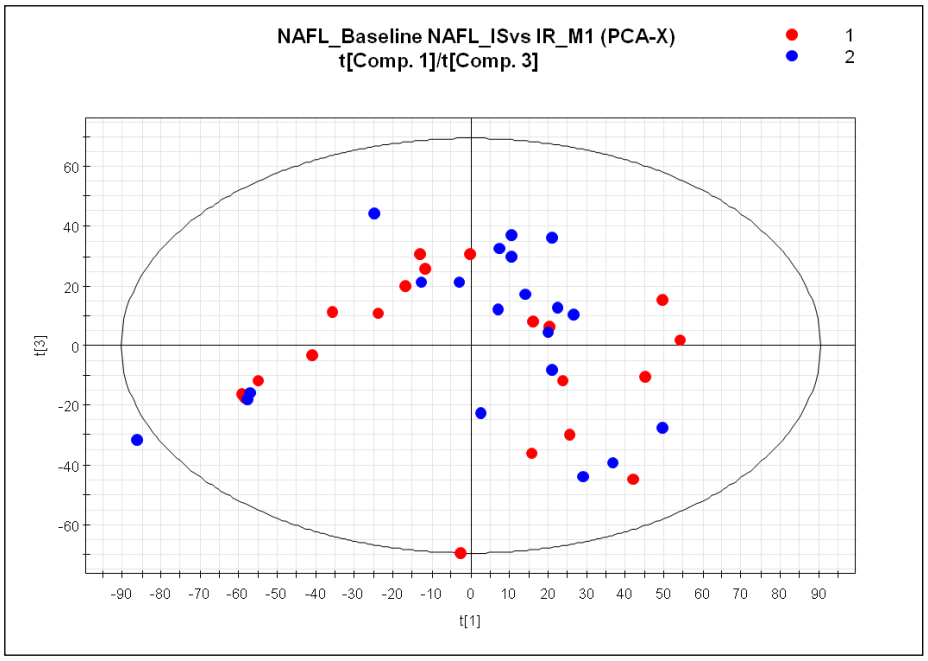


Figure 36: The score scatter plot of the PCA (colored following the classes). The first component absorbs 13 % of the variance the third component the 7%. (blu= IS and red=IR).

An OPLS/O2PLS model (**Fig. 37**) could differentiate the extremity of the two discriminative classes. Outliers were not considered for further analysis. PLS-modeling finds the direction through a normally distributed multivariate data cloud which explains most of its data variation and at the same time has high correlation to the classes of observations. The covariance and correlation between the metabolites and the modeled class designation was visualized by the S-Plot [208]. The extremal values in both directions give the most important m/z values the characterisation of the two different classes. The upper and lower 2.5% of the S-Plot curve signify the most discriminant masses. The intercept with the y-axis was taken as trend separator. Therefore cutting off 2.5% on each side of the plot leads to a slight assymmetry concerning the amount of found features. The longer tail marks the more significant group. 173 metabolites were classified as important. 94 features were up-regulated in the IS group, 79 features were down-regulated in regard of the same class (Appendix).

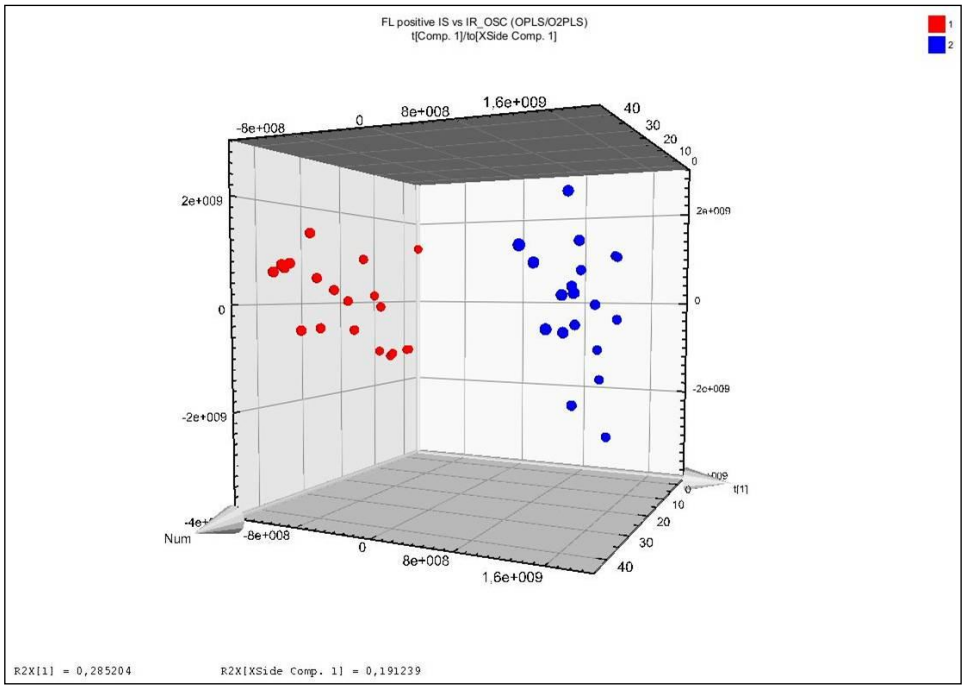


Figure 37: OPLS/O2PLS score scatter plot in which is visualized the two different group (blu= IS and red=IR), with Q^2(cum)=0.99 and R^2Y(cum)=0.99.

Of the 173 important masses 46 metabolites were assigned via MassTRIX and Netcalc; the sum formulas of 54 were annotated via NetCalc only; the rest remained classified under unknown (**Fig. 38**).

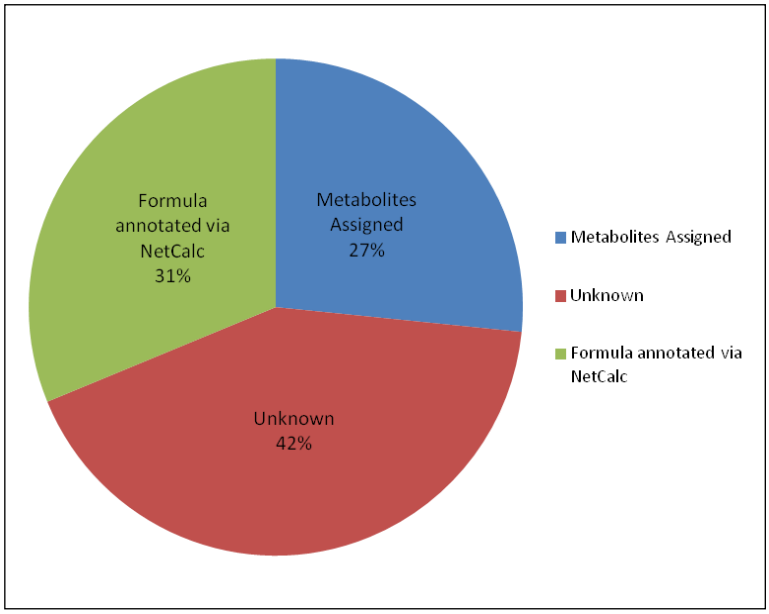


Figure 38: Percentage of annotated masses and unknowns among 173 marker candidates

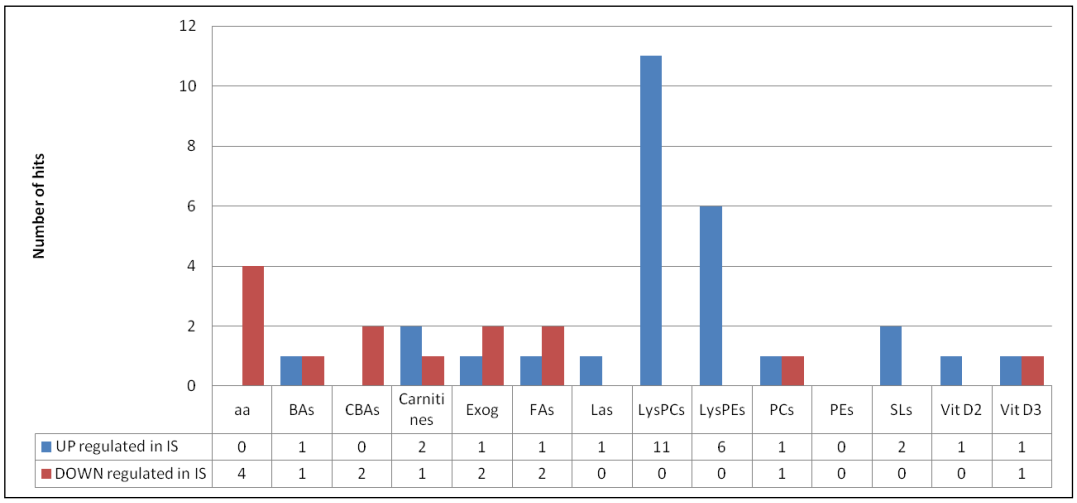


Figure 39: Important biomarkers in NAFL via FT-MS-ICR analysis.

In **Fig. 39** there is a description of the compounds up and down regulated in respect to the insulin sensitive class. A protagonist role in the insulin sensitive scenario is played by the class of phospholipids, especially LysoPhosphocholines (11 different formulae) and Lysophosphoethanolamines (6 different formulae). On the insulin sensitive down regulation horizon a group of amino acids is the prominent class, especially aromatic amino acids (Trp, PheAla; Tyr and Indoleacrylic acid that is Tryptophan metabolism related). Details on these classes of compounds along with their relation to clinical parameters such as fastening glucose, liver fat, CRP and ISI Matsuda will be discussed after the presentation of the LC-MS results

UPLC-Q-TOF-MS

Samples were analyzed with RP-UPLC-MS in randomized triplicates. For the sample preparation, chromatographic and mass spectrometric details see the Appendix. For the evaluation of retention time and mass accuracy drifts, samples were partially spiked with isotopically labeled standards. The use of a wide time range of standards is necessary but it is essential to consider possibly occurring suppression of low abundance matrix ions as well; therefore the number of standards used and their concentration has to be considered. The internal standards used during the study were aimed at a good retention time and mass coverage along the chromatographic gradient; the number of the aforementioned

standards comes from the decision of not overloading the matrix in order to avoid suppression effects in these highly valuable samples. In non-targeted metatabolmics the least number of treatment and passages prior analysis should always be preferred. Therefore a set of plasma from fatty liver subjects spiked with internal standards (Study 1) were analytically compared to a non-spiked plasma set (from the same subjects; Study 2). Quality control samples were evaluated during the entire study in order to verify the technical quality of the experiments in terms of mass accuracy, chromatographical robustness and intensity variability. The evaluation of the quality controls is an important step that reveals the goodness of an experimental set up [209, 65]. In the following table the standard compounds used during the experiments are listed.

Table 10: Isotopical labeled standard compounds used in NAFL study in positive ionization mode.

Nr.	Name	Formula	Neutral Monoisotopic Mass (m/z)	Possible ion adduct ions (m/z)		
				$[M+H]^+$	$[M+Na]^+$	$[M+K]^+$
1	[d_3]Acetyl-L-Carnitine	$C_9D_3H_{14}NO_4$	207.141865	207.14186	229.12381	245.09774
2	Nialamide	$C_{16}H_{18}N_4O_2$	298.142976	299.15052	321.13219	331.10613
3	[d_6]Sulfadimethoxine	$C_{12}D_6H_8N_4O_4S$	316.111236	317.11851	339.10045	355.07439
4	Reserpine	$C_{33}H_{40}N_2O_9$	608.273381	609.28065	631.26260	647.23653
5	[d_3]Hexadecanoyl-L-Carnitine	$C_{23}D_3H_{42}NO_4$	403.361514	403.36151	425.34345	441.31739

Extracted ion chromatograms of each internal standard mass were compared in order to evaluate the variation of RT (RSD %) the peak area, intensity and mass accuracy during the two different studies.

Table 11: Decription of averaged mass error of standard compounds in QC, during the positive LC-MS studies. The errors, calculated from the theoretical ion mass, are expressed in Da and in ppm.

Standard Compound	Theoretical Ion Mass (m/z) $[M+H]^+$	NAFLD_study1		NAFLD_study2	
		Exp.Error (Da)	Exp.Error (ppm)	Exp.Error (Da)	Exp.Error (ppm)
$[d_3]$Acetyl-L-carnitine	207.141865	-0.0049	-23.84	-0.0044	-21.12
Nialamide	299.15052	-0.0027	-9.03	-0.0047	-15.6
$[d_6]$Sulfadimethoxine	317.118513	-0.0013	-4.14	-0.001	-3.22
$[d_3]$Hexadecano-L-carnitine	403.361514	-0.0004	-0.87	0.0001	0.33
Reserpine	609.280657	0.0033	5.47	0.0032	5.28

Quality control samples were injected every ten acquisitions (See Appendix). For the evaluation of the characteristic parameters, only the quality controls after a conditioning period time of the column were considered. In the **Table 11** the experimental error (experimental m/z ion mass-theoretical ion mass) of the extracted ions of internal standards in plasma quality controls during the analysis of the fatty liver plasma samples in positive mode acquisition are described. The values of the masses shown derive from the average of the masses (each calculated from the averaged ion chromatogram) of the extracted standard compounds in the examined quality control samples; after column conditioning as already mentioned (from quality control 17 until quality control 28, see Apendix for the experimental injection set up). From these results it is possible to observe the typical TOF ionization dynamic mass accuracy range (**Fig. 40** and **Fig. 41**). In fact the experimental mass error displays larger error on smaller masses (207.14 m/z and299.15 m/z) and shows better mass accuracy in the larger mass range. In order to point out this trend, mass errors were calculated in ppm and Dalton.

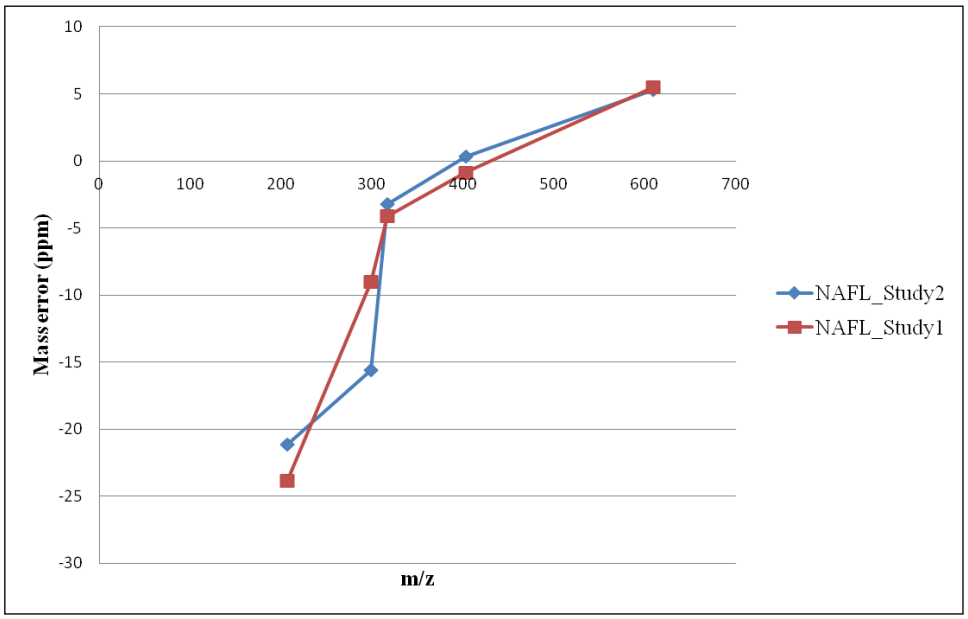


Figure 40: Dynamic Mass accuracy (ppm) of averaged standard compound measured via TOF-MS during NAFL experiment. The Study2 shows value slightly closer to the theoretical mass value.

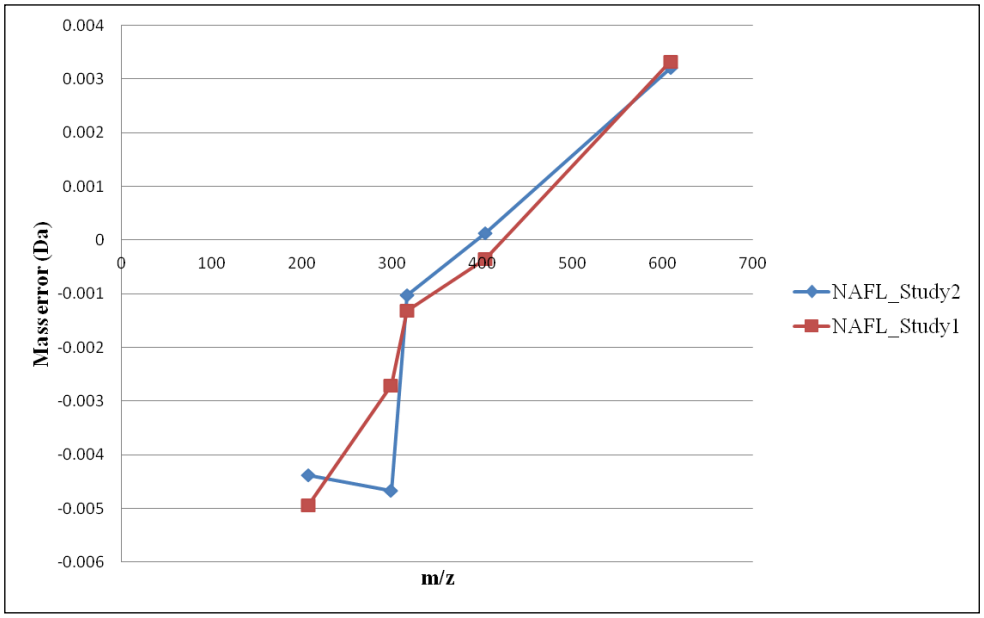


Figure 41: Dynamic Mass accuracy (Da) of averaged standard compound measured via TOF-MS during NAFL experiment. The Study2 shows value slightly closer to the theoretical mass value.

To evaluate the mass accuracy trend during the whole study, eight random plasma peaks were considered into the calculation of the mass drift (Appendix). The

averaged mass values of the extracted ion peaks lead to reasonable values of repeatability (RSD) expressed in percentage. The standard peaks and the randomly selected chromatographic peaks were tested as well in terms of chromatographic robustness, calculating the RT shift among the quality control samples (Appendix). Good values of RSD were reached (RSD < 1%). The calculation of the peak area RSD among the plasma and standards did not always display optimal repeatability. The reason for such behavior could lay in the variation of ion velocity and ion transfer, due to change in temperature. Variation of the room temperature as possible culprit of the variance in the ionization process is discussed in the chapter 4, where long time plasma measurements showed high variation in intensity signals along with a characteristic nocturnal and diurnal performance variation trend. Therefore the variation in intensity of the standards and randomly selected plasma peaks were studied. In **Fig. 42** it is possible to observe the variation of the mass intensity of the standard compounds during the study among quality control samples.

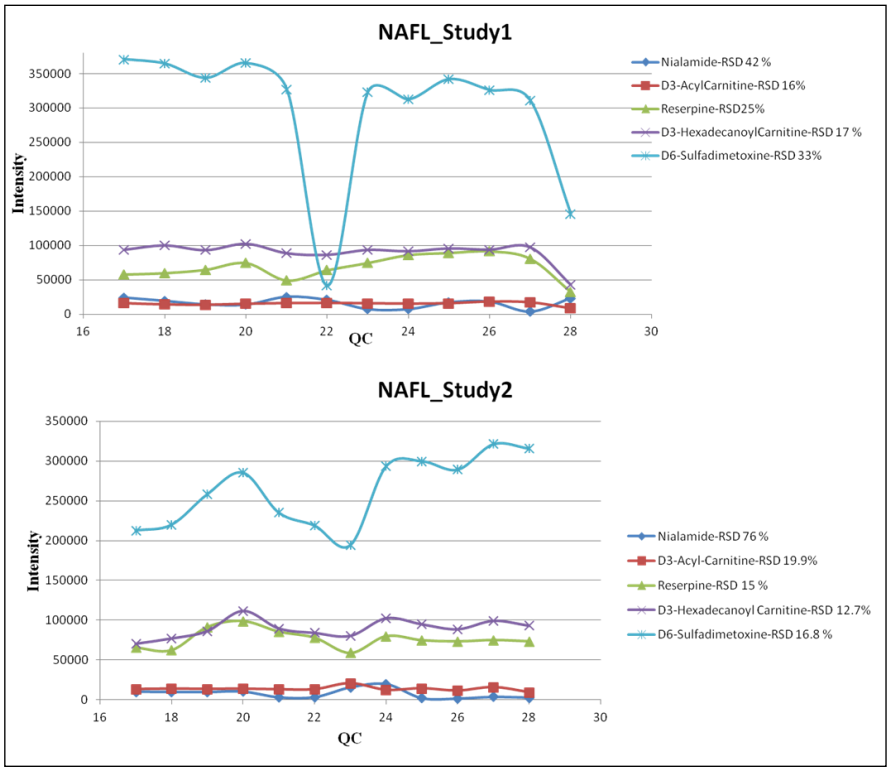


Figure 42: Variation of internal standards in QCs throughout the study.

Every standard is varies within the study according to an individual RSD value. Especially in study 2 it is possible to denote the same behaviour already observed in **Fig. 40** concerning the mass error: at small masses the variation of intensity is higher, showing better performance for the high masses range. This behavior could be explained by the characteristic performance of TOF-MS that is more accurate for high masses than smaller ones. High variation in the small mass intensity range can be an additional reason for the mass inaccuracy. The intensity variation among all the ions detected during the two fatty liver studies was followed.

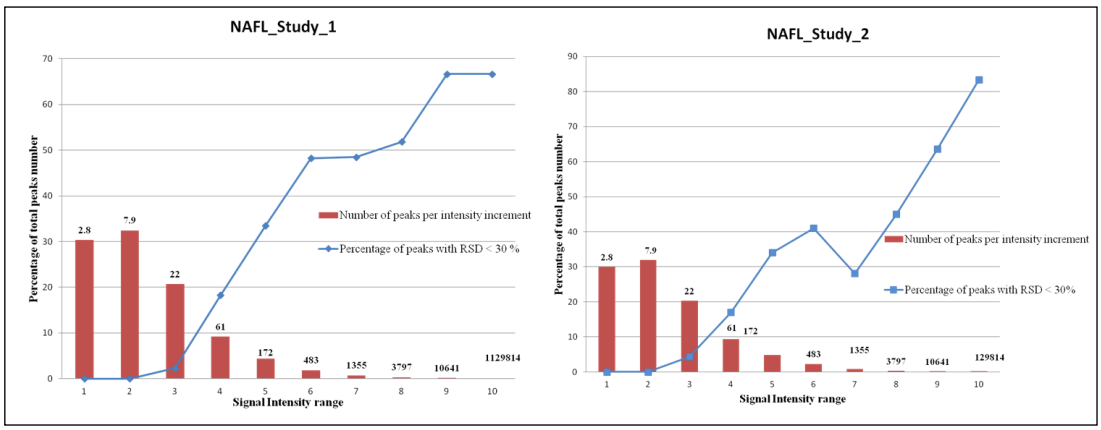


Figure 43: Partition of peaks with RSD < 30% (blue line) and intensity distribution (red bars).

Fig. 43 describes the impact of all ions detected during the two fatty liver studies in function of ten increments over the signal intensity range. The data were exported after elaboration with the software Markerlynx (Waters) and signals were filtered relative to the noise level. In the figure it is observable that the highest percentage of signals has the smallest intensities, and the fewest signals have high intensities. The RSD expressed in percentage and calculated over all the ion intensities (blue line) indicates less than 20 % of the peaks detected beneath 30 % RSD variability. The reason of such disappointing values can be explained by the fact that the majority of the ions detected lays in the region of the smallest intensity signals, the ones that are varying at most during long time measurements, especially after three experimental run days as it is the case of those two studies. Reference studies in literature display much lower abundances in the low intensity range. Whether the cases represented in literature are normality, whether the mass spectrometer used for this study is not sensitive

enough or whether the typical intensity profiles depend on the sample matrix can only be tested via cross validation studies with different mass spectrometers [209]. When the RSD of the ion intensities is being evaluated the causes of such variation have to be considered. The region of the smallest intensity is the most susceptible to the intensity variation, therefore in a non-targeted approach this would lead to discard small ion intensity molecules. Another factor, already mentioned above, is the room temperature change. In order to test the goodness of the experimental set up it is necessary to evaluate whether the quality control samples are clustering together relative to the other classes of interest. The data sets acquired were processed by the MarkerLynx software (MassLynx, Waters) using ApexTrack peak integration to detect chromatographic peaks (for details see Appendix). The masses were aligned and normalized with total area peak normalization to the sum of 10000. PCA models were evaluated for both fatty liver studies (**Fig. 44**).

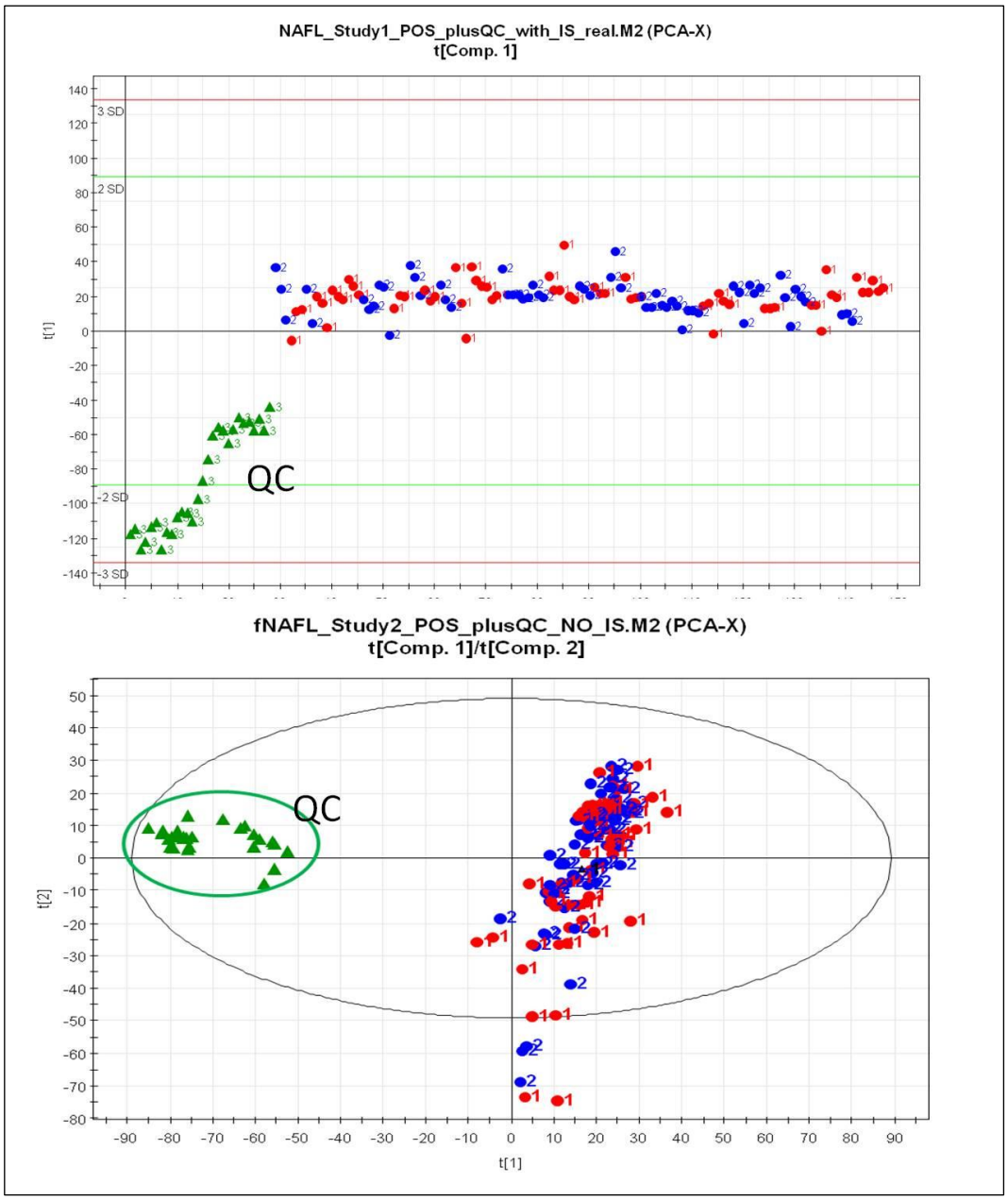


Figure 44: Differentiation of QC and samples. The PCA model on study 1 did not pass the cross validation and got only one component assigned. Study 2 was assigned with 4 components. Therefore the latter can be seen to be more robust.

In **Fig. 44** it is possible to evaluate the multivariate variability of the quality control samples (green dots) in a PCA calculated for each study. The variation among the quality control and the fatty liver samples (group1 and 2) is mainly expressed on the first component with regression cumulative value of R^2x (cum)=0.10 and prediction value of Q^2(cum)=0.07. The study2 shows a model that expresses the variability of the data with four components with a fitting value R^2x (cum)=0.15 and prediction value of Q^2(cum)=0.10 where the first component

separates the quality control and the fatty liver samples as well as in the first model. What we can remark is the fact that the quality control in the Study 2 are all included in the ellipse Hotelling's T^2 calculated with 0.05 at 95 % confidence, presenting a stronger homogeneity. Overall, multivariate reproducibility was acceptable. After the evaluation of the quality control samples, statistic evaluation was done on both fatty liver studies in order to discern significant masses. 20008 features were detected in Study 1; 13268 features were detected in Study 2. The data sets acquired were processed by MarkerLynx software (MassLynx, Waters) using ApexTrack peak integration to detect chromatographic peaks. The masses were aligned and normalized with total area peak normalization to the sum of 10000. Non-parametric Wilkoxon sum rank was applied defining only 5,5 % of masses significant. Further evaluations were done with multivariate statistics that is one of the most appropriate approaches for screening this complex human data set [50, 51]. In order to obtain first insights on the data structure apart from the PCA on the two classes of observation IR (class1) and IS(class2), PLS was performed on both studies. In each case the built models were not sufficiently able to separate the two groups of observations, insulin-sensitive and insulin resistant individuals, respectively. For this reason orthogonal signal correction prior to PLS modeling was applied in order to define features with discriminative power. The generated orthogonal-partial least squares (OPLS/O2PLS) model on the spiked fatty liver set(Study1) gave moderate model fit as well as moderate predictive variability with $R^2Y(cum)=0.79$ and $Q^2(cum)=0.674$. The (OPLS/O2PLS) model built on the non spiked fatty liver set (Study2) in turn gave very good model fit ($R^2Y(cum)=0.92$) and sufficient prodictiveness in cross-validation ($Q^2(cum)=0.74$). The difference in model quality can be observed in **Fig. 45**.

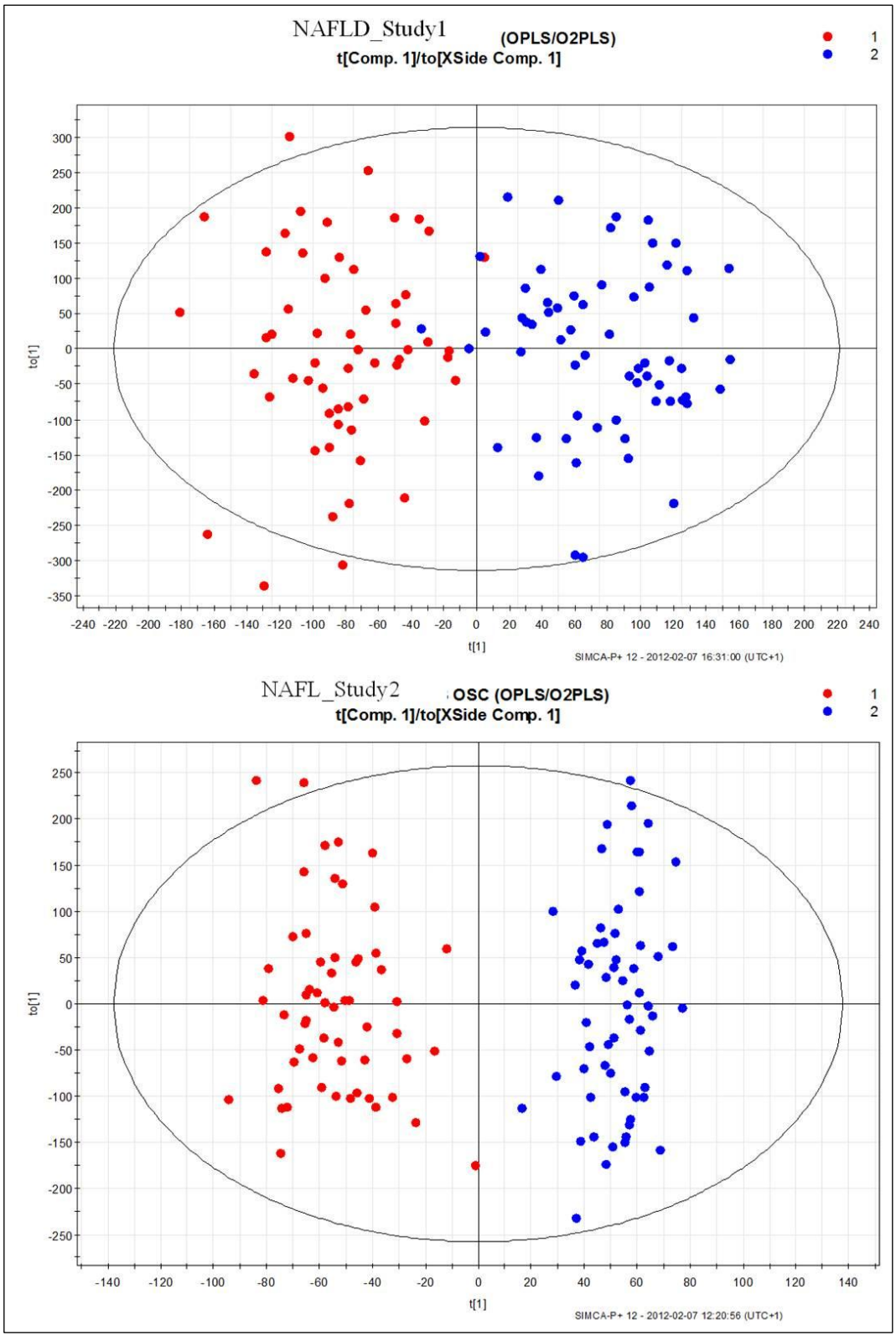


Figure 45: Comparison of OPLS/O2PLS scatter plot describing the separation among the two classes of interest (IR=1 and IS=2)of the population of fatty liver subjects in Study1 and Study2. NAFL Study 2 shows better model fit ($R^2Y(cum)=0.92$) and sufficient prouictiveness in cross-validation ($Q^2(cum)=0.74$).

As a consequence of apparently better data quality and separation among the fatty liver population when standard spiking was omitted, it was decided to prodeede in evaluating the discriminative compounds of Study 2.

Outliers were not considered for further analysis. Marker selection was performed as described for the ICR-FT/MS baseline. Therefore, cutting off 2.5% on each side of the plot led to. 338 metabolites, 172 classified as up-regulated in IS and 166 down-regulated in the IS. Further steps consisted of the mass assignment of the discriminative masses, for this issue the calibration of the masses is very important. The system was calibrated, prior to analysis, within an error window minor to 1 ppm. During the measurement, Lock mass solution (Leucin Enkephalin) was infused from the Lock spray in order to correct the mass drift during the acquisition. The lock mass correction is applied on the experimental masses following the calibration curve generated during the instrument calibration. The ppm error inferred from direct comparison of the experimental mass after lock mass correction and the theoretical ones could define an estimation of the limit of mass accuracy. The mass shift during the analysis (3 days) was leading to an averaged mass error shift ranging from 0.3 ppm for the highest masses until 20 ppm (**Table 11**). The reason, as already mentioned in the chapter preface, lays in the fact that only one mass correction was applied. Additionally, possible changes in room temperature between the day and the night could lead (as shown in Chapter 4) to stability variation of the mass spectrometer.

Due to the fact that post-mass calibration was not possible, the issue was to overcome the high tolerance mass window to apply in the inquiring chemical and biological databases (i.e ChemSpider, KEGG, HMDB). This passage could induce false mass annotation. The integration of ICR-FT/MS annotation (deriving from inquiries within 1 ppm error window) of the same data set of samples could offer a fast opportunity to annotate the common features between the two different platforms. In order to get a direction on the optimal mass difference window to apply for the alignment, random plasma peaks were selected and their mass variation among the study was followed (Appendix).

A graphic representation (**Fig. 46**) shows the ppm extremity in the Study 2 (blue) that ranges around (±15 ppm). The comparison with the Study1 highlights the same mass shift trend, even in different days of measurements.

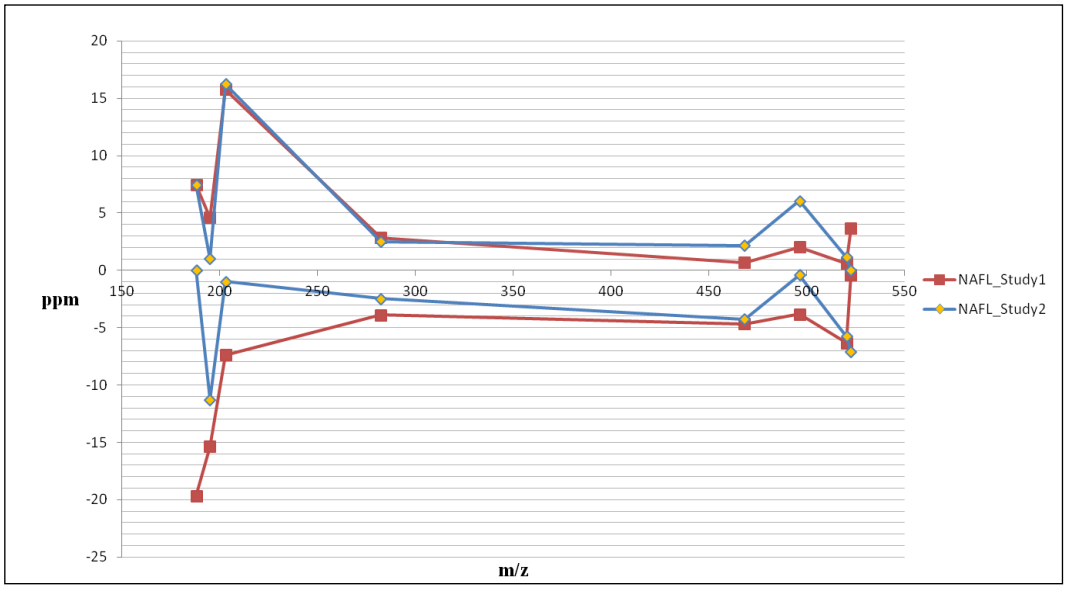


Figure 46: Upper and lower extremities of mass shift in the LC-studies NAFL_Study1 (red) and NAFL_Study2 (blue) as a function of mass. The high similarity in the maximal mass shift amplitude as well as the pronounced profile of the curves indicate instability regions of the TOF mass spectrometer which in the future may of use for calibration and LC-FT matching.

.

Different windows of different mass tolerance of the ICR-FT/MS data (from 1ppm until 50 ppm) were tested in order to verify whether the mass difference window mentioned above was a good compromise. Naturally 15 ppm are not sufficient for unambiguous formula annotation. Since the FWHM of m/z peaks of TOF-MS is roughly 20 times bigger than in ICR-FT/MS, one TOF-MS peak may potentially encompass 20 ICR-FT/MS peaks.

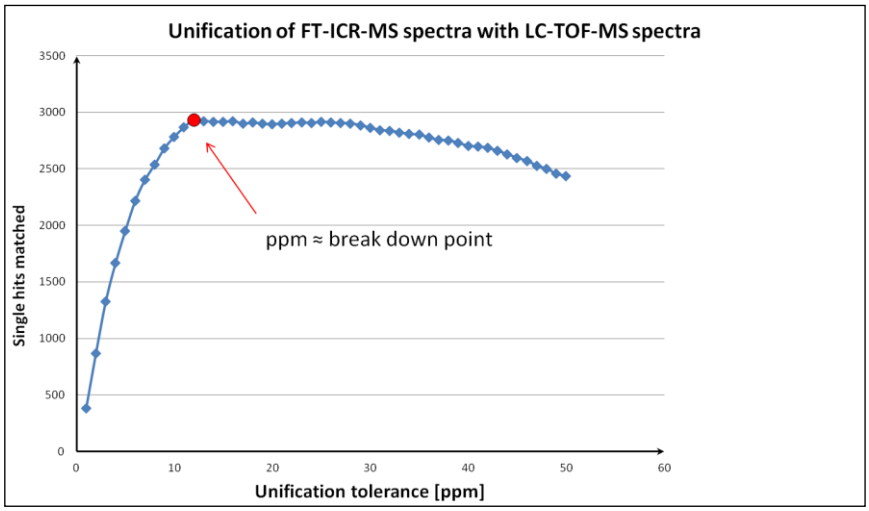


Figure 47: Alignment between LC-MS masses and FT-MS mass annotation. The red dot indicates the break down point where the singles hits matched slows down. This tolerance window (12 ppm) was selected for the annotations.

Therefore unification error point at which the number of single hits is superior to the number double hits needs to be found. Unifying LC-MS spectra and FT-MS spectra consequently is based on the search for a breakdown point in ppm error. The screen for this break down point was performed by consequent LC-FT-mass annotations at different unification windows and by recording the number of captured LC peaks and FT peaks for each window from 1 ppm to 50 ppm. The break down point or break down region was asses by plotting the number of singles hit matches between FT-MS and LC-MS over the unification windows giving the necessary limitation that only LC-MS masses which could be found in the respective FT spectra within a certain error range were to be regarded as valid marker candidates. The break down point (12ppm, mass tolerance), where it appears that the curve of the single hit matches started to decrease, was selected for the matching annotation, leading to the possible identification of 98 masses; 34 up regulated in respect to insulin sensitivity and 64 down regulated in respect to same criterion (Appendix) .

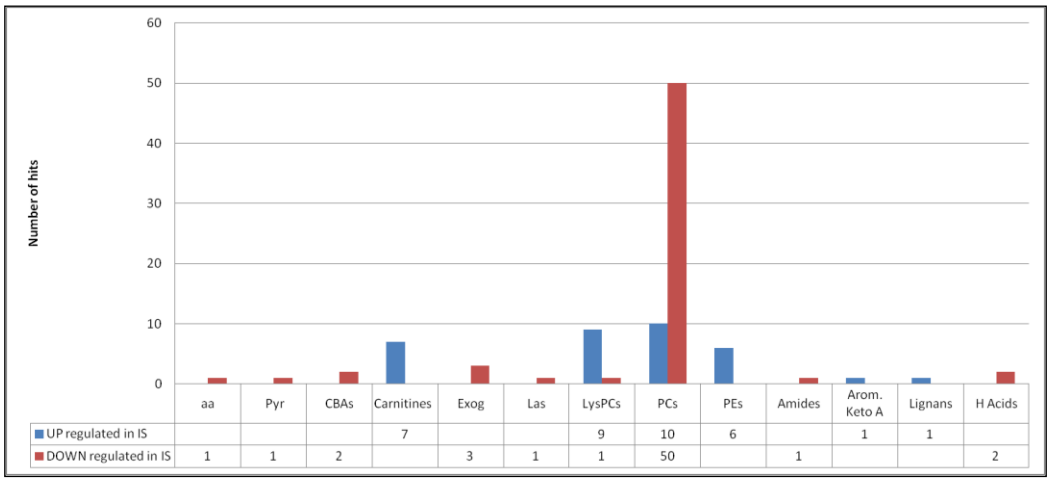


Figure 48: Possible mass annotation candidates in NAFL-LC-MS positive study.

In order to improve the matching algorithm, future developments will be done to refine the mass region of tolerance to each mass. The necessity of this further step is rising up due to the fact that the mass accuracy of TOF-MS is rather instable over time and m/z range relative to ICR-FT/MS and the resolution of ICR-FT/MS spectra at high masses is decreasing. Therefore the matching proposed above could lead to some false positive annotation. More accurately we might still go on increasing the error tolerance gaining new single hits while other masses may gain a second annotation in this process. If for each mass we will have an optimized break down point, the accuracy of the mass annotation will increase whereas the number of false positives may decrease making the annotation of the isomers easier.

The integration of the FOLLOW -UP study

Plasma samples from the same fatty liver subjects after nine months of life style intervention (diet and exercise) were analyzed via ICR-FT/MS in positive and negative ionization modes. Of the 21698 ions detected, the masses with a frequency below 10% were not considered for further statistical evaluation. Non-parametric statistical tests such as the Wilcoxon sum rank test was applied leading to 3% of the significance. Non-supervised parametric tests such as PCA were applied. The model has four components and shows goodness of fit as $R^2X(cum)$ = 0.72 and prediction coefficient as $Q^2X(cum)$=0.22 The score scatter plot of the PCA is shown in **Fig. 49** and the first component absorbed 32 % of the variance the fourth component absorbed 9%.

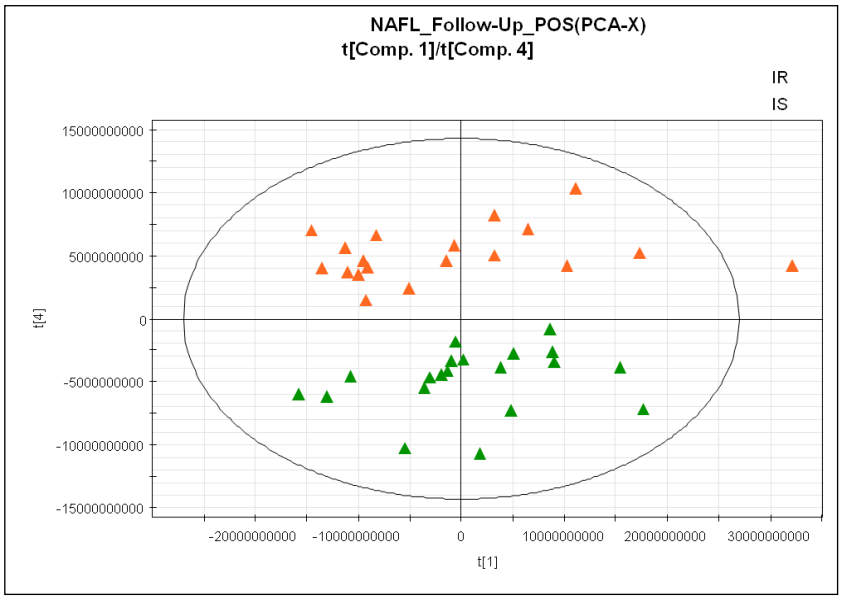


Figure 49: Scatter plot of the PCA model.The first component absorbs 32 % of the variance the third component the 9%.(orange= IS and green=IR).

Further multivariate statistics were conducted by the generation of a OPLS/O2PLS model with values of fitting ($R^2X(cum)$ = 0.74) and prediction coefficient ($Q^2X(cum)$=0.98) (**Fig. 50**).

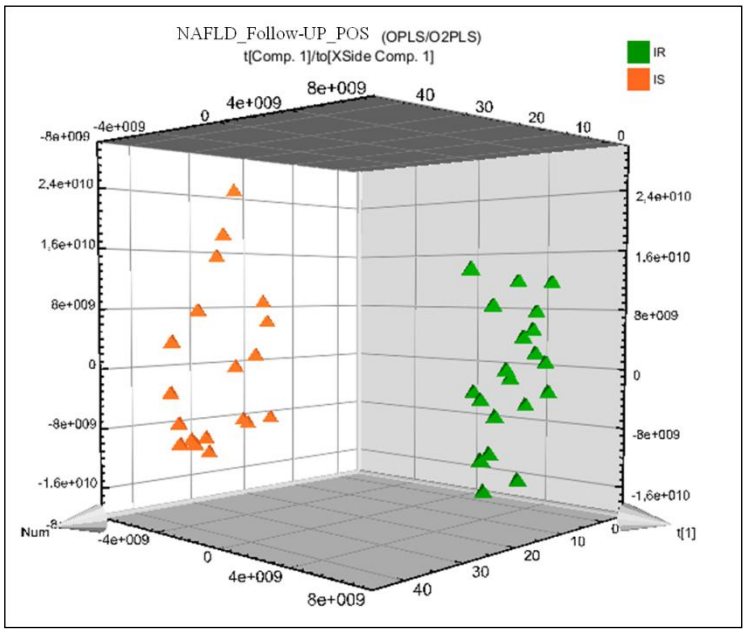


Figure 50: OPLS/O2PLS score scatter plot in which is visualized the two different group (orange= IS and green=IR), with Q^2(cum)=0.98 and R^2Y(cum)=0.74.

A cut off at 2.5% on each side of the S-plot was performed as already described in the case of the Baseline study. 518 significant masses were considered, with 265 down-regulated features in the IS group, and 253 up-regulated features in regard to the same class (Appendix). In **Fig. 51** the principal annotated compounds discriminative between the two classes of study are described.

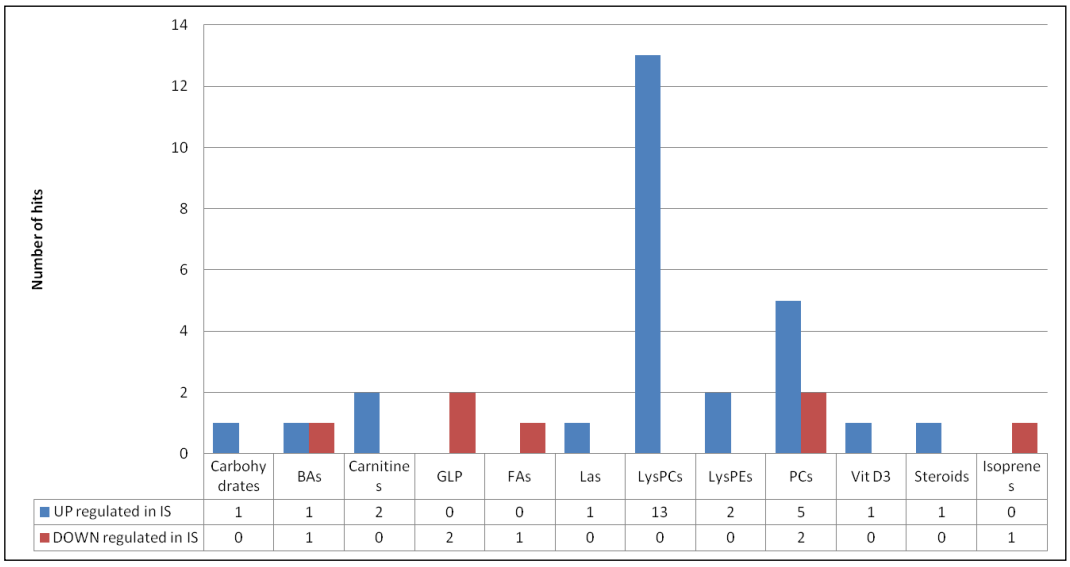


Figure 51: Classes of important compounds among UP-regulated insuin sensitive group and down reulated. As in the baseline study it is possible to denote the high presence of Lysophisphocholine.

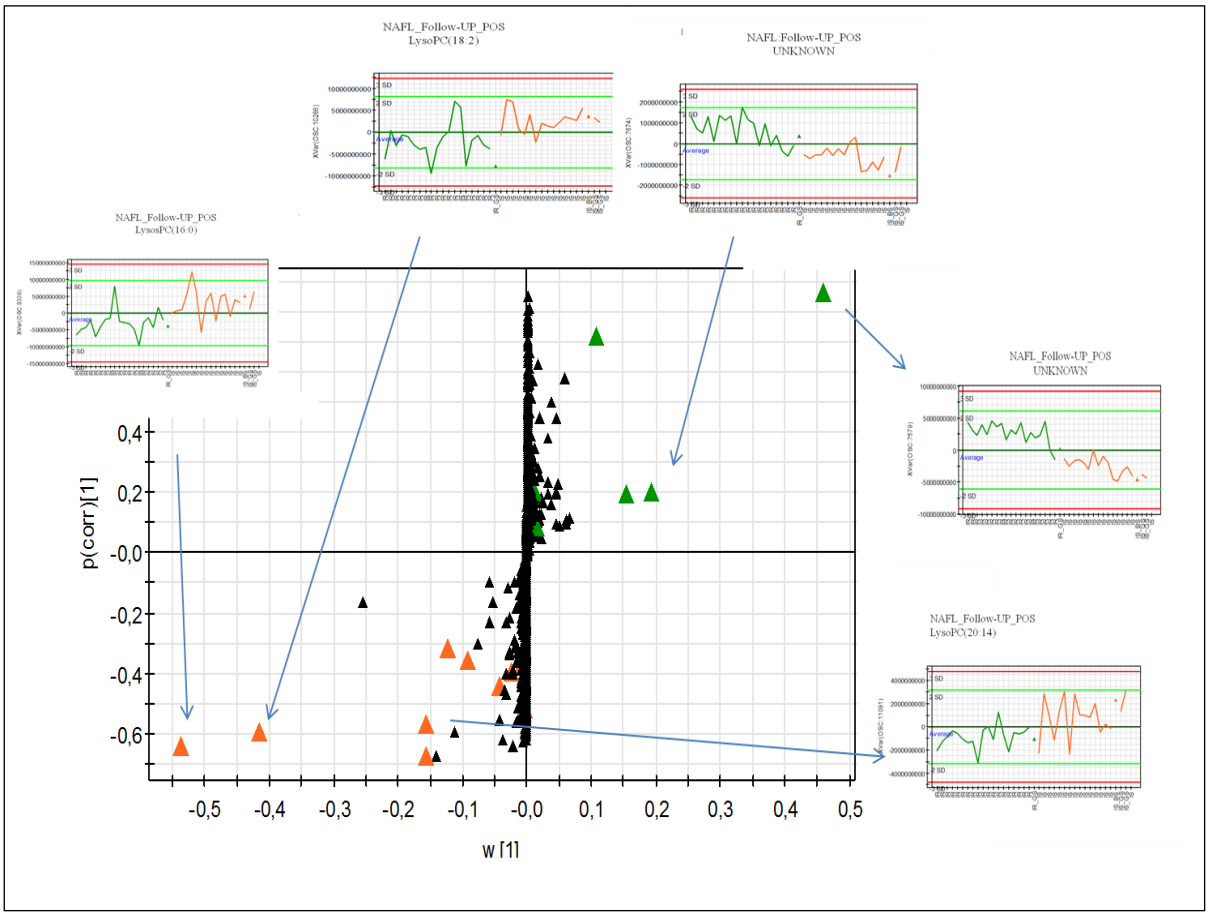


Figure 52: Characteristic S-lot showing the most significant masses. It is possible to denote the dominace of the phosphocholine class.

The most important markers in IR class (green) and IS class (orange) are shown in **Fig. 52**. The characteristic S-plot derived from the OPLS model shows the most extreme value, which are the most discriminative among the two classes.

Different Lysophosphocholines show their strong behaviour in the up- regulation of the insulin sensitive class. Finally Lyso-PC (16:0) was identified in plasma via ICR-FT/MS/MS with fragmentation in the linear ion trap. Tryptophan could only partially be identified. The fragmentation of a LysoPC(16:0) standard at 10eV resulted primarily in a loss of H2O from the glycerol moiety as well as in a loss of the C16-acyl group. Secondarily the rest of the glycerol moiety was fragmented leaving the phosphocholine [C5H14NO4P+H^+]$^+$ moiety behind. This moiety was observed whenever a phosphocholine was fragmented.

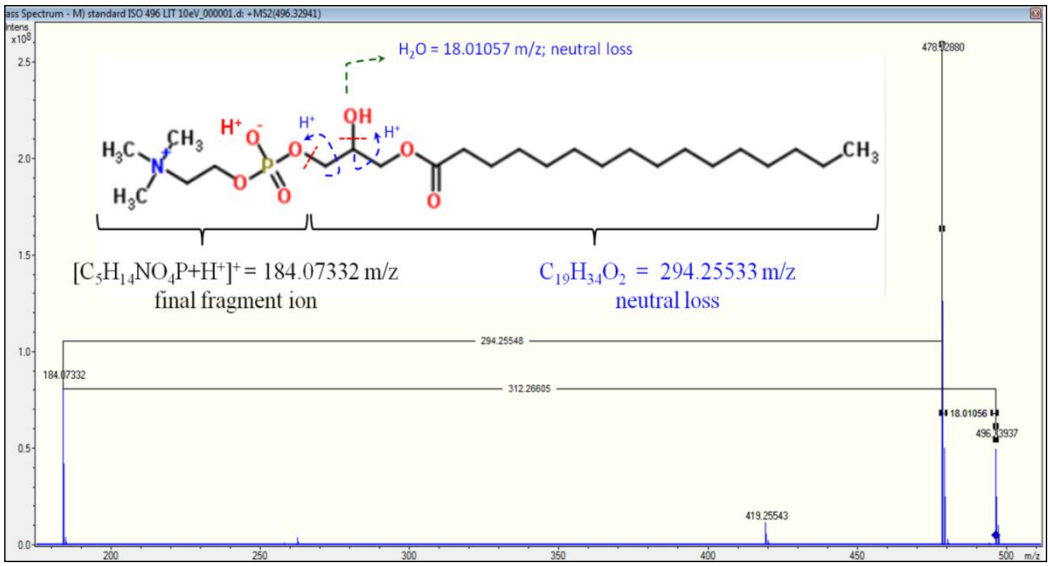


Figure 53: Typical Lyso-PC pattern yielded at 10 eV collision energy in the LIT partition of the FT-ICR-MS mass spectrometer; standard.

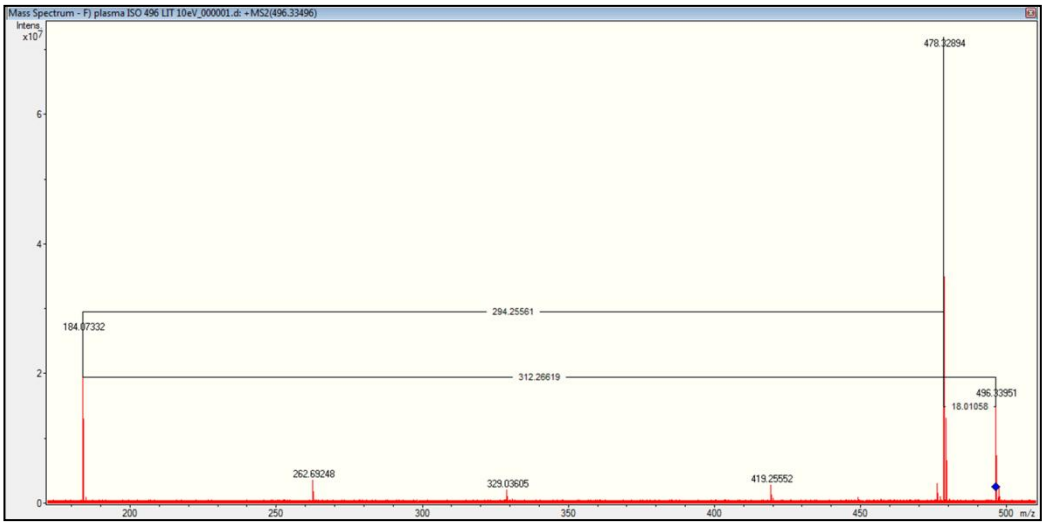


Figure 54: Fragment pattern for the same Lyso-PC (16:0) in plasma

Having attained the LysoPC(16:0) standard fragmentation pattern the same m/z was fragmented in a sample from the TULIP-cohort. The same fragmentation pattern could be confirmed. Therefore the same mass was identified to be a LysoPC(16:0).

In case of Tryptophan only the loss of NH3 relative to the parental Peak could be confirmed. Typical fragmentation patterns as documented in HMDB could not be confirmed, neither in the standard compound, nor in the sample.

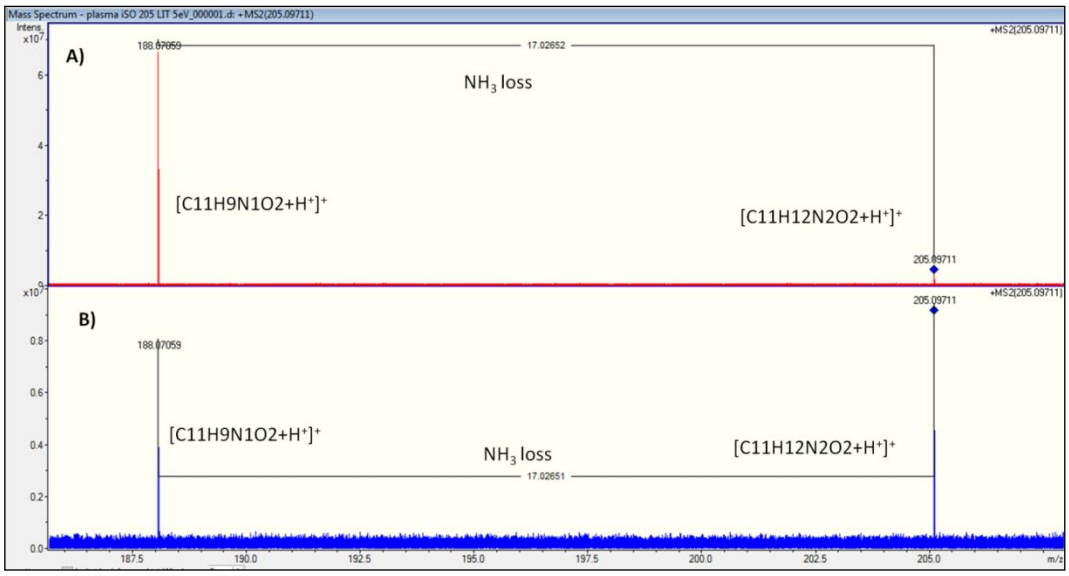


Figure 55: Ammonia loss of Tryptophan as a standard (red) and in Plasma (blue).

5.4.3 Discussion of the biological results

In this section the most prominent classes among the found discriminative features are discussed. The data of the different analytical platforms are discussed uniformly in order to gain a homogenous overview of the results. Present significant correlations with clinical parameters collected on the sampled individuals are underlined in order to emphasize the biochemical context of these findings.

Aminoacids

The aminoacids found to be discriminative are PhenylAlanine, Tryptophan,Tyrosine; additionally Indoleacrylic Acid that is linked with tryptophan metabolism and trihomomethionine connected with the biosynthesis of glucosinolate were found. All these compounds were only present as down-regulated in respect to the insulin sensitive group. The role of amino acids as promoter of insulin resistance was already reported in a long period study where branched chain amino acids (Leu , IsoLeu and Val) and aromatic amino acids (Tyr, PheAla and Trp) were found significant in diabetes prediction [210]. Moreover amino acid supplementation related studies in humans [211] and in animals [212] hypothesized that the circulation of amino acids can favour insulin resistance via breakdown of insulin signalling in the skeletal muscle. Alternatively the hypothesis that amino acids, especially branched chain amino acids, are involved in the debilitation of the pancreatic α-cells stimulating insulin secretion is considered. The question that rises up is the explanation of the reason why aromatic amino acids can be important in the development of insulin resistance (e.g. hippuric acid). Focusing on the case of tryptophan an elucidation can be attempted through its already known involvement with inflammation and from the other side through the gut-microbiota related world. In animal experiments on pigs it was demonstrated that dietary tryptophan surplus was reducing plasma stress hormon concentrations (cortisol, noradrenaine and epinephrine) after acute stress [213]. Additional experiments on pigs have shown that long term tryptophane dietary surplus induced insulin resistance. The hypothesis behind is that during an acute phase (3hours) the increase of stress hormones (noradrenaline

and adrenaline) induces insulin resistance whereas a chronic exposure (several days) to plasma catecholamine drives into insulin sensitive elevation [214]; therefore a chronic depicted concentration of stress hormones brings insulin sensitivity reduction. A supplementary interpretation of the phenomenon counted on a possible reduction in physical activity, as side effect of the Trp surplus diet [213], since it can reduce the concentration level of noradrenaline [215]. In the special scenario of this study where the concentration of tryptophan is decreased for the insulin sensitive class, as well as Indoleacrylic Acid (its metabolite derived form transamination reaction), a possible elucidation can be connected to the studies found in literature [214]. Additional findings in our work group point to Tryptophan as one of the key players in inflammation processes, which again connects this amino acid to insulin resistance.

Carboxylic acids

The citric acid cycle is one of the protagonists in metabolism. Its main purpose is to contribute directly to ATP production as well as to deliver metabolites for other pathways. For these purposes citric acid, the name-giving compound to the citrate cycle, is continuously shuffled between the mitochondrial matrix and the cytosol. Citric acid itself, which was down-regulated in Insulin sensitive patients, is produced via the glycolytic end-product phosphoenol-phosphate. Therefore citric acid concentrations in liver are as well tightly bound to gluconeogenesis and lactate concentrations. Several instances in literature indicate a strong but not necessarily causal relation between this cycle and diabetes Type2. Already in 1999 Large and Beylot showed, that diabetic rat livers had a higher pyruvate and lactate uptake and higher fluxes through gluconeogenesis and the citrate cycle [216]. In fact the commencing imbalance between effects of glucagon and insulin in insulin resistant liver leads to higher rates of gluconeogenesis [222]. Spégel et al. also found that insulin excretion in β-cells resistant to glucose triggering goes in parallel with limited citric acid cycle response [217]. In all described literature sources it becomes rather clear, that citrate is not a regulative entity but rather an indicator for insulin resistance involved processes. In the first example it is the

higher rate of glyconeogenesis and in the second a higher rate of glycolysis which determine the direction of citrate regulation.

Another compound found to be down-regulated in insulin-sensitive individuals is methylcitric acid. This compound is being produced once propionyl-CoA instead of acetyl-CoA reacts with oxaloacetate. Propionyl-CoA may be the final product of β-oxidation of odd chain fatty acids. Since these are predominantly produced by microbiota, one may conclude that the increase of methylcitric acid reflects the status of the intenstinal flora. Robert H. Allen et al. mentioned an increased production of propionyl-CoA due to deficiencies in the methionine-synthase and cobalamin metabolism. They have pointed out, that such deficiencies may lead to neuropsychiatric abnormalities [218]. In the recent years a connection between Diabetes *mellitus* and depression is discussed [219].

Bile acids

T. Claudel et al. described how bile acids may bind the Farnesoid X receptor (FXR) which would increase β-oxidation, decrease liponeogenesis as well as blood glucose levels. It furthermore inhibits bile acid synthesis. In our studies deoxycholic acid glycine coniugate was down-regulated in the insulin sensitive group whereas 12-ketodeoxycholic acid was found to be up-regulated. Since the deoxycholic acid glycine coniugate is a form created for intestinal reabsorption to be redirected to the liver, it is very well possible, that the FXR system works correctly in insulin sensitive individuals [220]. 12-ketodeoxycholic acid is a secondary bile acid produced in the intestine. C. Steiner et al. have shown that 7a-hydroxy-4-cholesten-3-one is significantly more abundant in type 2 diabetic patients than in controls. No specific role could be appointed in the given context. Further investigations may probably increase our knowledge in the involvement of FXR into insulin sensitivity among NAFLD patients [221].

Carnitines

Carnitine is necessary for the transport of fatty acids from the cytosol to the mitochondrial matrix, when fatty acids are needed for β-oxidation, i.e. when the body requires energy. They are as well needed to export the end-products of peroxysomal β-oxidation, medium-chain fatty acids, from the peroxysomes to the mitochondria. In normal, healthy persons peroxysomal β-oxidation makes up only 10% of whole β-oxidative activity [222]. Why is a second fatty acid oxidation mechanism next to mitochondrial oxidation required? Peroxysomal β-oxidation is needed for very long chain, branched chain and especially unsaturated fatty acids [223]. The initial oxidation step is performed via molecular O_2, i.e. no group specificity is necessary. This process usually ends at medium chain length of the fatty acid residue to be oxidized. In the present study one medium chain carnitine (octenoyl carnitine) was down-regulated but two poly-unsaturated long-chain carnitines up-regulated in insulin sensitive fatty liver patients. Octanoyl transferase, responsible for the formation of octanoyl carnitine is a peroxysomal enzyme. This pattern points towards regular mitochondrial β-oxidation in insulin sensitive and impaired mitochondrial β-oxidation in insulin resistant individuals.

Phospholipids

The compound class which exhibited the broadest impact in our studies are the Phospholipids. Phospholipids are an important class of membrane components; lipids that contain phosphorous and where both fatty acid and phosphoric acid are esterified to an alcohol. They are distinguished in glycerophospholipids and sphingolipids according as the alcohol is glycerol or sphingosine respectively (F**ig 56**).

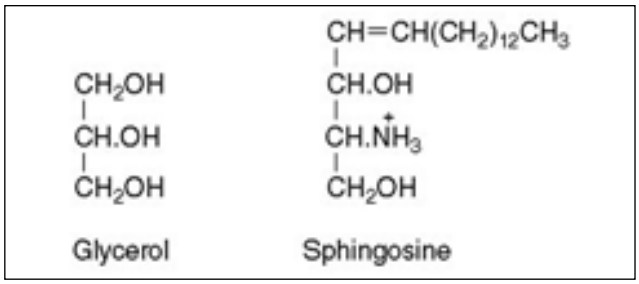


Figure 56: Chemical structure of glycerol and sphingosine.
Reprinted from *Functional Biochemistry in Health and Disease*,
Wiley-Blackwell, ISBN: 978-0-471-98820-5, Newsholme E., Leech A.,
Synthesis of Fatty Acids, Triacylglycerol, Phospholipids and Fatty Messengers: The Roles of Polyunsaturated Fatty Acids, 239, Copyright (2010), with permission from Elsevier.

In phosphoglycerides two long fatty acids are esterified with their carboxylic groups to the hydroxyl groups at position 1- and 2- of glycerol-3-phosphate; a base (ethanolamine, choline, insositol, serine) is attached to the phosphate of the latter one giving a positive charge contrasting the negative charge derived from the phosphate group (**Fig. 57**).

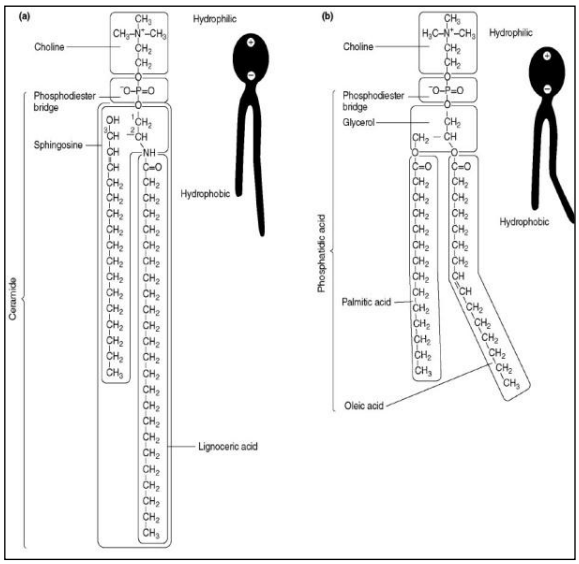


Figure 57: General lipid structures of Phospholipids.
Reprinted from *Functional Biochemistry in Health and Disease*, Wiley-Blackwell, ISBN: 978-0-471-98820-5, Newsholme E., Leech A., **Synthesis of Fatty Acids, Triacylglycerol, Phospholipids and Fatty Messengers: The Roles of Polyunsaturated Fatty Acids,** 240, Copyright (2010), with permission from Elsevier.

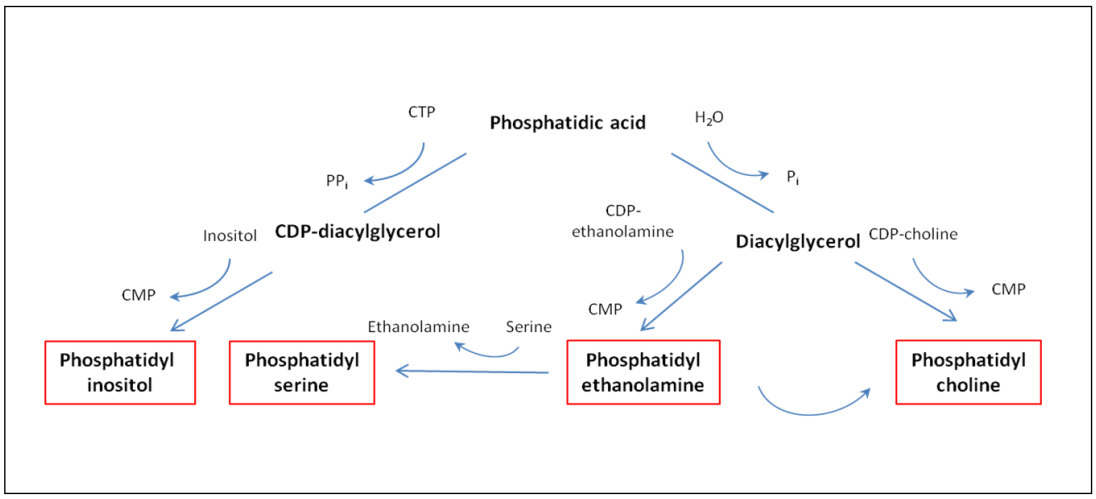


Figure 58: Glycerophospholipid synthesis.
Reprinted from *Functional Biochemistry in Health and Disease*, Wiley-Blackwell, ISBN: 978-0-471-98820-5, Newsholme E., Leech A., **Synthesis of Fatty Acids, Triacylglycerol, Phospholipids and Fatty Messengers: The Roles of Polyunsaturated Fatty Acids**, 241, Copyright (2010), adapted with permission from Elsevier.

The synthesis of glycerophospholipids (**Fig. 58**) generally starts with a formation of phosphatidate starting with a CoA-dependent acylation of glycerol-3-phosphate on the sn-2 position. The synthesis of phosphatidylserine, phosphatidylethanolamine and phosphatidylcholine goes on with a second

acylation and de-phosphorylation to give diacylglycerol. The fatty acid composition of the later glycerophospholipids is not yet decided in this process. In order to form the final glycerophospholipid the respective head groups which are bound as cytidine diphosphocholine or cytidine diphosphoethanolamine are being added to the remaining free hydroxyl group of the diacylglycerol. These reactions result in phosphatidylethanolamine (PE) and phosphatidylcholine (PC). The more versatile of these head groups is the ethanolamine group. It can be replaced by serine to give phosphatidylserine (PS) and it can be methylated to again give PC. Hence a part of PC synthesis is tetrahydrofolate and cobalamine dependent. As mentioned above, disorders regarding these two compounds may cause neuropsychiatric complications. All formed glycerophospholipids build the foundation of the cell membrane. Their hydrophilic head groups are important for cell-cell signalling. The phospholipid classes which differentiated between insulin-sensitive and insulin-resistant individuals across all experimental studies are Lyso-phospholipids. These compounds are more hydrophilic than the phospholipids. Many functions have been ascribed to the respective classes. Lyso-phospholipids may have pro- and anti-inflammatory effects. Their pro-inflammatory effects in diabetes type2 development are, as described later, by far more pronounced in literature than their anti-inflammatory action. These lipids are as well intermediates when oxidative damage in acyl-side chains of membrane lipids are repaired. Lysophospholipids are products of de-acylation via phospholipase A.

Phospholipase A consists of a group of enzymes responsible for the cleavage of one acyl residue from the glycerol backbone of a phospholipid, discerned in PLA1 ans PLA2 according to the stereochemistry of the reaction. Consequently a lyso-phospholipid is generated and free fatty acid is released [224] both with important messenger functions [225]. The concentration of lysophospholipids is documented as increased under inflammation disease status like atherosclerosis [226] lung infection [227] or rheumatoid arthritis [228]. The predominant mechanism involved in their synthesis is from the activation of PLA2 which is present for example in inflammation mediators like neutrophilic granulocytes that are as well releasing ROS species[229] responsible for the generation of lysophospholipids. The involved ROS species are $HO^{\cdot}$ Radicals, HOCl produced from the enzyme

myeloperoxidase (MPO), and $NO^{\cdot}$, NO_2 and O_2. The formation of LPL from PC in vivo is generally derived from the activity of LPA_2 in presence of HOCl [230]. In parallel it is known, that lysophospholipids occur intermittently after oxidative damage of cell membranes. Damaged acyl groups are being removed, under production of lysophospholipids, just to be re-acylated and re-integrated into the cell membrane.

The effects of the LPCs on the immune system are discussed in literature [231] considering especially the function of LPC and reporting their exotic characteristic having pro-inflammatory and anti-inflammatory effects. The anti-inflammatory effects are being documented from their inhibition action against lipopolysaccaride (LPS) inhibiting tissue factor (TF) activity for human monocytes [232] and TNF-α and IL-β from neutrophils [233]. Their inhibitory activity againt ROS generated in human neutrophils was demonstrated as well [234]. It is being argumented that the in vivo release of the unsaturated fatty acids like arachidonic acid could be a possible reason for the strong immuno-effect. LPC are being reported as mediator of insulin resistance [235] via C-Jun-N-Terminal kinase (JNK) activation and insulin receptor substrate (IRS-1) phosphorylation, which inhibits insulin signalling. This mechanism is specifically observed in the case of LPCs derived from saturated fatty acid like palmitic acid, in the case of unsaturated fatty acid like oleic acid weak activation of JNK was observed prefiguring different activation mechanisms [235]. In contrast to these findings among a slightly dominating group of unsaturated long chain Lyso-PCs specifically Lyso-PC(16:0), Lyso-PC(18:0) and Lyso-PC(20:0) were found to be up-regulated in insulin sensitive individuals (as it was confirmed in negative mode). Lyso-PEs and fully substituted phospholipids were greatly under-pronounced in this group when compared to Lyso-PCs. Parallel to the metabolomics data acquired, clinical data was collected for the TULIP-cohort individuals. The acquired parameters closely followed the example Kantartzis et al. [236]. The comparison of the metabolite profiles of statistically validated marker candidates with the clinical data suggested positive correlations of especially very long chain unsaturated Lyso-PCs with $ISI_{Matsuda}$. All Lyso-PCs showed significant anti-correlation to plasma CRP levels at baseline, irrespective of their acyl chain length. The discussed amino acids phenylalanine and tyrosine

as well as the carboxylic acid methylcitrate exhibited significant positive correlation to CRP.

The LC-MS data on the same samples only confirmed the ICR-FT/MS data in terms of Lsyo-PC (18:2) and Lyso-PC (20:2). Furthermore PCs were pronounced in LC-MS. Nonetheless, two apparent PC peaks at retention times 21.17 and 21.28 minutes each having the mass of 810.6014 m/z were found to be entirely contraindicative and may in the future be worth further investigation. As detected via LC-MS methylcitrate may additionally be anti-correlated to $ISI_{Matsuda}$. This, together with the predominance of Lyso-PC over Lyso-PE may indicate C_1-metabolism impairment in insulin resistance. Since citrate and methylcitrate were found down-regulated in IS via ICR-FT/MS and LC-MS analysis, respectively, targeted studies on the ratio of both compounds may give some explanations concerning the involvement of the C_1-metabolism. The LC study as well pointed at N6,N6,N6-trimethyllysin to be down-regulated in IS. Since this compound is a precursor to carnitine, it may be possible that some connection to the as well lower level of octenoylcarntitine and thus diminished peroxysomal fatty acid oxidation lays at hand.

In the follow-up study Lyso-PCs were still strongly dominant in insulin sensitive individuals. Correlation with clinical data pronounced the role of diacyl-phosphatidylcholines. Positive correlation with $ISI_{Matsuda}$ as well as almost entirely negative correlation with liver fat content was observed. Nonetheless, PCs were evenly distributed between insulin-sensitive and insulin-resistant individuals. The major part of correlation between metabolite levels and clinical parameters in the follow up is contributed by the insulin resistant individuals. This was not the case in the baseline study. One may therefore assume that the applied life-style intervention took effect.

The main discriminant between the two groups at baseline and at follow up overall remained the partition of Lyso-PCs.

Involvement of exogenous metabolites

Some compounds not originating from human anabolism like phytochemical compounds, alkaloids, phthalates were found down regulated in the insulin

sensitive group. A prominent position within this class is covered by caffeine and paraxanthin. The latter is a compound derived from the demethylation of caffeine after intake. Additionally, as possible annotation, other compounds belonging to the xanthine family (theobromine and theophylline) were found down-regulated in the insulin sensitive class. In the last decades many publications described the possible role of caffeine and paraxanthine in diabetes pathogenesis. The illation concerning the caffeine effect on insulin sensitivity is controversially debated between two different factions: the acute and the epidemiological (chronic) studies.

Acute Studies

Human studies have verified that oral caffeine intake 250 mg t.i.d. (equivalent of 5/7 cups of coffee per day) can produce plasma caffeine concentration of 40 µmol/l and paraxanthine, its principal metabolite, of 20 µmol/l. Caffeine and paraxanthine are both antagonist of the adenosine receptor [237]. These receptors are present in different kinds of tissue like fat, skeletal muscle, liver, brain and heart. A Study on healthy volunteers has found [238] that moderate caffeine intake (intravenous, plasma levels of 30 µmol/l) can lead to decreased insulin sensitivity of 15%. This phenomenon is accompanied by increased levels of catecholamines, FFAs and systolic and diastolic blood pressure [239]. It was already shown that infusion of the catecholamine epinephrine leads to a decrease in insulin sensitivity of a rate of 50 % [240]. In absence of significant levels of catecholamines, caffeine does not play a role on the level of glucose or insulin [241]. The increased level of FFAs can be the cause of the decreased insulin sensitivity and can fall off the glucose uptake in liver and the periphery [242]. The release of FFAs can be a consequence of the epinephrine mediated lipolysis or due to the inhibition of the adenosine-induced suppression of lipolysis [243]. Also other groups described decreased insulin sensitivity and higher levels of FFAs after caffeine consumption in the acute time scale, they concluded that insulin resistance is mediated by epinephrine antagonism of insulin action.

EpidemiologicalStudies (chronic)

Inverse associations between coffee consumption and insulin resistance were testified through different cohort studies performed in different countries between north Europe, south Europe, USA and Japan [244]. The data were always tried to be adjusted from possible confounders like sex, age, obesity, and other life style factors like smoking, physical activity or diet. It was found out that the kind of coffee preparation can also be incidental in the development of diabetes and CVD risks. A clear example is being provided from the different effects between filtered coffee and boiled coffee [245], since the latter would lead to cafestol production and increase of LDL levels. Another difference screened was the use of decaffeinated coffee; in Europe its use was too low [244] to bring tangible differences but in the USA high consumption of decaffeinated coffee was associated with low risk of insulin resistance [246]. In summary all the cohort studies are indicating the coffee consumers having a lower risk of diabetes2 until 28% (who drank 4-6 cups) and 35% (more than 6 cups), respectively. Different observations conclude that low risk of Type 2 Diabetes can derive from other beneficial components of coffee [246] like phenolchlorogenic acid or quinides. Some studies have shown that acute intake of caffeine can induce insulin resistance due to lower carbohydrates storage [239]. Anyway it has to be underlined that physiological effects of coffee can be different than the ones derived from caffeine consumption; intake of caffeine leads to elevation of catecholamines more evidently than for the same ammonut of caffeine from coffee [238]. A possible hypothesis involves the counteraction of quinides that increase the extracellular adenosine concentration [247]. Moreover it has to be considered that the acute effect of caffeine on insulin sensitivity can diminish as a function of time, due to the acquired tolerance.

The acute studies against the epidemiological studies are contributing to discordant conclusions but a certain hypothesis may be attempted. In animal studies it was already demonstrated that high levels of epinephrine lead to insulin resistance under short exposure, whilst it drove to insulin sensitivity amelioration

after long exposure; these observations could as well constitute a support of a possible explanation of caffeine action on the insulin.

This can point out a connection with tryptophan action. As discussed above, it was demonstrated that the high concentration of this amino acid can lead to low catecholamine concentrations and therefore insulin resistance under chronic exposure; the opposite circumstance is being met by caffeine where high concentrations are leading to insulin resistance in the acute exposure; therefore we could assume that also in this case long exposure will lead to amelioration of the insulin sensitivity. This assumption was also found from the epidemiological studies. Both cases are an example of how epinephrine plays an important role in the antagonism of adenosine receptor agonists and in the regulation of the insulin sensitivity where the exposure as function of time seems to be a key factor. In the up-regulated group only one phytochemical compound was prominent.

5.5 The puzzle composition: the description of the possible biochemical pathways involved

The same cohort samples analyzed in this thesis were analyzed by Prof. Lehmann et al. [257] with a targeted approach combining liquid chromatography, gas chromatography and direct infusion mass spectrometry. The study was based on the commercial Biocrates target kit, enabling the analysis of 265 metabolites. The targeted approach clearly pronounced the importance of LysoPCs in the insulin sensitive group. Especially the role of LysoPC (16:0) was confirmed. Furthermore the branched chain amino acids, one short chain acyl-carnitine and two long chain acyl-carnitines were shown to be up-regulated. Confirming the targeted study in this work especially Lyso-PCs were found to be up-regulated in insulin sensitive individuals. Lyso-PEs and fully substituted phospholipids were greatly under-pronounced, compared to the lysophospholipds, in this group at baseline. Parallel to the metabolomics data acquired a multitude of clinical data was collected for the TULIP-cohort individuals. The comparison of the metabolite profiles of statistically validated marker candidates with the clinical data suggested positive correlations of especially very long chain Lyso-PCs with $ISI_{Matsuda}$. These compounds showed significant anti-correlation to plasma CRP levels at baseline, suggesting an anit inflammatory role. The discussed amino acids phenylalanine and tyrosine as well as the carboxylic acid methylcitrate exhibited significant positive correlation to CRP, pointing out their pro-inflammatory nature.

In the follow-up study the class of Lyso-PCs was still strongly dominant in insulin sensitive individuals, suggesting a genetic prevalence, as confirmed in the targeted approach. However, correlation with clinical data pronounced the role of diacyl-phosphatidylcholines. Positive correlation with $ISI_{Matsuda}$ as well as almost entirely negative correlation with liver fat content was observed. Nonetheless, PCs were evenly distributed between insulin-sensitive and insulin-resistant individuals. The major part of correlation between metabolite levels and clinical parameters in the follow up is contributed by the insulin resistant individuals. This was not the case in the baseline study. It can therefore be assumed that the applied life-style intervention took effect. The main discriminant between the two groups at baseline and at follow up overall remained the partition of Lyso-PCs.

The compound the most discussed in literature in the context of insulin resistance is diacylglycerol (DAG) [248]. As described previously (section 5.4) the dissociation between NAFL and insulin resistance is hypothesized to be caused by the effective dislocation of DAG into triacylglycerol (TAG) or at least the elimination of DAG. Since Lyso-PCs are products of DAG it can be hypothesized, that the transformation of DAGs into Lyso-PCs and their consequent deportation may be an alternative way to rid the cell of DAG. In this context insulin resistant individuals may suffer from impaired PLA_1 and PLA_2 action or a deficiency of these enzymes. Another interesting result of the current study is the down-regulation of citrate and methyl citrate in insulin sensitive individuals. High concentrations of citric acid (aconitate in negative mode) may indicate a blockage in the TCA cycle. It is known that insulin resistant individuals are characterized by impaired oxidative phosphorylation at mitochondrial membranes while the rate of β-oxidation is high [248]. This phenomenon consequently causes an up-regulation of acetyl-CoA levels. According to our findings the impaired oxidative phosphorylation may be induced due to a blocked TCA activity. Citrate, as well as iso-aconitate were up-regulated in insulin resistant individuals which indicates inhibition of this cycle downstream to these compounds. Increased citrate concentrations were reported to activate liponeogenesis [249]. It was reported that insulin resistance especially coincides with undbalanced Acyl-CoA supply from β-oxidation and its catabolism [250]. The picture presented by the data indicates increased rates of β-oxidation in order to compensate “impaired” oxidative phosphorylation and liponeogenesis on the other side induced by elevated citrate levels. This scenario shows a complete dysregulation of fatty acid metabolism in a contradicting way, due to liponeogenesis, an increased formation of DAG. These findings are supported by an apparently high activity of peroxysomal β-oxidation as an answer to “impaired” oxidative phosphorylation as indicated by the up-regulation of C8-carnitine, a major product of peroxysomal β-oxidation [222]. A further link between the described cycles, DAG and Lyso-PCs is a poor activity of diacylglycerol kinase as observed in insulin resistant subjects [251]. This defect finally leads to an under-production of phosphatidic acid and consequently PCs, PEs, Lyso-PCs and Lyso-PEs [252]. Another factor for the insulin resistant cell to induce catabolic, potentially energy producing, pathways is the down-regulation

of BCAAs in the targeted study. These amino acids can be used for energy production as an alternative resort. This is though mostly used by muscle cells.
In summary Lyso-PCs seem to be the salvage pathway to rid the hepatic cells of DAGs. Especially the up-regulation of Lyso-PC(16:0) indicates PLA_2. Saturated acyl-chains are being ejected into the plasma while the more healthy unsaturated acyl-chains remain in the cell. Furthermore the strong pronunciation of diacyl-phosphatidylcholines and diacyl-phosphatidylethanolamines at the follow up may indicate intensive remodelling of the liver tissue, since especially the latter group is involved in such processes [248, 252]. The specific involvement of Lyso-PCs in inflammation is more and more becoming an important feature for future research, since their long-believed "bad" nature does not seem to hold in our context. Alternative routes for energy-supply as well as impaired citric-cycle-involvement in insulin resistant persons seem to be hypotheses to base future research on.

6. Conclusion

The ultimate aim of metabolomics to detect, identify and quantify all metabolites of a sample organsim or cell by means of one analytical technique seems still barely reachable due to the chemical diversity as well as the extreme dynamic range of abundances inherent to the metabolome. The approach addressed in this thesis, the combination of different analytical strategies into one platform is so far the most realistic way to achieve at least a partition of this goal. The optimization of UPLC-MS alone, in order to discern the human plasma metabolome and in order to find a golden standard applicable in all laboratories is mostly interfered due to hardware and software incoherence and instability. It was shown, that data quality from LC-TOF-MS greatly depends on good temperature insulation, the availability of stable ESI parameters and solvent setups, the choice of the column and sample preparation and none the least a freely adjustable software which enables to readjust undesired effects. Sampling, as well as sample preparation procedures, need to be subjected to constant quality checks. The effect of different anticoagulants in plasma sampling was, except for Li-heparin, negligible as compared to the impact of sample preparation. Where different SPE cartridges did

not vary greatly in affinity – as shown with ICR-FT/MS measurements – the quality of their materials are of higher importance. This was underlined by the apparent breakthrough in the load fraction of some SPE batches. A fast and effective alternative to SPE cartridges are OMIX Tips. MNPs exhibited very good retention for compounds detectable in positive mode but performed poorly for negatively ionizable compounds. Interestingly despite the breakthrough SPEs performed well in ICR-FT/MS measurements. Apparently these events were a good measure against overloading of the ICR cell. Unfortunately we do not yet know to which extent information got lost due to the breakthrough, since PPE did not prove itself to be a good reference approach in ICR-FT/MS. PPE produced mass spectra which were dominated by salt patterns, which do not play such a destructive role in LC-MS.

Concerning UPLC performance, mass accuracy and retention time stability stayed always in ranges typical for this technique. However, the mass spectrometer available had rather extreme shifts of ± 15 ppm over the time of all NAFLD studies analyzed. The amplitude of error was mass specific, which may allow the adaption correction algorithms in the future. There is a big discrepancy between data quality normatively published in the field and the actually achieved data quality in this work. Naturally lowly abundant metabolites will always be close to noise levels yielding logarithmic abundance distributions throughout the acquired intensity ranges. Such low abundance values will always have RSDs >> 30%. Literature references call for at least 80% of all detected peaks to have an RSD < 30%. Published intensity distributions showed not typical logarithmic intensity distributions. The acquired data did not satisfy these asked for normative. However, for the just mentioned reasons we have to doubt whether these normatives are realistic. At the moment it is not feasible to investigate whether this is the case or not, since we have found our TOF-performance to be extremely dependent on daytime, evidently due to insufficient heat insulation. Despite the described discrepancy, the performed studies on NAFLD yielded significant and non-over-fitted results. Both performed ICR-FT/MS studies validated each other, where the LC-MS study revealed not entirely the same information but were well in context. The reason for this may either be the different sensitivities of the instruments or the different sample preparations performed. Furthermore,

annotations of LC-MS peaks were performed by an attempt to match the LC-MS peaks to the ICR-FT/MS peaks which were acquired in parallel. This approach was performed, in order to counteract the relatively low resolution and mass accuracy of the TOF-MS used. The line of thought was to minimize the rate of false assignments. Conclusively, not necessarily the best method for metabolomics was found, but the major interfering effects were identified. The results attained on NAFLD suggest a regio/stereo specificity towards the composition of lipid acyl chain length and degree of unsaturation. The imbalance in abundance of LysoPCs and LysoPEs suggests high rates of methylation or at least a strong involvement of the C_1-metabolism. Where LysoPCs were upregulated in insulin sensitive persons, citric and methylcitric acid as well as N6,N6,N6-trimethyllysine, a carnitine precursor were down regulated. As Lyso-PC abundance negatively correlated with CRP and positively correlated with $ISI_{Matsuda}$ their protective effect against insulin resistance is evident. In the follow up study insulin sensitive individuals still had higher levels of Lyso-PCs. Insulin resistant individuals exhibited a dramatic increase in phosphatidylcholines of which many exhibited anti-correlation to CRP. The applied life-style intervention accordingly took effect. In conclusion, the combinatorial approach of LC-TOF-MS and ICR-FT/MS provides a promising basis for further specification of metabolic disease phenotypes especially in *pre-Diabetes mellitus Type2*.

6.1 Outlook

Future interest lies in the stabilization of the TOF instrument and in the validation of the analytical procedures specified. Regio- and stereo-specificity of Lyso-PC action may get further discerned using C_{30} columns, chiral columns and/or 2D-LC/MS. Algorithms for a more accurate matching of ICR-FT/MS and LC-MS data will be developed. Further follow up studies on the TULIP cohort will be analyzed in order to identify the mechanisms between IS and IR in NAFLD patients. This approach is specifically of importance, since the test groups are basically normalized for obesity which enables identification of patterns independent from this factor but significant for insulin action.

7. Appendix

Materials

Sample preparation

- Vortex-Genie2 Model G-560 E, Scientific Industries
- Oven Nabertherm GmbH, Model: N 30/65 HA, T Max: 650 ºC
- Speed Vac Concentration System:
    1. Speed Vac Concentrator SAVANT SPD 121 P, Thermo Scientific
    2. Refrigerator Vapor Trap RVT 400, Thermo Scientific
    3. Vacuum Pump OF P400, Thermo Scientific
- Centrifuge 5804 R, Max 20800 g, Eppendorf
- Ultrasonic Bath: RK 102P, Bandelin Sonorex
- Pipettes: 2-200µ, 50-200 µl, 100-1000 µl, Eppendorf Reference, Eppendorf (the date of the last calibration must be documented. Calibration need to be performed at least once per year)
- Eppendorf tubes 1.5 ml, Part. number: 27249, Eppendorf
- Eppendorf tubes 2 ml, Part. number: 31399, Eppendorf
- Milli-Q Plus system, Serial number: F3SM59643H, Millipore S.A.

UPLC-MS Analysis

- Aquity UPLC system, Waters
- Synapt HDMS ao-Q-TOF mass spectrometry, Waters
- Glass Syringes: 1010 LTN 10 ML (22/51/3), Part. number: 81616, Hamilton
- UPLC/MS Vials: Total Recovery Vial, Screw cap with bonded pre-slit, PTFE-Silicone Septa, Part. number: 600000671CV, Waters

ICR-FT/MS Analysis

-Brucker FT-ICR_MS

-TRiVersa Nanomate chip electrospray ionization system (Advion Ithaca, USA).

- Autosampler, Gilson

Chemicals

LC-MS Analysis

- Water, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)
- Acetonitrile, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich))
- SodiumHydroxide (≥98%), Roth
- Leucine Enkephalin solution (400 ng/µl), Waters
- Formic Acid, ULC/MS grade, Biosolve (The Netherlands)

Preparation of standard solutions

- Water, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)
- Formic Acid, ULC/MS grade, Biosolve (The Netherlands)
- Internal standards

Nr.	**Name**	**Company**
1	[d_3]Acetyl-L-carnitine.HCl	Dr. Herman J. ten Brink, VU medical center, Netherlands
2	[d_{10}]Adipic acid	Sigma-Aldrich
3	Nialamide	Sigma-Aldrich
4	[d_6]Sulfadimethoxine	Sigma-Aldrich
5	Reserpine	Sigma-Aldrich
6	[d_4]Cholic acid	Sigma-Aldrich
7	Decanoic Acid-C13	Sigma-Aldrich
8	[16,16,16-d_3]hexadecanoyl-L-carnitine.HCl	Dr. Herman J. ten Brink, VU medical center, Netherlands

Solutions

MS

- Calibration Solution

Prepare a solution of sodium formate mixing 10 %Vol. of formic acid in water (1 part), 0.1 M sodium hydroxide in water (1 part) and acetonitrile (8 part)

- Lock Mass Solution

Add 0.2 % Vol. of Leucine Enkephaline solution and 0.1 % Vol. of formic acid to a solution of MeOH/H_2O:1/1 (LC-MS grade) to a final concentration of 400 µg/l.

UPLC

- Mobile Phases

All the solvents are ULC/MS

A_1: H_2O/AcN : 9/1 (weak wash)

A_2: H_2O/AcN : 95/5 + 0.1 % formic acid

B_1: AcN/ MeOH/2-propanol : 6/3/1 (strong wash)

B_2: AcN

- -Blank

H_2O/AcN: 80/20.

- Plasma Blank

Water collected in plasma Sarstedt tubes and treated as plasma; then reconstituted in H_2O/AcN : 80/20 at volume of 50 µl.

- Plasma QC

Pool of plasma from different donors treated as plasma and reconstitute in H_2O/AcN : 80/20 at volume of 50 µl.

- Internal Standards

All the internal standards stock solutions are added, at the following concentrations to the plasma aliquot before protein precipitation extraction (In the table are described as well the final concentrations of internal standards after reconstitution of the dried sample in 50 µl of H_2O/AcN: 80/20 solution

Nr.	Name	STOCK SOLUTION 1 mg/ml standard solved in	SAMPLE PREPARATION Concentration in 80µl of plasma mg/L	RECONSTITUTION Final Concentration (in a final volume of 50 µl of H_2O/AcN:8/2) mg/L
1	[d_3]Acetyl-L-carnitine HCl	Methanol	0.625	1
2	[d_{10}]Adipic acid	Acetonitrile	6.25	10
3	Nialamide	Acetonitrile	0.625	1
4	[d_6]Sulfadimethoxine	Acetonitrile	0.625	1
5	Reserpine	Acetonitrile	0.625	1
6	[d_4]Cholic acid	Methanol	0.625	1
7	Decanoic Acid-C13	Acetonitrile	6.25	10
8	[16,16,16d_3]hexadecanoyl-L carnitine HCl	Methanol	0.625	1

Sample preparation (protein precipitation) prior UPLC-MS analysis

This protocol was used for the sample preparation oprimization, testing different kinds of organic solvents. The same protocol, using acetonitrile as organic solvent, was used for the sample preparation of plasma from fatty liver individuals.

All the samples need to be thrown on ice and vortex mixed for 30 seconds.

All the standards need to be added before the protein precipitation and let the samples stand on ice for at least 3 minutes; then vortex mixing before using of the sample.

The protein precipitation extraction procedure is performed adding Acetonitrile (320 µl) to a plasma aliquot (80 µl) in Eppendorf tubes of 1.5ml.

After the addition of the organic solvent, the samples are vortex mixed for 30 seconds at RT and then are centrifuged at 15294 g for 10 minutes at 4ºC.
The recovered supernatant is dried in Eppendorf tube of 1.5 ml.
In the protein precipitation tests using different kinds of organic solvents, the same ratio among plasma and organic solvent (1/4) was maintained. The total volume was always set at 400 µl. The amount of plasma was always kept at 80 µl.

Sample preparation prior FT-MS analysis

The fllowing protocol was used in the sample preparation of plasma from NAFL individuals. An additional pre- step as protein precipitation, was performed .

Materials:

-Omix C18 100µl tips, Varian
-MeoH, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)
-Water, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)
-H_3PO_4 85 % , Merck
- Formic Acid, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)
Solutions:
2% Formic acid in water
2% H_3PO_4 in water

Procedure:
For the conditioning, washing and elution steps set the pipettor to the max tip volume (100µl)
Pretreat the plasma (50µl in 100 µl of 2%H_3PO_4) and vortex for
30 seconds

- Conitioning: Aspirate and discard MeoH solution for 10 times
- Equilibration: Aspirate and discard 2% Formic acid solution for 10 times
- Loading: Aspirate and release the pre-treated sample 5 times. Repeated the procedure 10 cycles.

- Washing: Aspirate and discard 2% Formic acid solution for 10 times
- Elution: Aspirate MeoH solution and dispense into a vial. Aspirate and release it for 10 times.

Additional sample preparation

Additional protocols used in sample preparation

Solid Phase Extraction Method

Materials:

-HLB Oasis cartridges (1cc, 30 mg) , Waters, Milford
-C1, C2, C8, C18, Ph and CN Varian Bond Elute (1cc, 100 mg), Varian
-C18 Varian Bond Elute (6cc, 1 g), Varian
-Zic-Hilic cartridges (1cc, 25 mg and 20 cc/1g), Sequant
-HLB Oasis cartridges (20cc, 1g)
-Omix C18 100µl tips, Varian
-MeoH, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)
-Water, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)
-H_3PO_4 85 % , Merck
- Formic Acid, LC/MS grade, Chromosolv®, Fluka Analytical (Sigma-Aldrich)

Solutions:
2% Formic acid in water
2% H_3PO_4 in water

Procedure:
The solid phase extraction procedure was performed using the different kinds of cartridges selected. The vacuum pressure applied was maintained at 13 mmHg. The cartridges were conditionated with 1 ml of MeOH and 1 ml of H_2O in the case of HLB Oasis (2 ml of each step when 1 g sorbent was applied), 1ml of MeOH and 1 ml of formic acid 2% in the case of Ph, CN, C18, C8 and C2(2 ml of each step when 1 g sorbent was applied); 1ml

H_2O and 1 ml AcN in the case of Hilic application. The plasma was acidified with 2 % of H_3PO_4 at a dilution rate 1/1, then it was loaded onto the cartridge. In the case of Hilic application the samples was diluted in AcN, therefore protein precipitated and the supernatant was load on the cartridge. The wash step within HLB protocol was performed by 5% MeOH (1 ml) and the elution one using MeOH. 2% formic acid in the washing step and MeOH in the elution step in the case of the other sorbents. For Hilic application the washing step was performed using acetonitrile, the elution using a small volume of water.

- For all the methods different elution volumes were tested reaching 3 ml as maximum level. In the case of protocol optimization different kinds of solvents were testes (IPA, MeoH/AcN, AcN/IPA, EtOac.)
- During every protocols load fraction and elute fraction were collected and evaporated.

  The residue was reconstituted in 50 µl of H_2O/AcN: 8/2 for UPLC-MS applications. The eluate was diluted 1/50 in MeoH /H_2O :8/2 prior analysis via ICR-FT/MS.

- At the same time solid phase extraction of aqueous solution, collected in plasma Sarstedt tubes, was conducted following the procedure above mentioned in order to evaluate possible interferences from the cartridges.

Magnetic Nanoparticle application

Production :

The solvothermal reduction method to obtain magnetite (Fe_3O_4) nanoparticles was set in one simple step based on the method described by Hong Deng et.al. [253]. In an effort to deposit magnetic nanoparticles on the surface of adsorbent micro material for free-solution solid phase extraction, Si-C_{18} particles were introduced into the solvothermal system. An amount of 0.540g of $FeCl_3 \cdot 6H_2O$ (Santa Cruz Biotechnology, Heidelberg- Germany) was dissolved in 16 mL ethylenglycol PA (Serva Electrophoresis, Heidelberg-Germany) until it was obtained an

homogenous solution, followed by addition of 1,44g of sodium acetate crystal (Merck, Darmstadt-Germany) and 0,200g of Si-C_{18} (Fluka) and HLB (Oasis, Waters). The mixture was sealed in Teflon-lined stainless steel autoclave and was maintained at 200 ºC during 12 h.The shape and the size of our microparticles were verified by using electron microscopy

Application:

Prior to application, the MNPs were carefylly washed with Methanol.

Every step, from conditioning until elution, was maintained as in the protocol applied in the cartridges application (using C18 and HLB). The amount of plasma was kept at 80 µl and it was pre-treatreated with orthophosphoric acid 2% at dilution rate 1/1. Elution was performed, in both cases (C18 and HLB), with methanol (500 µl, volume). The eluate was dilutetd 1/50 in MeoH /H_2O :8/2 prior analysis via ICR-FT/MS.

Analysis via UPLC-TOF-MS

Columns

TECHNICAL CHARACTERISTICS	RP
Brand	Alltech Grom GmbH, Germany
Type	VisionHT C18 HL
Dimension	2 x 150 mm
Particle Size	1.5 µm
Pre-column	-
Handling	
Pre-equilibration of NEW Columns	60 minutes AcN ; 0.1 ml/min
Equilibration	30 min B_2 /A_2 :50/50, flow: 0.2 ml/min 30 min B_2 / A_2 :80/20, flow: 0.25 ml/min 20 min A_2, flow: 0.3 ml/min
Temperature	40º C
Wash	30 min A_1; flow: 0.2 ml/min 30 min $B_{1;}$ flow: 0.2 ml/min
Storage	20 min AcN /H2O :50/50, flow: 0.2 ml/min 20 min AcN /H2O :80/20, flow: 0.2 ml/min

Injecton parameters

Loop volume:	10 µl
Injection volume:	5 µl
Injection type:	PL
Weak wash:	H2O/AcN: 9/1
V_{weak}:	3000 µl
Strong wash	AcN / MeOH/ 2-propanol: 6/3/1
V_{strong}:	1000 µl
Applied advanced parameters	
Syringe draw rate:	100 µl/min
Needle placment (from the bottom)	3 mm
Air Gaps:	
pre-aspirate:	4 µl
post-aspirate:	4 µl

Gradient

Time	**Flow (ml/min)**	**A_2 %**	**B_2 %**	**Slope**
0	0.3	100	0	6
1.12	0.3	100	0	6
22.27	0.3	0	100	6
29.49	0.3	0	100	6
29.56	0.3	100	0	6
35.00	0.3	100	0	6

Detection

- The Lock mass is measured every 15s with scan time of 1s and flow rate of 5 µl/min
- The acquisition in negative mode require flow splitting.
  Peek tubes lenght: 22.5 cm. Flow rate into the MS system: 100 µl/min

Detection Parameters	**Polarity**	**Polarity**
	ES^+	**ES^-**
Capillary Voltage (kV)	3.10	2.30
Sample Cone (V)	30.00	40
Extraction cone	4.00	5
Source Temperature	120	
Desolvatation temperature (ºC)	300	
Desolvatation Flow (L/h)	800	
Cone(L/h)	50	
Detector (V)	1800	

Sample analysis order

The sample set was analyzed in the following way:

- 3 x blanks at the beginning of the study
- 16 x quality control plasma (QC)
- 3 x internal standard mix diluted in blank solution (after QC plasma column conditioning, in the middle of the study and at the end)
- Samples are measured in random triplicates within three different batches, where one batch include all samples measured one time
- 1 x quality control plasma (QC) every 10 measurements
- 4 x Plasma Blank randomly in order to examine possible carryover

The samples sequence in every batch it is randomized applying the function RAND of Microsoft Excel program

Analysis via ICR-FT/MS

The analyses of the baseline fatty liver samples were acquired on a Brucker APEX Qe FT-ICR_MS with a TRiVersa Nanomate chip electrospray ionization system (Advion Ithaca, USA). The resolving power was 120000 at m/z 400 in positive mode and 400000 at m/z 400 in the negative mode. The ionization was applied in both modes within a mass range of 147 m/z to 2000 m/z and with a time domain of 1 MW in positive mode and 4 MW in negative. 1024 scans were acquired for spectra in both modes .The spectra were externally calibrated on clusters of arginine (1mg/ml in methanol /water: 80/20) and internally calibrated on fatty acid masses in the case of the negative mode. The signal to noise ratio threshold was set to S/N = 4. Due to the high density of ions found in positive mode ionization, it was necessary to use a different calibration approach. This calibration was based on data base matching using MassTRIX. All the input masses exported from the original raw data (S/N ratio= 4) were assigned to the closest reference mass found in the database, within 1ppm error. Consequently a new calibration list was created performing a regression through the error distribution of these MassTRIX annotations.

Intermittent hardware updates (i.e. ion optics, electronics, pumps) and software updates (new FT algorithm) necessitated the development of new methods within ICR-FT/MS applications. The methods used to acquire the follow up samples (from NAFL subjects) used a time domain of 2MW, over 400 scans. This yielded a resolution of 220000 at m/z 400 as well as more populated mass spectra. The follow up samples were analyzed by automated infusion via a Gilson auto-sampler. The mass range acquired was set from 122 until 1000 m/z. The negative mode spectra were internally calibrated based on fatty acids and then exported with a signal to noise ration of S/N = 4. In positive mode acquisition, like in the above citied case, the adoption of an extra methodology for the calibration of the spectra was necessary. The combination of higher resolution and higher density, derived from the ionization of samples, required a network application. Prior to the calibration of all mass spectra, a screening for symmetric peak shape, low abundance at mass defect 0.6 to 0.8 at m/z < 600 and maximum noise amplitude was implemented. Afterwards, spectra were filtered for S/N > 4, a minimum peak

height of $1.5*10^6$ counts as well as a minimum relative intensity of 10^{-7}%, respectively. At first, the mass spectrum 1087_AD was chosen as a reference spectrum since it exhibited large peak counts. Consequently, 1087_AD was processes by the NetCalc annotation tool [133]. Choosing palmitic acid as a starting mass, a list of 6111 putatively annotated formulas was created based on a network reconstruction margin of 0.25 ppm. Consequently all mass spectra as well as the annotated mass list were aligned using the in-house software 'MatrixGenerator' at a ppm error of 2 ppm. Masses matching the annotated list and throughout more than 95% of all samples were extracted and stored as a new calibration list. Further filtering led to a final size of 594 calibration masses ranging from $200 < m/z < 1100$. This amount of calibration masses enabled direct adjustment of spectra at hand of a well resolved error distribution and consequently to a comparable calibration with exquisite peak alignment over a wide mass range. The standard deviation of mass relative to the calibration masses within the m/z range of interest was always below 100 ppb. The peaks of all exported spectra were aligned through in-house software 'MatrixGenerator' using a 1 ppm error window. The total frequency threshold of m/z value occurrence across the samples was set to 10 %.

Alignment and Annotation

The exported FT-MS spectra were aligned through the in-house software 'MatrixGenerator' and the data were combined into a matrix within 1 ppm error. The data acquired through UPLC-Q-TOF were aligned using MarkerLynx software (Waters, Mildfors) within a mass range of 0.02 Da and RT window of 0.1 minutes. All the masses were submitted to MassTRIX for the assignment through KEGG (http://www.genome.jp/kegg/), LipidMaps (http://www.lipidmaps.org/) and HMDB (http://www.hmdb.ca/). Integration of the annotation was done as well using the formula calculation derived from the network approach.

Table 12: Characteristic clinical parameters of the NAFL subjecties studied via non-targeted approach. The population is divided according to their insulin sensitivity. Base line and follow up (after 9 months of life style ntervention) are recorded.

Characteristics	Insulin sensitive subjects		Insulin resistant subjects				
	Baseline (N=20)	Follow-Up (N=20)	Baseline (N=20)	Follow-Up (N=20)	p-value at baseline	p-value at Follow-Up	
Insulin Sensitivity [arb. U]	14.7 ± 1.1	18.4 ± 1.72	7.3 ± 0.52	7.9 ± 0.75	<0.0001	<0.0001	
Liver Fat MRS [%]	8.6 ± 0.38	8.6 ± 0.85*	8.5 ± 0.51	5.6 ± 0.65*	0.72	0.71	
Gender [males/females]	15/5		7/13		0.01		
Age [years]	52 ± 2		44 ± 2		0.016		
Body Mass Index [kg*m^(-2)]	30.1 ± 0.76	28.9 ± 0.93*	32.4 ± 0.81	31.8 ± 0.81*	0.05	0.02	
Waist Circumference [cm]	103.5 ± 1.99	98.1 ± 2.63*	105.1 ± 2.12	101.9 ± 2.14*	0.61	0.25	
Total Body Fat MRT [Kg]	24.4 ± 1.82	20.5 ± 2.27	32 ± 1.5	28.1 ± 2.06*	0.002	0.09	
Visceral Fat MRT [Kg]	4.5 ± 0.41	3.8 ± 0.54*	3.7 ± 0.35	3.6 ± 0.50*	0.14	0.83	
AST [U/L]	26.6 ± 1.46	26.6 ± 3.44	27.1 ± 2.58	22 ± 1.51*	0.89	0.18	
ALT [U/L]	27.3 ± 2.33	26.8 ± 4.58	29.4 ± 2.2	26.5 ± 2.99	0.78	0.95	
Hs-CRP [mg/dL]	0.2 ± 0.05	0.17 ± 0.04*	0.36 ± 0.09	0.16 ± 0.04*	0.32	0.97	
Free Fatty Acids [µmol/L]	650 ± 45	555 ± 43*	645 ± 54	591 ± 43	0.77	0.51	
Fasting Glucose [mM]	5.3 ± 0.15	5.1 ± 0.15	5.3 ± 0.12	5.2 ± 0.12	0.96	0.86	
2h Glucose [mM]	7.5 ± 0.37	6.9 ± 0.50*	7.3 ± 0.43	6.4 ± 0.32*	0.58	0.51	
PAI-1 [ng/mL]	4.5 ± 0.51	4.6 ± 0.43	5.5 ± 0.53	6.5 ± 0.79	0.1	0.07	
Fetuin-A [µg/mL]	237 ± 9	±	293 ± 14	±	0.002		

Values are means ± SEM. * $p < 0.05$ for change within each group, AT: adipose tissue; MRT: magnetic resonance tomography; MRS: magnetic resonance spectroscopy; hs-CRP: high sensitivity C-reactive protein; PAI-1: plasminogen activator inhibitor 1; at follow-up data from MR were only available in 28 and for insulin sensitivity in 38 subjects.

Table 13: Mass deviation (expressed as RSD %) of labelled standard compounds used for spiking plasma quality control (QC) samples in NAFL studies. Analyses performed via UPLC-Q-TOF-MS, positive mode ionization. Comparison between Study 1 and Study 2.

	NAFL_Study_1					NAFL_Study_2				
Theoretical ion mass (m/z)	**207.1419**	**299.1505**	**317.1185**	**403.3615**	**609.2807**	**207.1419**	**299.1505**	**317.1185**	**403.3615**	**609.2807**
	St. 1(m/z)	St. 2(m/z)	St. 3(m/z)	St. 4(m/z)	St. 5(m/z)	St. 1(m/z)	St. 2(m/z)	St. 3(m/z)	St. 4(m/z)	St. 5(m/z)
QC17	207.137	299.149	317.117	403.3618	609.283	207.1381	299.1454	317.1188	403.3622	609.2816
QC18	207.1369	299.1486	317.1174	403.3604	609.2834	207.1386	299.1477	317.1177	403.3625	609.2824
QC19	207.1365	299.1479	317.1174	403.3616	609.2833	207.1358	299.1467	317.1175	403.3605	609.2836
QC20	207.136	299.147	317.119	403.3619	609.281	207.138	299.1479	317.1193	403.3489	609.2665
QC21	207.1366	299.1487	317.1169	403.3608	609.284	207.1368	299.1466	317.1181	403.3623	609.2836
QC22	207.136	299.148	317.117	403.3613	609.286	207.1385	299.1465	317.1168	403.3607	609.2822
QC23	207.1385	299.1476	317.1172	403.3619	609.2836	207.1372	299.1469	317.1172	403.3614	609.2861
QC24	207.136	299.147	317.118	403.3617	609.283	207.1385	299.149	317.1177	403.3618	609.2847
QC25	207.1387	299.1459	317.1184	403.3595	609.2837	207.1372	299.1427	317.1181	403.3615	609.2852
QC26	207.137	299.148	317.117	403.3616	609.284	207.1377	299.1435	317.1173	403.3633	609.2847
QC27	207.138	299.1463	317.1179	403.3625	609.2859	207.1364	299.1431	317.1159	403.3612	609.2867
QC28	207.135	299.149	317.115	403.3597	609.284	207.1376	299.1463	317.1173	403.3607	609.2818
STD	0.001099	0.001026	0.000985	0.000931	0.001311	0.000886	0.001986	0.00089	0.003772	0.005286
AV	207.1369	299.1478	317.1174	403.3612	609.2837	207.1375	299.146	317.1176	403.3606	609.2824
RSD%	0.000531	0.000343	0.000311	0.000231	0.000215	0.000428	0.000664	0.000281	0.000935	0.000868

Table 14: Mass deviation (expressed as RSD %) of selected ion peaks present in plasma quality control (QC) samples in NAFL studies. Analyses performed via UPLC-Q-TOF-MS, positive mode ionization. Comparison between Study 1 and Study 2.

	NAFL_Study_1								NAFL_Study_2							
	Peak 1 (m/z)	Peak 2 (m/z)	Peak 3 (m/z)	Peak 4 (m/z)	Peak 5 (m/z)	Peak 6 (m/z)	Peak 7 (m/z)	Peak 8 (m/z)	Peak 1 (m/z)	Peak 2 (m/z)	Peak 3 (m/z)	Peak 4 (m/z)	Peak 5 (m/z)	Peak 6 (m/z)	Peak 7 (m/z)	Peak 8 (m/z)
QC17	188.0728	195.0887	203.0505	282.2798	468.3104	496.3417	520.3428	522.3574	188.0704	195.0887	203.0496	282.2789	468.309	496.3398	520.3426	522.3598
QC18	188.0736	195.088	203.049	282.28	468.3094	496.3409	520.3409	522.358	188.0718	195.0868	203.0494	282.2795	468.307	496.3396	520.343	522.3574
QC19	188.0723	195.0889	203.0506	282.2791	468.3094	496.341	520.3431	522.3593	188.0706	195.0889	203.0518	282.2789	468.3091	496.3417	520.3432	522.3568
QC20	188.0719	195.0887	203.0505	282.2798	468.3096	496.3414	520.3433	522.3588	188.0716	195.0887	203.0521	282.2797	468.295	496.3257	520.3267	522.3428
QC21	188.0719	195.0876	203.0507	282.2791	468.309	496.3408	520.3428	522.3575	188.0713	195.0885	203.0529	282.2792	468.3079	496.3409	520.3416	522.3571
QC22	188.0713	195.0879	203.0499	282.2799	468.3093	496.3403	520.3413	522.3591	188.0715	195.0875	203.0527	282.2782	468.31	496.3428	520.3417	522.3571
QC23	188.0719	195.0881	203.0498	282.2795	468.3099	496.3412	520.3411	522.358	188.0712	195.0871	203.0528	282.2785	468.3083	496.3419	520.3396	522.3575
QC24	188.0715	195.0896	203.0506	282.2805	468.3097	496.3398	520.3422	522.3576	188.0716	195.0881	203.0516	282.2786	468.31	496.3404	520.342	522.3586
QC25	188.0742	195.0874	203.0537	282.2806	468.3107	496.3416	520.3398	522.3578	188.0713	195.0885	203.0519	282.2789	468.3073	496.3418	520.3412	522.358
QC26	188.073	195.0882	203.0497	282.2799	468.3095	496.341	520.3417	522.3586	188.0711	195.0865	203.0515	282.2796	468.3085	496.3401	520.3422	522.3561
QC27	188.0719	195.087	203.0505	282.279	468.3106	496.3398	520.3423	522.3572	188.0713	195.0881	203.0517	282.2793	468.3077	496.3404	520.3431	522.3566
QC28	188.0691	195.0857	203.0505	282.2787	468.3082	496.3427	520.3395	522.3584	188.0709	195.0865	203.0495	282.2788	468.3085	496.3404	520.3414	522.3583
STDEV	0.001285	0.00101	0.00113	0.000593	0.000703	0.000818	0.001249	0.000692	0.000415	0.000902	0.00127	0.000462	0.004005	0.004491	0.004519	0.004379
AVERAGE	188.0721	195.088	203.0505	282.2797	468.3096	496.341	520.3417	522.3581	188.0712	195.0878	203.0515	282.279	468.3074	496.3396	520.3407	522.3563
RSD%	0.000683	0.000518	0.000556	0.00021	0.00015	0.000165	0.00024	0.000132	0.000221	0.000462	0.000626	0.000164	0.000855	0.000905	0.000868	0.000838

Table 15: Retention Time shift (expressed as RSD %) of labelled standard compounds used for spiking plasma quality control (QC) samples in NAFL studies. Analyses performed via UPLC-Q-TOF-MS, positive mode ionization. Chromatographic separation via : C18 Vision HT-HL UPLC column (2 x 150 mm, 1.5 µm, Alltech Grom GmbH, Germany). Comparison between Study 1 and Study 2.

	NAFL_Study_1					NAFL_Study_2				
Theoretical ion mass (m/z)	207.1419	299.1505	317.1185	609.2807	403.3615	207.1419	299.1505	317.1185	609.2807	403.3615
	St. 1(RT)	St. 2(RT)	St. 3(RT)	St. 4(RT)	St. 5(RT)	St. 1(RT)	St. 2(RT)	St. 3(RT)	St. 4(RT)	St. 5(RT)
QC17	1.21	5.37	8	10.27	15.87	1.24	5.32	8	10.34	15.93
QC18	1.22	5.37	8	10.27	15.86	1.23	5.37	8.01	10.29	15.86
QC19	1.22	5.37	8	10.27	15.86	1.23	5.49	8.02	10.21	15.71
QC20	1.22	5.37	8.01	10.27	15.86	1.22	5.58	8.03	10.15	15.62
QC21	1.22	5.37	8	10.26	15.85	1.23	5.5	8.03	10.22	15.74
QC22	1.23	5.37	8	10.24	15.84	1.23	5.5	8.03	10.23	15.74
QC23	1.23	5.37	8	10.25	15.8	1.23	5.5	8.04	10.23	15.74
QC24	1.22	5.37	8	10.25	15.82	1.23	5.5	8.03	10.23	15.75
QC25	1.23	5.38	8	10.25	15.81	1.23	5.5	8.03	10.23	15.74
QC26	1.23	5.37	7.99	10.25	15.81	1.23	5.5	8.03	10.23	15.74
QC27	1.23	5.37	8	10.25	15.84	1.23	5.5	8.03	10.24	15.76
QC28	1.23	5.43	8.04	10.31	15.88	1.23	5.5	8.03	10.23	15.76
STDEV	0.006686	0.017299	0.012309	0.018505	0.026227	0.004264	0.068091	0.010836	0.045017	0.075694
AVERAGE	1.224167	5.375833	8.003333	10.26167	15.84167	1.23	5.48	8.025833	10.23583	15.7575
RSD %	0.546133	0.321785	0.1538	0.180328	0.16556	0.346668	1.242534	0.135017	0.439796	0.480367

Table 16: Retention Time shift (expressed as RSD %) of selected ion peaks present in plasma quality control (QC) samples in NAFL studies. Analyses performed via UPLC-Q-TOF-MS, positive mode ionization. Chromatographic separation via: C18 Vision HT-HL UPLC column (2 x 150 mm, 1.5 µm, Alltech Grom GmbH, Germany). Comparison between Study 1 and Study 2.

	NAFL_Study_1								NAFL_Study_2							
Peak(m/z)	203.0496	188.0704	195.0887	468.309	520.3426	496.3398	522.3598	282.2789	203.0496	188.0704	195.0887	468.309	520.3426	496.3398	522.3598	282.2789
	Peak 1 (RT)	Peak 2 (RT)	Peak 3 (RT)	Peak 4 (RT)	Peak 5 (RT)	Peak 6 (RT)	Peak 7 (RT)	Peak 8 (RT)	Peak 1 (RT)	Peak 2 (RT)	Peak 3 (RT)	Peak 4 (RT)	Peak 5 (RT)	Peak 6 (RT)	Peak 7 (RT)	Peak 8 (RT)
QC17	1.06	3.92	4.69	14.76	16.07	17.34	17.82	20.15	1.07	4.04	4.74	14.84	16.16	17.46	17.95	20.13
QC18	1.06	3.92	4.69	14.76	16.07	17.32	17.82	20.15	1.07	4	4.74	14.83	16.16	17.45	17.96	20.12
QC19	1.06	3.92	4.69	14.76	16.06	17.31	17.82	20.15	1.09	3.93	4.74	14.83	16.16	17.46	17.97	20.13
QC20	1.06	3.92	4.69	14.76	16.06	17.32	17.82	20.15	1.09	3.89	4.75	14.83	16.16	17.47	17.96	20.13
QC21	1.06	3.92	4.69	14.76	16.07	17.32	17.82	20.15	1.09	3.95	4.76	14.84	16.16	17.46	17.97	20.13
QC22	1.07	3.93	4.7	14.76	16.07	17.33	17.83	20.14	1.09	3.94	4.76	14.85	16.16	17.47	17.98	20.13
QC23	1.07	3.93	4.7	14.76	16.07	17.33	17.83	20.15	1.09	3.95	4.76	14.85	16.17	17.48	17.97	20.13
QC24	1.07	3.93	4.7	14.76	16.07	17.33	17.83	20.15	1.09	3.95	4.76	14.85	16.17	17.48	17.97	20.13
QC25	1.07	3.94	4.7	14.76	16.07	17.34	17.84	20.15	1.09	3.95	4.76	14.84	16.16	17.47	17.97	20.12
QC26	1.06	3.93	4.7	14.76	16.08	17.33	17.82	20.14	1.09	3.95	4.76	14.85	16.17	17.48	17.97	20.12
QC27	1.07	3.93	4.7	14.76	16.08	17.33	17.83	20.14	1.09	3.95	4.78	14.85	16.17	17.49	17.98	20.13
QC28	1.07	4	4.7	14.83	16.15	17.42	17.94	20.15	1.09	3.95	4.76	14.85	16.17	17.49	17.99	20.13
STDEV	0.005222	0.022208	0.005149	0.020207	0.023868	0.028123	0.03371	0.004523	0.007785	0.036296	0.011645	0.00866	0.005149	0.012673	0.010445	0.004523
AVERAGE	1.065	3.9325	4.695833	14.76583	16.07667	17.335	17.835	20.1475	1.086667	3.954167	4.755833	14.8425	16.16417	17.47167	17.97	20.1275
RSD%	0.49036	0.564722	0.109657	0.136851	0.148466	0.162233	0.18901	0.022448	0.71641	0.917926	0.244857	0.058348	0.031856	0.072535	0.058123	0.02247

Table 17: Peak Area shift (expressed as RSD %) of labelled standard compounds used for spiking plasma quality control (QC) samples in NAFL studies. Analyses performed via UPLC-Q-TOF-MS, positive mode ionization. Chromatographic separation via : C18 Vision HT-HL UPLC column (2 x 150 mm, 1.5 µm, Alltech Grom GmbH, Germany). Comparison between Study 1 and Study 2.

	NAFL_Study_1					NAFL_Study_2				
Theoretical ion mass (m/z)	207.1419	299.1505	317.1185	403.3615	609.2807	207.1419	299.1505	317.1185	403.3615	609.2807
	St. 1(Area)	St. 2(Area)	St. 3(Area)	St. 4(Area)	St. 5(Area)	St. 1(Area)	St. 2(Area)	St. 3(Area)	St. 4(Area)	St. 5(Area)
QC17	37.018	62.1	342.24	1427.535	248.24	45.439	31.51	262.635	818.755	250.009
QC18	34.922	68.673	358.186	1433.103	244.529	47.567	32.398	273.97	857.476	263.91
QC19	35.24	44.115	339.964	1369.438	261.481	43.256	32.283	308.488	957.867	337.358
QC20	38.018	32.46	317.56	1426.723	327.793	36.185	83.45	408.267	1087.716	370.206
QC21	44.029	87.051	318.888	1281.478	202.595	40.725	9.772	316.229	887.715	316.67
QC22	36.037	69.26	303.41	159.128	264.49	35.128	10.458	304.285	842.934	289.994
QC23	45.9	24.715	338.312	1274.125	304.465	46.995	47.73	286.609	729.449	216.122
QC24	30.067	10.88	296.79	1280.133	365.18	32.929	59.256	386.114	1115.45	316.316
QC25	42.018	60.817	343.355	1351.548	351.388	43.65	6.67	335.231	1125.068	276.305
QC26	45.377	37.169	341.37	1329.321	1390.13	27.99	4.83	337.265	1111.08	273.112
QC27	45.568	14.072	348.355	1257.636	319.732	58.125	11.801	349.086	1216.13	280.83
QC28	25.749	67.82	169.04	554.461	147.73	20.542	8.146	342.181	1178.431	282.358
STDEV	6.468325	24.50331	50.50703	398.0052	327.5989	9.998642	24.82459	43.20754	163.0772	40.93111
AVERAGE	38.32858	48.261	318.1225	1178.719	368.9794	39.87758	28.192	325.8633	994.0059	289.4325
RSD%	16.87598	50.77248	15.8766	33.76591	88.78513	25.07334	88.05546	13.25941	16.40605	14.14185

Table 18: Paeak Area shift (expressed as RSD %) of selected ion peaks present in plasma quality control samples in NAFL studies. Analyses performed via UPLC-Q-TOF-MS, positive mode ionization. Chromatographic separation via: C18 Vision HT-HL UPLC column (2 x 150 mm, 1.5 µm, Alltech Grom GmbH, Germany). Comparison between Study 1 and Study 2**.**

	NAFL_Study_1								NAFL_Study_2							
Peak(m/z)	203.0496	188.0704	195.0887	468.309	520.3426	496.3398	522.3598	282.2789	203.0496	188.0704	195.0887	468.309	520.3426	496.3398	522.3598	282.2789
	Peak 1 (Area)	Peak 2 (Area)	Peak 3 (Area)	Peak 4 (Area)	Peak 5 (Area)	Peak 6 (Area)	Peak 7 (Area)	Peak 8 (Area)	Peak 1 (Area)	Peak 2 (Area)	Peak 3 (Area)	Peak 4 Area)	Peak 5 (Area)	Peak 6 (Area)	Peak 7 (Area)	Peak 8 (Area)
QC17	790.47	343.142	45.209	1666.51	88.474	5578.977	2953.451	2088.425	668.416	226.883	34.451	2680.62	145.022	5098.834	2888.827	1996.571
QC18	818.064	335.212	70.156	1792.244	101.572	5489.114	2910.791	2073.778	702.681	227.322	46.876	2552.361	151.437	5256.27	2860.484	2016.165
QC19	991.456	414.886	87.326	3526.018	165.898	7335.31	3991.421	2972.875	811.687	345.479	63.675	2249.478	161.242	5298.052	2997.781	2080.144
QC20	953.244	343.142	45.209	3584.452	180.29	7402.536	4085.336	3036.309	1040.299	430.842	64.424	2011.093	169.238	5052.958	3073.535	1895.811
QC21	839.39	377.034	55.84	2719.258	152.871	6772.418	3624.591	2622.07	889.977	410.756	77.229	2033.901	163.302	5198.473	3054.045	2004.243
QC22	851.096	368.12	50.418	2618.406	144.735	6893.026	3736.381	2707.911	858.028	396.013	43.029	2369.941	156.997	5043.805	2875.501	1953.422
QC23	901.775	395.06	80.998	2618.022	155.981	7148.976	3995.064	2861.239	754.837	290.167	69.827	1851.711	131.934	4542.342	2502.095	1713.787
QC24	913.143	388.99	46.93	2774.585	172.427	7305.221	3958.603	2760.812	945.231	319.976	75.714	3315.906	170.587	6565.374	3587.641	2285.595
QC25	918.834	375.024	99.759	2998.563	147.597	6884.27	3836.595	2748.468	665.509	356.57	93.625	3141.095	179.348	6840.15	3769.466	2441.572
QC26	875.315	389.856	67.041	2805.69	151.759	6874.42	3831.588	2689.081	855.113	332.581	77.818	3228.584	171.005	6725.134	3677.136	1238.216
QC27	923.388	398.292	96.396	3937.365	190.03	7816.375	4307.196	3061.293	926.519	270.506	96.048	2967.936	175.128	6569.518	3698.329	2309.124
QC28	396.394	160.287	28.597	2058.27	52.822	3965.358	2006.152	1485.131	913.813	280.418	78.161	3096.601	167.478	6552.371	3625.716	2314.168
STDEV	153.2896	66.74907	22.77862	699.2485	40.64231	1084.885	656.135	472.8296	118.2394	67.70502	19.1069	521.0766	13.6231	838.4676	427.5455	323.0004
AVERAGE	847.7141	357.4204	64.48992	2758.282	142.038	6622.167	3603.097	2592.283	836.0092	323.9594	68.40642	2624.936	161.8932	5728.607	3217.546	2020.735
RSD%	18.0827	18.67523	35.32122	25.35087	28.61369	16.38263	18.21031	18.23989	14.14332	20.89923	27.93144	19.85102	8.414871	14.6365	13.28794	15.98431

Table 19: Significant biomarkers UP-regulated respect to the insulin sensitive group at base line. Experimental analyses via direct infusion with ICR-FT/MS in positive mode ionization. The p-val indicated derives from Wilcoxon test.

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
1	↑	0.576	0.007	40	496.33966	$C_{24}H_{50}NO_7P$	$[M+H^+]^+$	-0.22	LysoPC(16:0)	Lysophosphatidylcholines
2	↑	0.318	NS	40	520.33968	$C_{26}H_{50}NO_7P$	$[M+H^+]^+$	-0.17	LysoPC(18:2)	Lysophosphatidylcholines
3	↑	0.251	0.004	40	522.35532	$C_{26}H_{52}NO_7P$	$[M+H^+]^+$	-0.19	LysoPC(18:1(9Z))	Lysophosphatidylcholines
4	↑	0.251	0.004	40	522.35532	$C_{31}H_{49}NO_4$	$[M+Na^+]^+$	-0.11	unknown	Unknown
5	↑	0.122	0.014	40	524.37104	$C_{26}H_{54}NO_7P$	$[M+H^+]^+$	-0.05	LysoPC(18:0)	Lysophosphatidylcholines
6	↑	0.094	0.027	40	544.33982	$C_{28}H_{50}NO_7P$	$[M+H^+]^+$	0.1	LysoPC(20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylcholines
7	↑	0.069	NS	40	518.3216	$C_{24}H_{50}NO_7P$	$[M+Na^+]^+$	-0.21	LysoPC(16:0)	Lysophosphatidylcholines
8	↑	0.056	NS	14	474.92365	Unknown	-	-	unknown	Unknown
9	↑	0.041	0.031	40	338.34173	Unknown	-	-	unknown	Unknown
10	↑	0.04	NS	13	475.25834	$C_{25}H_{41}O_5P$	$[M+H^+]^+$	-0.1	unknown	Unknown
11	↑	0.038	NS	40	545.42338	$C_{31}H_{60}O_5S$	$[M+H^+]^+$	-0.08	unknown	Unknown
12	↑	0.036	NS	39	542.32161	$C_{26}H_{50}NO_7P$	$[M+Na^+]^+$	-0.18	LysoPC(18:2)	Lysophosphatidylcholines
13	↑	0.03	NS	30	496.33033	$C_{24}H_{49}NO_7S$	$[M+H^+]^+$	0.15	unknown	unknown
14	↑	0.025	NS	18	520.33268	unknown	-	-	unknown	unknown
15	↑	0.024	NS	9	711.88213	unknown	-	-	unknown	unknown
16	↑	0.023	NS	40	546.35547	$C_{28}H_{52}NO_7P$	$[M+H^+]^+$	0.1	LysoPC(20:3(5Z,8Z,11Z))	Lysophosphatidylcholines
17	↑	0.022	NS	7	470.58625	unknown	-	-	unknown	unknown
18	↑	0.019	NS	40	494.32417	$C_{29}H_{45}NO_4$	$[M+Na^+]^+$	0.17	Cervonyl carnitine	Long chain acyl carnitines
19	↑	0.019	NS	40	494.32417	$C_{24}H_{48}NO_7P$	$[M+H^+]^+$	0.1	LysoPC(16:1(9Z))	Lysophosphatidylcholines
20	↑	0.018	NS	7	712.38372	$C_{35}H_{57}N_3O_{10}S$	$[M+H^+]^+$	-0.03	unknown	unknown
21	↑	0.018	0.023	40	568.33986	$C_{30}H_{50}NO_7P$	$[M+H^+]^+$	0.16	LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Lysophosphatidylcholines
22	↑	0.016	0.007	40	498.34694	unknown	-	-	unknown	unknown
23	↑	0.015	NS	7	704.87415	unknown	-	-	unknown	unknown
24	↑	0.015	NS	40	478.29286	$C_{28}H_{41}NO_4P$	$[M+H^+]^+$	0.17	LysoPE(0:0/18:2(9Z,12Z))	Phosphatidylethanolamines

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
25	↑	0.015	NS	40	502.29289	$C_{25}H_{44}NO_7P$	$[M+H^+]^+$	0.15	LysoPE(0:0/20:4(5Z,8Z,11Z,14Z))	Phosphatidylethanolamines
26	↑	0.015	NS	7	475.59268	unknown	-	-	unknown	unknown
27	↑	0.014	NS	31	496.34905	unknown	-	-	unknown	unknown
28	↑	0.013	0.028	40	526.2928	$C_{27}H_{44}NO_7P$	$[M+H^+]^+$	-0.04	LysoPE(0:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylethanolamines
29	↑	0.012	NS	39	546.4268	unknown	-	-	unknown	unknown
30	↑	0.011	NS	25	524.36428	unknown	-	-	unknown	unknown
31	↑	0.011	NS	4	705.37566	unknown	-	-	unknown	unknown
32	↑	0.011	NS	40	286.12916	unknown	-	-	unknown	unknown
33	↑	0.011	NS	40	391.28431	$C_{24}H_{38}O_4$	$[M+H^+]^+$	0.05	12-Ketodeoxycholic acid	BA
34	↑	0.011	NS	40	566.32174	$C_{28}H_{50}NO_7P$	$[M+Na^+]^+$	0.06	LysoPC(20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylcholines
35	↑	0.01	0.03	40	298.04205	unknown	-	-	unknown	unknown
36	↑	0.01	NS	40	259.255	unknown	-	-	unknown	unknown
37	↑	0.01	0.031	40	339.3451	unknown	-	-	unknown	unknown
38	↑	0.009	0.009	40	510.35547	$C_{25}H_{52}NO_7P$	$[M+H^+]^+$	0.1	LysoPE(0:0/20:0)	Phosphatidylethanolamines
39	↑	0.009	0.017	40	534.29579	unknown	-	-	unknown	unknown
40	↑	0.009	NS	40	468.30852	$C_{27}H_{43}NO_4$	$[M+Na^+]^+$	-	1alpha,25-dihydroxy-24-oxo-23-azavitamin D2 / 1alpha,25-dihydroxy-24-oxo-23-azaergocalciferol	Vitamin D2 derivate
41	↑	0.009	NS	40	468.30852	$C_{22}H_{46}NO_7P$	$[M+H^+]^+$	0.11	LysoPC(14:0)	Lysophosphatidylcholines
42	↑	0.008	NS	5	470.92088	unknown	-	-	unknown	unknown
43	↑	0.007	0.025	29	497.35225	unknown	-	-	unknown	unknown
44	↑	0.007	NS	29	497.3337	unknown	-	-	unknown	unknown
45	↑	0.007	NS	40	288.28971	$C_{17}H_{37}NO_2$	$[M+H^+]^+$	-	C17 Sphinganine	Sphingolipids
46	↑	0.007	NS	40	369.3516	$C_{27}H_{44}$	$[M+Na^+]^+$	-	3-Deoxyvitamin D3	Vitamin D3
47	↑	0.007	0.015	40	570.35546	$C_{30}H_{52}NO_7P$	$[M+H^+]^+$	0.08	LysoPC(22:5(4Z,7Z,10Z,13Z,16Z))	Lysophosphatidylcholines
48	↑	0.007	0.033	40	454.29287	$C_{21}H_{44}NO_7P$	$[M+H^+]^+$	0.11	LysoPE(0:0/16:0)	Phosphatidylethanolamines

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
49	↑	0.007	NS	4	465.57989	unknown	-	-	unknown	Unknown
50	↑	0.007	0.009	40	363.66052	unknown	-	-	unknown	Unknown
51	↑	0.007	NS	3	712.88557	unknown	-	-	unknown	Unknown
52	↑	0.007	NS	40	703.57509	$C_{39}H_{79}N_2O_6P$	$[M+H^+]^+$	0.34	SM(d18:0/16:1(9Z))	SM
53	↑	0.006	0.039	33	496.35549	unknown	-	-	unknown	Unknown
54	↑	0.006	0.009	16	638.60832	$C_{40}H_{79}NO_4$	$[M+H^+]^+$	0.21	unknown	Unknown
55	↑	0.005	0.013	40	360.3237	$C_{22}H_{43}NO$	$[M+H^+]^+$	0.03	unknown	Unknown
56	↑	0.005	NS	40	482.32416	$C_{23}H_{48}NO_7P$	$[M+H^+]^+$	0.09	LysoPC(15:0)	Lysophosphatidylcholines
57	↑	0.005	NS	40	282.27913	$C_{18}H_{35}NO$	$[M+H^+]^+$	-0.05	Oleamide	Lipoamide
58	↑	0.005	0.024	17	400.37856	$C_{24}H_{49}NO_3$	$[M+H^+]^+$	0.11	unknown	Unknown
59	↑	0.005	0.001	40	267.647	unknown	-	-	unknown	Unknown
60	↑	0.005	NS	37	430.38906	$C_{25}H_{51}NO_4$	$[M+H^+]^+$	-0.05	unknown	Unknown
61	↑	0.005	NS	4	465.91444	unknown	-	-	unknown	Unknown
62	↑	0.005	0.01	30	497.35559	unknown	-	-	unknown	Unknown
63	↑	0.005	NS	40	480.3085	$C_{23}H_{46}NO_7P$	$[M+H^+]^+$	0.07	LysoPE(0:0/18:1(11Z))	Phosphatidylethanolamines
64	↑	0.005	NS	25	496.31771	unknown	-	-	unknown	Unknown
65	↑	0.005	NS	40	413.26625	$C_{24}H_{38}O_4$	$[M+Na^+]^+$	0.05	12-Ketodeoxycholic acid	BA
66	↑	0.004	NS	37	402.35778	$C_{23}H_{47}NO_4$	$[M+H^+]^+$	0	unknown	Unknown
67	↑	0.004	0.04	39	520.32231	unknown	-	-	unknown	Unknown
68	↑	0.004	NS	40	263.089	$C_{12}H_{16}O_5$	$[M+Na^+]^+$	0.02	unknown	Unknown
69	↑	0.004	NS	23	520.34635	unknown	-	-	unknown	Unknown
70	↑	0.004	0.001	40	279.647	unknown	-	-	unknown	Unknown
71	↑	0.004	NS	33	496.3588	unknown	-	-	unknown	Unknown
72	↑	0.004	NS	40	304.26108	$C_{18}H_{35}NO$	$[M+Na^+]^+$	-0.03	Oleamide	Lipoamide
73	↑	0.004	0.03	40	478.32925	$C_{24}H_{48}NO_6P$	$[M+Na^+]^+$	0.1	LysoPC(16:2)	Lysophosphatidylcholines
74	↑	0.004	0.03	40	478.32925	$C_{24}H_{48}NO_6P$	$[M+Na^+]^+$	0.1	unknown	Unknown

Enty	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
75	↑	0.004	NS	3	705.87781	unknown	-	-	unknown	Unknown
76	↑	0.004	0.013	40	524.35722	$C_{28}H_{47}N_5O_3$	$[M+Na^+]^+$	0.21	unknown	Unknown
77	↑	0.004	NS	3	474.92937	unknown	-	-	unknown	Unknown
78	↑	0.004	NS	19	521.34968	unknown	-	-	unknown	Unknown
79	↑	0.004	0.012	40	526.37804	unknown	-	-	unknown	Unknown
80	↑	0.004	NS	40	760.58583	$C_{42}H_{82}NO_8P$	$[M+H^+]^+$	0.98	PC(14:0/20:1(11Z))	Phosphatidylcholines
81	↑	0.004	NS	28	496.36222	unknown	-	-	unknown	Unknown
82	↑	0.004	NS	3	697.86648	unknown	-	-	unknown	Unknown
83	↑	0.004	NS	3	475.92728	unknown	-	-	unknown	Unknown
84	↑	0.004	NS	3	474.91789	unknown	-	-	unknown	Unknown
85	↑	0.004	NS	40	516.30613	$C_{24}H_{48}NO_7P$	$[M+Na^+]^+$	0.14	LysoPC(16:1(9Z))	Lysophosphatidylcholine
86	↑	0.004	NS	40	244.26348	$C_{15}H_{33}NO$	$[M+H^+]^+$	-0.03	unknown	Unknown
87	↑	0.004	0.025	40	375.66051	unknown	-	-	unknown	Unknown
88	↑	0.004	NS	40	558.29578	$C_{25}H_{43}N_5O_7S$	$[M+H^+]^+$	0.32	unknown	Unknown
89	↑	0.003	0.014	24	372.34722	$C_{22}H_{45}NO_3$	$[M+H^+]^+$	0.01	unknown	Unknown
90	↑	0.003	0.00007	36	527.31289	$C_{33}H_{44}O_4$	$[M+Na^+]^+$	-	Eudesobovatol A	Exogenous
91	↑	0.003	0.031	12	666.63948	$C_{42}H_{83}NO_4$	$[M+H^+]^+$	-0.01	unknown	Unknown
92	↑	0.003	0.015	32	338.33884	unknown	-	-	unknown	Unknown
93	↑	0.003	NS	40	316.321	$C_{19}H_{41}NO_2$	$[M+H^+]^+$	-0.02	unknown	Unknown
94	↑	0.003	0.0003	39	515.31327	$C_32H_{44}O_4$	$[M+H^+]^+$	0.18	unknown	Unknown
95	↑	0.003	0.027	10	610.57684	$C_{38}H_{75}NO_4$	$[M+H^+]^+$	-0.08	Tetracosahexaenoic acid	Fatty acids

Table 20: Significant biomarkers DOWN-regulated respect to the insulin sensitive group at base line. Experimental analyses via direct infusion ICR-FT/MS. The p-val indicated derive from Wilcoxon test.

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
1	↓	-0.143	0.033	40	215.01619	C_6H_8O7	$[M+Na^+]^+$	-0.16	Citric acid	Carboxylic acid
2	↓	-0.098	NS	21	214.09071	unknown	-	-	Unknown	unknown
3	↓	-0.076	0.03	40	332.9643	unknown	-	-	Unknown	unknown
4	↓	-0.074	0.045	40	215.09295	unknown	-	-	Unknown	unknown
5	↓	-0.068	NS	40	236.07154	$C_{10}H_{15}NO_2S$	$[M+Na^+]^+$	-0.12	Unknown	unknown
6	↓	-0.057	0.013	40	407.04325	$C_{12}H_{16}O_{14}$	$[M+Na^+]^+$	0.04	Unknown	unknown
7	↓	-0.04	NS	25	214.08841	unknown	$[M+Na^+]^+$	-	Unknown	unknown
8	↓	-0.034	NS	40	515.41285	$C_{30}H_{58}O_4S$	$[M+H^+]^+$	-0.01	Unknown	unknown
9	↓	-0.028	NS	40	216.0854	unknown	-	-	Unknown	unknown
10	↓	-0.025	NS	40	217.13689	$C_{10}H_{20}N_2OS$	$[M+H^+]^+$	-0.1	Unknown	unknown
11	↓	-0.024	NS	40	187.12634	$C_9H_{18}N_2S$	$[M+H^+]^+$	-	Unknown	unknown
12	↓	-0.019	NS	25	214.08781	unknown	-	-	Unknown	unknown
13	↓	-0.018	0.04	40	429.0252	unknown	-	-	Unknown	unknown
14	↓	-0.017	NS	25	214.08722	$C_8H_{17}NO_2S$	$[M+Na^+]^+$	-	Unknown	unknown
15	↓	-0.016	NS	40	188.07061	$C_{11}H_9NO_2$	$[M+H^+]^+$	0.21	Indoleacrylic acid	Amino acid
16	↓	-0.014	0.028	39	334.96012	unknown	-	-	Unknown	unknown
17	↓	-0.014	NS	40	198.18523	$C_{12}H_{23}NO$	$[M+H^+]^+$	-0.03	Unknown	unknown
18	↓	-0.014	NS	40	409.16218	$C_{22}H_{26}O_6$	$[M+Na^+]^+$	0.06	Burseran	Exogenous
19	↓	-0.014	NS	40	680.49207	unknown	-	-	Unknown	unknown
20	↓	-0.014	NS	40	205.09715	$C_{11}H_{12}N_2O_2$	$[M+H^+]^+$	-0.02	L-Tryptophan	Amino acid
21	↓	-0.014	NS	26	214.0913	unknown	-	-	Unknown	unknown
22	↓	-0.013	0.033	40	285.29003	$C_{17}H_{36}N_2O$	$[M+H^+]^+$	-0.01	Unknown	unknown
23	↓	-0.012	NS	40	719.503	$C_{41}H_{70}N_2O_6S$	$[M+H^+]^+$	0.37	Unknown	unknown

Enty	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
24	↓	-0.012	NS	40	646.50758	unknown	-	-	unknown	unknown
25	↓	-0.012	NS	40	241.17328	$C_{13}H_{24}N_2S$	$[M+H^+]^+$	-0.06	Unknown	unknown
26	↓	-0.011	NS	40	516.41624	unknown	-	-	Unknown	unknown
27	↓	-0.011	0.033	39	333.96766	unknown	-	-	Unknown	unknown
28	↓	-0.01	NS	21	215.01733	unknown	-	-	Unknown	unknown
29	↓	-0.01	NS	40	225.19611	$C_{13}H_{24}N_2O$	$[M+H^+]^+$	-0.12	Anapheline	Exogenous
30	↓	-0.009	NS	40	214.09187	unknown	-	-	Unknown	unknown
31	↓	-0.008	0.013	40	408.04661	unknown	-	-	Unknown	unknown
32	↓	-0.008	NS	32	419.31567	$C_{26}H_{42}O_4$	$[M+H^+]^+$	0.19	(24R)-1alpha,24-dihydroxy-22-oxavitamin D3	Vitamin D3 derivate
33	↓	-0.008	NS	40	518.27488	$C_{28}H_{39}NO_8$	$[M+H^+]^+$	0.06	Unknown	unknown
34	↓	-0.008	NS	39	429.24016	$C_{27}H_{34}O_3$	$[M+Na^+]^+$	-0.33	Unknown	unknown
35	↓	-0.008	NS	40	280.26349	$C_{18}H_{33}NO$	$[M+H^+]^+$	0.003	Unknown	unknown
36	↓	-0.008	NS	40	198.00418	unknown	-	-	Unknown	unknown
37	↓	-0.007	NS	40	230.99014	$C_9H_4O_6$	$[M+Na^+]^+$	-0.5	Unknown	Amino acid
38	↓	-0.007	NS	40	758.56973	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.4	PC(14:0/20:2(11Z, 14Z))	Phosphatidylcholines
39	↓	-0.007	NS	40	196.16959	$C_{12}H_{21}NO$	$[M+H^+]^+$	0.005	Unknown	unknown
40	↓	-0.007	NS	40	158.02703	$C_6H_7NO_2S$	$[M+H^+]^+$	0.03	Unknown	unknown
41	↓	-0.007	NS	40	237.07491	unknown	-	-	Unknown	unknown
42	↓	-0.007	NS	40	166.08626	$C_9H_{11}NO_2$	$[M+H^+]^+$	0.04	L-PhenylAlanine	Amino acid
43	↓	-0.006	NS	40	171.14919	$C_9H_{18}N_2O$	$[M+H^+]^+$	0	Unknown	unknown
44	↓	-0.006	NS	40	614.51775	$C_{36}H_{71}NO_4S$	$[M+H^+]^+$	0.15	Unknown	unknown
45	↓	-0.006	NS	33	214.08655	unknown	-	-	Unknown	unknown
46	↓	-0.006	NS	40	681.4954	unknown	-	-	Unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
47	↓	-0.006	NS	40	309.20362	$C_{16}H_{30}O_4$	$[M+H^+]^+$	-0.04	Hexadecanedioic acid	Fatty acid
48	↓	-0.006	NS	40	486.28508	$C_{28}H_{39}NO_6$	$[M+H^+]^+$	-0.13	Unknown	unknown
49	↓	-0.006	NS	40	450.32148	$C_{26}H_{43}NO_5$	$[M+H^+]^+$	0.17	Deoxycholic acid glycine conjugate	Bile acid
50	↓	-0.005	0.016	40	203.12126	$C_9H_{18}N_2OS$	$[M+H^+]^+$	-0.02	Unknown	unknown
51	↓	-0.005	NS	30	215.01501	unknown	-	-	Unknown	unknown
52	↓	-0.005	NS	39	352.09703	unknown	-	-	Unknown	unknown
53	↓	-0.005	NS	31	214.09305	unknown	-	-	Unknown	unknown
54	↓	-0.005	NS	40	1051.8011	unknown	-	-	Unknown	unknown
55	↓	-0.005	NS	40	720.50644	unknown	-	-	Unknown	unknown
56	↓	-0.005	NS	40	449.15396	unknown	-	-	Unknown	unknown
57	↓	-0.005	0.005	40	397.12418	$C_{17}H_{20}N_2O_9$	$[M+Na^+]^+$	0.04	Unknown	unknown
58	↓	-0.004	NS	40	229.03186	$C_7H_{10}O_7$	$[M+Na^+]^+$	-0.06	2-Methylcitric acid	Carboxylic acid
59	↓	-0.004	NS	40	228.23218	$C_{14}H_{29}NO$	$[M+H^+]^+$	-0.06	Unknown	unknown
60	↓	-0.004	0.023	40	621.05227	unknown	-	-	Unknown	unknown
61	↓	-0.004	0.048	40	181.07201	$C_7H_8N_4O_2$	$[M+H^+]^+$	0.02	Paraxanthine	Exogenous
62	↓	-0.004	NS	29	215.09415	unknown	-	-	Unknown	unknown
63	↓	-0.004	NS	40	254.24784	$C_{16}H_{31}NO$	$[M+H^+]^+$	0.003	Unknown	unknown
64	↓	-0.004	NS	40	286.20129	$C_{15}H_{27}NO_4$	$[M+H^+]^+$	0.01	2-Octenoylcarnitine	Carnitine
65	↓	-0.004	0.011	38	214.08483	unknown	-	-	Unknown	unknown
66	↓	-0.004	0.14	14	332.96707	unknown	-	-	Unknown	unknown
67	↓	-0.004	NS	4	760.57536	unknown	-	-	Unknown	unknown
68	↓	-0.004	NS	40	271.18386	$C_{14}H_{26}N_2OS$	$[M+H^+]^+$	0.01	Unknown	unknown
69	↓	-0.004	NS	40	305.24509	$C_{18}H_{34}O_2$	$[M+Na^+]^+$	-0.02	Oleic acid	Fatty acid
70	↓	-0.004	NS	9	332.96158	unknown	-	-	Unknown	unknown
71	↓	-0.003	NS	40	357.27882	$C_{24}H_{36}O_2$	$[M+H^+]^+$	0.05	Unknown	unknown
72	↓	-0.003	NS	40	216.09631	unknown	-	-	Unknown	unknown

Enty	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
73	↓	-0.003	NS	22	215.01793	unknown	-	-	Unknown	unknown
74	↓	-0.003	NS	12	407.03932	unknown	-	-	Unknown	unknown
75	↓	-0.003	0.033	40	214.09367	unknown	-	-	Unknown	unknown
76	↓	-0.003	0.079	40	182.08118	$C_9H_{11}NO_3$	$[M+Na^+]^+$	0.03	L-Tyrosine	Amino acid
77	↓	-0.003	NS	40	1052.8044	unknown	-	-	Unknown	unknown

Table 21: Significant biomarkers UP-regulated respect to the insulin sensitive group at base line. Experimental analyses via UPLC-MS-Q-TOF in positive mode ionization. The possible isomers and isobars masses annotated are grouped together. The same entry number indicates that they belong to the same isomers/isobar assignment group.

Entry	Insulin sensitive (IS)	Regression coefficient	Freq	LC-m/z	FT-m/z	RT	logD	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compounds class
1	↑	0.511	18	496.3398	496.33966	17.6444	3.18	$C_{24}H_{50}NO_7P$	$[M+H^+]^+$	-0.21	LysoPC(16:0)	Lysophosphatidylcholines
1	↑	0.511	18	496.3398	496.33966	17.6444	2.99	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	-0.11	Docosa-4,7,10,13,16-pentaenoyl carnitine	Long chain acyl carnitines
1	↑	0.511	18	496.3398	496.33966	17.6444	2.99	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	-0.11	Clupanodonyl carnitine	Long chain acyl carnitines
2	↑	0.395	15	522.3559	522.35532	17.9925	3.71	$C_{26}H_{52}NO_7P$	$[M+H^+]^+$	-0.19	LysoPC(18:1(9Z))	Lysophosphatidylcholines
2	↑	0.395	15	522.3559	522.35532	17.9925	3.71	$C_{26}H_{52}NO_7P$	$[M+H^+]^+$	-0.19	LysoPC(18:1(11Z))	Lysophosphatidylcholines
3	↑	0.314	25	496.3392	496.33966	17.491	3.18	$C_{24}H_{50}NO_7P$	$[M+H^+]^+$	-0.21	LysoPC(16:0)	Lysophosphatidylcholines
3	↑	0.314	25	496.3392	496.33966	17.491	2.99	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	-0.11	Docosa-4,7,10,13,16-pentaenoyl carnitine	Long chain acyl carnitines
3	↑	0.314	25	496.3392	496.33966	17.491	2.99	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	-0.11	Clupanodonyl carnitine	Long chain acyl carnitines
4	↑	0.282	19	520.3399	520.33968	16.1693	3.34	$C_{26}H_{50}NO_7P$	$[M+H^+]^+$	-0.22	LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines
5	↑	0.168	22	544.3403	544.33982	16.0197	3.51	$C_{28}H_{50}NO_7P$	$[M+H^+]^+$	0.05	LysoPC(20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylcholines
5	↑	0.168	22	544.3403	544.33982	16.0197	3.51	$C_{28}H_{50}NO_7P$	$[M+H^+]^+$	0.05	LysoPC(20:4(8Z,11Z,14Z,17Z))	Lysophosphatidylcholines
6	↑	0.09	27	494.3233	494.32417	15.4048	2.81	$C_{24}H_{48}NO_7P$	$[M+H^+]^+$	0.11	LysoPC(16:1(9Z))	Lysophosphatidylcholines
6	↑	0.09	27	494.3233	494.32417	15.4048	2.62	$C_{29}H_{45}NO_4$	$[M+Na^+]^+$	0.15	Cervonyl carnitine	Long chain acyl carnitines
7	↑	0.087	9	496.34	496.33966	16.9194	3.18	$C_{24}H_{50}NO_7P$	$[M+H^+]^+$	-0.21	LysoPC(16:0)	Lysophosphatidylcholines
7	↑	0.087	9	496.34	496.33966	16.9194	2.99	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	-0.11	Docosa-4,7,10,13,16-pentaenoyl carnitine	Long chain acyl carnitines
7	↑	0.087	9	496.34	496.33966	16.9194	2.99	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	-0.11	Clupanodonyl carnitine	Long chain acyl carnitines
8	↑	0.063	26	758.5724	758.56973	21.2847	11.93	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PE(15:0/22:2(13Z,16Z))	Phosphatidylethanolamines
8	↑	0.063	26	758.5724	758.56973	21.2847	11.93	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PE(22:2(13Z,16Z)/15:0)	Phosphatidylethanolamines
10	↑	0.063	26	758.5724	758.56973	21.2847	11.71	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PE-NMe(18:1(9Z)/18:1(9Z))	Phosphatidylethanolamines
11	↑	0.062	31	281.1171	281.11722	13.0025	3.04	$C_{18}H_{18}O_4$	$[M+H^+-H_2O]^+$	-0.02	7C-aglycone	Aromatic keto-acids

Entry	Insulin sensitive (IS)	Regression coefficient	Freq	LC-m/z	FT-m/z	RT	logD	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compounds class
12	↑	0.062	31	281.1171	281.11722	13.0025	3.61	$C_{18}H_{18}O_4$	$[M+H^+-H_2O]^+$	-0.02	Enterolactone	Lignans
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(14:0/20:2(11Z,14Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(14:1(9Z)/20:1(11Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(16:0/18:2(9Z,12Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(16:1(9Z)/18:1(11Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(16:1(9Z)/18:1(9Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(18:1(11Z)/16:1(9Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(18:1(9Z)/16:1(9Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(18:2(9Z,12Z)/16:0)	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(20:1(11Z)/14:1(9Z))	Phosphatidylcholines
13	↑	0.062	12	758.5712	758.56973	17.3548	10.27	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PC(20:2(11Z,14Z)/14:0)	Phosphatidylcholines
14	↑	0.045	11	758.572	758.56973	19.9776	11.92	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PE(15:0/22:2(13Z,16Z))	Phosphatidylethanolamines
14	↑	0.045	11	758.572	758.56973	19.9776	11.93	$C_42H_{80}NO_8P$	$[M+H^+]^+$	0.42	PE(22:2(13Z,16Z)/15:0)	Phosphatidylethanolamines
14	↑	0.045	11	758.572	758.56973	19.9776	11.7	$C_{42}H_{80}NO_8P$	$[M+H^+]^+$	0.42	PE-NMe(18:1(9Z)/18:1(9Z))	Phosphatidylethanolamines

Table 22: Significant biomarkers DOWN-regulated respect to the insulin sensitive group at base line. Experimental analyses via UPLC-MS-Q-TOF in positve mode ionization. The possible isomers and isobars masses annotated are grouped together. The same entry number indicates that they belong to the same isomers/isobar assignment group.

Entry	Insulin sensitive (IS)	Regression coefficient	Freq	LC_m/z	FT_m/z	RT	logD	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compounds class
1	↓	-0.139	36	282.2785	282.27913	20.1252	5.98	$C_{18}H_{35}NO$	$[M+H^+]^+$	-0.08	Oleamide	Lipoamide
1	↓	-0.139	36	282.2785	282.27913	20.1252	4.98	$C_{18}H_{37}NO_2$	$[M+H^+-H_2O]^+$	-0.08	Palmitoylethanolamide	Aliphatic amides
2	↓	-0.101	28	195.0859	195.08765	4.7502	-0.55	$C_8H_{10}N_4O_2$	$[M+H^+]^+$	-0.04	Caffeine	Exogenous
3	↓	-0.096	37	524.3709	524.37104	15.5215	4.07	$C_{26}H_{54}NO_7P$	$[M+H^+]^+$	-0.01	LysoPC(18:0)	Lysophosphatidylcholines
4	↓	-0.084	5	181.0709	181.07201	3.381	-2.29	$C_7H_{10}N_4O_3$	$[M+H^+-H_2O]^+$	0.04	5-Acetylamino-6-amino-3-methyluracil	Methyl pyrimidines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_46H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(16:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(16:1(9Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(16:1(9Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:1(11Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:1(9Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:2(9Z,12Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:2(9Z,12Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:3(6Z,9Z,12Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:3(6Z,9Z,12Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:3(9Z,12Z,15Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:3(9Z,12Z,15Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(18:4(6Z,9Z,12Z,15Z)/20:2(11Z,14Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:2(11Z,14Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:3(5Z,8Z,11Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:3(5Z,8Z,11Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:3(8Z,11Z,14Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	Freq	LC_m/z	FT_m/z	RT	logD	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compounds class
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:4(5Z,8Z,11Z,14Z)/18:2(9Z,12Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:4(8Z,11Z,14Z,17Z)/18:2(9Z,12Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(11Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(9Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(22:5(4Z,7Z,10Z,13Z,16Z)/16:1(9Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(22:5(7Z,10Z,13Z,16Z,19Z)/16:1(9Z))	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:0)	Phosphatidylcholines
5	↓	-0.084	24	806.5694	806.56966	21.2767	10.61	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.31	PC(20:3(8Z,11Z,14Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
6	↓	-0.074	38	215.0157	215.01619	1.2043	-1.47	$C_6H_8O_7$	$[M+Na^+]^+$	-0.18	Citric acid	Short chain tricarboxylic acids
6	↓	-0.074	38	215.0157	215.01619	1.2043	-1.59	$C_6H_8O_7$	$[M+Na^+]^+$	-0.18	Isocitric acid	Short chain hydroxy acids
6	↓	-0.074	38	215.0157	215.01619	1.2043	-1.59	$C_6H_8O_7$	$[M+Na^+]^+$	-0.18	D-threo-Isocitric acid	Short chain tricarboxylic acids
6	↓	-0.074	38	215.0157	215.01619	1.2043	-2.04	$C_6H_8O_7$	$[M+Na^+]^+$	-0.18	Diketogulonic acid	Short chain hydroxy acids
6	↓	-0.074	38	215.0157	215.01619	1.2043	-2.04	$C_6H_8O_7$	$[M+Na^+]^+$	-0.18	2,3-Diketo-L-gulonate	Simple alcohols
7	↓	-0.071	24	181.0708	181.07201	3.9468	-0.77	$C_7H_8N_4O_2$	$[M+H^+]^+$	0.01	Theobromine	Exogenous
7	↓	-0.071	24	181.0708	181.07201	3.9468	0.24	$C_7H_8N_4O_2$	$[M+H^+]^+$	0.01	Theophylline	Exogenous
7	↓	-0.071	24	181.0708	181.07201	3.9468	0.24	$C_7H_8N_4O_2$	$[M+H^+]^+$	0.01	Paraxanthine	Exogenous
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(16:0/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(16:0/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(16:1(9Z)/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:0/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:1(11Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:1(11Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:1(9Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:1(9Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:2(9Z,12Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	Freq	LC_m/z	FT_m/z	RT	logD	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compounds class
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:2(9Z,12Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:3(6Z,9Z,12Z)/20:2(11Z,14Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:3(9Z,12Z,15Z)/20:2(11Z,14Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(18:4(6Z,9Z,12Z,15Z)/20:1(11Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:1(11Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:2(11Z,14Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:2(11Z,14Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:3(5Z,8Z,11Z)/18:2(9Z,12Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:3(8Z,11Z,14Z)/18:2(9Z,12Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:4(5Z,8Z,11Z,14Z)/18:1(11Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:4(5Z,8Z,11Z,14Z)/18:1(9Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:4(8Z,11Z,14Z,17Z)/18:1(11Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:4(8Z,11Z,14Z,17Z)/18:1(9Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:0)	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(22:4(7Z,10Z,13Z,16Z)/16:1(9Z))	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(22:5(4Z,7Z,10Z,13Z,16Z)/16:0)	Phosphatidylcholines
8	↓	-0.054	17	808.5882	808.58581	21.2871	10.97	$C_{46}H_{82}NO_8P$	$[M+H^+]^+$	0.89	PC(22:5(7Z,10Z,13Z,16Z,19Z)/16:0)	Phosphatidylcholines
9	↓	-0.052	39	171.1473	171.14919	7.7579	-6.46	$C_9H_{20}N_2O_2$	$[M+H^+-H_2O]^+$	0.02	N6,N6,N6-Trimethyl-L-lysine	Methylated aminoacids

Table 23: Significant biomarkers UP-regulated respect to the insulin sensitive group at follow-up. Experimental analyses via direct infusion with ICR-FT/MS in positve mode ionization. The possible isomers and isobars masses annotated, are grouped together. The same entry number indicates that they belong to the same isomers/isobar assignment group.

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
1	↑	-0.536	0.023	39	518.3219	$C_{24}H_{50}NO_7P$	$[M+Na^+]^+$	0.31	LysoPC(16:0)	Lysophosphatidylcholines
1	↑	-0.416	NS	39	542.3219	$C_{26}H_{50}NO_7P$	$[M+Na^+]^+$	0.29	LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines
1	↑	-0.254	NS	39	409.1621	$C_{17}H_{29}O_9P$	$[M+H^+]^+$	-0.25	unknown	unknown
2	↑	-0.157	0.023	39	544.3378	$C_{26}H_{52}NO_7P$	$[M+Na^+]^+$	0.71	LysoPC(18:1(9Z))	Lysophosphatidylcholines
2	↑	-0.157	0.023	39	544.3378	$C_{26}H_{52}NO_7P$	$[M+Na^+]^+$	0.71	LysoPC(18:1(11Z))	Lysophosphatidylcholines
3	↑	-0.156	0.043	39	566.3222	$C_{28}H_{50}NO_7P$	$[M+Na^+]^+$	0.75	LysoPC(20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylcholines
3	↑	-0.156	0.043	39	566.3222	$C_{28}H_{50}NO_7P$	$[M+Na^+]^+$	0.75	LysoPC(20:4(8Z,11Z,14Z,17Z))	Lysophosphatidylcholines
4	↑	-0.14	0.028	39	519.3254	unknown	-	-	unknown	unknown
5	↑	-0.124	0.04	39	496.3398	$C_{24}H_{50}NO_7P$	$[M+H^+]^+$	0.06	LysoPC(16:0)	Lysophosphatidylcholines
5	↑	-0.124	0.04	39	496.3398	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	0.16	Docosa-4,7,10,13,16-pentaenoyl carnitine	carnitines
5	↑	-0.124	0.04	39	496.3398	$C_{29}H_{47}NO_4$	$[M+Na^+]^+$	0.16	Clupanodonyl carnitine	carnitines
6	↑	-0.113	NS	39	543.3254	unknown	-	-	unknown	unknown
7	↑	-0.091	0.025	39	520.34	$C_{26}H_{50}NO_7P$	$[M+H^+]^+$	0.39	LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines
8	↑	-0.077	NS	39	387.1805	$C_{19}H_{29}N_2O_3P$	$[M+Na^+]^+$	-0.89	unknown	unknown
9	↑	-0.059	NS	39	795.337	$C_{35}H_{57}N_4O_{11}SP$	$[M+Na^+]^+$	-0.55	unknown	unknown
10	↑	-0.059	NS	39	609.2813	unknown	-	-	unknown	unknown
11	↑	-0.053	NS	39	410.1656	unknown	-	-	unknown	unknown
12	↑	-0.043	0.016	39	516.3064	$C_{24}H_{48}NO_7P$	$[M+Na^+]^+$	0.65	LysoPC(16:1(9Z))	Lysophosphatidylcholines
13	↑	-0.042	0.043	39	567.3256	unknown	-	-	unknown	unknown
14	↑	-0.038	0.028	39	545.3412	unknown	-	-	unknown	unknown
15	↑	-0.034	NS	39	590.3225	$C_{34}H_{43}N_3O_6$	$[M+H^+]^+$	0.02	unknown	unknown
16	↑	-0.033	0.031	39	522.356	unknown	-	-	unknown	unknown
17	↑	-0.031	NS	39	411.0938	unknown	-	-	unknown	unknown
18	↑	-0.031	NS	39	546.3535	$C_{30}H_{47}N_3O_6$	$[M+H^+]^+$	-0.4	unknown	unknown
19	↑	-0.029	NS	39	425.1364	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
20	↑	-0.027	NS	39	796.3408	unknown	-	-	unknown	unknown
21	↑	-0.027	NS	39	904.4979	$C_{43}H_{80}NO_{13}SP$	$[M+Na^+]^+$	-0.18	unknown	unknown
22	↑	-0.026	0.04	39	497.3436	unknown	-	-	unknown	unknown
23	↑	-0.024	NS	39	490.2902	$C_{22}H_{46}NO_7P$	$[M+Na^+]^+$	0.42	LysoPC(14:0)	Lysophosphatidylcholines
24	↑	-0.023	0.038	39	564.307	$C_{29}H_{44}N_5O_3P$	$[M+Na^+]^+$	-0.72	unknown	unknown
25	↑	-0.021	0.026	39	518.3151	unknown	-	-	unknown	unknown
26	↑	-0.021	0.031	39	521.3436	unknown	-	-	unknown	unknown
27	↑	-0.019	NS	26	544.3413	$C_{27}H_{49}N_3O_6S$	$[M+H^+]^+$	-0.29	unknown	unknown
28	↑	-0.019	NS	39	610.2847	unknown	-	-	unknown	unknown
29	↑	-0.018	NS	39	337.0751	$C_{11}H_{22}O_6S_2$	$[M+Na^+]^+$	0.24	unknown	unknown
30	↑	-0.017	NS	39	542.3143	unknown	-	-	unknown	unknown
31	↑	-0.016	NS	39	659.2887	$C_{29}H_{48}O_{15}$	$[M+Na^+]^+$	0.19	unknown	unknown
32	↑	-0.016	NS	39	388.1838	unknown	-	-	unknown	unknown
32	↑	-0.016	NS	39	568.338	$C_{32}H_{45}N_3O_6$	$[M+H^+]^+$	-0.26	unknown	unknown
33	↑	-0.015	NS	39	639.2923	$C_{29}H_{48}N_2O_{10}S$	$[M+Na^+]^+$	0.19	unknown	unknown
34	↑	-0.015	NS	37	518.3269	$C_{28}H_{49}NO_4S$	$[M+Na^+]^+$	-0.99	unknown	unknown
35	↑	-0.013	NS	39	905.5012	unknown	-	-	unknown	unknown
36	↑	-0.013	NS	39	771.4886	$C_{40}H_{78}O_6SP_2$	$[M+Na^+]^+$	-0.11	unknown	unknown
37	↑	-0.013	NS	39	413.2666	$C_{24}H_{38}O_4$	$[M+Na^+]^+$	0.77	12-Ketodeoxycholic acid	Bile Acids
37	↑	-0.013	NS	39	413.2666	$C_{24}H_{38}O_4$	$[M+Na^+]^+$	0.77	7-Hydroxy-3-oxocholanoic acid	Bile Acids
37	↑	-0.013	NS	39	413.2666	$C_{24}H_{38}O_4$	$[M+Na^+]^+$	0.77	Nutriacholic acid	Bile Acids
37	↑	-0.013	NS	39	413.2666	$C_{24}H_{38}O_4$	$[M+Na^+]^+$	0.77	7a-Hydroxy-3-oxo-5b-cholanoic acid	Bile Acids
38	↑	-0.013	NS	36	542.3275	$C_{30}H_{49}NO_4S$	$[M+Na^+]^+$	0.02	unknown	unknown
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(16:1(9Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(18:2(9Z,12Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(18:3(6Z,9Z,12Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(18:3(6Z,9Z,12Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(18:3(9Z,12Z,15Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(18:3(9Z,12Z,15Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(18:4(6Z,9Z,12Z,15Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(18:4(6Z,9Z,12Z,15Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(20:3(5Z,8Z,11Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(20:3(8Z,11Z,14Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(20:4(5Z,8Z,11Z,14Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(20:4(5Z,8Z,11Z,14Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(20:4(8Z,11Z,14Z,17Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(20:4(8Z,11Z,14Z,17Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:2(9Z,12Z))	Phosphatidylcholines
39	↑	-0.012	NS	39	804.5538	$C_{46}H_{78}NO_8P$	$[M+H^+]^+$	0.002	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:1(9Z))	Phosphatidylcholines
40	↑	-0.012	0.04	39	534.3171	$C_{28}H_{43}N_3O_7$	$[M+H^+]^+$	-0.57	unknown	unknown
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Melibiose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Cellobiose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	D-Maltose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Alpha-Lactose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Sucrose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Lactulose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Trehalose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{2}2O_{11}$	$[M+Na^+]^+$	0.07	Isomaltose	Disaccharides

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Galactinol	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	3-b-Galactopyranosyl glucose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Epimelibiose	Disaccharides
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{1}1$	$[M+Na^+]^+$	0.07	Turanose	Carbohydrates
41	↑	-0.012	NS	39	365.1055	$C_{12}H_{22}O_{11}$	$[M+Na^+]^+$	0.07	Kojibiose	Carbohydrates
42	↑	-0.011	NS	39	727.4622	$C_{38}H_{64}N_4O_8$	$[M+Na^+]^+$	0.71	unknown	unknown
43	↑	-0.011	NS	39	435.3021	$C_{28}H_{44}O$	$[M+K^+]^+$	0.67	Ergosterol	Vitamin D3 derivatives
43	↑	-0.011	NS	39	435.3021	$C_{28}H_{44}O$	$[M+K^+]^+$	0.67	Ergocalciferol	Steroids and Steroid Derivatives
43	↑	-0.011	NS	39	435.3021	$C_{28}H_{44}O$	$[M+K^+]^+$	0.67	4a-Methyl-5a-cholesta-8,24-dien-3-one	Steroids and Steroid Derivatives
43	↑	-0.011	NS	39	435.3021	$C_{28}H_{44}O$	$[M+K^+]^+$	0.67	3-Keto-4-methylzymosterol	Steroids and Steroid Derivatives
43	↑	-0.011	NS	39	435.3021	$C_{28}H_{44}O$	$[M+K^+]^+$	0.67	5-Dehydroepisterol	Steroids and Steroid Derivatives
44	↑	-0.011	NS	39	520.3292	unknown	-	-	unknown	unknown
45	↑	-0.01	NS	39	518.3106	unknown	-	-	unknown	unknown
46	↑	-0.01	NS	39	928.4978	$C_{45}H_{80}NO_{13}SP$	$[M+Na^+]^+$	-0.27	unknown	unknown
47	↑	-0.01	0.038	39	494.3241	$C_{24}H_{48}NO_7P$	$[M+H^+]^+$	0.07	LysoPC(16:1(9Z))	Lysophosphatidylcholines
47	↑	-0.01	0.038	39	494.3241	$C_{29}H_{45}NO_4$	$[M+Na^+]^+$	0.11	Cervonyl carnitine	Carnitines
48	↑	-0.01	NS	39	404.207	$C_{22}H_{29}NO_6$	$[M+H^+]^+$	0.55	unknown	unknown
49	↑	-0.01	0.043	39	544.3308	$C_{25}H_{52}N_3O_4SP$	$[M+Na^+]^+$	-0.07	unknown	unknown
50	↑	-0.01	NS	24	303.0989	unknown	-	-	unknown	unknown
51	↑	-0.009	NS	16	467.1048	$C_{15}H_{29}N_2O_7S_2P$	$[M+Na^+]^+$	0.38	unknown	unknown
52	↑	-0.009	NS	39	591.326	unknown	-	-	unknown	unknown
53	↑	-0.009	NS	39	409.1584	$C_{17}H_{26}N_2O_8$	$[M+Na^+]^+$	0.67	unknown	unknown
54	↑	-0.009	0.031	39	517.3098	unknown	-	-	unknown	unknown
55	↑	-0.008	NS	39	815.5149	unknown	-	-	unknown	unknown
56	↑	-0.008	0.04	32	545.3448	unknown	-	-	unknown	unknown
57	↑	-0.008	NS	8	795.33	$C_{41}H_{57}O_{10}SP$	$[M+Na^+]^+$	-0.34	unknown	unknown
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(18:3(6Z,9Z,12Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(18:3(9Z,12Z,15Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(18:4(6Z,9Z,12Z,15Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(18:4(6Z,9Z,12Z,15Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(20:4(5Z,8Z,11Z,14Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(20:4(8Z,11Z,14Z,17Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
58	↑	-0.008	NS	39	828.5537	$C_{48}H_{78}NO_8P$	$[M+H^+]^+$	0.16	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
59	↑	-0.008	NS	39	393.0835	$C_{13}H_{24}O_8P_2$	$[M+Na^+]^+$	-0.93	unknown	unknown
60	↑	-0.008	NS	39	647.0779	unknown	-	-	unknown	unknown
61	↑	-0.008	0.025	39	523.3591	unknown	-	-	unknown	unknown
62	↑	-0.008	NS	39	504.3066	$C_{27}H_{41}N_3O_6$	$[M+H^+]^+$	-0.42	unknown	unknown
63	↑	-0.007	NS	39	766.5329	unknown	-	-	unknown	unknown
64	↑	-0.007	NS	39	586.3101	$C_{29}H_{45}N_3O_8$	$[M+Na^+]^+$	0.41	unknown	unknown
65	↑	-0.007	NS	39	412.0936	$C_{18}H_{19}N_3O_5S$	$[M+Na^+]^+$	-0.41	unknown	unknown
66	↑	-0.007	NS	39	722.5072	unknown	-	-	unknown	unknown
67	↑	-0.007	0.028	39	519.318	unknown	-	-	unknown	unknown
68	↑	-0.007	NS	39	566.3133	$C_{29}H_{45}N_5O_3S$	$[M+Na^+]^+$	-0.45	unknown	unknown
69	↑	-0.007	NS	20	519.3288	unknown	-	-	unknown	unknown
70	↑	-0.007	NS	39	542.3096	unknown	-	-	unknown	unknown
71	↑	-0.007	0.023	28	568.3417	$C_{26}H_{52}N_5O_3SP$	$[M+Na^+]^+$	-0.58	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
72	↑	-0.007	NS	39	524.3714	$C_{26}H_{54}NO_7P$	$[M+H^+]^+$	0.74	LysoPC(18:0)	Lysophosphatidylcholines
72	↑	-0.007	NS	39	524.3714	$C_{26}H_{54}NO_7P$	$[M+H^+]^+$	0.74	LysoPC(0:0/18:0)	Lysophosphatidylcholines
73	↑	-0.007	0.049	39	565.3103	unknown	-	-	unknown	unknown
74	↑	-0.007	0.009	39	592.3384	$C_{34}H_{45}N_3O_6$	$[M+H^+]^+$	0.52	unknown	unknown
75	↑	-0.007	0.038	39	472.3038	$C_{24}H_{47}N_3SP_2$	$[M+H^+]^+$	-0.1	unknown	unknown
76	↑	-0.007	NS	39	540.3071	$C_{30}H_{41}N_3O_6$	$[M+H^+]^+$	0.5	unknown	unknown
77	↑	-0.007	NS	39	683.4356	$C_{36}H_{60}N_4O_7$	$[M+Na^+]^+$	0.23	unknown	unknown
78	↑	-0.007	NS	21	543.3292	$C_{30}H_{48}O_7$	$[M+Na^+]^+$	-0.14	unknown	unknown
79	↑	-0.006	NS	39	485.1123	$C_{16}H_{30}O_11S_2$	$[M+Na^+]^+$	0.2	unknown	unknown
80	↑	-0.006	NS	39	547.357	unknown	-	-	unknown	unknown
81	↑	-0.006	NS	39	527.3139	$C_{25}H_{50}N_2O_4P_2$	$[M+Na^+]^+$	0.14	unknown	unknown
82	↑	-0.006	NS	39	524.2752	$C_{25}H_{44}NO_7P$	$[M+Na^+]^+$	0.76	LysoPE(0:0/20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylethanolamines
82	↑	-0.006	NS	39	524.2752	$C_{25}H_{44}NO_7P$	$[M+Na^+]^+$	0.76	LysoPE(0:0/20:4(8Z,11Z,14Z,17Z))	Lysophosphatidylethanolamines
82	↑	-0.006	NS	39	524.2752	$C_{25}H_{44}NO_7P$	$[M+Na^+]^+$	0.76	LysoPE(20:4(5Z,8Z,11Z,14Z)/0:0)	Lysophosphatidylethanolamines
82	↑	-0.006	NS	39	524.2752	$C_{25}H_{44}NO_7P$	$[M+Na^+]^+$	0.76	LysoPE(20:4(8Z,11Z,14Z,17Z)/0:0)	Lysophosphatidylethanolamines
83	↑	-0.006	NS	39	543.3173	unknown	-	-	unknown	unknown
84	↑	-0.006	NS	29	496.3446	unknown	-	-	unknown	unknown
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(14:0/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(16:0/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(16:0/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(16:1(9Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(16:1(9Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:0/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:1(11Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:1(11Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:1(9Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:1(9Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:2(9Z,12Z)/18:2(9Z,12Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:3(6Z,9Z,12Z)/18:1(11Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:3(6Z,9Z,12Z)/18:1(9Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_80NO_8P$	$[M+H^+]^+$	0.99	PC(18:3(9Z,12Z,15Z)/18:1(11Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:3(9Z,12Z,15Z)/18:1(9Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(18:4(6Z,9Z,12Z,15Z)/18:0)	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(20:3(5Z,8Z,11Z)/16:1(9Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(20:3(8Z,11Z,14Z)/16:1(9Z))	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(20:4(5Z,8Z,11Z,14Z)/16:0)	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(20:4(8Z,11Z,14Z,17Z)/16:0)	Phosphatidylcholines
85	↑	-0.006	NS	39	782.5702	$C_{44}H_{80}NO_8P$	$[M+H^+]^+$	0.99	PC(22:4(7Z,10Z,13Z,16Z)/14:0)	Phosphatidylcholines
86	↑	-0.006	NS	39	426.1393	$C_{19}H_{23}NO_{10}$	$[M+H^+]^+$	-0.51	unknown	unknown
87	↑	-0.006	NS	17	544.3434	unknown	-	-	unknown	unknown
88	↑	-0.006	NS	39	496.3333	unknown	-	-	unknown	unknown
89	↑	-0.006	0.004	39	558.3174	$C_{30}H_{43}N_3O_7$	$[M+H^+]^+$	0.11	unknown	unknown
90	↑	-0.006	NS	9	804.5465	unknown	-	-	unknown	unknown
91	↑	-0.005	NS	39	663.4551	unknown	-	-	unknown	unknown
92	↑	-0.005	NS	39	569.3413	unknown	-	-	unknown	unknown
93	↑	-0.005	NS	22	795.3446	$C_{37}H_{61}N_2O_9S_2P$	$[M+Na^+]^+$	-0.27	unknown	unknown
94	↑	-0.005	NS	39	805.5571	unknown	-	-	unknown	unknown
95	↑	-0.005	NS	39	929.5009	unknown	-	-	unknown	unknown
96	↑	-0.005	NS	38	539.314	$C_{24}H_{53}O_7SP$	$[M+Na^+]^+$	-0.42	unknown	unknown
97	↑	-0.005	NS	39	607.266	$C_{28}H_{44}N_2O_9S$	$[M+Na^+]^+$	0.01	unknown	unknown
98	↑	-0.005	0.049	39	558.2965	$C_{25}H_{53}NO_4S_2P_2$	$[M+H^+]^+$	0.17	unknown	unknown
99	↑	-0.005	NS	31	487.2672	$C_{24}H_{44}N_2O_2SP_2$	$[M+H^+]^+$	0.12	unknown	unknown
100	↑	-0.005	NS	39	413.0904	unknown	-	-	unknown	unknown
101	↑	-0.005	NS	39	387.1764	$C_{17}H_{26}N_2O_8$	$[M+H^+]^+$	0.61	unknown	unknown
102	↑	-0.005	NS	15	467.0982	$C_{19}H_{24}O_{10}S$	$[M+Na^+]^+$	-0.07	unknown	unknown
103	↑	-0.005	NS	36	518.3305	$C_{27}H_{50}N_3OSP$	$[M+Na^+]^+$	0.15	unknown	unknown
104	↑	-0.005	NS	29	566.3282	$C_{24}H_{55}N_3O_4SP2$	$[M+Na^+]^+$	0.29	unknown	unknown
105	↑	-0.005	NS	39	625.2767	$C_{23}H_{49}N_2O_{13}SP$	$[M+H^+]^+$	0.19	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
106	↑	-0.005	NS	39	930.6373	unknown	-	-	unknown	unknown
107	↑	-0.005	NS	39	797.3441	$C_{32}H_{59}N_2O_{17}P$	$[M+Na^+]^+$	-0.38	unknown	unknown
108	↑	-0.005	NS	7	796.3336	unknown	-	-	unknown	unknown
109	↑	-0.005	NS	39	660.2921	unknown	-	-	unknown	unknown
110	↑	-0.005	NS	39	548.2755	$C_{28}H_{48}NO_2S_2P$	$[M+Na^+]^+$	-0.26	unknown	unknown
111	↑	-0.005	NS	39	371.101	unknown	-	-	unknown	unknown
112	↑	-0.005	NS	39	491.2945	$C_{24}H_{45}N_4OSP$	$[M+Na^+]^+$	0.16	unknown	unknown
113	↑	-0.005	NS	39	640.2959	unknown	-	-	unknown	unknown
114	↑	-0.005	NS	39	678.4801	unknown	-	-	unknown	unknown
115	↑	-0.005	NS	39	489.2831	unknown	-	-	unknown	unknown
116	↑	-0.004	NS	39	859.5414	unknown	-	-	unknown	unknown
117	↑	-0.004	NS	38	241.1774	$C_{12}H_{26}O_3$	$[M+Na^+]^+$	0.02	unknown	unknown
118	↑	-0.004	NS	39	338.0745	$C_{15}H_{13}N_3O_5$	$[M+Na^+]^+$	-0.71	unknown	unknown
119	↑	-0.004	NS	33	518.334	$C_{27}H_{43}N_5O_5$	$[M+H^+]^+$	0.57	unknown	unknown
120	↑	-0.004	NS	39	544.3251	$C_{27}H_{51}N_3O_2SP_2$	$[M+H^+]^+$	0.25	unknown	unknown
121	↑	-0.004	NS	39	693.2898	$C_{30}H_{48}N_2O_{14}S$	$[M+H^+]^+$	-0.16	unknown	unknown
122	↑	-0.004	NS	39	772.4918	unknown	-	-	unknown	unknown
123	↑	-0.004	0.013	39	1015.676	unknown	-	-	unknown	unknown
124	↑	-0.004	NS	39	468.3086	$C_{22}H_{46}NO_7P$	$[M+H^+]^+$	0.28	LysoPC(14:0)	Lysophosphatidylcholines
125	↑	-0.004	NS	39	728.4654	unknown	-	-	unknown	unknown
126	↑	-0.004	NS	31	546.3571	$C_{27}H_{51}N_3O_6S$	$[M+H^+]^+$	0	unknown	unknown
127	↑	-0.004	NS	39	811.3113	$C_{37}H_{59}O_{12}P_3$	$[M+Na^+]^+$	0.19	unknown	unknown
128	↑	-0.004	NS	39	437.1933	$C_{19}H_{33}O_9P$	$[M+H^+]^+$	-0.36	unknown	unknown
129	↑	-0.004	NS	39	634.4536	unknown	-	-	unknown	unknown
130	↑	-0.004	0.04	39	520.3253	unknown	-	-	unknown	unknown
131	↑	-0.004	NS	39	615.1406	$C_{22}H_{30}O_{20}$	$[M+H^+]^+$	0.5	unknown	unknown
132	↑	-0.004	NS	39	411.1687	unknown	-	-	unknown	unknown
133	↑	-0.004	NS	29	519.3304	unknown	-	-	unknown	unknown
134	↑	-0.004	NS	39	482.2164	unknown	-	-	unknown	unknown
135	↑	-0.004	0.033	39	566.3082	unknown	-	-	unknown	unknown
136	↑	-0.004	0.04	25	520.3351	unknown	-	-	unknown	unknown
137	↑	-0.004	NS	39	425.11	$C_{14}H_{28}O_9P_2$	$[M+Na^+]^+$	-0.22	unknown	unknown
138	↑	-0.004	NS	38	829.5577	unknown	-	0	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
139	↑	-0.004	NS	39	500.2748	$C_{23}H_{44}NO_7P$	$[M+Na^+]^+$	0.12	LysoPE(0:0/18:2(9Z,12Z))	Lysophosphatidylethanolamines
139	↑	-0.004	NS	39	500.2748	$C_{23}H_{44}NO_7P$	[M+Na+]+	0.12	LysoPE(18:2(9Z,12Z)/0:0)	Lysophosphatidylethanolamines
140	↑	-0.004	NS	39	543.1181	unknown	-	-	unknown	unknown
141	↑	-0.004	NS	17	566.3261	$C_{26}H_{50}N_5O_3SP$	$[M+Na^+]^+$	-0.54	unknown	unknown
142	↑	-0.004	0.026	39	518.3323	unknown	-	-	unknown	unknown
143	↑	-0.004	NS	35	542.3314	$C_{22}H_{48}N_5O_8P$	$[M+H^+]^+$	0.06	unknown	unknown
144	↑	-0.004	NS	39	930.5135	unknown	-	-	unknown	unknown
145	↑	-0.004	NS	39	532.3381	$C_{29}H_{45}N_3O_6$	$[M+H^+]^+$	-0.08	unknown	unknown
146	↑	-0.004	NS	39	810.5595	unknown	-	-	unknown	unknown
147	↑	-0.004	NS	39	803.1005	unknown	-	-	unknown	unknown
148	↑	-0.003	NS	39	821.4775	$C_39H_{74}N_4O_8SP_2$	$[M+H^+]^+$	-0.05	unknown	unknown
149	↑	-0.003	NS	39	405.2099	$C_{17}H_{34}O_9$	$[M+Na^+]^+$	0.97	unknown	unknown
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(18:2(9Z,12Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(18:3(6Z,9Z,12Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(18:3(6Z,9Z,12Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(18:3(9Z,12Z,15Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(18:3(9Z,12Z,15Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(18:4(6Z,9Z,12Z,15Z)/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:3(5Z,8Z,11Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:3(8Z,11Z,14Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:4(5Z,8Z,11Z,14Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:4(5Z,8Z,11Z,14Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:4(8Z,11Z,14Z,17Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:4(8Z,11Z,14Z,17Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(22:4(7Z,10Z,13Z,16Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
150	↑	-0.003	NS	39	830.5698	$C_{48}H_{80}NO_8P$	$[M+H^+]^+$	0.47	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:2(9Z,12Z))	Phosphatidylcholines
151	↑	-0.003	NS	16	518.3287	unknown	-	-	unknown	unknown
152	↑	-0.003	NS	39	387.1739	$C_{15}H_{28}N_2O_8$	$[M+Na^+]^+$	0.22	unknown	unknown
153	↑	-0.003	NS	36	409.1697	$C_{16}H_{26}N_4O_7$	$[M+Na^+]^+$	0.8	unknown	unknown
154	↑	-0.003	NS	35	167.1042	$C_8H_{16}O_2$	$[M+Na^+]^+$	-0.07	unknown	unknown
155	↑	-0.003	NS	38	529.3715	$C_{27}H_{54}O_8$	$[M+Na^+]^+$	0.77	unknown	unknown
156	↑	-0.003	NS	32	542.3351	$C_{32}H_{45}N_3O_3$	$[M+Na^+]^+$	-0.4	unknown	unknown
157	↑	-0.003	NS	7	496.3354	$C_{24}H_{47}N_3O_6$	$[M+Na^+]^+$	-0.56	unknown	unknown
158	↑	-0.003	NS	39	351.0904	$C_{12}H_{24}O_6S_2$	$[M+Na^+]^+$	-0.62	unknown	unknown
159	↑	-0.003	NS	37	638.2815	$C_{28}H_{46}N_3O_{10}P$	$[[M+Na^+]^+$	0.23	unknown	unknown
160	↑	-0.003	NS	38	539.8155	unknown	-	-	unknown	unknown
161	↑	-0.003	0.038	26	461.2507	$C_{24}H_{38}O_7$	$[M+Na^+]^+$	-0.55	unknown	unknown
162	↑	-0.003	NS	39	684.439	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
163	↑	-0.003	NS	39	263.0561	$C_8H_{16}O_6S$	$[M+Na^+]^+$	0.31	unknown	unknown
164	↑	-0.003	NS	39	595.3825	$C_{29}H_{60}N_2O_4SP_2$	$[M+H^+]^+$	0.55	unknown	unknown
165	↑	-0.003	NS	37	610.3102	unknown	-	-	unknown	unknown
166	↑	-0.003	NS	39	932.5295	unknown	-	-	unknown	unknown
167	↑	-0.003	NS	21	524.2318	$C_{18}H_{43}N_3O_8SP_2$	$[M+H^+]^+$	-0.21	unknown	unknown
168	↑	-0.003	0.026	39	498.3467	unknown	-	-	unknown	unknown
169	↑	-0.003	NS	39	952.4982	$C_{47}H_{80}NO_{13}SP$	$[M+Na^+]^+$	0.14	unknown	unknown
170	↑	-0.003	NS	39	1128.618	unknown	-	-	unknown	unknown
171	↑	-0.003	NS	39	188.7751	unknown	-	-	unknown	unknown
172	↑	-0.003	NS	39	304.2611	$C_{18}H_{35}NO$	$[M+Na^+]^+$	0.06	Oleamide	Lipoamide
173	↑	-0.003	NS	34	546.3455	$C_{28}H_{51}NO_7S$	$[M+H^+]^+$	-0.77	unknown	unknown
174	↑	-0.003	NS	29	520.3452	unknown	-	-	unknown	unknown
175	↑	-0.003	0.028	32	459.2355	$C_{19}H_{39}O_{10}P$	$[M+H^+]^+$	0.38	unknown	unknown
176	↑	-0.003	NS	39	445.256	$C_{19}H_{41}O_9P$	$[M+H^+]^+$	-0.21	unknown	unknown
177	↑	-0.003	0.04	34	475.2441	$C_{25}H_{34}N_2O_7$	$[M+H^+]^+$	0.49	unknown	unknown
178	↑	-0.003	NS	39	441.298	$C_{19}H_{42}N_6O_2S$	$[M+Na^+]^+$	-0.47	unknown	unknown
179	↑	-0.003	NS	39	389.1121	$C_{13}H_{26}O_9P_2$	$[M+H^+]^+$	-0.98	unknown	unknown
180	↑	-0.003	NS	39	931.6408	unknown	-	-	unknown	unknown
181	↑	-0.003	0.008	39	1037.658	unknown	-	-	unknown	unknown
182	↑	-0.003	NS	39	339.0541	$C_{10}H_{20}O_7S_2$	$[M+Na^+]^+$	-0.4	unknown	unknown
183	↑	-0.003	NS	13	590.3267	$C_{26}H_{53}N_3O_6S_2$	$[M+Na^+]^+$	-0.18	unknown	unknown
184	↑	-0.003	0.043	37	555.2937	unknown	-	-	unknown	unknown
185	↑	0.003	NS	30	795.3491	$C_{31}H_{66}O_{15}SP_2$	$[M+Na^+]^+$	0.08	unknown	unknown
186	↑	-0.003	NS	37	795.3207	$C_{32}H_{58}N_2O_{15}P_2$	$[M+Na^+]^+$	0.34	unknown	unknown
187	↑	-0.003	NS	39	410.1613	unknown	-	-	unknown	unknown
188	↑	-0.003	NS	39	1104.617	unknown	-	-	unknown	unknown
189	↑	-0.003	0.023	39	272.1687	unknown	-	-	unknown	unknown
190	↑	-0.003	NS	39	906.5041	unknown	-	-	unknown	unknown
191	↑	-0.003	0.013	37	522.3478	unknown	-	-	unknown	unknown
192	↑	-0.003	NS	37	473.2879	$C_{24}H_{46}N_2OSP_2$	$[M+H^+]^+$	0.05	unknown	unknown
193	↑	-0.003	NS	25	410.1626	unknown	-	-	unknown	unknown
194	↑	-0.003	NS	39	767.5363	unknown	-	-	unknown	unknown
195	↑	-0.003	NS	39	527.8155	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
196	↑	-0.003	NS	39	271.161	$C_{18}H_{23}P$	$[M+H^+]^+$	-0.13	unknown	unknown
197	↑	-0.003	NS	39	414.2694	unknown	-	-	unknown	unknown
198	↑	-0.003	NS	39	639.409	$C_{34}H_{56}N_4O_6$	$[M+Na^+]^+$	-0.35	unknown	unknown
199	↑	-0.003	0.04	39	530.3224	$C_{29}H_{43}N_3O_6$	$[M+H^+]^+$	-0.07	unknown	unknown
200	↑	-0.003	NS	39	472.734	unknown	-	-	unknown	unknown
201	↑	-0.003	0.002	39	1039.676	unknown	-	-	unknown	unknown
202	↑	-0.003	0.049	39	542.3333	unknown	-	-	unknown	unknown
203	↑	-0.003	NS	39	954.6374	unknown	-	-	unknown	unknown
204	↑	-0.003	NS	9	496.3429	$C_{26}H_51NO_4S$	$[M+Na^+]^+$	-0.38	unknown	unknown
205	↑	-0.003	NS	37	478.3228	unknown	-	-	unknown	unknown
206	↑	-0.003	0.043	38	568.3303	$C_{24}H_{51}NO_{12}$	$[M+Na^+]^+$	-0.01	unknown	unknown
207	↑	-0.003	NS	39	446.2289	$C_{20}H_{34}N_5O_3P$	$[M+Na^+]^+$	-0.48	unknown	unknown
208	↑	-0.003	NS	39	802.5379	$C_{46}H_{76}NO_8P$	$[M+H^+]^+$	0.36	PC(18:3(6Z,9Z,12Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
208	↑	-0.003	NS	39	802.5379	$C_{46}H_{76}NO_8P$	$[M+H^+]^+$	0.36	PC(18:3(9Z,12Z,15Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
208	↑	-0.003	NS	39	802.5379	$C_{46}H_{76}NO_8P$	$[M+H^+]^+$	0.36	PC(18:4(6Z,9Z,12Z,15Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
208	↑	-0.003	NS	39	802.5379	$C_{46}H_{76}NO_8P$	$[M+H^+]^+$	0.36	PC(18:4(6Z,9Z,12Z,15Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
208	↑	-0.003	NS	39	802.5379	$C_{46}H_{76}NO_8P$	$[M+H^+]^+$	0.36	PC(20:4(5Z,8Z,11Z,14Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
208	↑	-0.003	NS	39	802.5379	$C_{46}H_{76}NO_8P$	$[M+H^+]^+$	0.36	PC(20:4(8Z,11Z,14Z,17Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
208	↑	-0.003	NS	39	802.5379	$C_{46}H_{76}NO_8P$	$[M+H^+]^+$	0.36	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
208	↑	-0.003	NS	39	802.5379	$C_{46}H_76NO_8P$	$[M+H^+]^+$	0.36	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
209	↑	-0.003	0.02	39	534.2962	$C_{24}H_{50}NO_7P$	$[M+K^+]^+$	0.99	LysoPC(16:0)	Lysophosphatidylcholines
210	↑	-0.003	NS	39	501.3405	$C_{26}H_{54}O_3S_2$	$[M+Na^+]^+$	-0.42	unknown	unknown
211	↑	-0.003	0.028	39	991.6759	unknown	-	-	unknown	unknown
212	↑	-0.003	NS	39	478.3291	$C_{24}H_{50}NO_7P$	$[M+H^+-H_2O]^+$	0.24	LysoPC(16:0)	Lysophosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
213	↑	-0.003	NS	39	816.5186	unknown	-	-	unknown	unknown
214	↑	-0.003	NS	26	796.3481	unknown	-	-	unknown	unknown
215	↑	-0.003	NS	24	543.3309	unknown	-	-	unknown	unknown
216	↑	-0.003	NS	39	473.2513	$C_{20}H_{41}O_{10}P$	$[M+H^+]^+$	0.57	unknown	unknown
217	↑	-0.002	NS	39	637.3066	$C_{29}H_{48}O_{15}$	$[M+H^+]^+$	-	unknown	unknown
218	↑	-0.002	NS	39	812.3149	unknown	Unknown	-	unknown	unknown
219	↑	-0.002	0.017	30	569.3452	unknown	Unknown	-	unknown	unknown
220	↑	-0.002	0.01	39	1013.658	unknown	Unknown	-	unknown	unknown
221	↑	-0.002	NS	39	311.2193	unknown	Unknown	-	unknown	unknown
222	↑	-0.002	NS	8	567.3297	unknown	Unknown	-	unknown	unknown
223	↑	-0.002	NS	39	515.3138	unknown	Unknown	-	unknown	unknown
224	↑	-0.002	NS	39	307.1152	$C_{14}H_{20}O_6$	$[M+Na^+]^+$	0.05	unknown	unknown
225	↑	-0.002	NS	39	410.1586	unknown	Unknown	-	unknown	unknown
226	↑	-0.002	0.018	31	503.2618	$C_{21}H_{43}O_{11}P$	$[M+H^+]^+$	0.45	unknown	unknown
227	↑	-0.002	NS	39	590.4271	unknown	Unknown	-	unknown	unknown
228	↑	-0.002	NS	10	831.567	unknown	Unknown	-	unknown	unknown
229	↑	-0.002	NS	39	436.3051	unknown	Unknown	-	unknown	unknown
230	↑	-0.002	NS	32	627.3516	$C_{28}H_{54}N_2O_{11}S$	$[M+H^+]^+$	-0.86	unknown	unknown
231	↑	-0.002	NS	29	567.3314	$C_{34}H_{46}O_7$	$[M+H^+]^+$	-0.37	unknown	unknown
232	↑	-0.002	0.006	39	593.3417	unknown	Unknown	-	unknown	unknown
233	↑	-0.002	NS	39	1045.465	$C_{45}H_{78}N_2O_{21}P_2$	$[M+H^+]^+$	0.06	unknown	unknown
234	↑	-0.002	NS	37	551.3141	$C_{25}H_{53}O_7SP$	$[M+Na^+]^+$	-0.08	unknown	unknown
235	↑	-0.002	NS	39	625.0956	$C_{31}H_{22}O_{13}$	$[M+Na^+]^+$	0.56	unknown	unknown
236	↑	-0.002	NS	39	640.4126	unknown	Unknown	-	unknown	unknown
237	↑	-0.002	NS	37	412.0975	unknown	Unknown	-	unknown	unknown
238	↑	-0.002	NS	36	411.1661	unknown	Unknown	-	unknown	unknown
239	↑	-0.002	NS	39	445.1119	$C_{14}H_{24}N_2O_{12}S$	$[M+H^+]^+$	-0.87	unknown	unknown
240	↑	-0.002	NS	39	519.3136	unknown	Unknown	-	unknown	unknown
241	↑	-0.002	NS	7	542.3126	$C_{26}H_{49}NO_7S$	$[M+H^+]^+$	0.66	unknown	unknown
242	↑	-0.002	NS	39	723.5106	unknown	Unknown	-	unknown	unknown
243	↑	-0.002	0.033	39	495.3279	unknown	Unknown	-	unknown	unknown
244	↑	-0.002	NS	39	931.5169	unknown	Unknown	-	unknown	unknown
245	↑	-0.002	NS	39	339.0716	unknown	Unknown	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-val	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
246	↑	-0.002	NS	39	823.3689	$C_{37}H_{61}N_4O_{11}SP$	$[M+Na^+]^+$	0.14	unknown	unknown
247	↑	-0.002	NS	39	854.5861	unknown	Unknown	-	unknown	unknown
248	↑	-0.002	NS	39	535.3204	unknown	Unknown	-	unknown	unknown

Table 24: Significant biomarkers DOWN-regulated respect to the insulin sensitive group at follow-up. Experimental analyses via direct infusion with ICR-FT/MS in positve mode ionization. The possible isomers and isobars masses annotated are grouped together. The same entry number indicates that they belong to the same isomers/isobar assignment group.

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
1	↓	0.459	NS	39	467.10203	$C_{16}H_{20}N_4O_{11}$	$[M+Na^+]^+$	-0.12	unknown	unknown
2	↓	0.194	NS	39	360.32353	unknown	-	-	unknown	unknown
3	↓	0.154	NS	39	536.16556	$C_{23}H_{32}NO_{10}P$	$[M+Na^+]^+$	-0.09	unknown	unknown
4	↓	0.107	0.049	39	468.10189	unknown	-	-	unknown	unknown
5	↓	0.066	NS	39	408.30876	$C_{22}H_{43}NO_4$	$[M+Na^+]^+$	0.81	unknown	unknown
6	↓	0.062	NS	39	445.12027	$C_{13}H_{31}N_2O_7S_2P$	$[M+Na^+]^+$	0.05	unknown	unknown
7	↓	0.06	NS	39	462.14689	$C_{16}H_{23}N_5O_{11}$	$[M+H^+]^+$	0.44	unknown	unknown
8	↓	0.058	NS	39	469.09876	unknown	-	-	unknown	unknown
9	↓	0.051	NS	39	425.21461	$C_{20}H_{34}O_8$	$[M+Na^+]^+$	0.04	unknown	unknown
10	↓	0.047	0.028	39	1022.77433	unknown	-	-	unknown	unknown
11	↓	0.046	0.003	39	722.52889	unknown	-	-	unknown	unknown
12	↓	0.046	NS	39	537.16523	$C_{17}H_{30}N_4O_{14}$	$[M+Na^+]^+$	0.29	unknown	unknown
13	↓	0.045	NS	39	393.29745	unknown	-	-	unknown	unknown
14	↓	0.044	NS	39	361.32685	unknown	-	-	unknown	unknown
15	↓	0.039	0.021	39	468.10556	unknown	-	-	unknown	unknown
16	↓	0.037	NS	39	685.4369	unknown	-	-	unknown	unknown
17	↓	0.032	NS	39	538.16268	unknown	-	-	unknown	unknown
18	↓	0.032	0.023	39	1023.77765	unknown	-	-	unknown	unknown
19	↓	0.032	0.011	33	467.1062	$C_{21}H_{20}N_2O_9$	$[M+Na^+]^+$	0.21	unknown	unknown
20	↓	0.026	NS	39	794.48343	$C_{38}H_{71}N_3O_{12}S$	$[M+H^+]^+$	0.39	unknown	unknown
21	↓	0.023	NS	39	736.5447	unknown	-	-	unknown	unknown
22	↓	0.023	NS	39	541.12113	$C_{18}H_{32}O_{13}P_2$	$[M+Na^+]^+$	0.17	unknown	unknown
23	↓	0.021	NS	39	763.60796	unknown	-	-	unknown	unknown
24	↓	0.021	0.003	39	723.53211	unknown	-	-	unknown	unknown
25	↓	0.02	NS	39	697.65979	unknown	-	-	unknown	unknown
26	↓	0.02	NS	39	446.25361	$C_{25}H_{35}NO_6$	$[M+H^+]^+$	-0.23	unknown	unknown
27	↓	0.019	NS	32	409.16531	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
28	↓	0.018	NS	39	686.4405	$C_{35}H_{63}N_3O_8S$	$[M+H^+]^+$	-0.53	unknown	unknown
29	↓	0.017	NS	39	446.11996	unknown	-	-	unknown	unknown
30	↓	0.017	0.038	39	467.09662	$C_{16}H_{22}N_2O_{12}S$	$[M+H^+]^+$	-0.01	unknown	unknown
31	↓	0.017	NS	36	369.15184	$C_{16}H_{26}O_8$	$[M+Na^+]^+$	-0.41	unknown	unknown
32	↓	0.016	NS	39	460.26932	unknown	-	-	unknown	unknown
33	↓	0.015	NS	31	505.03135	$C_{22}H_{14}N_2O_9S$	$[M+Na^+]^+$	0.25	unknown	unknown
34	↓	0.015	NS	39	463.14676	unknown	-	-	unknown	unknown
35	↓	0.014	NS	39	409.3116	unknown	-	-	unknown	unknown
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(14:0/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}N_{O8}P$	$[M+H^+]^+$	0.51	PC(14:0/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(14:1(9Z)/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(16:0/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(16:1(9Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(16:1(9Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:1(11Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:1(9Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:2(9Z,12Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:2(9Z,12Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:3(6Z,9Z,12Z)/18:2(9Z,12Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:3(9Z,12Z,15Z)/18:2(9Z,12Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:4(6Z,9Z,12Z,15Z)/18:1(11Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(18:4(6Z,9Z,12Z,15Z)/18:1(9Z))	Phosphatidylcholines
50	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(20:4(5Z,8Z,11Z,14Z)/16:1(9Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_78NO_8P$	$[M+H^+]^+$	0.51	PC(20:4(8Z,11Z,14Z,17Z)/16:1(9Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(20:5(5Z,8Z,11Z,14Z,17Z)/16:0)	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(22:4(7Z,10Z,13Z,16Z)/14:1(9Z))	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(22:5(4Z,7Z,10Z,13Z,16Z)/14:0)	Phosphatidylcholines
36	↓	0.013	NS	39	780.55339	$C_{44}H_{78}NO_8P$	$[M+H^+]^+$	0.51	PC(22:5(7Z,10Z,13Z,16Z,19Z)/14:0)	Phosphatidylcholines
37	↓	0.013	NS	37	793.62978	unknown	-	-	unknown	unknown
38	↓	0.013	NS	39	675.67765	unknown	-	-	unknown	unknown
39	↓	0.013	NS	39	403.36079	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
49	↓	0.012	0.046	38	542.12525	unknown	-	-	unknown	unknown
41	↓	0.012	NS	39	1071.61228	unknown	-	-	unknown	unknown
42	↓	0.012	0.046	39	746.49834	$C_{38}H_{71}N_3O_9S$	$[M+H^+]^+$	-0.05	unknown	unknown
43	↓	0.012	NS	39	795.48696	unknown	-	-	unknown	unknown
44	↓	0.012	NS	39	1070.75866	unknown	-	-	unknown	unknown
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(18:4(6Z,9Z,12Z,15Z)/24:1(15Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:0/22:5(4Z,7Z,10Z,13Z,16Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:0/22:5(7Z,10Z,13Z,16Z,19Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:1(11Z)/22:4(7Z,10Z,13Z,16Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:3(5Z,8Z,11Z)/22:2(13Z,16Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:3(8Z,11Z,14Z)/22:2(13Z,16Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:4(5Z,8Z,11Z,14Z)/22:1(13Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:4(8Z,11Z,14Z,17Z)/22:1(13Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(20:5(5Z,8Z,11Z,14Z,17Z)/22:0/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:0/20:5(5Z,8Z,11Z,14Z,17Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:1(13Z)/20:4(5Z,8Z,11Z,14Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:1(13Z)/20:4(8Z,11Z,14Z,17Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:2(13Z,16Z)/20:3(5Z,8Z,11Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:2(13Z,16Z)/20:3(8Z,11Z,14Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:4(7Z,10Z,13Z,16Z)/20:1(11Z)/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:5(4Z,7Z,10Z,13Z,16Z)/20:0/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(22:5(7Z,10Z,13Z,16Z,19Z)/20:0/0:0)	Diacylglycerols
45	↓	0.012	NS	39	737.54786	$C_{45}H_{78}O_5$	$[M+K^+]^+$	0.27	DG(24:1(15Z)/18:4(6Z,9Z,12Z,15Z)/0:0)	Diacylglycerols
46	↓	0.011	NS	16	542.32558	unknown	-	-	unknown	unknown
47	↓	0.011	NS	39	394.3008	unknown	-	-	unknown	unknown
48	↓	0.011	0.028	39	1024.78078	unknown	-	-	unknown	unknown
49	↓	0.01	NS	39	447.11718	unknown	-	-	unknown	unknown
50	↓	0.01	NS	39	470.09833	unknown	-	-	unknown	unknown
51	↓	0.01	NS	39	426.21788	unknown	-	-	unknown	unknown
52	↓	0.01	NS	27	469.10504	unknown	-	-	unknown	unknown
53	↓	0.01	NS	16	467.10763	unknown	-	-	unknown	unknown
54	↓	0.01	NS	39	469.10231	$C_{22}H_{23}O_8P$	$[M+Na^+]^+$	0.08	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
55	↓	0.009	NS	39	698.66297	unknown	-	-	unknown	unknown
56	↓	0.009	NS	39	764.61127	unknown	-	-	unknown	unknown
57	↓	0.009	NS	39	464.14375	unknown	-	-	unknown	unknown
58	↓	0.009	NS	37	536.17099	unknown	-	-	unknown	unknown
59	↓	0.009	NS	39	1072.61596	unknown	-	-	unknown	unknown
60	↓	0.009	NS	39	360.32029	unknown	-	-	unknown	unknown
61	↓	0.008	NS	39	961.80768	unknown	-	-	unknown	unknown
62	↓	0.008	NS	39	1071.7622	unknown	-	-	unknown	unknown
63	↓	0.008	NS	39	779.47245	$C_{38}H_{70}N_2O_{12}S$	$[M+H^+]^+$	0.29	unknown	unknown
64	↓	0.008	NS	39	468.09605	unknown	-	-	unknown	unknown
65	↓	0.008	NS	18	536.17278	$C_{17}H_{32}N_5O_{11}P$	$[M+Na^+]^+$	-0.08	unknown	unknown
66	↓	0.007	NS	18	409.16638	$C_{20}H_{34}OS_3$	$[M+Na^+]^+$	-0.04	unknown	unknown
67	↓	0.007	NS	39	638.60953	unknown	-	-	unknown	unknown
68	↓	0.007	NS	39	537.16964	unknown	-	-	unknown	unknown
69	↓	0.007	NS	39	750.60267	unknown	-	-	unknown	unknown
70	↓	0.007	NS	39	536.15812	$C_{19}H_{31}NO_{15}$	$[M+Na^+]^+$	-0.88	unknown	unknown
71	↓	0.007	NS	39	781.55695	unknown	-	-	unknown	unknown
72	↓	0.007	NS	39	668.62013	unknown	-	-	unknown	unknown
73	↓	0.007	NS	39	730.6339	unknown	-	-	unknown	unknown
74	↓	0.007	NS	18	518.31696	unknown	-	-	unknown	Unknown
75	↓	0.006	NS	39	745.64485	unknown	-	-	unknown	Unknown
76	↓	0.006	NS	28	462.15084	$C_{21}H_{23}N_3O_9$	$[M+H^+]^+$	0.28	unknown	Unknown
77	↓	0.006	0.046	39	467.09296	$C_{13}H_{29}N_2O_8SP_3$	$[M+H^+]^+$	-0.15	unknown	unknown
78	↓	0.006	0.035	39	747.5019	unknown	-	-	unknown	unknown
79	↓	0.006	NS	39	539.16216	$C_{19}H_{37}N_2O_8S_2P$	$[M+Na^+]^+$	0.08	unknown	unknown
80	↓	0.006	NS	37	470.10215	unknown	-	-	unknown	unknown
81	↓	0.006	0.018	11	483.07913	$C_{15}H_{26}O_{12}P_2$	$[M+Na^+]^+$	-0.09	unknown	unknown
82	↓	0.006	NS	29	794.63307	unknown	-	-	unknown	unknown
83	↓	0.006	NS	39	356.35251	$C_{22}H_{45}NO_2$	$[M+H^+]^+$	0.58	unknown	unknown
84	↓	0.006	NS	39	640.58875	unknown	-	-	unknown	unknown
85	↓	0.005	NS	12	462.1494	$C_{15}H_{32}N_3O_7S_2P$	$[M+H^+]^+$	0.41	unknown	unknown
86	↓	0.005	NS	39	447.25695	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
87	↓	0.005	NS	39	676.68084	unknown	-	-	unknown	unknown
88	↓	0.005	NS	39	1000.79208	unknown	-	-	unknown	unknown
89	↓	0.005	NS	25	360.32602	$C_{24}H_{41}NO$	$[M+H^+]^+$	-0.2	unknown	unknown
90	↓	0.005	0.002	39	724.53572	unknown	-	-	unknown	unknown
91	↓	0.005	NS	39	1271.73182	unknown	-	-	unknown	unknown
92	↓	0.005	NS	39	962.81082	unknown	-	-	unknown	unknown
93	↓	0.005	NS	39	610.18501	$C_{26}H_{38}NO_{10}SP$	$[M+Na^+]^+$	0.63	unknown	unknown
94	↓	0.005	NS	39	463.15064	unknown	-	-	unknown	unknown
95	↓	0.005	NS	39	362.33037	unknown	-	-	unknown	unknown
96	↓	0.005	NS	39	429.24005	$C_{24}H_{38}O_4$	$[M+K^+]^+$	0.27	12-Ketodeoxycholic acid	Bile Acids
96	↓	0.005	NS	39	429.24005	$C_{24}H_{38}O_4$	$[M+K^+]^+$	0.27	7-Hydroxy-3-oxocholanoic acid	Bile Acids
96	↓	0.005	NS	39	429.24005	$C_{24}H_{38}O_4$	$[M+K^+]^+$	0.27	Nutriacholic acid	Bile Acids
96	↓	0.005	NS	39	429.24005	$C_{24}H_{38}O_4$	$[M+K^+]^+$	0.27	7a-Hydroxy-3-oxo-5b-cholanoic acid	Bile Acids
97	↓	0.005	NS	39	446.12384	unknown	-	-	unknown	unknown
98	↓	0.005	NS	39	701.41111	$C_{33}H_{63}N_2O_{10}P$	$[M+Na^+]^+$	-0.21	unknown	unknown
99	↓	0.005	NS	9	537.15996	$C_{17}H_{32}N_2O_{15}S$	$[M+H^+]^+$	0.63	unknown	unknown
100	↓	0.005	0.031	38	467.1091	$C_{18}H_{24}N_2O_9S$	$[M+Na^+]^+$	-0.8	unknown	unknown
101	↓	0.005	NS	24	536.16024	$C_{20}H_{27}N_5O_{11}$	$[M+Na^+]^+$	0.57	unknown	unknown
102	↓	0.005	NS	39	469.09316	$C_{17}H_{25}O_{11}SP$	$[M+H^+]^+$	0.77	unknown	unknown
103	↓	0.005	0.033	28	1022.7868	unknown	-	-	unknown	unknown
104	↓	0.005	NS	39	483.07648	$C_{12}H_{22}N_2O_{16}S$	$[M+H^+]^+$	0.41	unknown	unknown
105	↓	0.004	NS	39	461.27247	unknown	-	-	unknown	unknown
106	↓	0.004	NS	18	410.16764	unknown	-	-	unknown	unknown
107	↓	0.004	0.015	27	467.11332	$C_{23}H_{24}O_7S$	$[M+Na^+]^+$	-0.37	unknown	unknown
108	↓	0.004	NS	38	467.11052	$C_{26}H_{20}O_7$	$[M+Na^+]^+$	0.84	unknown	unknown
109	↓	0.004	NS	39	332.29247	$C_{20}H_{39}NO$	$[M+Na^+]^+$	0.26	unknown	unknown
110	↓	0.004	NS	39	666.64088	unknown	-	-	unknown	unknown
111	↓	0.004	NS	8	360.32516	unknown	-	-	unknown	unknown
112	↓	0.004	NS	17	541.12846	$C_{21}H_{30}N_2O_9S_2$	$[M+Na^+]^+$	-0.07	unknown	unknown
113	↓	0.004	NS	39	376.97571	unknown	-	-	unknown	unknown
114	↓	0.004	NS	39	360.31793	unknown	-	-	unknown	unknown
115	↓	0.004	NS	39	1272.73499	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
116	↓	0.004	NS	39	534.37103	$C_{31}H_{52}NO_4P$	$[M+H^+]^+$	0.67	unknown	unknown
117	↓	0.004	NS	39	1055.74765	unknown	-	-	unknown	unknown
118	↓	0.004	NS	39	408.30421	unknown	-	-	unknown	unknown
119	↓	0.004	NS	16	445.12261	$C_{20}H_{26}N_2O_4S_2$	$[M+Na^+]^+$	-0.02	unknown	unknown
120	↓	0.004	NS	20	445.11662	$C_{16}H_{27}N_2O_7SP$	$[M+Na^+]^+$	-0.59	unknown	unknown
121	↓	0.004	NS	30	541.11564	$C_{19}H_{28}N_2O_{12}S_2$	$[M+H^+]^+$	-0.01	unknown	unknown
122	↓	0.004	NS	39	687.44381	unknown	-	-	unknown	unknown
123	↓	0.004	0.04	39	330.33671	$C_{20}H_{43}NO_2$	$[M+H^+]^+$	0.16	unknown	unknown
124	↓	0.004	NS	39	610.57796	unknown	-	-	unknown	unknown
125	↓	0.004	NS	20	541.12479	$C_{24}H_{26}N_2O_9S$	$[M+Na^+]^+$	-0.63	unknown	unknown
126	↓	0.004	NS	39	1199.77741	unknown	-	-	unknown	unknown
127	↓	0.004	NS	39	751.60605	unknown	-	-	unknown	unknown
128	↓	0.004	NS	20	469.09495	unknown	-	-	unknown	unknown
129	↓	0.004	NS	39	304.2999	$C_{21}H_{37}N$	$[M+H^+]^+$	0.09	unknown	unknown
130	↓	0.003	NS	39	541.11327	$C_{17}H_{30}N_2O_{12}S_2$	$[M+Na^+]^+$	0.05	unknown	unknown
131	↓	0.003	NS	14	412.32087	$C_{27}H_{41}NO_2$	$[M+H^+]^+$	-0.32	unknown	unknown
132	↓	0.003	NS	39	780.47604	unknown	-	-	unknown	unknown
133	↓	0.003	NS	39	388.35493	unknown	-	-	unknown	unknown
134	↓	0.003	NS	36	541.12664	$C_{22}H_{31}O_{10}SP$	$[M+Na^+]^+$	-0.25	unknown	unknown
135	↓	0.003	NS	12	697.65405	unknown	-	-	unknown	unknown
136	↓	0.003	NS	39	404.36436	unknown	-	-	unknown	unknown
137	↓	0.003	NS	39	553.45992	unknown	-	-	unknown	unknown
138	↓	0.003	NS	39	1073.61954	unknown	-	-	unknown	unknown
139	↓	0.003	NS	39	1001.79525	unknown	-	-	unknown	unknown
140	↓	0.003	NS	39	400.37873	$C_{24}H_{49}NO_3$	$[M+H^+]^+$	0.54	unknown	unknown
141	↓	0.003	0.043	38	468.1099	unknown	-	-	unknown	unknown
142	↓	0.003	NS	26	722.5351	unknown	-	-	unknown	unknown
143	↓	0.003	NS	19	462.14294	$C_{19}H_{27}NO_{10}S$	$[M+H^+]^+$	0.21	unknown	unknown
144	↓	0.003	NS	39	462.14121	$C_{20}H_{30}N_3O_2S_2P$	$[M+Na^+]^+$	0.61	unknown	unknown
145	↓	0.003	NS	37	746.64793	unknown	-	-	unknown	unknown
146	↓	0.003	NS	39	1200.78084	unknown	-	-	unknown	unknown
147	↓	0.003	NS	10	1022.7621	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
148	↓	0.003	NS	39	1072.76577	unknown	-	-	unknown	unknown
149	↓	0.003	NS	39	731.63705	unknown	-	-	unknown	unknown
150	↓	0.003	NS	39	537.15752	unknown	-	-	unknown	unknown
151	↓	0.003	NS	39	639.61275	unknown	-	-	unknown	unknown
152	↓	0.003	NS	38	796.49051	$C_{38}H_{7}4N_{3}O_{10}SP$	$[M+H^{+}]^{+}$	-0.03	unknown	unknown
153	↓	0.003	NS	39	605.42467	$C_{31}H_{58}N_{4}O_{6}$	$[M+Na^{+}]^{+}$	-0.31	unknown	unknown
154	↓	0.003	NS	39	332.33127	unknown	-	-	unknown	unknown
155	↓	0.003	NS	39	649.45111	unknown	-	-	unknown	unknown
156	↓	0.003	NS	21	408.3118	unknown	-	-	unknown	unknown
157	↓	0.003	NS	39	669.62335	unknown	-	-	unknown	unknown
158	↓	0.003	NS	39	738.55143	unknown	-	-	unknown	unknown
159	↓	0.003	NS	39	561.3982	$C_{29}H_{54}N_{4}O_{5}$	$[M+Na^{+}]^{+}$	-0.78	unknown	unknown
160	↓	0.003	NS	29	445.12397	$C_{15}H_{31}N_{2}O_{5}SP_{3}$	$[M+H^{+}]^{+}$	0.1	unknown	unknown
161	↓	0.003	NS	39	468.09252	unknown	-	-	unknown	unknown
162	↓	0.003	NS	39	471.09607	unknown	-	-	unknown	unknown
163	↓	0.003	NS	37	383.20388	$C_{18}H_{32}O_{7}$	$[M+Na^{+}]^{+}$	-0.38	unknown	unknown
164	↓	0.003	NS	39	425.44678	$C_{27}H_{56}N_{2}O$	$[M+H^{+}]^{+}$	0.56	unknown	unknown
165	↓	0.003	NS	39	542.12075	$C_{23}H_{25}N_{3}O_{9}S$	$[M+Na^{+}]^{+}$	0.7	unknown	unknown
166	↓	0.003	NS	9	462.15238	$C_{24}H_{25}NO_{7}$	$[M+Na^{+}]^{+}$	0.13	unknown	unknown
167	↓	0.003	NS	38	538.16926	unknown	-	-	unknown	unknown
168	↓	0.003	NS	7	361.32852	unknown	-	-	unknown	unknown
169	↓	0.003	NS	22	468.0977	unknown	-	-	unknown	unknown
170	↓	0.003	NS	39	664.62521	unknown	-	-	unknown	unknown
171	↓	0.003	NS	39	1056.75116	unknown	-	-	unknown	unknown
172	↓	0.003	NS	8	780.54646	unknown	-	-	unknown	unknown
173	↓	0.003	NS	16	360.32682	unknown	-	-	unknown	unknown
174	↓	0.003	NS	39	484.07566	unknown	-	-	unknown	unknown
175	↓	0.003	NS	39	519.13942	unknown	-	-	unknown	unknown
176	↓	0.003	NS	20	507.02722	$C_{12}H_{22}O_{16}P_{2}$	$[M+Na^{+}]^{+}$	-0.63	unknown	unknown
177	↓	0.002	NS	36	778.61877	unknown	-	-	unknown	unknown
178	↓	0.002	NS	8	403.23283	unknown	-	-	unknown	unknown
179	↓	0.002	NS	26	543.12182	$C_{2}3H_{24}N_{2}O_{12}$	$[M+Na^{+}]^{+}$	-0.61	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
180	↓	0.002	NS	39	517.37157	$C_{27}H_{5}8O_{3}S_{2}$	$[M+Na^{+}]^{+}$	-0.74	unknown	unknown
181	↓	0.002	NS	12	425.2099	$C_{17}H_{39}O_{6}SP$	$[M+Na^{+}]^{+}$	0.43	unknown	unknown
182	↓	0.002	NS	39	362.30308	unknown	-	-	unknown	unknown
183	↓	0.002	NS	13	410.16983	unknown	-	-	unknown	unknown
184	↓	0.002	NS	39	369.38378	$C_{23}H_{48}N_{2}O$	$[M+H^{+}]^{+}$	-0.43	unknown	unknown
185	↓	0.002	NS	35	562.31713	$C_{29}H_{49}NO_{6}S$	$[M+Na^{+}]^{+}$	-0.26	unknown	unknown
186	↓	0.002	NS	39	646.62579	unknown	-	-	unknown	unknown
187	↓	0.002	NS	8	403.36283	unknown	-	-	unknown	unknown
188	↓	0.002	NS	33	536.1748	unknown	-	-	unknown	unknown
189	↓	0.002	0.031	39	1025.78371	unknown	-	-	unknown	unknown
190	↓	0.002	NS	38	551.33049	$C_{27}H_{48}N_{2}O_{8}$	$[M+Na^{+}]^{+}$	0.36	unknown	unknown
191	↓	0.002	NS	39	445.11506	$C_{17}H_{28}O_{8}P_{2}$	$[M+Na^{+}]^{+}$	-0.25	unknown	unknown
192	↓	0.002	NS	39	707.49333	$C_{36}H_{68}N_{4}O_{8}$	$[M+Na^{+}]^{+}$	0.56	unknown	unknown
193	↓	0.002	NS	39	361.32355	unknown	-	-	unknown	unknown
194	↓	0.002	NS	7	370.1555	unknown	-	-	unknown	unknown
195	↓	0.002	NS	38	620.41165	$C_{30}H_{57}N_{3}O_{10}$	$[M+H^{+}]^{+}$	-0.04	unknown	unknown
196	↓	0.002	NS	39	378.97088	unknown	-	-	unknown	unknown
197	↓	0.002	NS	39	699.66657	unknown	-	-	unknown	unknown
198	↓	0.002	NS	33	467.1119	$C_{20}H_{22}N_{2}O_{9}S$	$[M+H^{+}]^{+}$	0.04	unknown	unknown
199	↓	0.002	NS	5	413.26847	unknown	-	-	unknown	unknown
200	↓	0.002	NS	39	696.65172	unknown	-	-	unknown	unknown
201	↓	0.002	NS	39	421.23265	$C_{18}H_{39}O_{7}P$	$[M+Na^{+}]^{+}$	0.2	unknown	unknown
202	↓	0.002	NS	38	598.3985	$C_{29}H_{59}NO_{9}S$	$[M+H^{+}]^{+}$	0.28	unknown	unknown
203	↓	0.002	NS	36	573.3438	unknown	-	-	unknown	unknown
204	↓	0.002	NS	39	448.11685	$C_{16}H_{28}NO_{8}SP$	$[M+Na^{+}]^{+}$	0.68	unknown	unknown
205	↓	0.002	NS	9	873.70303	unknown	-	-	unknown	unknown
206	↓	0.002	NS	39	693.47762	$C_{35}H_{66}N_{4}O_{8}$	$[M+Na^{+}]^{+}$	0.47	unknown	unknown
207	↓	0.002	NS	38	576.38531	$C_{28}H_{53}N_{3}O_{9}$	$[M+H^{+}]^{+}$	-0.26	unknown	unknown
208	↓	0.002	NS	39	641.59176	unknown	-	-	unknown	unknown
209	↓	0.002	NS	39	548.30138	unknown	-	-	unknown	unknown
210	↓	0.002	NS	39	485.07255	unknown	-	-	unknown	unknown
211	↓	0.002	NS	38	422.2941	unknown	-	-	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
212	↓	0.002	0.009	35	467.11897	$C_{13}H_{28}N_2O_{12}P_2$	$[M+H^+]^+$	-0.12	unknown	unknown
213	↓	0.002	NS	15	835.59056	unknown	-	-	unknown	unknown
214	↓	0.002	NS	39	465.14284	$C_{16}H_{31}N_2O_8SP$	$[M+Na^+]^+$	-0.54	unknown	unknown
215	↓	0.002	NS	39	663.46733	unknown	-	-	unknown	unknown
216	↓	0.002	NS	21	736.55108	unknown	-	-	unknown	unknown
217	↓	0.002	NS	29	1023.79013	unknown	-	-	unknown	unknown
218	↓	0.002	NS	39	619.44035	$C_{32}H_{60}N_4O_6$	$[M+Na^+]^+$	-0.25	unknown	unknown
219	↓	0.002	NS	39	402.3579	$C_{23}H_{47}NO_4$	$[M+H^+]^+$	0.29	unknown	unknown
220	↓	0.002	NS	38	642.42498	$C_{31}H_{63}NO_{10}S$	$[M+H^+]^+$	0.68	unknown	unknown
221	↓	0.002	NS	27	685.4426	$C3_9H_{60}N_2O_8$	$[M+H^+]^+$	0.52	unknown	unknown
222	↓	0.002	NS	39	358.36816	$C_{22}H_{47}NO_2$	$[M+H^+]^+$	0.56	unknown	unknown
223	↓	0.002	NS	13	327.14151	$C_{14}H_{24}O_7$	$[M+Na^+]^+$	0.27	unknown	unknown
224	↓	0.002	NS	39	430.38912	$C_{25}H_{51}NO_4$	$[M+H^+]^+$	0.07	unknown	unknown
225	↓	0.002	NS	39	609.36143	$C_{31}H_{54}O_{10}$	$[M+Na^+]^+$	0.83	unknown	unknown
226	↓	0.002	NS	39	536.15336	$C_{16}H_{36}NO_{13}SP$	$[M+Na^+]^+$	-0.68	unknown	unknown
227	↓	0.002	NS	39	481.15472	$C_{18}H_{36}O_7SP_2$	$[M+Na^+]^+$	-0.42	unknown	unknown
228	↓	0.002	NS	39	636.59385	unknown	-	-	unknown	unknown
229	↓	0.002	NS	39	667.64437	unknown	-	-	unknown	unknown
230	↓	0.002	NS	39	702.41459	unknown	-	-	unknown	unknown
231	↓	0.002	NS	34	465.14684	unknown	-	-	unknown	unknown
232	↓	0.002	NS	36	697.41401	$C_{33}H_{66}N_2O_7SP_2$	$M+H^+]^+$	0.19	unknown	unknown
233	↓	0.002	NS	39	305.24513	$C_{18}H_{34}O_2$	$[M+Na^+]^+$	0.07	Oleic acid	Fatty Acids
233	↓	0.002	NS	39	305.24513	$C_{18}H_{34}O_2$	$[M+Na^+]^+$	0.07	Elaidic acid	Fatty Acids
233	↓	0.002	NS	39	305.24513	$C_{18}H_{34}O_2$	$[M+Na^+]^+$	0.07	Vaccenic acid	Fatty Acids
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(16:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(16:1(9Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(16:1(9Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:1(11Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:1(9Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:2(9Z,12Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:2(9Z,12Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:3(6Z,9Z,12Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:3(6Z,9Z,12Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:3(9Z,12Z,15Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:3(9Z,12Z,15Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(18:4(6Z,9Z,12Z,15Z)/20:2(11Z,14Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:2(11Z,14Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:3(5Z,8Z,11Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:3(5Z,8Z,11Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:3(8Z,11Z,14Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:3(8Z,11Z,14Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:4(5Z,8Z,11Z,14Z)/18:2(9Z,12Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:4(8Z,11Z,14Z,17Z)/18:2(9Z,12Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(11Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(9Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(22:5(4Z,7Z,10Z,13Z,16Z)/16:1(9Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(22:5(7Z,10Z,13Z,16Z,19Z)/16:1(9Z))	Phosphatidylcholines
234	↓	0.002	NS	39	806.56947	$C_{46}H_{80}NO_8P$	$[M+H^+]^+$	0.08	PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:0)	Phosphatidylcholines
235	↓	0.002	NS	38	554.37212	$C_{27}H_{55}NO_8S$	$[M+H^+]^+$	0.01	unknown	unknown
236	↓	0.002	0.038	35	1022.79442	unknown	-	-	unknown	unknown
237	↓	0.002	NS	38	584.38272	$C_{28}H_{57}NO_9S$	$[M+H^+]^+$	0.07	unknown	unknown
238	↓	0.002	NS	39	569.31516	$C_{30}H_{52}O_7P_2$	$[M+H^+-H_2O]^+$	0.7	Presqualene diphosphate	Isopranes
239	↓	0.002	NS	39	569.31516	$C_{30}H_{52}O_7P_2$	$[M+H^+-H_2O]^+$	0.7	All-trans-hexaprenyl diphosphate	Isopranes
240	↓	0.002	NS	39	540.15957	unknown	-	-	unknown	unknown
241	↓	0.002	NS	39	376.31852	$C_{22}H_{43}NO_2$	$[M+Na^+]^+$	-0.22	unknown	unknown
242	↓	0.002	NS	16	794.49089	$C_{50}H_{68}NO_5P$	$[M+H^+]^+$	0.13	unknown	unknown
243	↓	0.002	NS	38	538.15475	unknown	-	-	unknown	unknown
244	↓	0.002	NS	9	462.30574	unknown	-	-	unknown	unknown
245	↓	0.002	0.038	37	467.08667	$C_{22}H_21O_8P$	$[M+Na^+]^+$	0.1	unknown	unknown
246	↓	0.002	NS	39	474.28477	unknown	-	-	unknown	unknown
247	↓	0.002	NS	39	557.09538	$C_{22}H_{22}N_4O_{10}S$	$[M+Na^+]^+$	0.88	unknown	unknown
248	↓	0.002	NS	14	653.38819	$C_{34}H_{62}O_6S_2$	$[M+Na^+]^+$	0.29	unknown	unknown

Entry	Insulin sensitive (IS)	Regression coefficient	p-value	Freq	Raw mass m/z	Formula	Ion Type	Error (ppm)	Metabolite assignment	Compound class
249	↓	0.002	NS	37	606.39605	$C_{29}H_{55}N_3O_{10}$	$[M+H^+]^+$	0.05	unknown	unknown
250	↓	0.002	0.046	39	469.1097	$C_{14}H_{26}N_2O_{12}S$	$[M+Na^+]^+$	-0.36	unknown	unknown
251	↓	0.002	NS	38	945.81235	unknown	-	-	unknown	unknown
252	↓	0.002	NS	11	931.7448	unknown	-	-	unknown	unknown
253	↓	0.002	NS	38	539.16651	$C_{22}H_{32}N_2O_{10}S$	$[M+Na^+]^+$	-0.9	unknown	unknown
254	↓	0.002	NS	39	406.2926	unknown	-	-	unknown	unknown
255	↓	0.002	NS	25	1215.75298	unknown	-	-	unknown	unknown
256	↓	0.002	NS	10	811.39003	$C_{33}H_{65}N_4O_{13}SP$	$[M+Na^+]^+$	0.2	unknown	unknown
257	↓	0.002	NS	39	764.57617	unknown	-	-	unknown	unknown
258	↓	0.002	NS	35	693.45645	unknown	-	-	unknown	unknown
259	↓	0.002	NS	39	613.34167	$C_{27}H_{50}N_4O_{10}$	$[M+Na^+]^+$	-0.4	unknown	unknown

Table 25: Correlation between discriminative markers and Fasting glucose at base line. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Fasting Glc		Fasting Glc		Fasting Gluc	
				Corr	p-val	Corr	p-val	Corr	p-val
Unknown	unknown	286.1292	↑	0.387983	0.0134	0.4952	0.0264	-	-
LysoPC(16:1(9Z))	Lysophosphatidylcholines	494.3242	↑	-0.31218	0.0499	-	-	-	-
Unknown	unknown	496.3555	↑	-	-	-0.4651	0.0388	-	-
Unknown	unknown	496.3256	↑	-0.32261	0.0423	-0.4901	0.0282	-	-
Unknown	unknown	216.0963	↓	-	-	-	-	0.6001	0.0052
Unknown	unknown	216.0854	↓	-	-	-	-	0.596592	0.005491
Unknown	unknown	198.1852	↓	-	-	-	-	0.583479	0.00692
Unknown	unknown	215.093	↓	-	-	-	-	0.58274	0.007009
Unknown	unknown	196.1696	↓	-	-	-	-	0.544564	0.013042
Unknown	unknown	214.0919	↓	-	-	-	-	0.530183	0.016188
Unknown	unknown	214.0937	↓	-	-	-	-	0.5098	0.0217
Unknown	unknown	158.027	↓	-	-	-	-	0.479094	0.032576
Unknown	unknown	214.0848	↓	-	-	-	-	0.479	0.0326
Unknown	unknown	214.0865	↓	-	-	-	-	0.462809	0.039894
Unknown	unknown	397.1242	↓	-	-	-0.4484	0.0474	-	-

Table 26 : Correlation between discriminative markers and Liver fat at base line. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	RAW m/z	Insulin sensitive (IS)	IS and IR		IS only		IR only	
				Liver fat		Liver fat		Liver fat	
				Corr	p-val	Corr	p-val	Corr	p-val
Unknown	Unknown	298.042	↑	-	-	-	-	0.517	0.0196
Unknown	Unknown	369.352	↑	-	-	-	-	0.4669	0.0379
Unknown	Unknown	711.882	↑	-	-	-	-	0.4645	0.0391
Unknown	Unknown	470.252	↑	-	-	-	-	0.4524	0.0452
Unknown	unknown	520.322	↑	-0.37	0.0188	-	-	-0.4738	0.0348
LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines	520.34	↑	-0.3929	0.0121	-	-	-0.5106	0.0214
L-Phenylalanine	Amino acids	166.086	↓	0.45639	0.0031	0.4895	0.0285	0.47122	0.03597
Indoleacrylic acid	Amino acids	188.071	↓	0.50167	0.001		-	0.59369	0.00578
L-Tyrosine	Amino acids	182.081	↓	0.39553	0.0115	-	-	0.4555	0.0436
Unknown	Unknown	285.29	↓	-	-	0.4518	0.0455	-	-
L-Tryptophan	Amino acids	205.097	↓	0.45912	0.0029	-	-	-	-
Unknown	unknown	720.506	↓	0.3239	0.0415	-	-	-	-
Unknown	unknown	719.503	↓	0.32136	0.0432	-	-	-	-

Table 27: Correlation between discriminative markers and ISI Matsuda index at base. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	RAW m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				ISI-Matsuda		ISI-Matsuda		ISI-Matsuda	
				Corr	p-val	Corr	p-val	Corr	p-val
Unknown	unknown	430.3891	↑	0.523959	0.0005	0.6157	0.0039	-	-
Unknown	unknown	402.3578	↑	0.518368	0.0006	0.5983	0.0053	-	-
Oleamide	Lipoamides	304.2611	↑	0.366358	0.0201	0.7268	0.0003	-	-
LysoPC(22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Lysophosphatidylcholines	568.3399	↑	0.494596	0.0012	0.6042	0.0048	-	-
LysoPE(0:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Lysophosphatidylethanolamines	526.2928	↑	0.437624	0.0047	0.5762	0.0078	-	-
LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines	520.3397	↑	0.315422	0.0474	-	-	0.6615	0.0015
LysoPC(18:1(9Z))	Lysophosphatidylcholines	522.3553	↑	0.384338	0.0143	-	-	0.5793	0.0074
Unknown	unknown	520.3223	↑	0.327996	0.0388	-	-	0.5907	0.0061
Unknown	unknown	524.3572	↑	0.366733	0.0199	-	-	0.5088	0.022
Oleamide	Lipoamides	282.2791	↑	-	-	0.4653	0.0387	-	-
LysoPC(20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylcholines	544.3398	↑	0.481446	0.0017	-	-	-	-
LysoPC(20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylcholines	566.3217	↑	0.441271	0.0044	-	-	-	-
LysoPC(22:5(4Z,7Z,10Z,13Z,16Z))	Lysophosphatidylcholines	570.3555	↑	0.41457	0.0078	-	-	-	-
Unknown	unknown	497.3556	↑	0.389835	0.0129	-	-	-	-
Unknown	unknown	497.3522	↑	0.383825	0.0145	-	-	-	-
Unknown	unknown	478.3292	↑	0.375781	0.0169	-	-	-	-
Unknown	unknown	503.2962	↑	0.371759	0.0182	-	-	-	-
LysoPE(0:0/20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylethanolamines	502.2929	↑	0.362182	0.0216	-	-	-	-
LysoPE(0:0/20:0)	Lysophosphatidylethanolamines	510.3555	↑	0.351004	0.0264	-	-	-	-
unknown	unknown	279.647	↑	0.349084	0.0273	-	-	-	-
unknown	unknown	527.3129	↑	0.337489	0.0332	-	-	-	-
LysoPC(16:0)	Lysophosphatidylcholines	496.3397	↑	0.328887	0.0383	-	-	-	-
LysoPC(18:0)	Lysophosphatidylcholines	524.371	↑	0.326002	0.0401	-	-	-	-

Metabolite assignment	Compounds class	RAW m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				ISI-Matsuda		ISI-Matsuda		ISI-Matsuda	
				Corr	p-val	Corr	p-val	Corr	p-val
unknown	unknown	526.378	↑	0.32395	0.0414	-	-	-	-
unknown	unknown	498.3469	↑	0.323345	0.0418	-	-	-	-
unknown	unknown	534.2958	↑	0.319496	0.0445	-	-	-	-
SM(d18:0/16:1(9Z))	Sphingolipids	703.5751	↑	0.317749	0.0457	-	-	-	-
LysoPE(0:0/16:0)	unknown	454.2929	↑	0.31622	0.0468	-	-	-	-
unknown	unknown	558.2958	↑	0.313928	0.0485	-	-	-	-
unknown	unknown	520.3464	↑	-	-	-0.4481	0.0475	-	-
Oleic acid	Fatty acids	305.2451	↓	0.359029	0.0229	-	-	-	-
Paraxanthine	exogenous	181.072	↓	-	-	0.5321	0.0157	-	-
unknown	unknown	407.0393	↓	-0.34426	0.0296	-	-	-	-

Table 28: Correlation between discriminative markers and CRP at base line study. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	RAW m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				CRP		CRP		CRP	
				Corr	p-val	Corr	p-val	Corr	p-val
unknown	unknown	545.4234	↑	-	-	0.6349	0.0026	-	-
unknown	unknown	546.4268	↑	-	-	0.6287	0.003	-	-
unknown	unknown	375.6605	↑	-	-	0.5374	0.0145	-	-
LysoPC(20:4(5Z,8Z,11Z,14Z))	Lysophosphatidylcholines	544.3398	↑	-0.33126	0.0368	-	-	-	-
LysoPC(18:1(9Z))	Lysophosphatidylcholines	522.3553	↑	-0.33457	0.0348	-	-	-	-
LysoPC(15:0)	Lysophosphatidylcholines	482.3242	↑	-0.3475	0.028	-	-	-	-
LysoPC(14:0)	Lysophosphatidylcholines	468.3085	↑	-0.35726	0.0236	-	-	-	-
LysoPC(20:3(5Z,8Z,11Z))	Lysophosphatidylcholines	546.3555	↑	-0.37079	0.0185	-	-	-	-
unknown	unknown	524.3643	↑	-0.37119	0.0184	-	-	-	-
LysoPE(0:0/20:0)	Lysophosphatidylethanolamines	510.3555	↑	-0.38249	0.0149	-	-	-	-
unknown	unknown	496.3622	↑	-0.38672	0.0137	-	-	-	-
unknown	unknown	498.3469	↑	-0.40719	0.0091	-	-	-	-
LysoPC(16:0)	Lysophosphatidylcholines	496.3397	↑	-0.41009	0.0086	-	-	-	-
LysoPC(22:5(4Z,7Z,10Z,13Z,16Z))	Lysophosphatidylcholines	570.3555	↑	-0.40818	0.0089	-	-	-0.5477	0.0124
unknown	unknown	215.0179	↓	0.37808	0.0162	0.4531	0.0448	-	-
unknown	unknown	203.1213	↓	0.360455	0.0223	0.4471	0.0481	-	-
unknown	unknown	646.5076	↓	-	-	0.6158	0.0038	-	-
unknown	unknown	647.5111	↓	-	-	0.6076	0.0045	-	-
unknown	unknown	230.9901	↓	-	-	0.5886	0.0063	-	-
unknown	unknown	614.5177	↓	-	-	0.5761	0.0079	-	-
unknown	unknown	1051.801	↓	-	-	0.5293	0.0164	-	-
unknown	unknown	1052.804	↓	-	-	0.5271	0.0169	-	-
unknown	unknown	486.2851	↓	-	-	0.5196	0.0189	-	-
unknown	unknown	217.1369	↓	-	-	0.4927	0.0273	-	-
Hexadecanedioic acid	Dicarboxylic fatty acids	309.2036	↓	-	-	0.4701	0.0365	-	-
unknown	unknown	352.097	↓	-	-	0.4568	0.0429	-	-

Metabolite assignment	Compounds class	RAW m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				CRP		CRP		CRP	
				Corr	p-val	Corr	p-val	Corr	p-val
unknown	unknown	238.0673	↓	0.418441	0.0072	-	-	-	-
unknown	unknown	237.0749	↓	0.410638	0.0085	-	-	-	-
unknown	unknown	236.0715	↓	0.41046	0.0085	-	-	-	-
unknown	unknown	236.9981	↓	0.398332	0.0109	-	-	-	-
unknown	unknown	215.0173	↓	0.387064	0.0136	-	-	-	-
unknown	unknown	449.154	↓	0.375124	0.0171	-	-	-	-
Citric acid	Carboxylic acids	215.0162	↓	0.347591	0.028	-	-	-	-
unknown	unknown	429.0252	↓	0.341925	0.0308	-	-	-	-
L-Phenylalanine	Amino acids	166.0863	↓	0.329585	0.0378	-	-	-	-
L-Tyrosine	Amino acids	182.0812	↓	0.326934	0.0395	-	-	-	-
2-Methylcitric acid	Carboxylic acids	229.0319	↓	0.316849	0.0464	-	-	-	-

Table 29: Correlation between discriminative markers and Fasting glucose at base line. Experiment performed via UPLC-Q-TOF-MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		IS		IR	
					Fasting Glc		Fasting Glc		Fasting Glc	
					Corr	p_val	Corr	p-Val	Corr	p-Val
LysoPC(18:0)	Lysophosphatidylcholines	524.371	15.4316	↑	-	-	0.4795	0.0002	-	-
LysoPC(0:0/18:0)	Lysophosphatidylcholines	524.371	15.4316	↑	-	-	0.4795	0.0002	-	-
PC(14:0/20:3(5Z,8Z,11Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(14:0/20:3(8Z,11Z,14Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(14:1(9Z)/20:2(11Z,14Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(16:0/18:3(6Z,9Z,12Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(16:0/18:3(9Z,12Z,15Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(16:1(9Z)/18:2(9Z,12Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(18:2(9Z,12Z)/16:1(9Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(18:3(6Z,9Z,12Z)/16:0)	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(18:3(9Z,12Z,15Z)/16:0)	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(20:2(11Z,14Z)/14:1(9Z))	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(20:3(5Z,8Z,11Z)/14:0)	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
PC(20:3(8Z,11Z,14Z)/14:0)	Phosphatidylcholines	756.5541	21.1687	↑	-	-	0.3756	0.004	-	-
2-Octenoylcarnitine	-	268.1907	6.1552	↑	-	-	-	-	0.3057	0.0012
Metoprolol	Secondary amino alcohols	268.1907	6.1552	↑	-	-	-	-	0.3057	0.0012
LysoPC(18:0)	Lysophosphatidylcholines	524.371	15.6214	↓	-	-	0.4795	0.0002	-	-
LysoPC(P-18:1(9Z))	Lysophosphatidylcholines	506.3605	19.0248	↓	-	-	0.3942	0.0024	-	-
LysoPC(18:0)	Lysophosphatidylcholines	506.3605	19.0248	↓	-	-	0.3942	0.0024	-	-
LysoPC(0:0/18:0)	Lysophosphatidylcholines	506.3605	19.0248	↓	-	-	0.3942	0.0024	-	-
PC(14:0/20:3(5Z,8Z,11Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(14:0/20:3(8Z,11Z,14Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(14:1(9Z)/20:2(11Z,14Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		IS		IR	
					Fasting Glc		Fasting Glc		Fasting Glc	
					Corr	p_val	Corr	p-Val	Corr	p-Val
PC(16:0/18:3(6Z,9Z,12Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(16:0/18:3(9Z,12Z,15Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(16:1(9Z)/18:2(9Z,12Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(18:2(9Z,12Z)/16:1(9Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(18:3(6Z,9Z,12Z)/16:0)	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(18:3(9Z,12Z,15Z)/16:0)	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(20:2(11Z,14Z)/14:1(9Z))	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(20:3(5Z,8Z,11Z)/14:0)	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-
PC(20:3(8Z,11Z,14Z)/14:0)	Phosphatidylcholines	756.5541	21.2813	↓	-	-	0.3756	0.004	-	-

Table 30: Correlation between discriminative markers and Liver fat at base line. Experiment performed via UPLC-Q-TOF-MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		IS		IR	
					Liver fat		Liver fat		Liver fat	
					Corr	p-val	Corr	p-Val	Corr	p-Val
MG(0:0/24:6(6Z,9Z,12Z,15Z,18Z,21Z)/0:0)	Monoacylglycerols	413.3051	19.6662	↑	-	-	0.3221	0.0145	-	-
MG(24:6(6Z,9Z,12Z,15Z,18Z,21Z)/0:0/0:0)	Monoacylglycerols	413.3051	19.6662	↑	-	-	0.3221	0.0145	-	-
7 alpha-Hydroxy-3-oxo-4-cholestenoate	Bile acids	413.3051	19.6662	↑	-	-	0.3221	0.0145	-	-
LysoPE(0:0/18:2(9Z,12Z))	Lysophosphatidylethanolamines	478.2929	15.4993	↑	-	-	0.316	0.0166	-	-
LysoPE(18:2(9Z,12Z)/0:0)	Lysophosphatidylethanolamines	478.2929	15.4993	↑	-	-	0.316	0.0166	-	-
Palmitoleoyl Ethanolamide	endocannabinoids	320.256	16.5459	↓	-	-	-0.3002	0.0233	-	-

Table 31: Correlation between discriminative markers and ISI Matsuda index at base line. Experiment performed via UPLC-Q-TOF-MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		IS		IR	
					ISI Matsuda		ISI Matsuda		ISI Matsuda	
					Corr	p_val	Corr	p-Val	Corr	p-Val
LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines	520.3397	16.1693	↑	-	-	-	-	0.4867	0.001
PC(16:0/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:0/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:0/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:1(11Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:1(11Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:1(9Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:1(9Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:2(9Z,12Z)/20:2(11Z,14Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:3(6Z,9Z,12Z)/20:1(11Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:3(9Z,12Z,15Z)/20:1(11Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(18:4(6Z,9Z,12Z,15Z)/20:0)	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:0/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:1(11Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:1(11Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:2(11Z,14Z)/18:2(9Z,12Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:3(5Z,8Z,11Z)/18:1(11Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:3(5Z,8Z,11Z)/18:1(9Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:3(8Z,11Z,14Z)/18:1(11Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:3(8Z,11Z,14Z)/18:1(9Z))	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:4(5Z,8Z,11Z,14Z)/18:0)	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(20:4(8Z,11Z,14Z,17Z)/18:0)	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
PC(22:4(7Z,10Z,13Z,16Z)/16:0)	Phosphatidylcholines	810.6014	21.1706	↑	-	-	-0.3034	0.0218	-	-
LysoPE(0:0/18:2(9Z,12Z))	Lysophosphatidylethanolamines	478.2929	15.4993	↑	-	-	-0.3618	0.0057	-	-
LysoPE(18:2(9Z,12Z)/0:0)	Lysophosphatidylethanolamines	478.2929	15.4993	↑	-	-	-0.3618	0.0057	-	-
9-cis-Retinoic acid	Carotenoids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
13-cis-Retinoic acid	Long chain fatty acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		IS		IR	
					ISI Matsuda		ISI Matsuda		ISI Matsuda	
					Corr	p-val	Corr	p-Val	Corr	p-Val
4-Oxoretinol	Carotenoids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
4-OH-Retinal	Carotenoids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
Leukotriene A4	Leukotrienes	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
5-HEPE	Hydroxyeicosapolyenoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
12-HEPE	Hydroxyeicosapolyenoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
14,15-EpETE	Epoxyeicosatrienoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
15-HEPE	Hydroxyeicosapolyenoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
15-KETE	Oxoeicosapolyenoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
17,18-EpETE	Epoxyeicosatrienoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
5-KETE	Oxoeicosapolyenoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
12-KETE	Oxoeicosapolyenoic acids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
Oxymesterone	Ketosteroids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
Retinoic acid	Retinoids	301.2161	13.5133	↓	-	-	0.4087	0.0016	-	-
PC(18:1(11Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(18:1(9Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(18:2(9Z,12Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(18:2(9Z,12Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(18:3(6Z,9Z,12Z)/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(18:3(9Z,12Z,15Z)/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:2(11Z,14Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:3(5Z,8Z,11Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:3(5Z,8Z,11Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:3(8Z,11Z,14Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:3(8Z,11Z,14Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:4(5Z,8Z,11Z,14Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:4(5Z,8Z,11Z,14Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:4(8Z,11Z,14Z,17Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:4(8Z,11Z,14Z,17Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:2(11Z,14Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		IS		IR	
					ISI Matsuda		ISI Matsuda		ISI Matsuda	
					Corr	p-val	Corr	p-Val	Corr	p-Val
PC(22:4(7Z,10Z,13Z,16Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(22:4(7Z,10Z,13Z,16Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:2(9Z,12Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:2(9Z,12Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:1(11Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:1(9Z))	Phosphatidylcholines	832.5845	21.2876	↓	-	-	0.3929	0.0025	-	-
Bovinic acid	Long chain fatty acids	303.2295	16.5974	↓	-	-	0.3082	0.0197	-	-
9E,11E-Octadecadienoic acid	Long chain fatty acids	303.2295	16.5974	↓	-	-	0.3082	0.0197	-	-
10E,12Z-Octadecadienoic acid	Long chain fatty acids	303.2295	16.5974	↓	-	-	0.3082	0.0197	-	-
Linoelaidic acid	Long chain fatty acids	303.2295	16.5974	↓	-	-	0.3082	0.0197	-	-
Linoleic acid	Long chain fatty acids	303.2295	16.5974	↓	-	-	0.3082	0.0197	-	-
PC(16:0/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:0/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:0/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:1(11Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:1(11Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:1(9Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:1(9Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:2(9Z,12Z)/20:2(11Z,14Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:3(6Z,9Z,12Z)/20:1(11Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:3(9Z,12Z,15Z)/20:1(11Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(18:4(6Z,9Z,12Z,15Z)/20:0)	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:0/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:1(11Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		IS		IR	
					ISI Matsuda		ISI Matsuda		ISI Matsuda	
					Corr	p-val	Corr	p-Val	Corr	p-Val
PC(20:1(11Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:2(11Z,14Z)/18:2(9Z,12Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:3(5Z,8Z,11Z)/18:1(11Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:3(5Z,8Z,11Z)/18:1(9Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:3(8Z,11Z,14Z)/18:1(11Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:3(8Z,11Z,14Z)/18:1(9Z))	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:4(5Z,8Z,11Z,14Z)/18:0)	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(20:4(8Z,11Z,14Z,17Z)/18:0)	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
PC(22:4(7Z,10Z,13Z,16Z)/16:0)	Phosphatidylcholines	810.6014	21.2768	↓	-	-	-0.3034	0.0218	-	-
Citric acid	Short chain tricarboxylic acids	215.0162	1.2043	↓	-	-	-	-	-0.461	0.01
Isocitric acid	Short chain hydroxy acids	215.0162	1.2043	↓	-	-	-	-	-0.461	0.01
D-threo-Isocitric acid	Short chain tricarboxylic acids	215.0162	1.2043	↓	-	-	-	-	-0.461	0.01
Diketogulonic acid	Short chain hydroxy acids	215.0162	1.2043	↓	-	-	-	-	-0.461	0.01
2,3-Diketo-L-gulonate	Simple alcohols	215.0162	1.2043	↓	-	-	-	-	-0.461	0.01

Table 32: Correlation between discriminative markers and CRP at base line. Experiment performed via UPLC-Q-TOF-MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	RT	Insulin sensitive (IS)	IS and IR		only IS		only IR	
					CRP		CRP		CRP	
					Corr	p_val	Corr	p-val	Corr	p-val
LysoPC(20:2(11Z,14Z))	Lysophosphatidylcholines	548.371	18.8032	↑	-	-	-0.4205	0.0011	-	-
LysoPC(P-18:1(9Z))	Lysophosphatidylcholines	506.3605	19.0248	↓	-	-	0.3977	0.0022	-	-

Table 33: Correlation between discriminative markers and Fasting Glucose at follow-up. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Fasting Glc		Fasting Glc		Fasting Glc	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	520.3452	↑	0.544852	0.000335	-	-	0.6632	0.0014
unknown	unknown	487.2672	↑	-	-	0.4841	0.0357	0.6341	0.0027
unknown	unknown	522.3478	↑	0.490275	0.001533	-	-	0.5393	0.0141
LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines	520.34	↑	0.458007	0.003368	-	-	0.5484	0.0123
unknown	unknown	545.3448	↑	0.474155	0.002293	-	-	0.4618	0.0404
unknown	unknown	640.4126	↑	0.388106	0.01463	-	-	0.5505	0.0119
unknown	unknown	1039.676	↑	0.38721	0.01488	0.4641	0.0453	-	-
unknown	unknown	568.3303	↑	0.343699	0.032168	-	-	0.4676	0.0376
unknown	unknown	485.1123	↑	-	-	-	-	0.671	0.0012
unknown	unknown	409.1621	↑	-	-	-	-	0.6595	0.0016
unknown	unknown	489.2831	↑	-	-	-	-	0.613	0.0041
unknown	unknown	412.0936	↑	-	-	-	-	0.611	0.0042
unknown	unknown	412.0975	↑	-	-	-	-	0.5942	0.0057
unknown	unknown	410.1613	↑	-	-	-	-	0.5573	0.0107
unknown	unknown	337.0751	↑	-	-	-	-	0.5444	0.0131
unknown	unknown	461.2507	↑	-	-	-	-	0.539	0.0142
unknown	unknown	566.3282	↑	-	-	-	-	0.5345	0.0152
unknown	unknown	543.3309	↑	-	-	-	-	0.5298	0.0163
unknown	unknown	425.11	↑	-	-	-	-	0.5292	0.0164
unknown	unknown	271.161	↑	-	-	-	0	0.5158	0.0199
unknown	unknown	389.1121	↑	-	-	-	0	0.5067	0.0226
unknown	unknown	542.3333	↑	-	-	-	-	0.4981	0.0254
LysoPC(18:2(9Z,12Z))	Lysophosphatidylcholines	542.3219	↑	-	-	-	-	0.4969	0.0258

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Fasting Glc		Fasting Glc		Fasting Glc	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	543.3291	↑	-	-	0.4873	0.0343	-	-
unknown	unknown	684.439	↑	-	-	-	-	0.4943	0.0267
unknown	unknown	473.2879	↑	-	-	-	-	0.4821	0.0314
unknown	unknown	795.337	↑	-	-	-	-	0.4819	0.0314
unknown	unknown	595.3825	↑	-	-	-	-	0.4806	0.032
unknown	unknown	683.4356	↑	-	-	-	-	0.48	0.0322
unknown	unknown	797.3441	↑	-	-	-	-	0.48	0.0322
unknown	unknown	338.0745	↑	-	-	-	-	0.4778	0.0331
unknown	unknown	542.3275	↑	-	-	-	-	0.476	0.0339
unknown	unknown	796.3408	↑	-	-	-	-	0.4739	0.0348
unknown	unknown	566.3261	↑	-	-	0.4605	0.0473	-	-
unknown	unknown	445.1119	↑	-	-	0.4589	0.0481	-	-
unknown	unknown	534.3171	↑	-	-	0.4589	0.0481	-	-
unknown	unknown	812.3149	↑	-	-	0.4589	0.0481	-	-
unknown	unknown	727.4622	↑		-	-	-	0.468	0.0374
Ergosterol	Vitamin D3 derivates	435.3021	↑		-	-	-	0.4643	0.0392
Ergocalciferol	Steroids	435.3021	↑	-	-	-	-	0.4643	0.0392
4a-Methyl-5a-cholesta-8,24-dien-3-one	Steroids	435.3021	↑	-	-	-	-	0.4643	0.0392
3-Keto-4-methylzymosterol	Steroids	435.3021	↑	-	-	-	-	0.4643	0.0392
5-Dehydroepisterol	Steroids	435.3021	↑	-	-	-	-	0.4643	0.0392
unknown	unknown	409.1584	↑		-	-	-	0.4629	0.0399
unknown	unknown	728.4654	↑	-	-	-	-	0.4621	0.0402
unknown	unknown	795.3207	↑	-	-	-	-	0.4532	0.0448
unknown	unknown	639.409	↑	-	-	-	-	0.4487	0.0472
unknown	unknown	459.2355	↑	-	-	-	-	0.4482	0.0475

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Fasting Glc		Fasting Glc		Fasting Glc	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	544.3308	↑	-	-	-	-	0.4468	0.0482
unknown	unknown	1015.676	↑	0.425518	0.006921	-	-	-	-
unknown	unknown	678.4801	↑	0.416309	0.008386	-	-	-	-
unknown	unknown	723.5106	↑	0.415829	0.008469	-	-	-	-
unknown	unknown	722.5072	↑	0.407967	0.009934	-	-	-	-
unknown	unknown	496.3446	↑	0.375377	0.018533	-	-	-	-
unknown	unknown	766.5329	↑	0.37428	0.018906	-	-	-	-
unknown	unknown	634.4536	↑	0.366925	0.021579	-	-	-	-
unknown	unknown	767.5363	↑	0.364606	0.022485	-	-	-	-
unknown	unknown	522.356	↑	0.362814	0.023206		-	-	-
LysoPC(14:0)	Lysophosphatidylcholines	468.3086	↑	0.359177	0.024729	-	-	-	-
unknown	unknown	569.3452	↑	0.350693	0.028607	-	-	-	-
unknown	unknown	810.5595	↑	0.349044	0.029416	-	-	-	-
unknown	unknown	496.3354	↑	0.348998	0.029439	-	-	-	-
unknown	unknown	991.6759	↑	0.335184	0.036989	-	-	-	-
unknown	unknown	339.0541	↑	0.328996	0.04085	-	-	-	-
unknown	unknown	637.3066	↑	0.323613	0.044468	-	-	-	-
LysoPC(16:0)	Lysophosphatidylcholines	496.3398	↑	0.319081	0.047712	-	-	-	-
Docosa-4,7,10,13,16-pentaenoyl carnitine	Carnitines	496.3398	↑	0.319081	0.047712	-	-	-	-
Clupanodonyl carnitine	Carnitines	496.3398	↑	0.319081	0.047712	-	-	-	-
unknown	unknown	167.1042	↑	0.317311	0.04903	-	-	-	-
unknown	unknown	546.3571	↑	0.316652	0.049528	-	-	-	-
unknown	unknown	854.5861	↑	0.316269	0.049819	-	-	-	-
unknown	unknown	409.1653	↓	-	-	-	-	0.666	0.0013
unknown	unknown	746.6479	↓	-	-	0.5278	0.0202	-	-

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Fasting Glc		Fasting Glc		Fasting Glc	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	697.414	↓	-	-	0.5174	0.0233	-	-
unknown	unknown	641.5918	↓	-	-	0.4873	0.0343	-	-
unknown	unknown	646.6258	↓	-	-	-		0.4884	0.0289
12-Ketodeoxycholic acid	Bile acids	429.2401	↓	-	-	-	-	0.4555	0.0436
7-Hydroxy-3-oxocholanoic acid	Bile acids	429.2401	↓		-	-	-	0.4555	0.0436
Nutriacholic acid	Bile acids	429.2401	↓	-	-	-	-	0.4555	0.0436
7a-Hydroxy-3-oxo-5b-cholanoic acid	Bile acids	429.2401	↓		-	-	-	0.4555	0.0436
unknown	unknown	747.5019	↓	-	-	-	-	0.4509	0.046
unknown	unknown	811.39	↓	-0.34535	0.031298	-	-	-	-
unknown	unknown	370.1555	↓	-	-	-	-	-0.4459	0.0488
unknown	unknown	327.1415	↓	-	-	-	-	-0.4514	0.0457
unknown	unknown	369.1518	↓	-	-	-	-	-0.4525	0.0452
unknown	unknown	403.2328	↓	-	-	-	-	-0.5123	0.0209

Table 34: Correlation between discriminative markers and CRP at follow-up. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		only IS		only IR	
				CRP		CRP		CRP	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	615.1406	↑	-	-	-	-	0.6344	0.0027
unknown	unknown	592.3384	↑	-	-	0.5406	0.0169	-	-
unknown	unknown	543.1181	↑	-	-	-	-	0.5357	0.0149
unknown	unknown	446.2289	↑	0.372194	0.019635	-	-	-	-
unknown	unknown	473.2513	↑	0.347245	0.03032	-	-	-	-
unknown	unknown	241.1774	↑	-	-	-	-	-0.4626	0.04
unknown	unknown	763.608	↓	0.450453	0.004007	-	-	0.6164	0.0038
unknown	unknown	764.6113	↓	0.449733	0.004073	-	-	0.6157	0.0039
unknown	unknown	779.4724	↓	0.425929	0.006862	-	-	0.6082	0.0044
unknown	unknown	1056.751	↓	0.423836	0.007171	-	-	0.6017	0.005
unknown	unknown	780.476	↓	0.415632	0.008503	-	-	0.6058	0.0046
unknown	unknown	1055.748	↓	0.419342	0.007877	-	-	0.5943	0.0057
unknown	unknown	730.6339	↓	0.366793	0.021631	-	-	0.6061	0.0046
unknown	unknown	731.637	↓	0.368509	0.020979	-	-	0.6018	0.005
unknown	unknown	393.2974	↓	0.384478	0.015664	-	-	0.5498	0.012
unknown	unknown	542.1208	↓	-	-	-	-	0.5584	0.0105
unknown	unknown	609.3614	↓	-	-	-	-	0.5511	0.0118
unknown	unknown	605.4247	↓	-	-	0.5406	0.0169	-	-
unknown	unknown	541.1211	↓	-	-	-	-	0.541	0.0138
unknown	unknown	697.414	↓	-	-	-	-	0.5345	0.0152
unknown	unknown	653.3882	↓	-	-	-	-	0.5237	0.0178
unknown	unknown	467.1105	↓	-	-	-	-	0.5136	0.0205
unknown	unknown	470.0983	↓	-	-	-	-	0.5124	0.0209
unknown	unknown	467.093	↓	-	-	-	-	0.5123	0.0209
unknown	unknown	541.1266	↓	-	-	-	-	0.507	0.0225
unknown	unknown	470.1022	↓	-	-	-	-	0.4949	0.0265
unknown	unknown	467.102	↓	-	-	-	-	0.4915	0.0277
unknown	unknown	468.1019	↓	-	-	-	-	0.4883	0.0289
unknown	unknown	467.1091	↓	-	-	-	-	0.486	0.0298
unknown	unknown	467.1062	↓	-	-	-	-	0.4858	0.0299

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		only IS		only IR	
				CRP		CRP		CRP	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	541.1133	↓	-	-	-	-	0.485	0.0302
unknown	unknown	541.1285	↓	-	-	-	-	0.4821	0.0314
unknown	unknown	471.0961	↓	-	-	-	-	0.4793	0.0325
unknown	unknown	469.095	↓	-	-	-	-	0.4772	0.0334
unknown	unknown	467.119	↓	-	-	-	-	0.4688	0.0371
unknown	unknown	542.1252	↓	-	-	-	-	0.4679	0.0375
unknown	unknown	468.1056	↓	-	-	-	-	0.4631	0.0398
unknown	unknown	778.6188	↓	-	-	-	-	0.4556	0.0435
unknown	unknown	483.0791	↓	-0.32813	0.041413	-	-	-	-

Table 35: Correlation between discriminative markers and Liver fat at follow-up. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Liver fat		Liver fat		Liver fat	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	473.2879	↑	-	-	0.5411	0.0167	-	-
unknown	unknown	446.2289	↑	-	-	0.4653	0.0447	-	-
unknown	unknown	461.2507	↑	-	-	0.4571	0.0491	-	-
unknown	unknown	501.3404	↑	0.393767	0.013131	-	-	-	-
unknown	unknown	529.3715	↑	0.378654	0.017453	-	-	-	-
PC(16:1(9Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(18:2(9Z,12Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(18:3(6Z,9Z,12Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(18:3(6Z,9Z,12Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(18:3(9Z,12Z,15Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(18:3(9Z,12Z,15Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(18:4(6Z,9Z,12Z,15Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(18:4(6Z,9Z,12Z,15Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(20:3(5Z,8Z,11Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(20:3(8Z,11Z,14Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(20:4(5Z,8Z,11Z,14Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(20:4(5Z,8Z,11Z,14Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(20:4(8Z,11Z,14Z,17Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(20:4(8Z,11Z,14Z,17Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:2(9Z,12Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:1(9Z))	Phosphatidylcholines	804.5538	↑	-0.34234	0.032901	-	-	-	-
unknown	unknown	660.2921	↑	-0.35795	0.025261	-	-	-	-
unknown	unknown	659.2887	↑	-0.35843	0.02505	-	-	-	-
PC(18:3(6Z,9Z,12Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024
PC(18:3(9Z,12Z,15Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024
PC(18:4(6Z,9Z,12Z,15Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024
PC(18:4(6Z,9Z,12Z,15Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024
PC(20:4(5Z,8Z,11Z,14Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024
PC(20:4(8Z,11Z,14Z,17Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Liver fat		Liver fat		Liver fat	
				Corr	p-val	Corr	p-val	corr	p-val
PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024
PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	802.5379	↑	-	-	-	-	-0.5024	0.024
PC(18:2(9Z,12Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(18:3(6Z,9Z,12Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(18:3(6Z,9Z,12Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(18:3(9Z,12Z,15Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(18:3(9Z,12Z,15Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(18:4(6Z,9Z,12Z,15Z)/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:3(5Z,8Z,11Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:3(8Z,11Z,14Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:4(5Z,8Z,11Z,14Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:4(5Z,8Z,11Z,14Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:4(8Z,11Z,14Z,17Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:4(8Z,11Z,14Z,17Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(22:4(7Z,10Z,13Z,16Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	830.5698	↑	-0.36929	0.020688	-	-	-0.5203	0.0187
unknown	unknown	403.2328	↓	-	-	-	-	0.4482	0.0475
unknown	unknown	764.5762	↓	-0.35231	0.027831	-	-	-	-
PC(16:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(16:1(9Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(16:1(9Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:1(11Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:1(9Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:2(9Z,12Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:2(9Z,12Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:3(6Z,9Z,12Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:3(6Z,9Z,12Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Liver fat		Liver fat		Liver fat	
				Corr	p-val	Corr	p-val	corr	p-val
PC(18:3(9Z,12Z,15Z)/20:3(5Z,8Z,11Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:3(9Z,12Z,15Z)/20:3(8Z,11Z,14Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(18:4(6Z,9Z,12Z,15Z)/20:2(11Z,14Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:2(11Z,14Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:3(5Z,8Z,11Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:3(5Z,8Z,11Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:3(8Z,11Z,14Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:3(8Z,11Z,14Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:4(5Z,8Z,11Z,14Z)/18:2(9Z,12Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:4(8Z,11Z,14Z,17Z)/18:2(9Z,12Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(11Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(9Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(22:5(4Z,7Z,10Z,13Z,16Z)/16:1(9Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(22:5(7Z,10Z,13Z,16Z,19Z)/16:1(9Z))	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:0)	Phosphatidylcholines	806.5695	↓	-0.38051	0.016864	-	-	-	-
PC(14:0/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(14:0/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(14:1(9Z)/22:4(7Z,10Z,13Z,16Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(16:0/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(16:1(9Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(16:1(9Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:1(11Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:1(9Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:2(9Z,12Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:2(9Z,12Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:3(6Z,9Z,12Z)/18:2(9Z,12Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:3(9Z,12Z,15Z)/18:2(9Z,12Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:4(6Z,9Z,12Z,15Z)/18:1(11Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(18:4(6Z,9Z,12Z,15Z)/18:1(9Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(20:4(5Z,8Z,11Z,14Z)/16:1(9Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(20:4(8Z,11Z,14Z,17Z)/16:1(9Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(20:5(5Z,8Z,11Z,14Z,17Z)/16:0)	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				Liver fat		Liver fat		Liver fat	
				Corr	p-val	Corr	p-val	corr	p-val
PC(22:4(7Z,10Z,13Z,16Z)/14:1(9Z))	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(22:5(4Z,7Z,10Z,13Z,16Z)/14:0)	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029
PC(22:5(7Z,10Z,13Z,16Z,19Z)/14:0)	Phosphatidylcholines	780.5534	↓	-0.40041	0.011541	-	-	-0.4881	0.029

Table 36: Correlation between discriminative markers and ISI Matsuda index at follow-up. Experiment performed via ICR-FT/MS in positive ionization mode.

Metabolite assignment	Compounds class	Raw m/z	Insulin sensitive (IS)	IS and IR		IS		IR	
				ISI-Matsuda		ISI Matsuda		ISI Matsuda	
				Corr	p-val	Corr	p-val	corr	p-val
unknown	unknown	564.307	↑	-	-	-	-	0.4997	0.0249
PC(18:3(6Z,9Z,12Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(18:3(9Z,12Z,15Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(18:4(6Z,9Z,12Z,15Z)/22:5(4Z,7Z,10Z,13Z,16Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(18:4(6Z,9Z,12Z,15Z)/22:5(7Z,10Z,13Z,16Z,19Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(20:4(5Z,8Z,11Z,14Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(20:4(8Z,11Z,14Z,17Z)/20:5(5Z,8Z,11Z,14Z,17Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:4(5Z,8Z,11Z,14Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(20:5(5Z,8Z,11Z,14Z,17Z)/20:4(8Z,11Z,14Z,17Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(22:5(4Z,7Z,10Z,13Z,16Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(22:5(7Z,10Z,13Z,16Z,19Z)/18:4(6Z,9Z,12Z,15Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:3(6Z,9Z,12Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
PC(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/18:3(9Z,12Z,15Z))	Phosphatidylcholines	828.5537	↑	-	-	-	-	0.4518	0.0455
unknown	unknown	590.3225	↑	-	-	-	-	0.4458	0.0488
unknown	unknown	568.3303	↑	-	-	-	-	0.4452	0.0492
unknown	unknown	649.4511	↓	-0.32697	0.042183	-	-	-	-
unknown	unknown	693.4776	↓	-0.33125	0.03941	-	-	-	-
unknown	unknown	675.6776	↓	-	-	-	-	-0.4451	0.0492
unknown	unknown	696.6517	↓	-	-	-	-	-0.4518	0.0455
unknown	unknown	676.6808	↓	-	-	-	-	-0.4556	0.0435
unknown	unknown	693.4564	↓	-	-	-0.4839	0.0358	-	-
Unknown	unknown	723.5321	↓	-	-	-	-	-0.5298	0.0163
Unknown	unknown	722.5289	↓	-	-	-	-	-0.5486	0.0123
Unknown	unknown	724.5357	↓	-	-	-	-	-0.5508	0.0118

8. Bibliography

1. Villas-Boas S.G., Roessner U., Hansen M. A. E., Smedsgaard J., Nielsen J., *Metabolome Analysis: An Introduction* 2007, Wiley, Hobocken, New Jersey, USA.
2. Nicholson J. K., *J. Mol.Syst.Biol.* 2006, 2, 52.
3. Loscalzo J., Kohane I., Barabasi A-L., *Mol.Syst.Biol*.2007, 3, doi:10.1038/msb4100163.
4. Oliver S. G., Winson M. K., Kell D. B., Baganz F., *Tibtech* 1998, 16, 373-378.
5. Nicholson J.K., Lindon J. C., Holmes E., *Xenobiotica* 1999, 29, 1181-1189.
6. Nicholson J.K., Buckingham M. J., Sadler P. J., *Biochem. J.* 1983, 211, 605-615.
7. Nicholson J.K., Higham D. P., Timbrell J. A., Sadler P. J., *Molecular Pharmacology* 1985, 36, 398-404.
8. Bales J.R., Higham D.P., Howe I., Nicholson J. K., Sadler P. J., Clinical Chemistry 1984, 30,426-432.
9. Gartland K.P.R., Bonner F. W., Nicholson J.K., *Molecular Pharmacology* 1989, 35, 242-251.
10. Nicholson J.K., Wilson I. D., *Prog. NMR Spectrosc*. 1989, 21, 449-501.
11. Moka D., Vorreuther R., Schicha H., Spraul M., Humpfer E., Lipinski M., Foxall P. J., Nicholson J.K., Lindon J. C., *Journal of Pharaceutical and Biomedical Analysis* 1998, 17, 125-132.
12. Raamsdonk L.M., Teusink B., Broadhurst D., Zhang N., Hayes A., Walsh M. C., Berden J. A., Brindle K.M. , Kell D. B., Rowland J. J., Westerhoff H.V., van Dam K., Oliver S.G., *Nature Biotechnology* 2001, 19, 45-50.
13. Fiehn O., *Plant Molecular biology* 2002, 48, 155-171.
14. Nicholson J.K., Lindon J. C., *Nature* 2008, 455, 1054-1056.
15. Nicholson J.K., Connelly J., Lindon J. C., Holmes E., *Nat. Rev. Nat. Discov.* 2002, 1, 153-161.
16. Gavaghan C. L., Holmes E., Lenz E., Wilson I. D., Nicholson J. K., *FEBS Lett.* 2000, 484, 169-174.
17. Daviglus M.L., Stamler J., Pirzada A., Yan L.L., Garside D.B., Liu K., Wang R., Dyer A.R., Lloyd-Jones D.M., Greenland P., *JAMA* 2004, 292, 1588-1592.
18. Li M., Wang B., Zhang M., Rantalainen M., Wang S., Zhou H., Zhang Y., Shen J., Pang X., Zhang M., Wei H., Chen Y., Lu H., Zuo J., Su M., Qiu Y., Jia W., Xiao C., Smith L. M., Yang S., Holmes E., Tang H., Zhao G., Nicholson J. K., Li L., Zhao L., *PNAS* 2008, 105, 2117-2122.
19. Holmes E., Wilson I. D., Nicholson J.K., *Cell* 2008, 134, 714-717.
20. Nicholson J.K., , Wilson I. D., *Nat. Rev. Nat.Discov*.2003, 2, 668-676.
21. Nicholson J.K., Holmes E., Wilson I.D., *Nat. Rev Microbiol.* 2005, 3, 431-438.
22. Zheng X., Xie G., Zhao A., Zhao L., Yao C., Chiu N. H., Zhou Z., Bao Y., Jia W., Nicholson J. K., Jia W., *J. Proteome Res.* 2011, 10, 5512-5522.

23. Dumas M. E., Barton R. H., Toye A., Cloarec O., Blancher C., Rothwell A., Fearnside J., Tatoud R., Blanc V., Lindon J.C., Mitchell S. C., Holmes E., McCarthy M. I., Scott J., Gauguier D., Nicholson J. K., *PNAS* 2006, 103, 12511-12516.
24. Dumas M. E., Maibaum E. C., Teague C., Ueshima H., Zhou B., Lindon J. C., Nicholson J. K., Stamler J., Elliott P., Chan Q., Holmes E., *Anal. Chem.* 2006, 78, 2199-2208.
25. Brugman S., Klatter F.A, Visser J.T., Wildeboer-Veloo A.C., Harmsen H.J., Rozing J., Bos N. A., *Diabetologia* 2006, 49, 2105-2108.
26. Kind T., Scholz M., Fiehn O., *PlosOne* 2009, 4(5), e5440.
27. Wishart D.S., Knox C., Guo A. C., Eisner R., Young N., Gautam B., Hau D. D., Psychogios N., Dong E., Bouatra S., Mandal R., Sinelnikov I., Xia J., Jia L., Cruz J. A., Lim E., Sobsey C.A., Shrivastava S., Huang P., Liu P., Fang L., Peng J., Fradette R., Cheng D., Tzur D., Clements M., Lewis A., De Souza A., Zuniga A., Dawe M., Xiong Y., Clive D., Greiner R., Nazyrova A., Shaykhutdinov R., Li L., Vogel H.J., Forsythe I., *Nucleic Acids Res.* 2009, 37, D603-D610.
28. Nagarajan N., Pop M., *Methods Mol. Biol.* 2010, 673, 1-17.
29. De Souza A. G., MacCormack T. J., Wang N., Li L., Gross G.G., *Zebrafish* 2009, 6, 229-238.
30. Chen J., Zhao X., Fritsche J., Yin P., Schmitt-Kopplin Ph., Wang W., Lu X., Häring H. U., Schleicher E. D., Lehmann R., Xu G., *Anal. Chem.* 2008, 80, 1280-1289.
31. Lenz E. M., Bright J., Knight R., Wilson I.D., Major H., *Analyst.* 2004, 129, 535-541.
32. Wishart D.S., Tzur D., Knox C., Eisner R., Guo A. C., Young N., Cheng D., Jewell K., Arndt D., Sawhney S., Fung C., Nikolai L., Lewis M., Coutouly M. A., Forsythe I., Tang P., Shrivastava S., Jeroncic K., Stothard P., Amegbey G., Block D., Hau D. D., Wagner J., Miniaci J., Clements M., Gebremedhin M., Guo N., Zhang Y., Duggan G. E., Macinnis G. D., Weljie A. M., Dowlatabadi R., Bamforth F., Clive D., Greiner R., Li L., Marrie T., Sykes B. D., Vogel H. J., Querengesser L., *Nucleic Acids Res.* 2007, 35, D521-D526.
33. Wishart D.S., Lewis M. J., Morrissey J. A., Flegel M. D., Jeroncic K., Xiong Y., Cheng D., Eisner R., Gautam B., Tzur D., Sawhney S., Bamforth F., Greiner R., Li L., *J. Chromatogr. B* 2008, 871, 164-173.
34. Want E.J., O'Maille G., Smith C. A., Brandon T. R., Uritboonthai W., Qin C., Trauger S. A., Siuzdak G., *Anal. Chem.* 2006, 78, 3, 743-752.
35. Villas-Boas, S. G., Mas S., Akesson M., Smedsgaard J., Nielsen J., *Mass. Spectrom. Rev.* 2005, 24, 613-646.
36. Wilson I. D., Plumb R., Granger J., Major H., Williams R., Lenz E. A., *J. Chromatogr. B* 2005, 817, 67-76.
37. Lenz E. M., Wilson I.D., *J. Proteome Res.* 2007, 6, 443-458.
38. Lu X., Zhao X., Bai C., Zhao C., Lu G., Xu G., *J. Chromatogr. B* 2008, 866, 64-76.
39. Theodoridis G., Gika H.G., Wilson I. D., *Trends Anal. Chem.* 2008, 27, 251-260.
40. Wu Z., Huang Z., Lehman R., Zhao C., Xu G., *Chromatographia* 2009, 69, S23-S32.

41. Wishart D.S., *Bioanalysis* 2011, 3, 15, 1769-1782.
42. Dunn W. B., Broadhurst D., Brown M., Baker P. N., Redman C. W., Kenny L. C., Kell D. B., *J. Chromatogr. B* 2008, 871, 288-298.
43. Dunn W. B., Bailey N. Y., Johnson H. E. , *Analyst.* 2005, 130, 606-625.
44. Lindon J. C., Nicholson J. K., Holmes E., *The Handbook of Metabonomics and Metabolomics* 2007, Chapter 1, Elsevier, Amsterdam, The Netherlands, 1st. ed.
45. Wilson I.D., Wade K. E., Nicholson J. K., *Trends Anal. Chem*. 1989, 8, 368-374.
46. Bhattacharya A., *Nature* 2010, 463, 605-606.
47. Duarte, I. F., Legido-Quigley C., Parker D. A., Swann J. R., Spraul M., Braumann U., Gil A. M., Holmes E., Nicholson J. K., Murphy G. M., Vilca-Melendez H., Heaton N., Lindon J. C., *Mol. Biosyst.* 2009, 5, 180-190.
48. King R., Bonfiglio R., Fernandez-Metzler C., Miller-Stein C., Olah T., *J. Am. Soc. Mass Spectrom.* 2000, 11, 942–950.
49. Sterner, J. L. Sterner J. L., Johnston M.V., Nicol G. R., Ridge D. P., *J. Mass Spectrom.* 2000, 35, 385-391.
50. Jansson J., Willing B., Lucio M., Fekete A., Dicksved J., Halfvarson J., Tysk C., Schmitt-Kopplin P. , *PlosOne* 2009, Vol 4, e6386.
51. Lucio M., Fekete A., Weigert C., Wägele B., Zhao X., Chen J., Fritsche A., Häring H. U, Schleicher E. D., Xu G., Schmitt-Kopplin P, Lehmann R., *PlosOne* 2010, 5, e13317.
52. Schmitt-Kopplin P., Gabelica Z., Gougeon R. D., Fekete A., Kanawati B., Harir M., Gebefuegi I., Eckel G., Hertkorn N. , *PNAS* 2010, 107, 2763-2768.
53. Gougeon R., Lucio M., Frommberger M., Peyron D., Chassagne D., Alexandre H, Feuillat F., Voilley A., Cayot P., Gebefügi I., Hertkorn N, Schmitt-Kopplin P., *PNAS* 2009, 16, 9174-9179.
54. Liger-Belair G., Cilindre C., Gougeon R. D., Lucio M., Gebefügi I., Jeandet P., Schmitt-Kopplin P., *PNAS* 2009, 16, 16545-16549.
55. Breitling R., Pitt A. R., Barrett M. P., *Trends Biotechnol.* 2006, 24, 543–548.
56. Boernsen, K. O., Gatzek S., Imbert G., *Anal.Chem.* 2005, **77,** 7255–7264.
57. Lin L., Yu Q., Yan X., Hang W., Zheng J., Xing J., Huang B., *Analyst* 2010, 135, 2970-2978.
58. Leon C., Rodriguez-Meizoso I., Lucio M., Garcia-Cañas V., Ibañez E., Schmitt-Kopplin P., Cifuentes A., *J. Chromatogr. A* 2009, 1216, 7314-7232.
59. Wagner, S., Scholz K., Donegan M., Burton L., Wingate J., Völkel W., *Anal. Chem.* 2006, 78, 1296-1305.
60. Guillarme D., Schappler J., Rudaz S., Veuthey J.-L., *TrAC Trends Anal*. Chem. 2010, 29, 15-27.
61. Wilson I.D., Nicholson J. K., Castro-Perez J., Granger J. H., Johnson K. A., Smith B. W., Plumb R. S. , *J. Proteome Res.* 2005, 4, 591-598.
62. Plumb R. S., Granger J. H., Stumpf C. L., Johnson K. A., Smith B. W., Gaulitz S., Wilson I. D., Castro-Perez J., *Analyst* 2005, 130, 844-849.

63. Idborg H., Zamani L., Edlund P. O., Schuppe-Koistinen I., Jacobsson S. P., *J. Chromatogr. B* 2005, 828, 9-13.
64. Cubbon S., Bradbury T., Wilson J., Thomas-Oates J., *Anal. Chem.* 2007, 79, 8911-8918.
65. Gika H. G., Theodoridis G. A., Wilson I. D., *J. Sep. Sci.* 2008, 31, 1598-1608.
66. Spagou K., Tsoukali H., Raikos N., Gika H., Wilson I. D., Theodoridis G., *J. Sep. Sci.* 2010, 33, 716-727.
67. Wang Y., Lehmann R., Lu X., Zhao X., Xu G., *J. Chromatogr. A* 2008, 1204, 28-34.
68. Plumb R.S., Rainville P., Smith B. W., Johnson K. A., Castro-Perez J., Wilson I. D., Nicholson J. K., *Anal. Chem.* 2006, 78, 7278-7283.
69. Tolstikov V. V., Lommen A., Nakanishi K., Tanaka N., Fiehn O., *Anal. Chem.* 2003, 75, 6737-6740.
70. Granger J., Plumb R., Castro-Perez J., Wilson I. D., *Chromatographia* 2005, 61, 375-380.
71. Ding J., Sorensen C. M., Zhang Q., Jiang H., Jaitly N., Livesay E. A., Shen Y., Smith R. D., Metz T. O., *Anal. Chem.* 2007, 79, 6081-6093.
72. Bottcher C., Roepenack-Lahaye E. V., Willscher E., Scheel D., Clemens S., *Anal. Chem.* 2007, 79, 1507-1513.
73. Sana, T. R., Waddell K., Fischer S. M., *J. Chromatogr. B* 2008, 871, 314-321.
74. Broeckling C. D., Huhman D. V., Farag M. A., Smith J. T., May G. D., Mendes P., Dixon R. A., Sumner L. W., *J. Exp. Bot.* 2005, 56, 323-336.
75. Begley P., Francis-McIntyre S., Dunn W. B., Broadhurst D. I., Halsall A., Tseng A., Knowles J., HUSERMET Consortium, Goodacre R., Kell D. B., *Anal. Chem.* 2009, 81, 7038-7046.
76. Brunelli C., Bicchi C., Di Stilo A., Salomone A., Vincenti M., *J. Sep. Sci.* 2006, 29, 2765-2771.
77. Hasegawa Y., Kimura M., Yang Y., Verma I. C., Yamaguchi S., *Brain Dev.* 2005, 27, 39-45.
78. Kuhara T., *Mass Spectrom. Rev.* 2005, 24, 817-827.
79. Martin D. M., Toub O., Chiang A., Lo B. C., Ohse S., Lund S. T., Bohlmann J., *PNAS* 2009, 106, 7245-7250.
80. Shellie R. A., Welthagen W., Zrostliková J., Spranger J., Ristow M., Fiehn O., Zimmermann R., *J. Chromatogr. A* 2005, 1086, 83-90.
81. Welthagen W., Shellie R. A., Spranger J., Ristow M., Zimmermann R., Fiehn O., *Metabolomics* 2005, 1, 65-73.
82. Almstetter M. F. Appel I. J., Gruber M. A., Lottaz C., Timischl B., Spang R., Dettmer K., Oefner P. J., *Anal. Chem.*2009, 81, 5731-5739.
83. Takats Z., Wiseman J. M., Cooks R. G., *J. Mass Spectrom.* 2005, 40, 1261-1275.
84. Allwood J. W., Erban A., de Koning S., Dunn W. B., Luedemann A., Lommen A., Kay L., Löscher R., Kopka J., Goodacre R., *Metabolomics* 2009, 5, 479-496.
85. *The Human body. Blood Physiology and circulation* 2011. Britannica educational Publishing.

86. Nelson D. L., Cox, M. M., *Leningher Principles of Biochemistry* 2002, W. H. Freeman and Company.
87. Reed S., *Essential Physiological Biochemistry* 2009, Wiley-BlackWell.
88. Papaspyros N. S., *The History of Diabetes Mellitus* 1964, Georg Thieme Verlag.
89. Medvei V.C., *The History of Clinical Endocrinology: A Comprehensive Account of Endocrinology from Earliest Times to the present Day* 1993, 34-37, The Parthenon Publishing Group.
90. Southgate M. T., De medicina. *JAMA* 1999, 282, 921. doi:10.1001/jama.282.10.921.
91. Medvei V.C., *The History of Clinical Endocrinology: A Comprehensive Account of Endocrinology from Earliest Times to the present Day* 1993, 55-56, The Parthenon Publishing Group.
92. Medvei V.C., The *History of Clinical Endocrinology: A Comprehensive Account of Endocrinology from Earliest Times to the present Day* 1993, 97, The Parthenon Publishing Group.
93. Zajac J., Shrestha, A., Patel, P., Poretsky, L., The Main Events in the History of Diabetes Mellitus, 2010, Springer-Verlag US.
94. Bliss M., *The Discovery of Insulin* 1982, 33-39, Chicago: The University of Chicago Press.
95. Bliss M., *The Discovery of Insulin* 1982, 104-1128. Chicago: The University of Chicago Press.
96. MacCracken J., Hoel D., *Postgrad. Med.* 1997, 101, 138-140, 143-145, 149-150.
97. www.gene.com, *First Successful Laboratory Production of Human Insulin Announced,* press release September 6 1978.
98. Alberti K., De Fronzo, R.A., Zimmet, P, *International textbook of Diabetes Mellitus* 1997, 9-23, Wiley.
99. Minkowski O., *Diabetes* 1989, 38, 1-6.
100. The Diabetes Control and Complications Trial research Group, *N. Engl. J. Med.* 1993, 329, 997-986.
101. Matsuda M., DeFronzo R. A., *Diabetes Care* 1999, 22, 1462-1470.
102. Froguel P., Velho G., *Recent. Prog. Horm. Res.* 2001, 56, 91-105.
103. Narayan K. M., Boyle J. P., Thompson T. J., Gregg E. W., Williamson D. F., *Diabetes Care* 2007, 30, 6, 1562-1566.
104. Grundy S., *Endocrine* 2000, 13, 155-165.
105. Groop L., *Diabetes Obes. Metab.* 1999, 1 (Suppl 1), S1-S7.
106. Clark D., *Diabetes Care* 1997, 20, 1176-1182.
107. Stern M., *J. Clin. Invest.* 2000, 106, 3, 323-327.
108. Kaprio J., Tuomilehto J., Koskenvuo M., Romanov K., Reunanen A., Eriksson J., Stengård J., Kesäniemi Y. A., *Diabetologia* 1992, 35, 1060-1067.
109. Stumvoll M., Goldstein B. J., van Haeften T. W., *Lancet* 2005, 365, 1333-1346.
110. International Diabetes Federation. *IDF Diabetes Atlas, 5th edn.* Brussels,

Belgium: International Diabetes Federation, 2011.

111. Marshall A., Hendrickson C. L., Jackson G. S., *Mass Spectrometry Review* 1998, 17, 1-35
112. Chernushevich I. V., Loboda A. V., Thomson B. A., *J. Mass Spectrom.* 2001, 36, 849-865
113. Li X., Fekete A., Englmann M., Frommberger M., Lv S., Chen G., Schmitt-Kopplin Ph., *Analytical and Bioanalytical Chemistry* 2007, 389, 5, 1439-1446
114. Rainville P.D., *Chromatography Today* February/March (2010) 12.
115. Guilhaus M., *J. Mass Spectrom.* 1995, 30, 1519-1532,
116. Mamyrin B.A., *International Journal of mass Spectromety* 2001, 206, 3, 251- 266.
117. Southam A., Payne T. G., Cooper H. J., Arvanitis T. N., Viant M. R., *Anal. Chem.* 2007, 79, 12, 4595-4560.
118. Cuadros-Inostroza A., Giavalisco P., Hummel J., Eckardt A., Willmitzer L., Peña-Cortés H., *Anal.Chem.* 2010, 82, 9, 3573-3580.
119. Skoog D. A., Leary J. J., *Principles of instrumental Analysis* 1992, Chapter24, University Saunders College Publishing.
120. Quinn H., *Int. Patent number* 1997, WO97/16724.
121. Swartz M. E., *Journal of Liquid Chromatography & Related Technologies* 2005, 28, 1253–1263.
122. Appelblad P., *A practical Guide to HILIC* 2005-2008, Merck.
123. Grumbach E. S., Wheat T. E., Kele M., Mazzeo J. R., *Separation Sience Redefined* 2005.
124. Hoffmann E., Stroobant V., *Mass Spectrometry* 2007, Chapter 5, Wiley.
125. Robb D. B., Covey T. R., Bruins A. P., *Anal Chem* 2000, 72, 15, 3653-3659.
126. Bruins A. P., *TrAC* 1994, 13, 1, 37-43.
127. Cech N. B., Enke C. G., *Mass Spectrometry Reviews* 2001, 20, 362-387.
128. Cole R. B., Harrata A. K., *J. Am. Soc. Mass Spectrom.* 1993, 4, 546-556.
129. Oss M., Kruve A., Herodes K., Leito I., *Anal. Chem.* 2010, 82, 2865-2872.
130. Henriksen T., Juhler R. K., Svensmark B., Cech N. B., *J. Am. Soc. Mass Spectrom.* 2005, 16, 446-455.
131. Bruce S. J., Tavazzi I., Parisod V., Rezzi S., Kochhar S., Guy P. A., *Anal. Chem* . 2009, 81, 3285-3296.
132. Polson C., Sarkar P., Incledon B., Raguvaran V., Grant R., *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* 2003, 785, 263-275.
133. Tziotis D., Hertkorn N., Schmitt-Kopplin Ph., *Eur. J. Mass. Spectrom. (Chichester, Eng).* 2011, 17, 4, 415- 21.
134. Weisberg S., *Applied linear Regression* 1985, New York Wiley.
135. Martens H., Naes T., *Spectrochimica Acta* 1990, 46, 1541.
136. Tuck M.K., Chan D. W., Chia D., Godwin A. K., Grizzle W. E., Krueger K. E., Rom W., Sanda M., Sorbara L., Stass S., Wang W., Brenner D. E., *Journal of Proteome Research* 2009, 8, 113-117.

137. Mitchell B., Yasui Y., Li C. I., Fitzpatrick A. L., Lampe P. D., *Cancer. Inf.* 2005, 1, 1, 98-104.

138. Rai A., Greening D. W., Barnes T. W., Lim J. W., Tauro B. J., Rai A., Xu R., Adda C., MathivananS., Zhao W., Xue Y., Zhu H. J., Simpson R. J., *Proteomics* 2005, 5, 13, 3262-3277.

139. Mischak H., Apweiler R., Banks R. E., Conaway M., Coon J., Dominiczak A., Ehrich J. H., Fliser D., Girolam M., Hermjakob H., Hochstrasser D., Jankowski J., Julian Ba, Kolch W., Massy Z. A., Neusuess C., Novak J., Peter K., Rossing K., Schanstra J., Semmes O. J., Theodorescu D., Thongboonkerd V., Weissinger E. M., Van Eylk J. E., Yamamoto T., *Proteomics Clin. Appl.* 2007, 1, 2, 148-156.

140. Luque-Garcia J. L., Neubert T. A., *J. Chromatography A* 2007, 1153, 259-276.

141. Drake S. K., Bowen R. A., Remaley A. T., Hortin G. L., *Clin. Chem.* 2004, 50, 2398-2401.

142. Wickus G. G., Mordan R. J., Mathews E. A., *Clin. Chem.* 1993, 38, 2347.

143. Sampson M., Ruddel M., Albright S., Elin R. J., *Clin. Chem.* 1997, 43,675.

144. Koziol J. A., Feng A. C., Yu J., Griffin N. M., Schnitzler J. E., 2008, *J. Proteomics Bioinform.*, 6, 287-292.

145. Bland J. M., Altman D. G., *Statist. Meth. Med Res.*1999, 8, 135-160.

146. Bland J. M., Altman D. G., *J. Biopharm. Statist*. 2007, 17, 571-582.

147. Michopoulos F., Lai L., Gika H., Theodoridis G., Wilson I., *Journal of Proteome Research* 2009, 8, 2114-2121.

148. Want E. J., O'Maille G., Smith C. A., Brandon T. R., Uritboonthai W., Qin C., Trauger S. A., Siuzdak G., *Anal. Chem.* 2006, 78, 743-752.

149. Preiss D., Sattar N., *Clinical Science(London, England:1979)* 2008, 115, 141-150.

150. Matteoni C. A., Younossi T. M., Gramlich T., Boparai N., Liu Y. C., McCullough A. J., *Gastroenterology1999*, 116, 1413-1419.

151. Maxwell J. D., Sanderson I., Butler W. H., Gazet J. C., Pilkington T. R., *Br. Med. J.* 1977, 2, 726-729.

152. Ludwig J., Viggiano T. R., McGill D. B., Oh B. J., *Mayo clin. Proc.* 1980, 55, 434-438.

153. Angulo P., *N. Eng. J. Med.* 2002, 346, 1221-1231.

154. Ruhl C. E., Everhart J.E., *Clin. Liver Dis.* 2004, 3, 501-519.

155. Browning J. D., Szczepaniak L. S., Dobbins R., Nuremberg P., Horton J. D., Cohen J. C., Grundy S. M., Hobbs H. H., *Hepatology* 2004, 40, 1387-1395.

156. Fassio E., Alvarez E., Dominguez N., Landeira G., Longo C., *Hepatology* 2004, 40, 820-826.

157. Edmison J., McCullough A. J., *Clin. Liver Dis.* 2007, 11, 75-104.

158. Bellentani S., Saccoccio G., Masutti F., Grocè L. S., Brandi G., Sasso F., Cristanini G., Tiribelli C., *Ann. Intern Med.* 2000, 132, 112-117.

159. Andersen T., Chrostpffersen P., Cluud C., *Int. J. Obes.* 1984, 8, 107-115.

160. Bacon B. R., Farahvash M. J., Janney C. G., Neuschwander-Tetri B. A., *Gastroenterology* 1994, 107, 1103-1109.
161. Cortez-Pinto H., Camilo M. E., *Best Pract. Res. Clin. Gastroenterol.* 2004, 18, 1089-1104.
162. Arun J., Clements R. H., Lazenby A. J., Leeth R. R., Abrams G. A., *Obes. Surg.* 2006, 16, 1351-1358.
163. Weston S. R., Leyden W., Murphy, R., Bass, N.M., Bell, B.P., Manos, M.M., Terrault, N.A., Hepatology 2005, 41, 372-379.
164. Sherry B., Mei Z., Scanlon K. S., Mokdad A. H., Grummer-Strawn L. M., *Arch. Pediatr. Adolesc.Med.* 2004, 158, 1116-1124.
165. Fan J. G., Saibara T., Chitturi S., Kim B. I., Sung J. J., Chutaputti A., Asia-Pacific Working Party for NAFLD, *J. Gastroenterol. Hepatol.* 2007, 22, 794-800.
166. Hsieh S. D., Yoshinaga H., Sakurai Y., Kosaka K., *Int. J. Obes. Relat. Metab. Disord.* 2000, 24, 358-362.
167. Deurenberg P., Deurenberg-Yap M., Guricci S., *Obes. Rev.* 2002, 3, 141-146.
168. Roden M., Price T. B., Perseghin G., Petersen K. F., Rothman D. L., Cline G. W., Shulman G. I., *J. Clin. Invest.* 1996, 97, 2859-2865.
169. Stranges S., Dorn J. M., Muti P., Freudenheim J. L., Farinaro E., Russell M., Nochajski T. H., Trevisan M., *Hepatology* 2004, 39, 754-763.
170. Dries D. R., Gallegos L. L., Newton A. C., *J. Biol. Chem.* 2007, 282, 826-830.
171. Angulo P., Keach J. C., Batt K. P., Lindor K. D., *Hepatology* 1999, 30, 1356-1362.
172. Sanyal A. J., *Gastroenterology* 2002, 123, 1705-1725.
173. Iijima H., Moriyasu F., Tsuchiya K., Suzuki S., Yoshida M., Shimizu M., Sasaki S., Nishiguchi S., Maeyama S., *Hepatol. Res.* 2007, 37, 722-730.
174. Campbell M. S., Reddy K. R., *Aliment. Pharmacol. Ther.* 2004, 20, 249-259.
175. Donnelly K. L., Smith C. I., Schwarzenberg S. J., Jessurun J., Boldt M. D., Parks E. J., *J. Clin. Invest.* 2005, 115, 1343-1351.
176. Dresner A., Laurent D., Marcucci M., Griffin M. E., Dufour S., Cline G. W., Slezak L. A., Andersen D. K., Hundal R. S., Rothman D. L., Petersen K. F., Shulman G. I., *J. Clin. Invest.*1999, 103, 253-259.
177. Hotamisligil G. S., *Nature* 2006, 444, 860-867.
178. Shoelson S. E., Lee J., Goldfine A. B., *J. Clin. Invest.* 2006, 116, 1793-1801.
179. Day C. P., James O. F., *Gastronenterology* 1998, 114, 842-845.
180. Roden M., *Nat. Clin. Pract. Endocrinol. Metab.* 2006, 2, 335-348.
181. Zhang K., Shen X., Wu J., Sakaki K., Saunders T., Rutkowski D. T., Back S. H., Kaufman R. J., *Cell* 2006, 124, 587-599.
182. Nielsen S., Guo Z., Johnson C. M., Hensur D. D., Jensen M. D., *J. Clin. Invest.* 2004, 113, 1582-1588.
183. Tilg H., Kaser A., *Nat. Clin. Pract. Gastroenterol. Hepatol.* 2005, 2, 148-155.
184. Despres J. P., Lemieux I., *Nature* 2006, 444, 881-887.

185. Iizuka K., Bruick R. K., Liang G., Horton J. D., Uyeda K., *PNAS* 2004, 101, 7281-7286.
186. Dentin R., Benhamed F., Hainault I., Fauveau V., Foufelle F., Dyck J. R., Girard J., Postic C., *Diabetes* 2006, 55, 2159-2170.
187. Cortez-Pinto H., Jesus L., Barros H., Lopes C., Moura M. C., Camilo M. E., *Clin. Nutr.* 2006, 25, 816-823.
188. Huang M. A., Greenson J. K., Chao C., Anderson L., Peterman D., Jacobson J., Emick D., Lok A. S., Conjeevaram H. S., *Am. J. Gastroenterology* 2005, 100, 1072-1081.
189. Ryan M. C., Abbaso F., Lamendola C., Carter S., McLaughlin T. L., *Diabetes Care* 2007, 30, 1075-1080.
190. Luyckx F. H., Lefebvre P. J., Scheen A. J., *Diabetes Metab.* 2000, 26, 98-106.
191. Ouyang X., Cirillo P., Sautin Y., McCall S., Bruchette, J. L., Diehl A. M., Johnson R. J., Abdelmalek M. F., *J. Hepatol.* 2008, 48, 993-999.
192. Nguyen-Duy T. B., Nichaman M. Z., Church T. S., Blair S. N., Ross R., *Am. J. Physiol. Endocrinol. Metab.* 2003, 284, E1065-E1071.
193. Wei Y., Rector R. S., Thyfault J. P., Ibdah J. A., *World J. Gastroenterol.* 2008, 14, 193-199.
194. Schmilovitz-Weiss H., *The 58th annual meeting of the American Association For the Study of the Liver* 2007, Boston, MA.
195. Lewis J. R., *Dig. Dis. Sci.* 2010, 55, 560-578.
196. Misbin R. I., Green L., Stadel B. V., Gueriquian J. L., Gubbi A., Fleming G. A., *N. Engl. J. Med.* 1998, 338, 265-266.
197. Chang C. Y., Argo C. K., Al-Osaimi A. M., Caldwell S. H., *J. Clin. Gastroenterol.* 2006, 40 (Suppl 1), S51- S60.
198. Molloy R., Komorowski R., Varma R. R., *Liver Transpl Surg.* 1997, 3, 177-178.
199. Kahn S. E., Hull R. L., Utzschneider K. M., *Nature* 2006, 444, 840-846.
200. Stefan N., Kantartzis K., Häring H. U., *Endocrine Reviews* 2008, 29, 7, 939-960.
201. Mathews S. T., Chellam N., Srinivas P. R., Cintron V. J., Leon M. A., Goustin A. S., Grunberger G., *Mol. Cell Endocrinol.* 2000, 164, 87-98.
202. Mathews S. T., Singh G. P., Ranalletta M., Cintron V. J., Qiang X., Goustin A. S., Jen K. L., Charron M. J., Jahnen-Dechent W., Grunberger G., *Diabetes* 2002, 51, 2450-2458.
203. Stefan N., Fritsche A., Weickert C., Boeing H., Joost H. G., Häring H. U., Schulze M. B., *Diabetes* 2008, 57, 2762-2767.
204. Choi C. S., Fillmore J. J., Kim J. K., Liu Z. X., Kim S., Collier E. F., Kulkarni A., Distefano A., Hwang Y. J., Kahn M., Chen Y., Yu C., Moore I. K., Reznick R. M., Higashimori T., Shulman G. I., *J. Clin. Invest.* 2007, 117, 1995-2003.
205. Gentile C. L., Pagliassotti M-J., *Journal of Nutritional Biochemistry* 2008, 19, 567-576.
206. Stefan N., Häring H. U., *Diabetes* 2011, 60, 2011-2017.
207. Trygg J., Wold S., *Journal of Chemometrics*, 2002; 16: 119-128.
208. Wiklund S., Johansson E., Sjöström L., Mellerowicz E. J., Edlund U., Shockcor J. P., Gottfries J., Moritz T., Trygg J., *Anal. Chem.* 2008, 80, 115-122.

209. Theodoridis G., Gika H. G., Wilson I. D., *Mass Spectrometry Reviews* 2011, 30, 5, 884-906.

210. 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., O'Donnell C. J., Carr S. A., Mootha V. K., Florez J. C., Souza A., Melander O., Clish C. B., Gerszten R. E., *Nat. Med.* 2011, 17, 4, 448-453.

211. Krebs M., Krassak M., Bernroider E., Anderwld C., Brehm A., Meyerspeer M., Nowotny P., Roth E., Waldhäusl W., Roden M., *Diabetes* 2002, 51, 599-605.

212. Newgard C. B., An J., Bain J. R., Muehlbauer M. J., Stevens R. D., Lien L. F., Hagg A. M., Shah S. H., Ariotto M., Sientz C. A., Rochon J., Gallup D., Ilkayeva O., Wenner B. R., Yancy W. S. Jr., Eisenson H., Musante G., Surwit R. S., Millington D. S., Butler M. D., Svetkey L. P ., *Cell. Metab.* 2009, 9, 311-326.

213. Koopmans S. J., Ruis M., Dekker R., van Diepen H., Korte M., Mroz Z., *Physiol. Behav. 2005, 85, 469-478.*

214. Budohoski L., Challiss R. A., Dubaniewicz A., Kaciuba-Uscilko H., Leighton B., Lozeman F. J., Naza K., Newshome E. A., Porta S., *Biochem. J.* 1987, 244, 655-660.

215. Goldstein D. S., *Stress, Catecholamines, and Cardiovascular Disease* 1995, New York University Press.

216. Large V., Beylot M., *Diabetes* 1999, 48, 6, 1251-1257.

217. Spégel P., Malmgren S., Sharoyko W., Newsholme P., Koeck T., Mulder H., *Biochem. J.*, 2011, 435, 277–284.

218. Allen R., Stabler S. P., Savage D. G., Lindenbaum J., *Metabolism* 1993, 42, 8, 978-988.

219. Hosoya T., Matsushima M., Nukariya K., Utsunomiya K., *Intern Med* 2012, 51, 263-269.

220. Claudel T. Staels B., Kuipers F., *Arterioscler Thromb Vasc Biol.* 2006, 25, 2020-2030.

221. Steiner C., Othman A., Saely C. H., Rhein P., Drexel H., von Eckardstein A., Rentsch K. M., *PLoS ONE* 2011, 6,11, e25006.

222. Newsholme E., Leech T., *Functional Biochemistry in Health ad Disease* 2009, Chapter 7, 138, Wiley-Blackwell, ISBN: 978-0-471-98820-5.

223. Reddy J. K., Hashimoto T., *Annu. Rev. Nutr.* 2001, 21,193-230.

224. Kaiser E., *Clin. Lab.Sci.* 1999, 36, 65-163.

225. Six D. A., Dennis E. A., *Biophys. Acta* 2000, 1488, 1-19.

226. Matsumoto T., Kobayashi T., Kamata K., *Curr. Med. Chem.* 2007, 14, 3209-3220.

227. Hammerschmidt S., Büchler N., Wahn H., *Chest.* 2002, 121, 573-581.

228. Fuchs B., Schiller J., Wagner U., Häntzschel H., Arnold L., *Clin. Biochem.* 2005, 38, 925-933.

229. Schiller J., Fuchs B., Arnhold J., Arnold K., *Curr. Med. Chem*. 2003, 10, 2123-2145.

230. Kougias P., Chai H., Lin P. H., Lumsden A. B., Yao Q., Chen C., *Med. Sci. Monit.* 2006, 12, RA5-16.

231. Hong C-W., Song D-K., *Biomol. Ther.* 2008, 16, 69-76.

232. Engelmann B., Zieseniss S., Brand K., Page S., Lentschat A., Ulmer A. J., Gerlach E., *Arterioscler. Thromb. Vasc.* Biol. 1999, 19, 1, 47-53.

233. Yan J. J., Jung J. S., Lee J. E., Huh S. O., Kim H. S., Jung K. C., Cho J. Y., Nam J. S., Suh H. W., Kim Y. H., Song D. K., *Nat. med.* 2004, 10, 2, 161-167.

234. Müller J., Petkovic M., Schiller J., Arnold K., Reichle S., Arnhold J., *Luminescence* 2002, 17, 141-149.

235. Han M. S., Lim Y. M., Quan W., Kim J. R., Chung K. W., Kang M., Kim S., Park S. Y., Han J. S., Park S. Y., Cheon H. G., Dal Rhee S., Park T. S., Lee M. S., *Journal of Lipid Research* 2011, 52, 1234-1246.

236. Kantartzis K., Thamer C., Peter A., Machann J., Schick F., Schraml C., Königsrainer A., Königsrainer I., Kröber S., Niess A., Fritsche A., Häring H. U., Stefan N., *Gut* 2009, 58, 1281-1288.

237. Biaggioni I., Davis S. N., *Diabetes Care* 2002, 25, 2, 399-400 (editorial).

238. Keijzers G. B., De Galan B. E., Tack C. J., Smits P ., *Diabetes Care* 2002, 25, 2, 364-369.

239. Greer F., Hudson R., Graham T., *Diabetes* 2001, 50, 2349-2354.

240. Deibert D. C., DeFronzo R. A., *J. Clin. Invest.* 1980, 65, 717-721.

241. Van Soeren M., Mohr T., Kjaer M., Graham T. E., *J. App. Physiol.(1985)* 1996, 80, 999-1005.

242. Andersson P., Lind L., Berne C., Berglund L., Lithell H. O., *J. Intern. Med.* 1999, 246, 529-537.

243. Fredholm B. B., *Acta Physiol. Scand.* 1978, 102, 191-198.

244. Van Dam R. M., Hu F. B., *JAMA* 2005, 294, 97-104.

245. Urgert R., Schulz A. G., Kata M. B., *The American Journal of Clinical Nutrition.* 1995, 61, 149-154.

246. Salazar-Martinez E., Willet W. C., Ascherio A., Manson J. E., Leitzmann M. F., Stampfer M. J., Hu F. B., *Ann. Intern. Med.* 2004, 140, 1-8.

247. De Paulis T., Schmidt D. E., Bruchery A. K., Kriby M. T., McDonald M. P., Commers P., Martin P. R., *Eur. J. Pharmacol.* 2002, 442, 215-223.

248. Samuel V. T., Petersen K. F., Shulman G. I., *Lancet* 2010, 375, 2267-2277.

249. Goodridge A. G., *The Journal of Biological Chemistry* 1973, 248, 12, 4318-4326.

250. Adams S. H., Hoppel C. L., Lok K. H., Zhao L., Wong S. W., Minkler P. E., Hwang D. H., Newman J. W., Garvey W. T., *The Journal of Nutrition* 2009, 139, 1073-1081.

251. Chibalin A. V., Leng Y., Vieira E., Krook A., Björnholm M., Long Y. C., Kotova O., Zhong Z., Sakane F., Steiler T., Nylén C., Wang J., Laasko M., Topham M. K., Gilbert M., Wallberg-Henricksson H., Zierath J. R., *Cell* 2008, 132, 375-386.

252. Fuchs B., Schiller J., *Mini-Reviews in Medicinal Chemistry* 2009, 9, 368-378.

253. Hong D., Li X., Peng Q., Wang X., Chen J., Li Y., *Angew. Chem. Int. Ed.* 2005, 44, 2782 –2785.

254. Moritz F., Forcisi S., Harir M., Kanawati B., Lucio M., Tziotis D., Schmitt-Kopplin P., *The potential of ultrahigh resolution MS [FTICR-MS] in Metabolomics Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data* 2013, 117-136, Wiley-VCH.
255. Forcisi S., Moritz F., Kanawati B., Tziotis D., Lehmann R., Schmitt-Kopplin P., *J Chromatogr. A* 2013, 1292, 51-65.
256. Yin P., Peter A., Franken H., Zhao X., Neukamm S. S., Rosenbaum L., Lucio M., Zell A., Häring H. U., Xu G., Lehmann R., *Clin. Chem.* 2013, 59, 833-845.
257. Lehmann R., Franken H., Dammeier S., Rosenbaum L., Kantartzis K., Peter A., Zell A., Adam P., Li J., Xu G., Königsrainer A., Machann J., Schick F., Hrabé de Angelis M., Schwab M., Staiger H., Schleicher E., Gastaldelli A., Fritsche A., Häring H. U., Stefan N., *Diabetes Care* 2013, 36, 2331-2338.

9. Acknowledgements

I would like to thank Dr.Philippe Schmitt-Kopplin, for having offered to me the great opportunity to go on board and start this great adventure, working in such a wonderful and stimulating ambiente transmitting enthusiasm and precious knowledge.

I would like to thank Dr.Rainer Lehman for the valued work in the consortium organization and for his helpful suggestions.

I would like to thank Dr. Agnese Fekete that followed me with patience and care during my first steps of my work

I would like to thank you all my collegues, everybody of you is really unique and made my working here special.

A special thanks to Marianna for her collaboration in the analysis of the data and for her precious advices

A special thanks to Dimitris and Franco for their collaboration in the networking analysis of the data

An additional special thanks to Franco for all we share together everyday

Un ringraziamento speciale ai miei genitori per avermi dato l'opportunità di studiare e seguire sempre i miei interessi. Grazie per i principi che mi avete trasmesso, per avermi insegnato cosa sono la tenacia, il coraggio e che tutte le esperienze, anche se negative, possono sempre essere fonte di forza.

The author thanks the German Federal Ministry of Education and Research (BMBF), the Kompetenznetz Diabetes mellitus (Competence Network for Diabetes mellitus) funded by the German Federal Ministry of Education and Research (FKZ 01GI0804).

Scientific Contributions

Sara Forcisi, Kilian Wörmann, Agnes Fekete, Rainer Lehmann, Philippe Schmitt Kopplin HPLC 2009 Conference, June 28-July 2009, Dresden, Germany
Poster presentation
Comparison of different anticoagulants in plasma sample preparation in non-targeted metabonomics studies using UPLC-TOF MS

Sara Forcisi
Visiting guest, October 2009, National Chromatographic R. & A. Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China
Oral introduction to my project

Sara Forcisi, Kilian Wörmann, Agnes Fekete, Rainer Lehmann, Holger Franken, Erhan Kenar, Guowang Xu, Philippe Schmitt-Kopplin
1st International Symposiun on Metabolomics and More, March 10-12, 2010, Freising-Weihenstephan, Germany
Poster presentation
Plasma sample preparation in non-targeted metabonomics studies by UPLC-Q-TOF MS

Wörmann K., Lucio M., Forcisi S., Heinzmann S.S., Kenar E., Franken H., Rosenbaum L., Schmitt-Kopplin P., Kohlbacher O., Zell A., Häring H.-U., and Lehmann R., **Metabolomics in der Diabetesforschung**, *Der Diabetologe* 2012, 1-5, Springer-Verlag, DOI:10.1007/s11428-011-0778-9.

Wörmann K., Walker A., Moritz F., Forcisi S., Tziotis D., Lucio M., Heinzmann S. S., Adamski J., Lehmann R., Häring H. U., Schmitt-Kopplin P., **Revolution in der Diabetesdiagnostik dank –omics. Biomarker mittels Metabolomics** Revolution in Diabetes Diagnostics - Metabolomics for Discovering Biomarkers *Diabetes aktuell* 2012, 10, 129-133, Georg Thieme Verlag, Stuttgart /New-York, DOI: 10.1055/s-0032-1320054.

Kenar E., Franken H., Rosenbaum L., Lehmann R., Forcisi S., Wörmann K., Lucio M., König A., Rahnenführer J., Schmitt-Kopplin P., Häring H.-U., Zell A., Kohlbacher O., **Mit bioinformatik zu biomarkern**, *Medizinische Welt* 2012, 63, 245-250.

Moritz F., Forcisi S., Harir M., Kanawati B., Lucio M., Tziotis D., Schmitt-Kopplin P., **The potential of ultrahigh resolution MS [FTICR-MS] in Metabolomics Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data** 2013, 117-136, Wiley-VCH.

Forcisi S., Moritz F., Kanawati B., Tziotis D., Lehmann R., Schmitt-Kopplin P., **Liquid chromatography-mass spectrometry in metabolomics research: mass analyzers in ultra high pressure liquid chromatography coupling**, *J. Chromatogr. A* 2013, 1292, 51-65.

Kenar E., Franken H., Forcisi S., Wörmann K., Häring H. U., Lehmann R., Schmitt-Kopplin P., Zell A., Kohlbacher O., **Automated Label.Free Quantification of Metabolites from LC-MS Data**, *Mol Cell Proteomics* 2013, M113.031278, Paper in Press.

Curriculum Vitae

Sara Forcisi

Persönliche Informationen

Alter:	28 Jahre
Gerburtsdatum:	7. August 1983 (Gorizia, Italien)
Nationalität:	Italienisch
Email:	sara.forcisi@helmholtz-muenchen.de

BILDUNG

Dezember 2008 – April 2012

Doktorandin am Helmholtz Zentrum München

Arbeitstitel: Chromatographie und Massenspektrometrie basierte Metabolomik für Typ2 Diabetes mellitus Studien.

Betreuung: Priv.-Doz. Dr. Philippe Schmitte-Kopplin (Technische Universität München, Lehrstuhl für Analytische Lebensmittelchemie, Alte Akademie, Freising, Deutschland)

17. Juli 2008

Titel der erworbenen Qualifikation:

Abschluss zweiten Grades (Msc) in industrieller Biotechnologie (Spezialisierung: Umwelt)

Arbeitstitel: Enzymatisch kinetische Auflösung von 1-(2-hydroxy-5-mathyl-phenyl)ethylaminen

Betreuung: Prof. Giulia Licini (Abteilung for chemische Wissenschaften, Universität von Padova, Italien)

Name der Bildungseinrichtung: Universität von Padova, Italien

September 2005

Titel der erworbenen Qualifikation:

Abschluss ersten Grades (BSc) in Biotechnologie (Spezialisierung: Chemie)

Arbeitstitel: Biokatalytische Prozesse für die industrielle Produktion von chemischen Substanzen

Betreuung: Prof. Fernando Formaggio (Abteilung für chemische Wissenschaften, Universität von Padova, Italien)

Name der Bildungseinrichtung: Universität von Padova, Italien

1997 - 2002

Titel der erworbenen Qualifikation:

Maturità Scientifica (wissenschaftliches Abitur)

Name der Bildungseinrichtung: Liceo Scientifico Duca degli Abruzzi Gorizia, Italien

ARBEITSERFAHRUNG

Dezember 2008 – jetzt:

(IÖC, jetzt BGA, TUM), UPLC/SYNAPT-QTOF für ungerichtete Metabolomik in Prä-Diabetes Forschung

Oktober 2009 – November 2009:

Bildungsbesuch gestiftet von der Sino-Deutschen Kooperation im Labor für sensitive Separation & Charakterisierung von Biomolekülen, dalian Institut für chemische Physik, Dalian, China

Februar 2007 – Juli 2008:

Arbeitsgruppe für organische Chemie, Universität von Padova, Italien

Juni 2007 – Dezember 2007:

Labor für angewandte und Computerisierte Biokatalyse, Universität von Trieste, Italien

TRAINING UND KURSE

11.-12. Oktober 2011:
Wissenschaftliches Management, Training, Workshop (ReMaT)
Gehalten von TuTech Innovation, Helmholtz Assoziation von Deutschen Wissenschaftszentren, Abteilung Hamburg und Brüssel

28. Juni – 02. Juli 2009
Kurzkurs: „Student-Setting-up“ und Optimierung von Multidimensionalen Separationen
HPLC Konferenz 2009, Dresden, Deutschland

21-22 April 2009
4. Internationaler HILIC-Day und „Advanced HILIC User Meeting“
Halterm am See, Deutschland

FÄHIGKEITEN UND KOMPETENZ

Sprache:
Italienisch: Muttersprache
Englisch: flüssig in Schrift und Sprache
Deutsch: Grundstufe

Technische Fähigkeiten und Kompetenzen:
Probenaufbereitung und Manipulation von biologischen Proben, chemische Analyse mittels UPLC-TOF-MS, ICR-FT/MS, Nanospray Ionisierung, NMR, IR, GC-MS, Multivariate Analyse sowie Interpretation biologischer Daten
Synthese und Charakterisierung organsicher Substanzen
Enzym-Manipulation und Biotransformation

Computer-Fähigkeiten und Kompetenzen:
Betriebssystem: Windows
Anwendungen: MS Office (Word, Excel, Powerpoint)

Software zur Datenanalyse:

MarkerLynx (Waters), Data Analysis (Bruker), Topspin (Bruker), mzmine, MeV, JChem, Chem Draw, Datenbanken zur Massenannotierung (MassTRIX, ChemSpider), Datenbanken zur bibliographischen Recherche