LC/MS-BASED TARGETED AND GLOBAL METABOLOMIC METHODOLOGIES AND THEIR APPLICATION TO BIOMARKER DISCOVERY

By

ESTELA SOLEDAD CERUTTI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2009

To my son and beloved family

ACKNOWLEDGMENTS

This work would not have been possible without the guidance and encouragement I received from my academic advisors, Dr. Richard A. Yost and Dr. David H. Powell, under whose supervision I chose this topic and began the research. I would like to sincerely express that I have enriched my life from this experience in both personal and professional aspects.

I also like to express my gratitude to Dr. Peggy Borum, whose thoughtful advice often served to give me a sense of direction during my PhD studies. I would like to extend my gratitude to Dr. Ben Smith and Dr. John Eyler for their assistance and for spending their valuable time serving on my committee. I also want to thank the members of Powell's (Julia, Jodie, Cris, Basri, Noelle, and Jacina) and Yost's groups for their friendship and helpful advice in situations related to science, research, and life.

I am deeply grateful to my academic institution in Argentina and the Fulbright Commission for the trust and support that they gave me in order to study in the United States. I am tempted to thank all of my friends individually; however, because the list might be too long and by fear of leaving someone out, I will simply say thank you very much to you all.

I give thanks from the bottom of my heart to my parents, Lila and Luis, for providing me with the values to discover and achieve my goals in life. My gratitude also goes to my beloved sister, Laura, for her continuous support and encouragement to find the positive side of everything. My family and family-to-be deserve so much of the credit for my success. I cannot end without thanking God for giving me the strength and the opportunity to achieve one of the most important dreams in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

LC/MS-BASED TARGETED AND GLOBAL METABOLOMIC METHODOLOGIES AND THEIR APPLICATION TO BIOMARKER DISCOVERY

By

Estela Soledad Cerutti

August 2009

Chair: Richard A. Yost
Cochair: David H. Powell
Major: Analytical Chemistry

Small-molecule profiling, termed metabolomics, is a valuable tool to study phenotype and changes in phenotype caused by environmental influences, disease, or changes in genotype. The metabolome can be defined as the complete complement of all small molecule (<1500 Da) metabolites found in a specific cell, organ or organism. The metabolome represents a vast number of components that belong to a wide variety of compound classes, and these compounds are very diverse in their physical and chemical properties and occur in a wide concentration range. Consequently, studying the metabolome is a major challenge to analytical chemistry.

Mass spectrometry (MS) is used in metabolomics to detect, quantify, and identify enzymatic substrates and byproducts from biological and clinical samples. MS-based metabolomics offers qualitative and quantitative analyses with high selectivity and sensitivity, wide dynamic range, and the ability to analyze biofluids with extreme molecular complexity. The combination of liquid chromatography with MS reduces the complexity of the mass spectra, decreases ion suppression, provides isobar resolution, and delivers information on the properties of the metabolites.

In this study, the hypothesis that detectable changes will occur in the blood plasma metabolic profile of healthy female and male adults before and after a ketogenic diet has been tested. In addition, changes in the plasma metabolome of piglets between days 2 through 8 of life have been evaluated.

Novel complementary chromatographic approaches–reversed phase and hydrophilic interaction liquid chromatography, directly coupled to a time-of-flight mass spectrometer operating under electrospray conditions in positive ion mode, have been developed and optimized. The performance/contribution of each separation strategy, identification of unique m/z features, and technical variability have been evaluated. The studies involved a large number of samples that required powerful data processing/analysis capabilities. In this sense, the raw data were processed using commercial instrument software. From the obtained chromatograms, features were extracted, aligned, normalized, filtered, and then analyzed by different statistical methods, including analysis of variance, principal component analysis, and volcano plots. Finally, using the accurate mass criterion of 2 ppm mass error, putative biomarkers responsible for the metabolic differences in the samples were identified using several databases.

CHAPTER 1
INTRODUCTION

Metabolomics

Background

Metabolomics is one of the most rapidly growing and changing fields of contemporary science and it is one of those relatively new 'omic' disciplines. This field is evolving toward the intention of extracting useful knowledge from metabolite pools [1, 2]. Metabolomics is complementary to the other omic technologies as it is downstream of genomics, transcriptomics, and proteomics and, is closest to cell's physiological state [3]. Consequently, metabolomic analysis holds the promise to contribute to a better understanding of the interactions from genes to phenotypes. The rising number of publications in the field (Figure 1-1) demonstrates that metabolomics is not just a new "omics" word, but a valuable emerging technology to study the phenotype (which refers to the observable physical or biochemical characteristics of an organism, as determined by both genetic makeup and environmental influences) and changes in the phenotype caused by environmental influences, disease, or changes in phenotype.

The metabolome was first described by Oliver et al. in 1998 [4] as the set of metabolites found in or produced by an organism. This definition has been limited to "the quantitative complement of all of the low molecular weight molecules (< 1500 Da) present in cells in a particular physiological or developmental state" [5]. Fiehn made distinctions between different metabolite analysis techniques and defined metabolomics as a comprehensive analysis in which all metabolites of a biological system were identified and quantified [6].

Metabolomic Approaches

Metabolomics defined as the unbiased identification and quantification of all the metabolites present in a specific biological sample cannot be carried out in its totality. Thus,

different analytical approaches can help to reduce the complexity of the analysis, these categories include target metabolomics, metabolite profiling, and metabolite fingerprinting. Target analysis is constrained to the qualitative and quantitative analysis of a particular set of metabolites of known identity. Metabolite profiling involves the identification and quantitation of a predefined set of metabolites of known or unknown identity and belonging to a class or a selected metabolic pathway. Metabolite fingerprinting aims to rapidly classify numerous samples using multivariate statistics, typically without differentiation of individual metabolites or their quantitation [5, 7-9].

The metabolome represents a vast number of components that belong to a wide variety of compound classes. These compounds are very diverse in their physical and chemical properties and occur in a wide concentration range [1, 10]. There have been many attempts to estimate the number of metabolites in a biological system. However, the size of the metabolome varies greatly, depending on the organism studied [11]. Current estimates put the mammalian metabolome at about 3,000 different compounds [12] whereas the yeast and *E. coli* metabolomes are believed to consist of between 600 to 800 compounds [13], and it is estimated that the plant kingdom may encode more than 200,000 metabolites [14]. These molecules can be ionic inorganic species, hydrophilic compounds or amphoteric analytes. Furthermore, the elemental composition, the order of atoms, and the stereochemical orientation may have to be elucidated *de novo* for metabolites [2]. Thus, there is no single technology platform that has the ability to profile the metabolome simultaneously and a combinatorial approach is used.

Analytical Instrument Platforms

Analytical variance is a major factor when using different technology platforms and refers to the coefficient of variance related to a particular experimental approach. Therefore, the knowledge and correction of analytical variance between experimental approaches is necessary. Numerous analytical techniques have been used in the field of metabolomics to monitor and

explore metabolic differences between biological samples, such as nuclear magnetic resonance (NMR) [15], Fourier transform-infrared spectroscopy (FT-IR) [16, 17] and mass spectrometry (MS) [18, 19] coupled to separation techniques or using direct flow injection. Magnetic resonance spectroscopy has several advantages relative to mass spectrometry: it does not require sample preparation and it produces signals that correlate directly and linearly with compounds' abundances. However, only medium to high abundance metabolites will be detected with this technique and the identification of individual metabolites is challenging in complex mixtures [18]. On the other hand, mass spectrometry offers quantitative analyses with high selectivity and sensitivity and the potential to identify metabolites [18]. However, mass spectrometry-based techniques usually require a sample preparation step, which can cause metabolite losses and, based on the sample introduction system and the ionization technique used, specific metabolite classes may be discriminated against.

There has been tremendous progress in mass spectrometry metabolomics in recent years, leaving researchers with a variety of choices for chromatographic separation, ionization, and mass spectrometry analysis. Separations may be achieved by gas chromatography (GC) [20], capillary electrophoresis (CE) [9], or liquid chromatography (LC) [21], with LC approaches continuously evolving. A comparison of different commonly used metabolic technologies is shown in Table 1-1; specificity, sensitivity and structural range of the different methods vary substantially. Later discussion will be focused on those specific technologies that were applied to this research work.

In order to overcome the drawbacks of directly injecting complex samples, liquid chromatography can be associated with the MS detector for metabolomic analysis. Liquid chromatography can reduce ion suppression caused by coeluting compounds, isobaric

interferences in the case of low-resolving mass analyzers, and often can separate isomers [1]. In addition, a good analytical separation will result in better detection limits and MS data quality due to reduced background noise. Important technological advances in liquid chromatography over the past few years have forged a new era of research. Recently, liquid chromatography has shifted from the standard high-performance liquid chromatography (HPLC) to the ultra-high pressure liquid chromatography (UHPLC) which can increase resolution sensitivity and peak capacity while decreasing sample volumes and mobile phases [22]. The downside of this approach is the high pressure (10,000–15,000 psi) needed to operate these columns, and thus special UHPLC systems are required.

Metabolomics deals with a great diversity of small molecules that differ greatly in their physical and chemical properties of size, polarity/hydrophobicity, and charge [18]. While no single chromatographic method is suitable for all classes of metabolites, different alternatives exist. Reversed-phase liquid chromatography (RPLC) is a standard tool for the separation of medium polar and non-polar analytes. However, very polar metabolites are not retained on classical reversed-phase stationary phases and elute with the void volume. An interesting alternative to RPLC is the use of hydrophilic interaction liquid chromatography (HILIC) for the separation of highly polar and hydrophilic compounds. HILIC is orthogonal to RPLC chromatography and uses polar stationary phase materials and low aqueous/high-organic solvent systems.

In summary, technological advances in NMR and MS have introduced metabolomics as an approach to study the metabolism and its regulation in relation to disease, genetic, and environmental factors. With regard to human health alone, multiple benefits of metabolomics investigations can be visualized. They can deliver new tools to diagnose disease or monitor the

success of nutritional, pharmacological, and therapeutic interventions. Metabolomics can also provide new biomarkers to assess human health, and over time a powerful list of diagnostic markers will likely be discovered, which can be measured using high-throughput assays. However, in order to proceed from the single biomarker concept to the global metabolome evaluation outside a research environment, the technology has yet to be developed to provide the clinician with the tools to assess entire wide classes of metabolites in biofluids and automatically process the data to evaluate the biochemical status of an individual. Many technical and methodological issues have to be addressed to create analytical platforms that readily answer biological questions efficiently.

In general, for every type of MS-based metabolomics experiment the following steps need to be addressed during method development and validation.

Sampling. Sample acquisition is driven by the experimental design and the experiment type. A sufficient number of samples are required in order to reduce the influence of biological variability. In particular when studying human samples, the influences of gender, diet, age, and genetic factors have to be considered. In addition, representative quality control samples (replicates) and blanks have to be analyzed.

Sample preparation. The main goal of the sample preparation procedure is to extract the analytes from the complex biological matrices, such as serum, plasma, whole blood, urine, tissue, saliva, cell pellets, etc.; and bring them into a format that it is compatible with the analytical technique used while removing matrix components that will interfere with the analysis. In terms of metabolic target analysis the sample preparation procedure can be tailored for the target metabolites, because the analytes are known and surrogate compounds or stable isotope-labeled standards can be utilized to optimize the extraction procedure and matrix

removal. On the other hand, if a global metabolomic analysis is pursued, the sample preparation step should be as simple and universal as possible. Sample preparation and sample introduction methods can include direct injection, liquid-liquid extraction (LLE), solid-phase extraction (SPE), supercritical fluid extraction, accelerated solvent extraction, microwave-assisted extraction, protein precipitation, and membrane methods, such as dialysis and ultracentrifugation.

Sample analysis. This point has already been discussed in the "Analytical Instrument Platforms" section, vide supra. Further details for the applied analytical strategies will be given later in the thesis.

Data processing and analysis. Besides the technological aspects, data processing is an important factor for making sense of the results and the steps taken depend on what question was initially asked (hypothesis approach). In all cases, statistical analysis must be performed in order to ensure analytical rigor [23]. How this is done remains a debate, since different methods could give varying results and no consensus approach has been selected by the research community. Table 1-2 shows some of the common methods used for data processing, all the steps taken before any statistical methodology is applied, and analysis or application of supervised or unsupervised statistical treatments. To deal with large datasets, it is very important to carry out the correct data processing steps as batch to batch differences in data become a clear problem. This process involves feature extraction (a feature is a molecular entity characterized by a specific mass and retention time), alignment (e.g. mass and/or retention time), normalization (e.g. to adjust the intensities within each sample run by reducing the systematic error), filtering (e.g. using the most intense peaks across the experiments), and quality control (e.g. clustering trees). On the other hand, data analysis uses multivariate techniques to reduce the dimensionality

of the measurement vector. This can be done by selecting a subset of variables directly as a representation of the total process, combining (transforming) the original measurements to a new set with fewer features. The probability distribution is estimated and this probability density function would help in the process of validation and testing the system against the available data. A classification system is determined during this process and now derives 'rules' that allocate new observations to pre-defined categories. Different methods include non-linear mapping (NLM), hierarchical clustering analysis (HCA), principal component analysis (PCA), k-means clustering and self-organization maps (SOM), independent component analysis (ICA), partial least squares-discriminant analysis (PLS-DA), multilevel component analysis (MCA); probabilistic neural networks (PNNs), artificial neural networks (ANNs), and orthogonal projection on latent structure-discriminant analysis (O-PLS-DA), among others. Once discriminant analysis is done, statistical significance can be determined by applying classical statistics like Student's t-test, analysis of variance (ANOVA), or multiple analyses of variance (MANOVA). Another way of dealing with data after the processing step is to detect for significant correlations of components, which is based on the covariance and/or correlations within a data matrix. Specific details of the applied data processing and statistical tools are given in Chapter 3.

Identification and biological interpretation. The identification of metabolites remains a major bottleneck for metabolomics. The unambiguous identification of a differentiating biomarker(s) is important from a biological point of view. A careful scrutiny on the biochemistry of differentiating metabolites is an essential part of the biomarker discovery process [24]. For the proper assignment of metabolites, whether in terms of identity or models of networks, bioinformatic tools and data bases are required [2]. There are few tools that can automatically

produce a list of possible metabolites from the mass signals at a particular retention time (MS) or chemical shift (NMR). In fact, the connection between experimental data (MS and NMR spectra, retention time, fragmentation pattern, chemical shift, coupling constant) and the available chemical databases (Table 1-3) is still weak, let alone automatic.

In summary, many technical and methodological issues need to be addressed to create analytical platforms that readily answer biological questions efficiently. In this sense, data analysis and visualization tools, libraries, and databases for metabolomics have yet to be developed. A major obstacle in this type of investigation is the high diversity and variability encountered in the metabolome. The physico-chemical properties of metabolites are so diverse that none of the currently available techniques can analyze all metabolites simultaneously. In addition, the quantification of the metabolites encountered is a challenge because signal intensity is not only a function of concentration or mass but also depends on the chemical structure of the analytes and can be influenced by matrix interference (especially when electrospray ionization associated with mass spectrometry is used). On the other hand, the concentration of the components in the metabolome varies over a tremendous dynamic range, which can be as high as 100,000. Although metabolomics is still in its infancy and the framework is still developing, the greater synergy between organisms will provide a much clearer picture of the function of cells, organs, and organisms, bringing us closer to understanding their roles in nature.

The principle aim of this study is to contribute to the development of global and targeted metabolomic strategies for the evaluation and identification of metabolomic changes in human and animal biological samples. Therefore, this work focuses on the development of chromatographic approaches for the separation of polar and nonpolar compounds, their coupling to mass spectrometric detection, and application to real metabolomic studies.

The remainder of Chapter 1 presents an overview of mass spectrometry, focusing on the time-of-flight mass spectrometer, a brief description of the type of ionization applied for this project- electrospray ionization, some general principles of liquid chromatography- focusing on C_{18}-monolithic and hydrophilic interaction liquid chromatographic approaches, a succinct explanation of the metabolic role of the targeted metabolites, and, finally, the scope of the present work.

Time-of-Flight Mass Spectrometer

Background

In a time-of-flight (TOF) analyzer, ions are separated on the basis of differences in their velocities as they move in a straight path toward a collector. The time-of-flight mass spectrometer (TOF-MS) is fast, capable of high resolving power and high accuracy, applicable to chromatographic separation, and it is used for the mass determination of small and large biomolecules [25]. William E. Stephens of the University of Pennsylvania proposed the concept of TOF-MS in 1946 [26]. Two years later, in 1948, Cameron and Eggers [27] reported the first time-of-flight mass spectra, demonstrating a mass resolution of around 5. In 1955, Wiley and McLaren [28] developed a technique to focus the spread of ionization position and initial ion energy, achieving a mass resolving power of 300 or higher. This modification provided TOF-MS with a practical mass resolving power, and the Bendix Corporation began distribution of commercial models [29]. In the 1970s, a reflectron TOF-MS [30] was developed, featuring a mass resolving power of a few thousand. In the late 1980s, Dawson and Guilhaus [31], and Dodonov [32] developed orthogonal acceleration (*oa*), which allowed an efficient combination of TOF-MS and continuous ionization sources.

Orthogonal Acceleration Time-of-Flight Mass Spectrometer

In the basic structure of an orthogonal acceleration time-of-flight mass spectrometer (*oa*-TOF-MS) the sample is continuously ionized by the ion source. Then, ions from the ion source pass into the flight tube, which is perpendicular to the incident beam direction. The perpendicular orientation of the source with respect to the flight tube reduces the kinetic energy dispersion in the direction of the flight tube improving the mass resolving power [33]. To initiate a pulse of ions into the TOF analyzer an electrostatic field is created in nanoseconds (typically) and as a consequence, the ions in this field experience a force. The direction of the force is strictly orthogonal to the ion beam axis. The resulting orthogonal acceleration imparts a new component of velocity to the sampled ions and this component is vectorially independent of the axial velocity of the ion beam. Vectorial decoupling of the velocity of ions in the ion beam and TOF directions is an important feature of *oa*-TOF-MS. Figure 1-2 shows a schematic of an *oa*-TOF-MS. The independence of the TOF and source axes leads to advantages such as reduction of the spread of the velocity distribution in the TOF analysis direction and the orthogonal acceleration region fill time can be matched to the longest flight of the ions in the TOF analysis region insuring that there are no ions wasted and thereby maximizing sensitivity.

Electrospray Ionization

Electrospray ionization (ESI) is a simple and elegant method that handles big molecules, operates at atmospheric pressure and at a moderate temperature, and is probably the most gentle ionization technique available for MS [34]. It has also become the most successful interface for LC/MS and CE/MS applications [1, 9]. Although the concept of electrospray was put forward by Malcolm Dole in 1968 [35], the development of ESI–MS is credited to John Fenn [36-39], who was awarded the 2002 Nobel Prize in Chemistry for that contribution. Electrospray applications account for most of the activity in *oa*-TOF-MS [40]. The coupling of electrospray with *oa*-TOF-

MS has important advantages: duty-cycle-related sensitivity, speed to facilitate MS and MS/MS with online liquid chromatography, high m/z capability, and excellent mass accuracy [41].

As the name implies, electrospray ionization is a process that produces a fine spray of highly charged droplets under the influence of an intense electric field. Evaporation of the solvent converts those charged droplets into gas-phase ions. The mechanism of ESI is a highly debated topic. It is generally believed that ionization in electrospray involves three different processes: droplet formation, droplet shrinkage, and desorption of gaseous ions [42].

Advantages of ESI are as follows: analysis of compounds with a molecular weight up to about 310,000 Da is possible [43]; ESI is very sensitive; typical detection limits range from low attomole to picomole levels [44]; ESI is a very mild ionization technique; usually only sample molecules carrying multiple protons are generated. It is possible to observe native biological complexes bound by noncovalent interactions. Fragmentation can be induced by increasing the potential difference in the ion source and on-line coupling with liquid chromatography or capillary electrophoresis (CE) equipment is possible because ESI is a solvent-based ionization technique.

Disadvantages of ESI are as follows: the presence of salts, buffers, detergents, and other additives reduces the sensitivity dramatically. If buffers are needed, volatile buffers, such as ammonium acetate or ammonium formate are preferred; and analysis of mixtures is difficult, because each compound gives rise to several signals corresponding to sample molecules carrying a range of protons [45]. In addition, signal suppression (the signal intensity of an analyte ion may decrease)/enhancement (the signal intensity of an analyte ion may increase) is common in ESI. Signal suppression/enhancement originates from a competition between an analyte and co-eluting species (both with the same charge) for placement in the surface excess charge layer

which ultimately transfers into the gas phase. Ion suppression or enhancement affects precision and accuracy of an assay and it is most commonly induced by matrix components.

Liquid Chromatography

Although chromatography has been around for more than a century and column chromatography for about half a century, new columns, new chemistry, new materials as well as new instrumentations are continuously introduced [11]. Highly efficient separations are the cornerstone of many analytical methods that deal with extremely complex mixtures. These new analytical tools need to be fast and of sufficient resolving power to detect minute qualitative and quantitative changes in a metabolic profile.

For the separation and identification of metabolites GC/MS, LC/MS, or CE/MS techniques have been primarily employed. Among these techniques, HPLC/MS is most widely applicable to metabolomics. HPLC separations are better suited for the analysis of labile and high molecular weight compounds and for the analysis of nonvolatile polar compounds in their natural form. However, since metabolomics deals with a diversity of small molecules, there is no single ideal chromatographic approach that can be applied to all classes of metabolites. We have found that two methods: one for nonpolar compounds and one for polar compounds (described below)-, provide a reasonable breadth of coverage.

Monolithic Columns

High performance liquid chromatography has traditionally been performed in columns packed with 3 or 5 μm particle diameters. The internal diameter of these columns is typically between 2 and 4.6 mm, although smaller column diameters are gaining popularity. High separation efficiency with a concomitant reduction in analysis time is achieved by reducing particle size [46]. Chromatographic packing materials with diameters in the range of 1-2 μm are

now commercially available and known as stationary phases for ultra-high pressure (performance) liquid chromatography. These smaller particles dramatically increase column efficiency, which in turn increases mass sensitivity, chromatographic resolution, and speed [47]. However, columns packed with such particles need a special LC system. The injection volume should be small as well as the volume of the detector cell, in case UV detection is applied [48]. In addition, smaller column particle size results in higher system pressures and requires components, such as pumps, with high pressure ratings (2,000 bar or 30,000 psi).

Emerging technology in separation sciences is enabling high separation efficiency and speed of analysis, surpassing the conventional particle-packed columns in HPLC. This includes the use of monolithic columns [48-50].

As compared to packed particle columns, monolithic columns consist of a single piece of porous cross-linked polymer or porous silica. Monoliths are made in different formats as porous rods, generated in thin capillaries or made as thin membranes or disks. As a result of their internal design, all the mobile phase flows through the stationary phase. The convective flow, which is the typical driving force for mass transfer in this type of columns, enables a substantial increase in the speed of separation [51]. Monolithic columns have high column permeability and small-sized skeletons that decrease the diffusion path length of molecules in the stationary phase, resulting in a reduced contribution to band broadening [52]; thus, van Deemter curves for some monolithic columns are much flatter at high flow rates compared to conventional columns. These features allow the operation of monolithic columns at very high flow rates for fast separations with no significant loss in efficiency [52-55].

Based on the nature of their construction materials, monolithic columns can be classified as organic polymer- or silica-based columns [49]. The organic polymer-based monoliths have

always done a better job of separating larger molecules; while silica-based monolithic columns enable fast separations of smaller molecules [49]. Monolithic silica columns were introduced much later than polymer monoliths. After conventional-size columns became commercially available in 2000, the publication rate on monolithic silica columns exceeded that on polymer monoliths [56]. In this research, a reversed-phase C_{18} silica-based monolithic column was used for the separation of the nonpolar metabolites in biological samples. Thus, the discussion will be briefly focused on this particular type of column material.

Monolithic silica columns: these columns contain a tailor-made bimodal pore structure with both macropores or through pores and mesopores [57]. The large macropores are responsible for a low flow resistance, and therefore allow the application of high eluent flow rate, while the small pores ensure sufficient surface area (300 m^2/g approximately) for separation efficiency [57]. One of the important features of these columns is their high permeability, which is nearly twice as high as that of packed columns [58]. Therefore, monolithic silica columns can be operated at high flow rates of up to 10 mL min^{-1}, thus allowing fast separations of various mixtures [59].

Monolithic silica columns are suitable for high throughput analysis, e.g. metabolomics analysis, as well as for two-dimensional HPLC methods [56, 60].

HILIC Columns

Reversed-phase mode is most often employed in HPLC, where chemically bonded stationary phases (C_8, C_{18}, C_{30}, etc.) have advantages in rapid equilibration with mobile phase, and high separation efficiency and high reproducibility in gradient separations, based on the hydrophobic properties of the stationary phases. However, the highest concentrations for endogenous metabolites in body fluids are found for very polar, small molecules [61, 62]; these compounds exhibit poor retention on reversed-phase columns, resulting in co-elution of many

polar compounds [63, 64]. These compounds are eluting early in the chromatographic run, under highly aqueous mobile phase conditions and without sufficient compound separation, and ionization efficiency of ESI is often poor [65].

The retention and separation of polar compounds is an on-going challenge for chromatographers [66-67]. Ion exchange or ion pairing, mobile phase pH manipulation, and reversed-phase chromatography with specially designed columns are techniques that traditionally have been used for retention of polar analytes. However, each of these techniques has certain drawbacks. Ion exchange or ion pairing work well only if the analytes of interest are ionizable. In addition, ion pairing is difficult with mass spectrometry because ion pairing reagents will cause signal suppression. Manipulation of mobile phase pH is a technique that also works for ionizable compounds, because the retention characteristics of ionizable compounds are a function of pH of the mobile phase; however, manipulation of mobile phase pH is not always successful in retaining analytes that are very polar. Also some compounds might not be stable outside a narrow pH range. Reversed-phase chromatography is versatile and able to retain and resolve many classes of compounds; however, the retention of polar analytes often requires a highly aqueous mobile phase to achieve retention, which can cause a number of issues such as dewetting of the stationary phase. In addition, the highly aqueous mobile phases that are required for polar retention are not ideal for mobile phase desolvation by ESI-MS, and thus result in poor sensitivity [68].

Hydrophilic interaction chromatography is a mode of chromatography that can address these issues. HILIC was first introduced by Alpert for the separation (amongst other polar compounds) of amino acids, which were eluted in opposite order to that found in reversed-phase chromatography [69]. HILIC is similar to normal phase liquid chromatography in that elution is

promoted by the use of polar modifiers (e.g. H_2O) in less polar mobile phases (e.g. acetonitrile). However, HILIC is unique in that the presence of water in the mobile phase is crucial for the establishment of a stagnant enriched aqueous layer on the surface of the stationary phase into which analytes may selectively partition. Alpert considered that dipole-dipole interactions (hydrogen bonds) might contribute to partitioning into the stationary phase layer. Hydrogen bonding, especially when using mobile phases of low water content, has also been proposed as a retention mechanism [70]. Clearly, the mechanism of HILIC separations is complex and different phenomena contribute to various degrees.

Despite the complexity of the mechanism, HILIC is simple and its general advantages can be summarized as follow: reasonable peak shapes obtained for bases; mass spectrometer sensitivity is enhanced due to the high organic content in the mobile phase and the high efficiency of spraying and desolvation techniques; direct injection can often be made of extracts eluted from C_{18} columns with solvents of high organic content; the order of elution is generally the opposite of that found in reversed phase separations, giving useful alternative selectivity; good retention of polar compounds is obtained in HILIC, whereas poor retention is often obtained in reversed-phase chromatography; and higher flow rates are possible due to the high organic content of typical mobile phases [70].

The HILIC separations of proteins [69], peptides [71-73], amino acids [69], oligonucleotides [69], carbohydrates [74], histones [75], and natural products [76, 77], among others compounds [78], have been reported. The use of HILIC chromatography as a tool for the evaluation of polar compounds in different metabolomic applications is increasing in popularity and acceptance [63, 76, 79-81]. One of the aims of this study has been to illustrate the utility of HILIC for global and targeted metabolomics.

Target Compounds

Carnitine and Acylcarnitines

Carnitine and acylcarnitines are endogenous metabolites present in most mammalian tissues. Carnitine, (3-hydroxy-4-(*N*,*N*,*N*-trimethylammonio) butanoate) is a small, water soluble, quaternary nitrogen-containing compound (see Figure 1-3) that is involved in the transport of activated fatty acids into the mitochondrial matrix, where they are metabolized via β-oxidation [82]. Fatty acid metabolism occurs in the mitochondrial matrix; however, the mitochondrial inner membrane is impermeable to fatty acids. Thus, carnitine plays a vital transport role in fatty acid metabolism and in cellular energy production.

In mammals, carnitine functions through the reversible esterification of its 3-hydroxyl group, with subsequent translocation of the acylcarnitines produced from one cellular compartment to another. Carnitine acyltransferases are the enzymes responsible for the production of acylcarnitines, which can have differing chain-lengths, commonly designated as short-, medium-, and long-chain length; and which vary by their cellular location and metabolic functions.

Carnitine is also implicated in the maintenance of the cellular pool of free coenzyme A (CoA) and in the elimination of potentially toxic acyl-CoA, originating from exposure to xenobiotics and/or from blockage of metabolic pathways. In contrast to acyl-CoAs, the corresponding acylcarnitines can be excreted via the urine [82, 83].

In healthy subjects, carnitine and acetylcarnitine are the dominant concentration carnitines of the body fluid and tissue carnitine pools [82, 84]. Skeletal muscle contains >95% of the total carnitine body stores, and the tissue concentrations are considerably higher than the concentration in plasma; transport systems ensure its widespread distribution from sites of absorption and synthesis throughout the body. The kidneys play a crucial role in carnitine homeostasis, since

they reabsorb >90% of the filtered carnitine [82], so that the plasma levels of free carnitine are maintained at 30–40 μmol L^{-1} [85, 86]. Reported plasma concentrations in healthy humans are 29-50 μmol L^{-1} for free carnitine, 2.5-8.6 μmol L^{-1} for acetylcarnitine, 0.18-0.6 μmol L^{-1} for propionylcarnitine, 0.03-0.17 μmol L^{-1} for isovalerylcarnitine, 0.02-0.05 μmol L^{-1} for hexanoylcarnitine, 0.01-0.13 μmol $L^{-1}$for octanoylcarnitine, and 2.2-4.9 μmol L^{-1} for the long-chain acylcarnitines [85]. The relative amounts of acylcarnitines are often expressed as a ratio of acylcarnitine to free carnitine. In plasma samples, a ratio greater than 0.4 is indicative of a carnitine deficiency [87].

In many metabolic disorders, carnitine metabolism is greatly disturbed, leading to a redistribution of the carnitine and acylcarnitine pools. The determination of individual acylcarnitines in biological fluids is a powerful means to diagnose and monitor these disorders [88-91].

Amino Acids

An amino acid is a molecule containing both amine and carboxyl functional groups. In the alpha amino acids ($H_2NCHRCOOH$, where R is an organic substituent), the amino and carboxylate groups are both attached to the α–carbon. The various alpha amino acids differ in which side chain (R group) is attached to their alpha carbon.

Twenty standard amino acids are used by cells in protein biosynthesis, and these are specified by the general genetic code. These 20 amino acids are biosynthesized from other molecules, but organisms differ in which ones they can synthesize and which ones must be provided in their diet. The eight amino acids (isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine) that cannot be synthesized by an organism are known as essential amino acids and must be obtained from food. However, the essentiality depends on many factors, e.g. cysteine, taurine, tyrosine, histidine and arginine are semi-essential

amino acids in children because the metabolic pathways that synthesize these amino acids are not fully developed. The amounts required also depend on the age and health of the individual [92, 93].

Depending on the polarity of the side chain, amino acids vary in their hydrophilic or hydrophobic character [94]. These properties are important in protein structure and protein-protein interactions. The importance of the physical properties of the side chains comes from the influence this has on the amino acid residues interactions with other structures, both within a single protein and between proteins. The distribution of hydrophilic and hydrophobic amino acids determines the tertiary structure of the protein, and their physical location on the outside structure of the proteins influences their quaternary structure. The analytical separation of amino acids is also influenced by their polar or nonpolar character. Chemical structures of carnitine, acylcarnitines, and amino acids are shown in Figure 1-3.

This research focused on the application of liquid chromatography with mass spectrometry to global and targeted metabolomic studies. Detailed description of the development and optimization of novel chromatographic approaches and their direct application to biological samples (plasma from animal and human sources) is presented in the following chapters. In the case of target metabolomic studies, special attention to carnitine, acylcarnitines (short-, medium-, and long-chain) and amino acids was given. Commercially available C_{18}-monolithic- and hydrophilic interaction liquid chromatography-columns for the separation of nonpolar and polar analytes were used. In addition, an *oa*-TOF mass spectrometer was utilized for the accurate mass analyses of the mentioned metabolites. Data were acquired in positive electrospray ionization mode.

There are two emerging column technologies that have attracted much attention in recent years, one is the monolithic column and the other is the use of hydrophilic interaction liquid chromatography. Herein, we explore these two chromatographic approaches. The analytical conditions for the development of a reversed-phase chromatographic strategy using two in-series C_{18}-monolithic columns for the efficient separation of nonpolar compounds is presented in Chapter 2. In addition, the steps corresponding to the optimization of a HILIC separation methodology for the chromatographic resolution of polar compounds are presented. A comparison of the selectivity of both chromatographic methods and application to real samples are detailed at the end of Chapter 2 as well.

The application of the optimized LC/MS parameters to the human plasma samples before and after a ketogenic diet therapy for global and targeted metabolomic studies is discussed in Chapter 3.

The application of the same LC/MS methodologies to piglet plasma samples to evaluate metabolome changes during days two to eight of life is described in Chapter 4.

Finally, Chapter 5 summarizes the work presented here, as well as some future research directions.

Table 1-1. Comparison of metabolic technologies (adapted from reference [7])

Criteria	NMR	GC/MS	LC/MS
Sensitivity	Poor	Good	Excellent
Limit of detection	10^{-6} mol	10^{-12} mol	10^{-15} mol
Metabolites detected	High concentration of organic compounds in a solution	Volatile, nonpolar compounds	Nonvolatile compounds in a solution
Robustness	Good	Reasonable	Reasonable
Speed	Rapid	Depending on chromatography	Depending on chromatography
Quantification	Good	Poor, standards needed	Poor, standards needed
Identification	By chemical shift calibration	By mass, fragmentation, and retention time	By mass, fragmentation, and retention time
Problems	Peak overlap	Volatility of metabolites	Ionization of metabolites
Disadvantages	-Exact adjustment of pH required after metabolite extraction -Different methods can circumvent the need for extraction, but result in a further loss of sensitivity	-Nonvolatile compounds have to be derivatized -Some metabolites cannot be made volatile even with derivatization	-Problem with ion supression and adduct formation -Different metabolites detected in positive and negative scanning mode -Lack of libraries for metabolite identification

Table 1-2. Methods for data processing and analysis (adapted from reference [2])

Data processing methods	Data analysis methods
Normalization	Coefficient of variation
Baseline correction, peak shifting, and noise removal	ANOVA[1] or MANOVA[2]
Missing value correction	PCA[3], ICA[4], and subtypes
Deconvolution of peak	Clustering-HCA[5], k-means
Data reduction	Self organizing maps
Limited data analysis to specified representative region of data	Fisher discriminant analysis
Exclude variables or sample outliers	Partial least squares (PLS)
	Neural networks (ANNs[6], PNNs[7])
	Genetic programming and algorithms

[1]ANOVA: analysis of variance; [2]MANOVA: multiple analysis of variance; [3]PCA: principal component analysis; [4]ICA: independent component analysis; [5]HCA: hierarchical clustering analysis; [6]ANNs: artificial neural networks; [7]PNNs: probabilistic neural networks

Table 1-3. MS, NMR, and metabolic pathway databases (DB)

Name	Source/URL
MS-based databases	
Golm Metabolome Database (GMDB@CSB.DB)	Max Planck Institute of Molecular Plant Physiology
Human Metabolome Database (HMDB)	Genome Alberta and Genome Canada
KNApSAcK (Comprehensive Species – Metabolite Relationship Database)	Nara Institute of Science and Technology (NAIST)
Metlin	The Scripps Research Institute
NIST/EPA/NIH Mass Spectral Library (NIST 0.5)	National Institute of Standards and Technology (NIST)
SpecInfo	Daresbury Laboratory
Spectral Database for Organic Compounds (SDBS)	National Institute of Advanced Industrial Science and Technology (AIST)
NMR-based databases	
ACD Databases	Advanced Chemistry Development, Inc.
Human Metabolome Database (HMDB)	Genome Alberta and Genome Canada
NMRShiftDB	University of Koeln
SpecInfo	Daresbury Laboratory
Spectral Database for Organic Compounds (SDBS)	National Institute of Advanced Industrial Science and Technology (AIST)
Standard Compounds on Biological Magnetic Resonance Bank (BMRB)	University of Wisconsin
Metabolic pathway databases	
BRENDA	BRaunschweig Enzyme Database (http://www.brenda.uni-koeln.de/)
Chemical Entities of Biological Interest (ChEBI)	http://www.ebi.ac.uk/chebi/
HumanCyc	Encyclopedia of Human Metabolic Pathways (http://humancyc.org/)
KEGG	Kyoto Encyclopedia of Genes and Genomes (http://www.genome,jp/kegg/)
Lipid Maps	http://www.lipidmaps.org/
MetaCyc	Encyclopedia of Metabolic Pathways (http://metacyc.org/)
Nicholson's Metabolic Minimaps	http://www.tcd.ie/Biochemistry/IUBMB-Nicholson/
PUMA2	Evolutionary Analysis of Metabolism (http://compbio.mcs.anl.gov/puma2/cgi-bin/index.cgi)
Reactome	A Curated Knowledgebase of Pathways (http://www.reactome.org/)
Roche Applied Sciences Biochemical Pathways Chart	http://www.expasy.org/cgi-bin/search-biochem-index

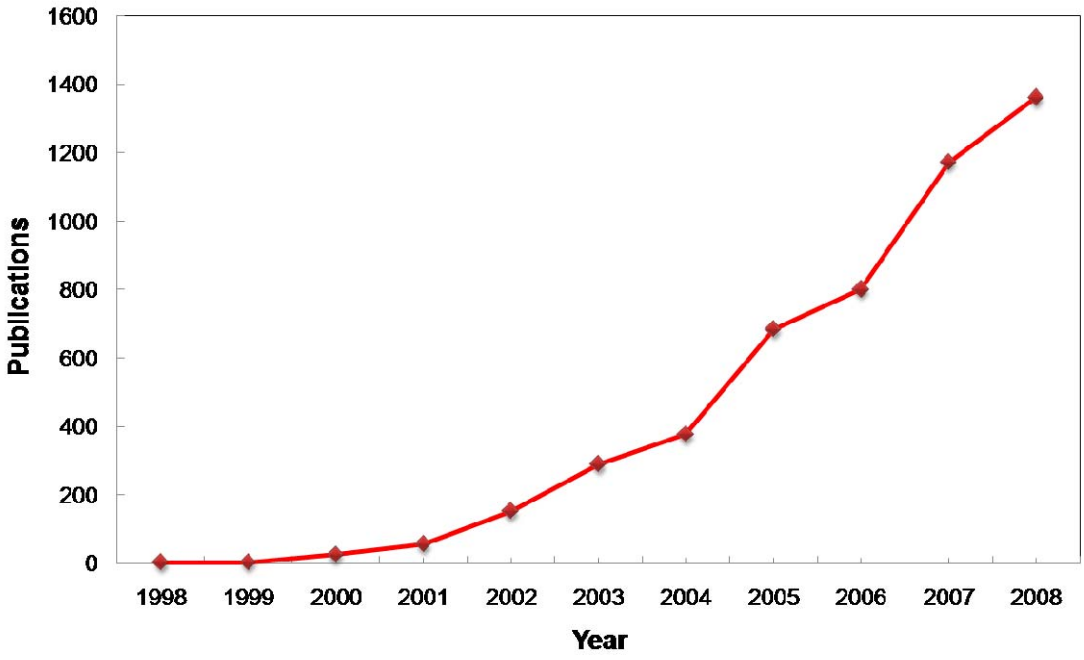


Figure 1-1. Representation of the number of metabolomic publications per year from 1998 to 2008 (obtained by searching " metabolomics" using SciFinder, Version 2007, (American Chemical Society)).

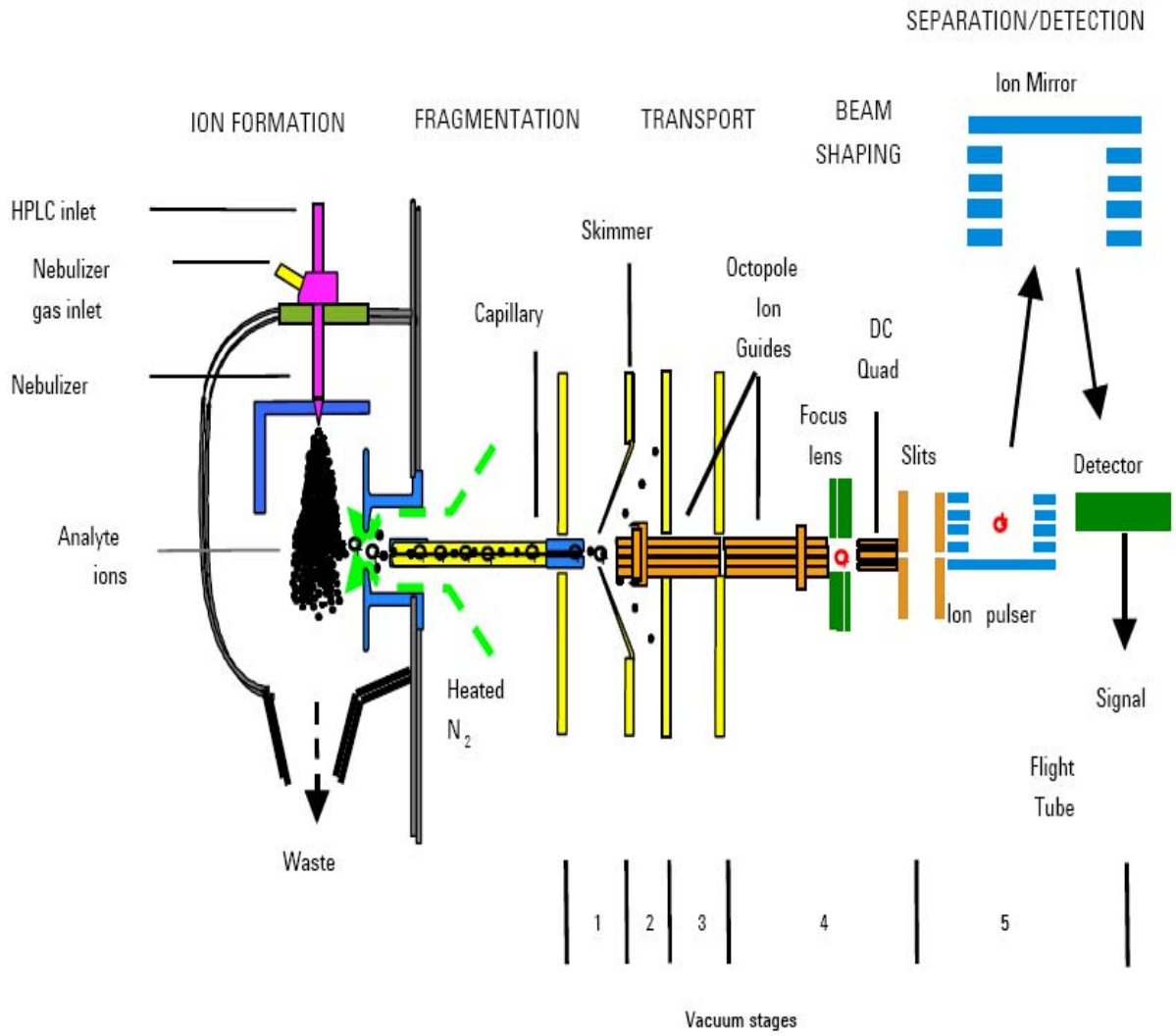
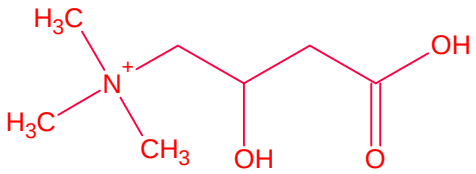
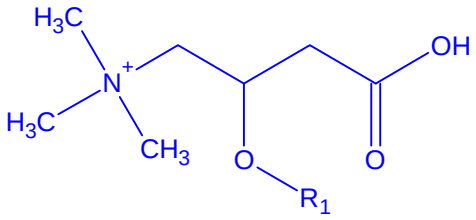


Figure 1-2. Main components of typical *oa*-TOF system with reflecting mass analyzer. The ion beam enters from a source at the left (ion formation region) and it is accelerated to enter the orthogonal accelerator (*oa* or ion pulser in this figure). The beam optics make the beam more parallel before it enters the oa. The beam "fills" the first stage of the oa until a bipolar push-out pulse pair is applied. A packet of ions is thus sampled and accelerated through grids to enter the drift region. Conventional reflecting TOF optics are used to bring the ions to a space-time focus on the detector. During the time that the ions are in the drift-region (and ion mirror) the oa is refilled with new beam. Figure adapted from www.chem.agilent.com.

Free carnitine



Acylcarnitines



Amino acids

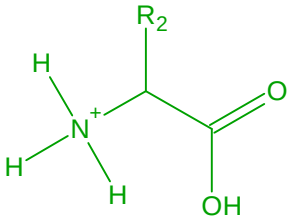


Figure 1-3. Chemical structures of free carnitine, acylcarnitines, and amino acids. The "R_1" group in the acylcarnitine structure is esterified by different chain lengths giving place to acylcarnitines with different chain lengths. The "R_2" group in the amino acid structure determines its identity.

CHAPTER 2
DEVELOPMENT AND COMPARISON OF TWO CHROMATOGRAPHIC APPROACHES

Liquid Chromatography

In liquid chromatography analytes are separated by virtue of differing partition coefficients between a liquid mobile phase and a solid stationary phase [95]. High-performance liquid chromatography (HPLC) is one mode of chromatography in which a liquid mobile phase is used to separate the analytes. Different steps take place during this separation process: the components are first dissolved in a solvent; they are forced to flow through a chromatographic column under a high pressure; and once in the column, the mixture is resolved into its components, which elute from the column sequentially over time

The interaction between the components and the column stationary phase directly determines the chromatographic resolution. The interaction of the solute with mobile and stationary phases can be manipulated through different choices of both solvents and stationary phases. As a result, HPLC acquires a high degree of versatility not found in other chromatographic systems and it has the ability to easily separate a wide variety of chemical mixtures [96, 97].

HPLC is characterized by small diameter (2 - 5 mm), reusable stainless steel columns, column packings with very small (1, 2, 3, 5 and 10 μm) particles and the continual development of new substances to be used as stationary phases, relatively high inlet pressures and controlled flow of the mobile phase, precise sample introduction without the need for large samples, special continuous flow detectors capable of handling small flow rates and detecting very small amounts, automated standardized instruments, rapid analysis, and high chromatographic resolution [95-97].

The high performance obtained with these chromatographic systems is the result of many factors: very small particles of narrow distribution range and uniform pore size and distribution, high pressure column slurry packing techniques, accurate low volume sample injectors, sensitive low volume detectors and, of course, good pumping systems. Naturally, pressure is needed to permit a given flow rate of the mobile phase; otherwise, pressure is a negative factor not contributing to the improvement in separation [95-97].

Today there is a competition between two means for fast LC analysis, namely, HPLC with monolithic phases and small particle phases (< 2 μm) used in ultra-high pressure liquid chromatography. With both phase types a substantially faster analysis, reduced solvent consumption, and increased mass sensitivity can be achieved.

It is important to understand the basic principles of a rapid resolution system; however, since the chromatographic technologies used in this work are not related to UHPLC, a further discussion will be avoided. Attention will instead be focused on monolithic and hydrophilic interaction chromatography columns.

Monolithic stationary phases are continuous separation media in the format that can be compared to a single "large" particle that does not contain interparticle voids. This type of material exhibits dual pore arrays named throughpores and mesopores. The throughpores are responsible for the free flow of the mobile phase through the column and the mesopores for the retention performance. Therefore, the efficiency of this column is controlled by the average sizes of the throughpores and mesopores. The major advantage of this new approach is that we might be able to choose these two sizes independently of each other, but it is still the beginning of a long period of research endeavors during which the immense number of possible approaches to design and preparation of these monolithic columns will be optimized.

Hydrophilic interaction liquid chromatography is a technique suitable for the separation of very polar and hydrophilic compounds. Therefore, solutes that have little or no retention on reversed-phase columns generally experience strong retention on HILIC columns. HILIC is a variation of normal-phase chromatography without the disadvantages of using solvents that are immiscible with water. Besides, HILIC enhances the chromatographic retention and separation of those analytes which are poorly retained by reversed-phase chromatography. In this type of column, the stationary phase consists of a polar material (e.g. silica, cyano, diol, amino, amide, polymeric coatings, etc.), the mobile phase is highly organic (> 80%) with a small amount of aqueous/polar solvent, and the order of elution goes from least to most polar.

Experimental

Mass Spectrometer

Positive ESI [(+)ESI] spectra were collected on an Agilent 6210 Time-of-Flight mass spectrometer (Agilent Technologies, Inc., Santa Clara, California) configured with a dual nebulizer electrospray source. With this source, ions were simultaneously generated from the LC stream by one capillary nebulizer and from mass calibrant solution, via a second capillary nebulizer. Data acquisition and processing were conducted using the Agilent MassHunter Workstation software.

TOF Calibration

An Agilent Technologies ES Tuning Mix (P/N G2421A) containing betaine (CAS #: 107-43-7, MW: 117.1), hexamethoxyphosphazene (CAS #: 957-13-1, MW: 321.0), hexakis(2,2-difluoroethoxy)phosphazene (CAS #: 186817-57-2, MW: 621.0), hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazene (CAS #: 58943-98-9, MW: 921.9), hexakis(1H, 1H, 5H-octafluoropentoxy)phosphazene (CAS #: 16059-16-8, MW: 1521.0), hexakis(1H, 1H, 7H-dodecafluoroheptoxy)phosphazene (CAS #: 3830-74-8, MW: 2121.4), and hexakis(1H, 1H, 9H-

perfluorononyloxy)phosphazene (CAS #: 186043-67-4, MW: 2720.1), provided mass reference ions for a baseline calibration of the mass axis.

HPLC System

An Agilent 1200 Series Rapid Resolution LC system equipped with a binary pump provided solvent delivery. For HILIC chromatography, a Luna HILIC column (Phenomenex, Torrance, California; 2.0 mm × 150 mm, with 3.0 μm particle diameter) was employed. In order to improve chromatographic resolution, two in-series C_{18}-monolithic columns (Phenomenex Onyx, each column was 100 mm × 4.6 mm) were used for conventional reversed-phase chromatographic separation.

Standard Solutions

Laboratory and commercial-made standards for carnitine, acylcarnitines, and amino acids were used for all the experiments. Standard stock solutions for each analyte were prepared in methanol and were stored at -4°C. Standard mixes were freshly prepared on the day of analysis by mixing known volumes of the different standard stock solutions to create a final solution with 0.5 mg L^{-1} of each analyte. Whenever necessary, dilutions of these stock solutions were used.

LC-MS grade acetonitrile (ACN), HPLC grade water, methanol, and isopropanol were obtained from Burdick and Jackson (Muskegon, Michigan). Acetic acid (glacial, TraceMetal grade), HPLC grade ammonium acetate, and certified ammonium formate were obtained from Fisher Chemical (Fisher Scientific Inc., Pittsburgh, Pennsylvania). Formic acid was purchased from Acros Organics (Fisher Scientific).

Mobile Phase Preparation

For HILIC chromatography, the mobile phase was prepared by first dissolving a known amount of ammonium acetate or ammonium formate in water (stock solution), and then mixing with the desired volume of acetonitrile. The salt concentration in the text and figure captions

refers to the final concentration in the mobile phase. The pH of the salt solution was not adjusted. The mobile phases A and B for C_{18}-monolithic chromatography were prepared by adding 1% (v/v) of acetic acid to water and acetonitrile.

Plasma Sample Preparation

Piglet plasma samples were treated according to the standard operating procedure established to provide standards that are used in the Metabolic Assessment Laboratory (MAL) as required by section 58.81 of the Good Laboratory Practice for the Non-clinical Laboratory [98]. Frozen plasma was gently thawed with swirling to maintain cold temperature throughout the sample. Plasma (100 µL) was pipetted into 1 mL of acetonitrile:methanol (3:1), mixed vigorously and frozen overnight. The plasma/solvent samples were thawed, mixed, centrifuged for 15 min at 12,000 rpm, the supernatant collected and the pellet discarded. The supernatant was centrifuged a second and a third time and evaporated to dryness using nitrogen gas. The sample replicates were all assayed together in order to reduce error. In order to solubilize polar and less polar compounds, the pellets resulting from the preparation of the plasma samples were dissolved in 100 µL solvent composed of 10.0 µL of water, 10.0 µL of isopropanol, 40.0 µL of methanol, and 40.0 µL of acetonitrile. Blank solutions were prepared in the same way. It is important to note that no derivatization procedure was used before or after column separation. This fact differentiates this procedure from other approaches reported in the literature for the separation and/or determination of carnitine and carnitine-based compounds in clinical samples [99].

MS and Chromatographic Conditions

The electrospray source was operated in positive ionization mode. The capillary was set to -4.0 kV, the nebulizer gas was operated at 50 psi, and the drying gas was set to 10 L min^{-1} at a temperature of 300 ºC. The capillary exit was set to 180 V with skimmer set to 60 V. The

octopole offset voltage was set to 250 V. The TOF provided resolving power and mass accuracy (less than three ppm) with high sensitivity, scanning time (40 full scan spectra from m/z 50 to 1650 per second), and dynamic range (3-4 orders of magnitude.) Signal was acquired using a fast analog-to-digital converter (ADC); the ADC did not require any dead time corrections and was configured to record up to 10,000 transients per second.

For HILIC chromatography, the mobile phase A was 7.5 mM ammonium formate in water and mobile phase B was 7.5 mM ammonium formate in 8:92 water: acetonitrile. HILIC chromatography was carried out with a flow rate of 0.3 mL min^{-1} and the following gradient program: 10% A held for 5 min, then increased linearly to 50% A over 20 min and then held at 50% A for 5 min. A return to the initial conditions was accomplished by 10-minute linear gradient from 50 to 10% A where it was held for additional 5 min. The column was held at a temperature of 30 ºC. The injection volumes for the standard mixture and plasma samples were 10 and 20 μL, respectively.

For C_{18}-monolithic chromatography, the mobile phase A was 1% acetic acid in water and mobile phase B was 1% acetic acid in acetonitrile. The C_{18} gradient at 1.0 mL min^{-1} was as follows: 95% A held for 6.5 min, 19 min linear gradient to 0% A, held at 0% A for 9.5 min. A return to the initial conditions was accomplished by 10-minute linear gradient to 95% A where it was held for additional 5 min. The column was held at a temperature of 30 ºC. The injection volume was 20 μL for the analysis of standards and plasma samples. The optimization of these parameters has previously been reported [100].

Analytes of Interest

Due to the zwitterionic nature of carnitine, short-chain acylcarnitines, and amino acids, conditions were used to promote protonation of the carboxylic group, which resulted in the

production of positively charged ions. Thus, the positive ESI mass spectra of these compounds were dominated by their $[M+H]^+$ ions with minimal fragmentation. The pertinent characteristics of the analytes under study as well as the m/z of the $[M+H]^+$ ions are given in Table 2-1.

Results and Discussion

Coupling of Two C_{18}-Monolithic Columns

It is important to note that the HPLC parameters, such as effect of the composition of the mobile phase, effect of mobile phase buffer, temperature, flow rate, and gradient elution profile related to this approach were optimized previously during my Master's degree research work [100]. As a summary, the previous results indicated that a chromatographic gradient system composed of water and acetonitrile as organic mobile phase, when adding 1% acetic acid, could sharpen peak shapes and improve analytical sensitivity and resolution for the HPLC analysis of carnitine and acylcarnitine compounds (Figure 2-1). The optimum gradient was summarized in the "MS and chromatographic conditions" section.

The novel aspect of this study was the coupling of two in-series C_{18}-monolithic columns with the purpose of improving the chromatographic resolution of the compounds. Therefore, experiments were designed to evaluate the system performance using carnitine, acylcarnitines, and amino acids as target analytes (amino acids were not included in the list of target compounds in the Master's degree research work).

As seen from Figure 2-1, the polarities and corresponding retention times of carnitine and acylcarnitines are quite different with the long-chain acylcarnitines being much less polar than free carnitine and short-chain (< hexanoylcarnitine, C_6) acylcarnitines. The non-polar character of the medium- (hexanoylcarnitine – decanoylcarnitine) and long-chain (> decanoylcarnitine) acylcarnitines due to their extended hydrocarbon chain results in strong retention on the C_{18}-functionalized stationary phase (Table 2-1). However, the short chain, highly polar

acylcarnitines, and most of the polar amino acids, are poorly retained on the C_{18} column. Thus, there is significant interest in developing a chromatographic approach for the separation of more polar compounds.

Consequently, experiments were designed to optimize the parameters that affect the retention efficiency on a HILIC column. There were three main objectives to these studies: the first was to optimize the experimental conditions to achieve an efficient separation of the target metabolites (carnitine, acylcarnitines, and amino acids), the second to compare the efficiency of both chromatographic methodologies, and third to demonstrate the applicability of the methodologies to the analysis of the targeted compounds in plasma samples.

HILIC Method Development and Optimization

Effect of acetonitrile content and change on retention

HILIC often employs acetonitrile and water mixtures as the mobile phase [101-103]. The choice of organic component and its concentration have a significant impact on the retention of polar compounds on HILIC [104]. The effect of the acetonitrile content on retention under isocratic separation was investigated by changing it from 50 to 97% (v/v) in 10% increments while keeping the ammonium formate concentration unchanged. All of the compounds exhibited typical HILIC behaviors of increasing retention between 9 and 17 minutes as the ACN content increased (Figure 2-2). A content of 90% of acetonitrile was chosen for further experiments.

Effect of salt type and concentration

When charged molecules are analyzed with HILIC, the buffer is an essential component in the mobile phase [104]. The buffer type (ammonium formate or acetate) and concentration (5, 7.5, and 10 mM) were varied for isocratic 90% acetonitrile HILIC separation of a mixture of carnitine, acylcarnitines, and amino acids. The use of ammonium formate led to improved peak shape and increased retention times compared to ammonium acetate. The use of either buffer

gave improved retention time, peak shape, and sensitivity compared to no buffer (Figures 2-3 and 2-4).

The effect of the ammonium formate concentration on retention was evaluated by varying its concentration from 5 to 10 mM in a mobile phase of 90% acetonitrile. The ammonium formate concentration could not be further increased due to its solubility limitations in the mobile phase. The retention time of all of the compounds increased as the concentration of ammonium formate increased from 5 to 7.5 mM; and then decreased as the concentration of the buffer increased to 10 mM (Figure 2-5**)**. As a result, a 7.5 mM concentration of ammonium formate was selected for all further experiments.

Effect of temperature: Van't Hoff plots

Column temperature is another important parameter that affects the retention of polar compounds in HILIC [104]. Operation at elevated temperatures or at sub-ambient temperatures has been shown to influence the retention mechanism between the solute and the stationary phase material, often resulting in selectivity changes [105]. A Van't Hoff plot of the natural logarithm of the retention factor, ln k, which is expressed as:

$$k = \frac{t_r - t_0}{t_0} \qquad (2\text{-}1)$$

where t_r and t_0 are the retention times of the solute and the void time; vs. the reciprocal temperature, 1/T, is often obtained when operating at various temperatures, with a non-linear plot indicating a shift in the retention mechanism involved [104].

Van't Hoff plots for carnitine, short-chain acylcarnitines and some model amino acids on the HILIC column over the temperature interval 20–60 °C, are displayed in Figure 2-6. Carnitine and acylcarnitines showed decreased retention with increasing column temperature. In contrast, amino acids demonstrated increased retention at higher temperature. The largest range of

retention times among all of the analytes was obtained when the temperature was fixed at 30 °C. This temperature was selected for further experiments.

Effect of flow rate: van Deemter plots

The effect of the mobile phase flow rate on the separation of carnitine, polar acylcarnitines, and amino acids was evaluated using van Deemter plots [52-55]. Ten microliters of the standard sample mixture was injected onto the HILIC system at varying flow rates, from 0.1 to 0.5 mL min^{-1} with a gradient separation. The height equivalent to a theoretical plate (H) was determined using the following equations for N (number of plates) and H:

$$H = \frac{L}{N} \tag{2-2}$$

$$N = 5.54 \left(\frac{t_r}{W_{1/2}}\right)^2 \tag{2-3}$$

where L is the column length (μm), t_r is the retention time for the analyte (min), and $W_{1/2}$ is the peak width at half height (min). The plate heights for twelve analytes at varying flow rates are shown in Figure 2-7. A flow rate of 0.3 mL min^{-1} gave the smallest H values for all analytes, and was the flow rate chosen for analytical separation.

Evaluation of ionization suppression

Since all the amino acids were very close in retention time, the potential for ionization suppression among them was investigated. Standard solutions containing individual compounds and a solution made of a mixture of the standards were analyzed without chromatographic separation. Each standard was at a concentration of 5 mg L^{-1}, except for phenylalanine, proline, valine, and tyrosine which were each at 2.5 mg L^{-1}. The ESI molar responses (peak area per moles injected) of each analyte in the individual solutions were compared to the values in the

standard mixture; at the level of concentration under evaluation, no ion suppression was observed.

Repeatability

The intra-day precision of the proposed method was tested with six repeated injections of standard solution mixture at the concentration level of 0.5 mg L^{-1}. The inter-day precision was studied by analyzing 0.5 mg L^{-1} standard solution mixture, with six injections randomly executed in a 30-day period. For retention time and molar response, the intra-day relative standard deviations were below 2 and 8%, respectively. Similar results were obtained for inter-day relative standard deviations.

Applicability of HILIC to the Separation of Endogenous Metabolites in Plasma Samples

Piglet plasma samples were prepared as described in the experimental section. The above-mentioned optimized parameters for the separation and determination of carnitine, short-chain acylcarnitines and amino acids by HILIC/ESI-TOF were applied. Total ion chromatograms for both methodologies are shown in Figure 2-8. Extracted ion chromatograms of the analytes under HILIC and C_{18}-monolithic conditions are shown in Figures 2-9 – 2-11. While the separation of the carnitine and the C_2-, C_3- and C_4- acylcarnitines was much improved over that obtained with C_{18} RPLC, the peak shapes of the acylcarnitines showed some peak splitting (C_2-acylcarnitine) and fronting (Figure 2-9). In contrast, most of the amino acids gave well-shaped chromatographic peaks with good separation (Figures 2-10 and 2-11). Proline, valine, methionine, and tyrosine did show some peak splitting (Figures 2-10 and 2-11.)

Specificity of HILIC vs. C_{18}-Monolithic Chromatography

The high sensitivity afforded by coupling HPLC with ESI-MS is subject to some limitations when RPLC is used: decreases in the ESI efficiency and stability as a consequence of

polar compound elution in a highly aqueous mobile phase, and ion suppression due to poor retention and co-elution of polar compounds.

In this study, retention efficiencies on both HILIC and RPLC columns were evaluated. As shown in Table 2-1, fourteen of the twenty-eight analytes were weakly retained under RPLC conditions and eluted near the solvent front (3.6 min). On the HILIC column, the same compounds were strongly retained and eluted in a high organic mobile phase (~90-65%), which facilitated analyte desolvation and resulted in enhanced (+)ESI-MS sensitivity. In general, the concentration of acetonitrile in the mobile phase resulted in a greater than 10 times increase in the MS signal when compared to the C_{18} phase.

From a global metabolomics perspective and using the C_{18}-column approach, 810 compounds were isolated by using the software Molecular Feature Extractor (MFE). Identification of the metabolites was performed by searching METLIN database with the accurate m/z of the detected ions. As a result, 8% of the metabolites matched compounds in the database agree within two ppm mass accuracy. Positive mass error was frequently observed, which was in agreement with the mass errors exhibited for the calibrant ions at the specific scan.

On the other hand, when the HILIC column information was analyzed, 105 compounds were isolated by the MFE software. A total of 54 (51%) compounds were tentatively identified in the METLIN database within two ppm mass accuracy. Negative mass error was most common, which was in agreement with the mass errors displayed for the calibrant ions at the specific scan.

In summary, 65 compounds were tentatively identified when using the monolithic C_{18}-HPLC-based method. On the other hand, 54 compounds were tentatively identified when performing HILIC-based chromatography. After comparing the results, 18 compounds were found by METLIN in both chromatographic approaches.

Conclusions

Two novel separation methodologies for the orthogonal resolution of metabolites in biological samples were developed and successfully applied. The two in-series C_{18}-monolithic columns demonstrated adequate performance for the separation of nonpolar compounds. On the other hand, a simple, relatively rapid, and accurate LC-MS methodology was successfully developed and validated for separating carnitine, short-chain acylcarnitines, and amino acids by hydrophilic interaction chromatography. In addition, the conditions required for the HILIC separation were favorable to ESI-MS detection. The optimized methodology showed applicability to biological samples. Based upon an extensive literature search using SciFinder Scholar and Web of Science, this is the first time that a HILIC separation approach has been applied to the simultaneous separation of carnitine, polar acylcarnitines, and amino acids.

Table 2-1. Chemical formula, RTs, and m/z of the $[M+H]^+$ ions of carnitine, acylcarnitines, and amino acids

ID	Formula	$[M+H]^+$,[1] m/z	Retention time (RT) (min) RPLC-C_{18} column	RRT (min) RPLC-C_{18} column[2]	Retention time (RT) (min) HILIC column	RRT (min) HILIC column[2]
Carnitine	$C_7H_{15}NO_3$	162.1125	4.3	0.7	18.2	15.4
Acetylcarnitine	$C_9H_{17}NO_4$	204.1230	4.6	1.0	16.4	13.6
Propionylcarnitine	$C_{10}H_{19}NO_4$	218.1387	4.8-6.1	1.2	14.5	11.7
Butyrylcarnitine	$C_{11}H_{21}NO_4$	232.1543	6.0-11.4	2.4	13.5	10.7
Hexanoylcarnitine	$C_{13}H_{25}NO_4$	260.1856	17.1	13.5	13.6	10.8
Octanoylcarnitine	$C_{15}H_{29}NO_4$	288.2169	19.2	15.6	12.7	9.9
Decanoylcarnitine	$C_{17}H_{33}NO_4$	316.2482	20.1	16.5	11.9	9.1
Lauroylcarnitine	$C_{19}H_{37}NO_4$	344.2795	22.4	18.8	11.1	8.3
Myristoylcarnitine	$C_{21}H_{41}NO_4$	372.3108	23.2	19.6	10.4	7.6
Palmitoylcarnitine	$C_{23}H_{45}NO_4$	400.3421	24.2	20.6	9.9	7.1
Stearoylcarnitine	$C_{25}H_{49}NO_4$	428.3734	26.1	22.5	9.3	6.5
Alanine	$C_3H_7NO_2$	90.0555	4.4	0.8	18.2	15.4
Aspartic acid	$C_4H_7NO_4$	134.0448	4.5	0.9	20.4	17.6
Asparagine	$C_4H_8N_2O_3$	133.0608	4.6	1.0	13.8	11.0
Taurine	$C_2H_7NO_3S$	126.0219	4.0	0.4	11.2	8.4
Glycine	$C_2H_5NO_2$	76.0393	4.6	1.0	13.1	10.3
Glutamic acid	$C_5H_9NO_4$	148.0604	4.5	0.9	20.8	18.0
Histidine	$C_6H_9N_3O_2$	156.0768	4.7	1.1	19.2	16.4
Isoleucine	$C_6H_{13}NO_2$	132.1019	4.7	1.1	14.5	11.7
Leucine	$C_6H_{13}NO_2$	132.1019	4.7	1.1	14.6	11.8
Methionine	$C_5H_{11}NO_2S$	150.0583	5.2	1.6	14.8	12.0
Phenylalanine	$C_9H_{11}NO_2$	166.0863	6.0	2.4	14.5	11.7
Proline	$C_5H_9NO_2$	116.0706	4.6	1.0	14.7	11.9
Threonine	$C_4H_9NO_3$	120.0655	4.7	1.1	18.2	15.4
Tyrosine	$C_9H_{11}NO_3$	182.0812	4.6	1.0	17.2	14.4
Valine	$C_5H_{11}NO_2$	118.0863	4.7	1.1	12.1	9.3
Tryptophan	$C_{11}H_{12}N_2O_2$	205.0972	11.1	7.5	14.7	11.9
Carnosine	$C_9H_{14}N_4O_3$	227.1139	4.5	0.9	21.2	18.4

[1]Monoisotopic mass/charge (m/z); [2]RRT (relative retention time) = RRT = RT(analyte) - RT(void)

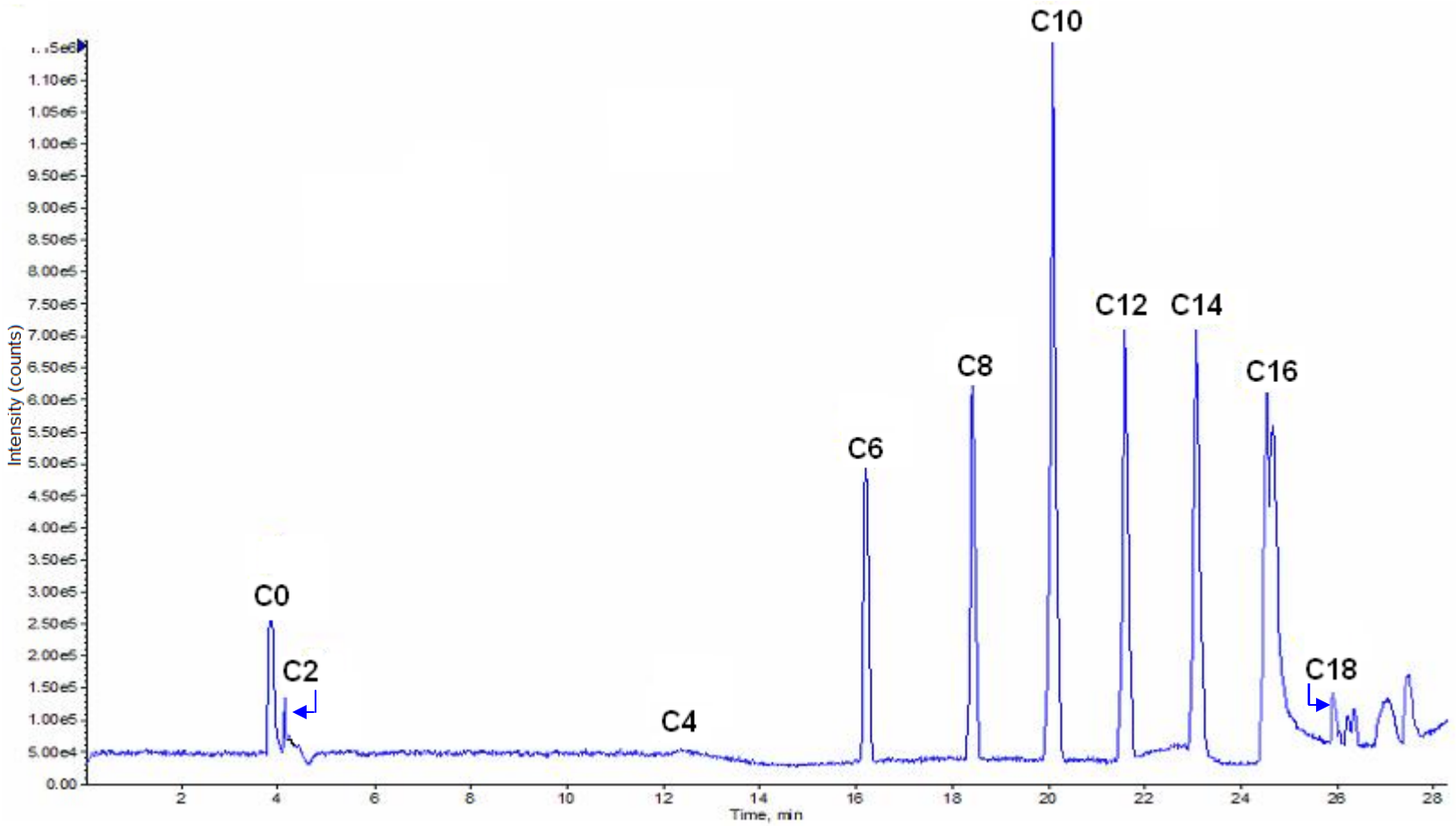


Figure 2-1. Base peak chromatogram obtained with a monolithic C_{18} column and acetonitrile as organic mobile phase; concentration standard mixture: 0.5 mg L^{-1}; injected volume: 15 µL; mobile phase flow rate: 1.0 mL min^{-1}; C_0: carnitine; C_2: acetylcarnitine; C_3: propionylcarnitine; C_4: butyrylcarnitine; C_6: hexanoylcarnitine; C_8: octanoylcarnitine; C_{10}: decanoylcarnitine; C_{12}: lauroylcarnitine; C_{14}: myristoylcarnitine; C_{16}: palmitoylcarnitine; C_{18}: stearoylcarnitine (adapted from reference [100].)

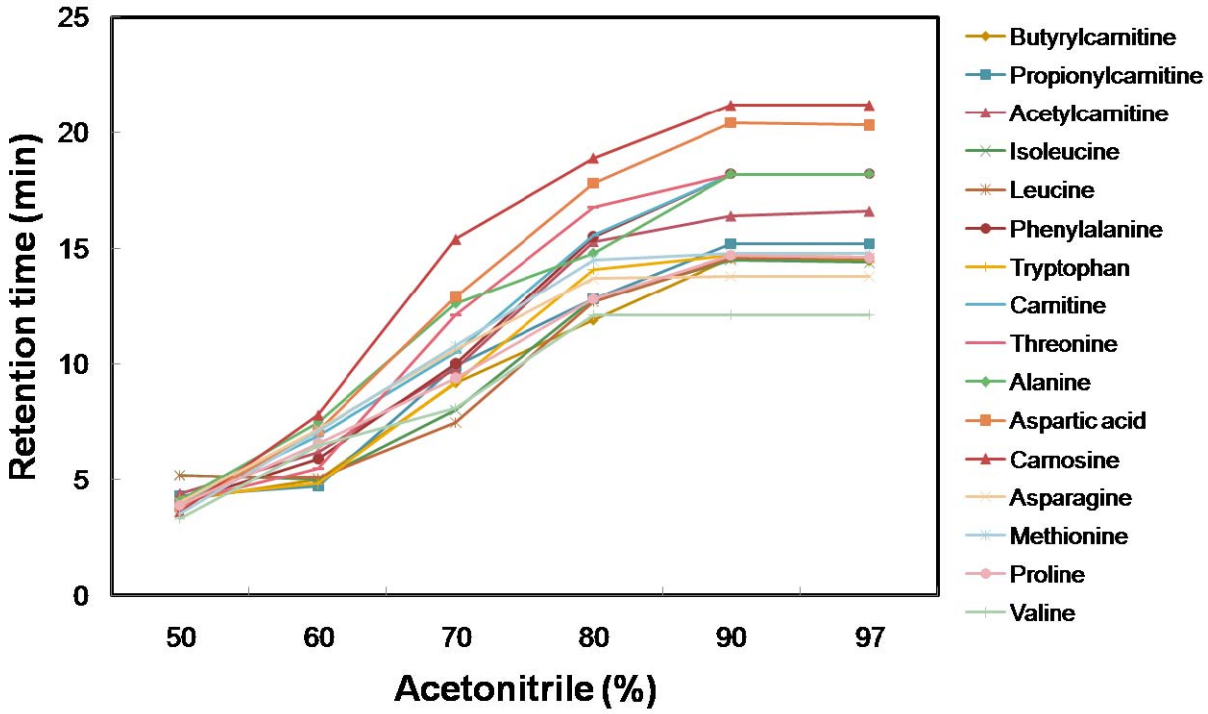


Figure 2-2. Effect of acetonitrile content on the isocratic HILIC of acylcarnitines and amino acids. Mobile phase: ACN:H_2O; Flow rate: 0.3 mL min^{-1}; Temperature: 30 ˚C; Concentration of standard mixture: 0.5 mg mL^{-1}.

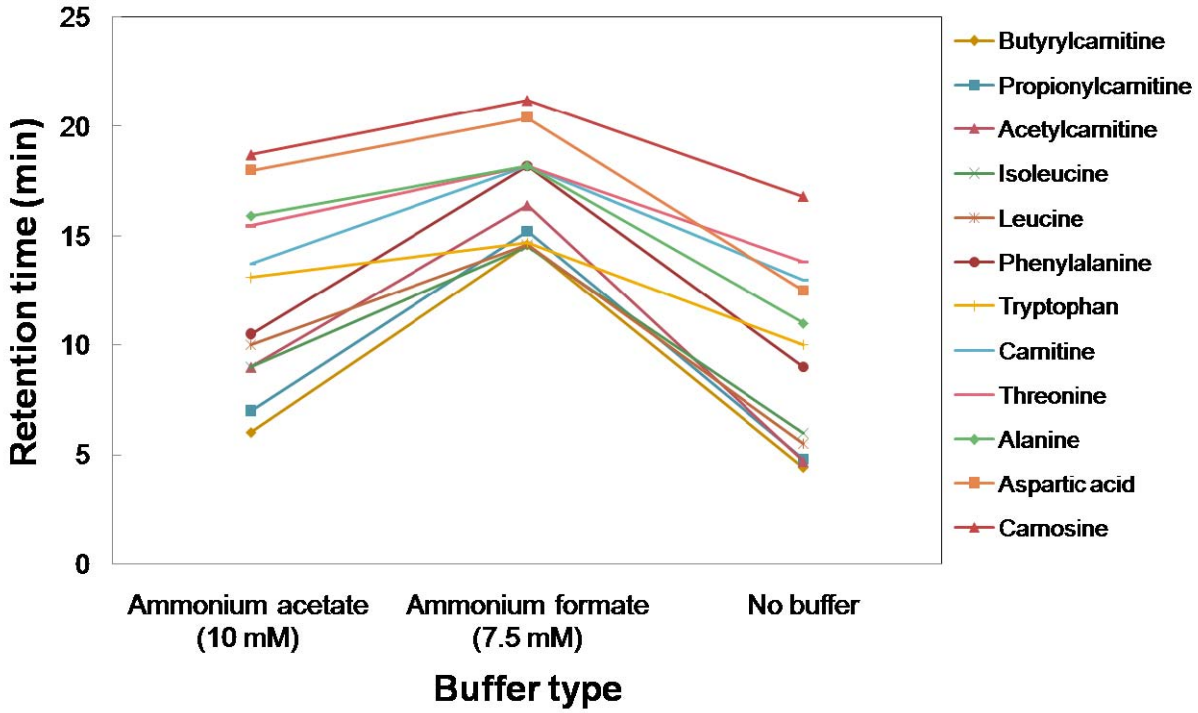


Figure 2-3. Effect of buffer addition to the mobile phase for HILIC separation. Mobile phase: ACN:H_2O; Flow rate: 0.3 mL min^{-1}; Temperature: 30 ˚C; Concentration of standard mixture: 0.5 mg mL^{-1}.

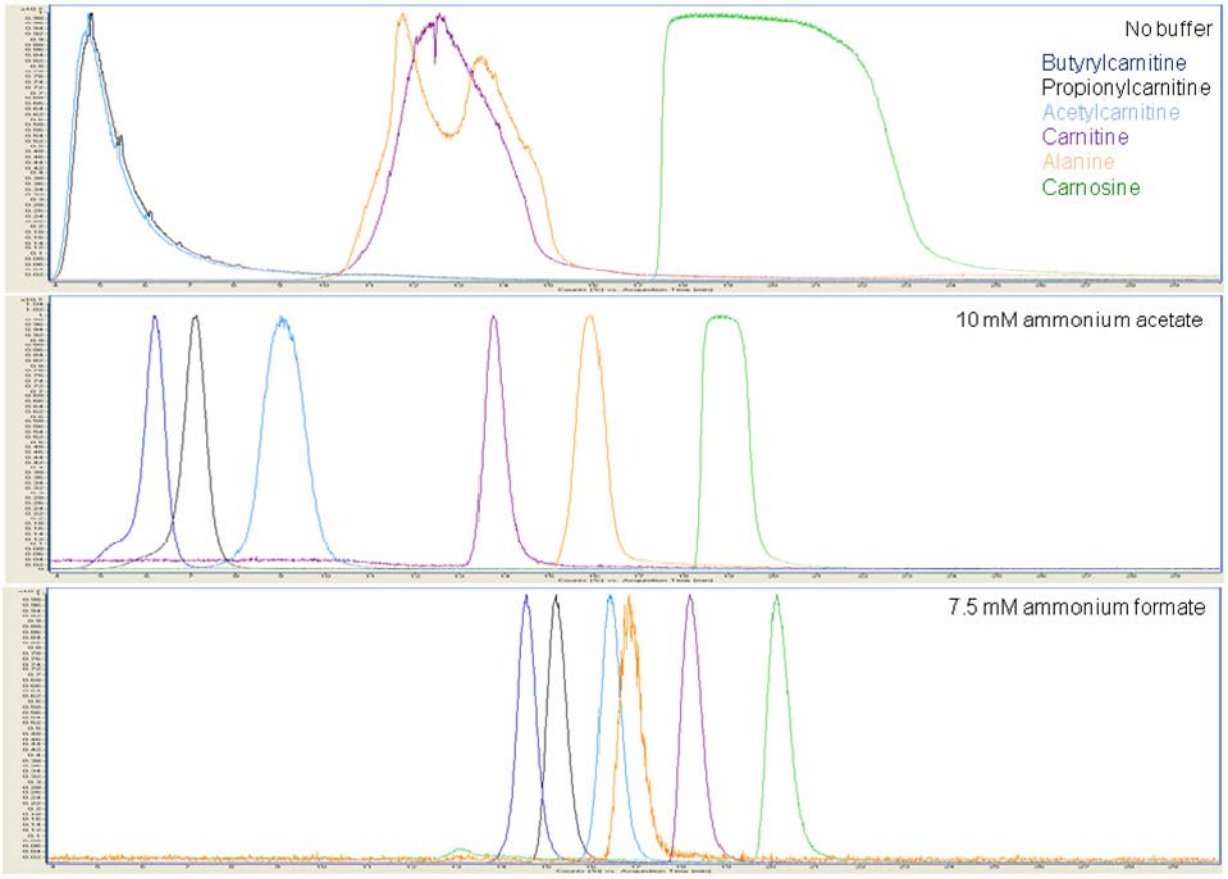


Figure 2-4. Extracted ion chromatograms for some model compounds showing the effect of buffer addition to the mobile phase for HILIC separation. Mobile phase: ACN:H_2O; Flow rate: 0.3 mL min^{-1}; Temperature: 30 ˚C; Concentration of standard mixture: 0.5 mg mL^{-1}.

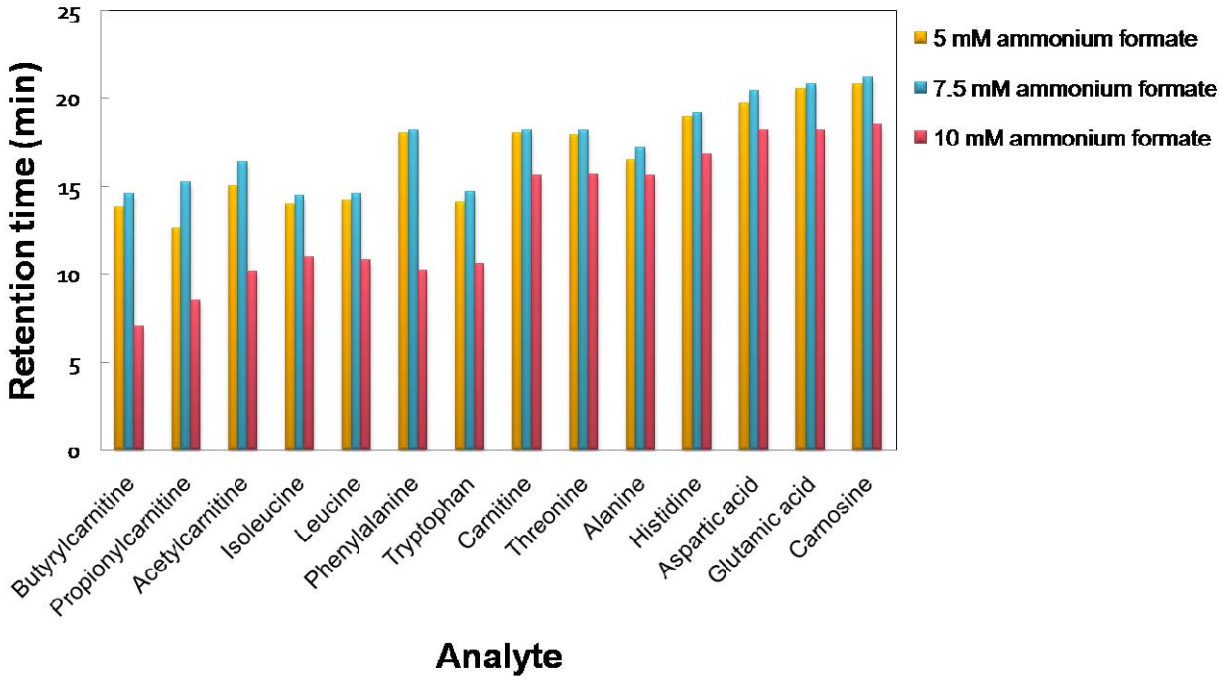


Figure 2-5. Effect of concentration of buffer on the retention time of some representative acylcarnitines and amino acids. Mobile phase: ACN:H_2O; Buffer: ammonium formate; Flow rate: 0.3 mL min^{-1}; Temperature: 30 ˚C; Concentration of standard mixture: 0.5 mg mL^{-1}.

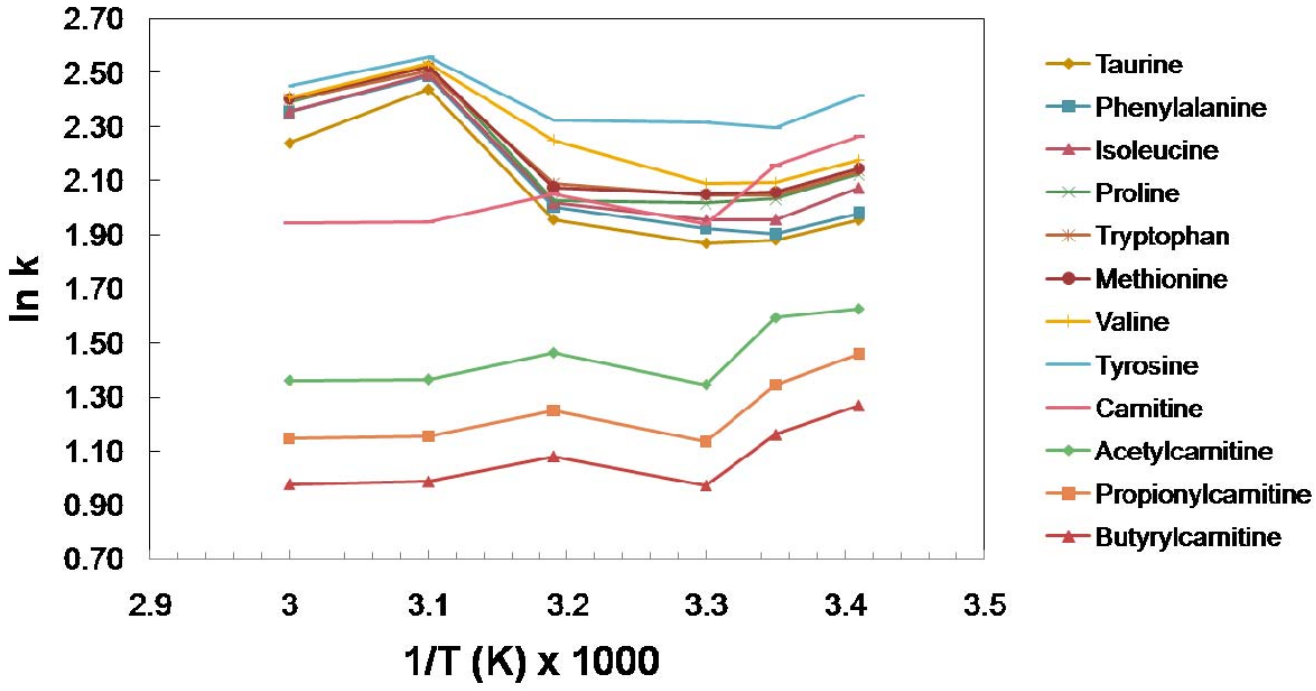


Figure 2-6. Van't Hoff plots for evaluating the effect of the temperature on the retention efficiency. Mobile phase: ACN:H_2O; Buffer: ammonium formate; Flow rate: 0.3 mL min^{-1}; Concentration of standard mixture: 0.5 mg mL^{-1}.

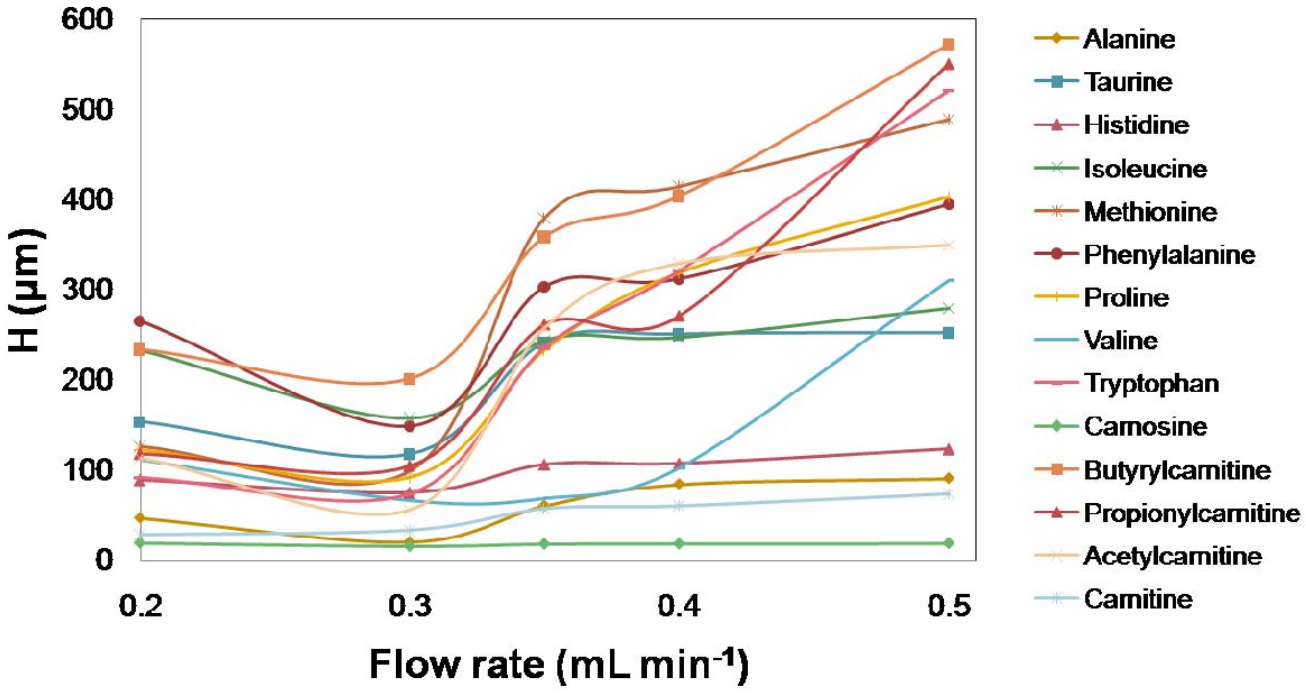


Figure 2-7. van Deemter plots for evaluating the effect of the mobile phase flow rate on the retention efficiency. Mobile phase: ACN:H_2O; Buffer: ammonium formate; Temperature: 30 ˚C; Concentration of standard mixture: 0.5 mg mL^{-1}.

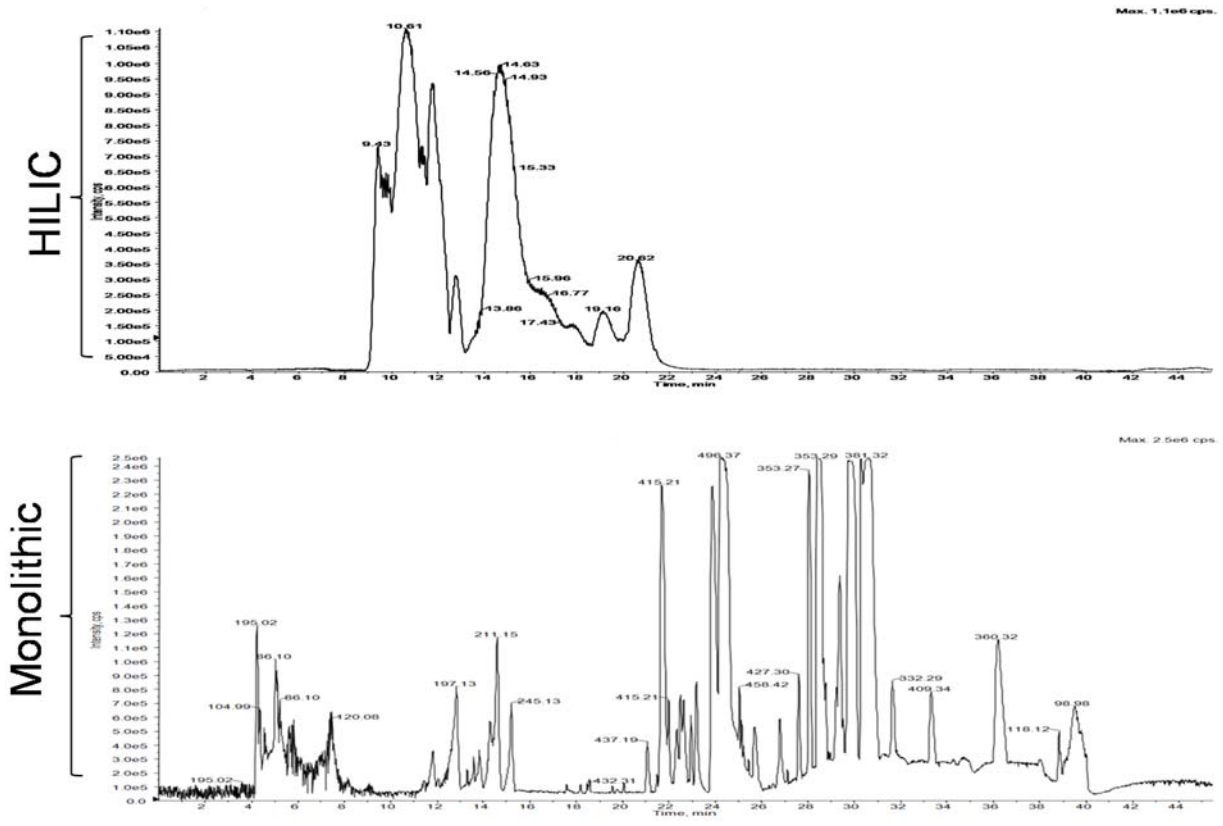


Figure 2-8. Total ion chromatograms for the application of C_{18}-monolithic/ and HILIC/MS on piglet plasma samples.

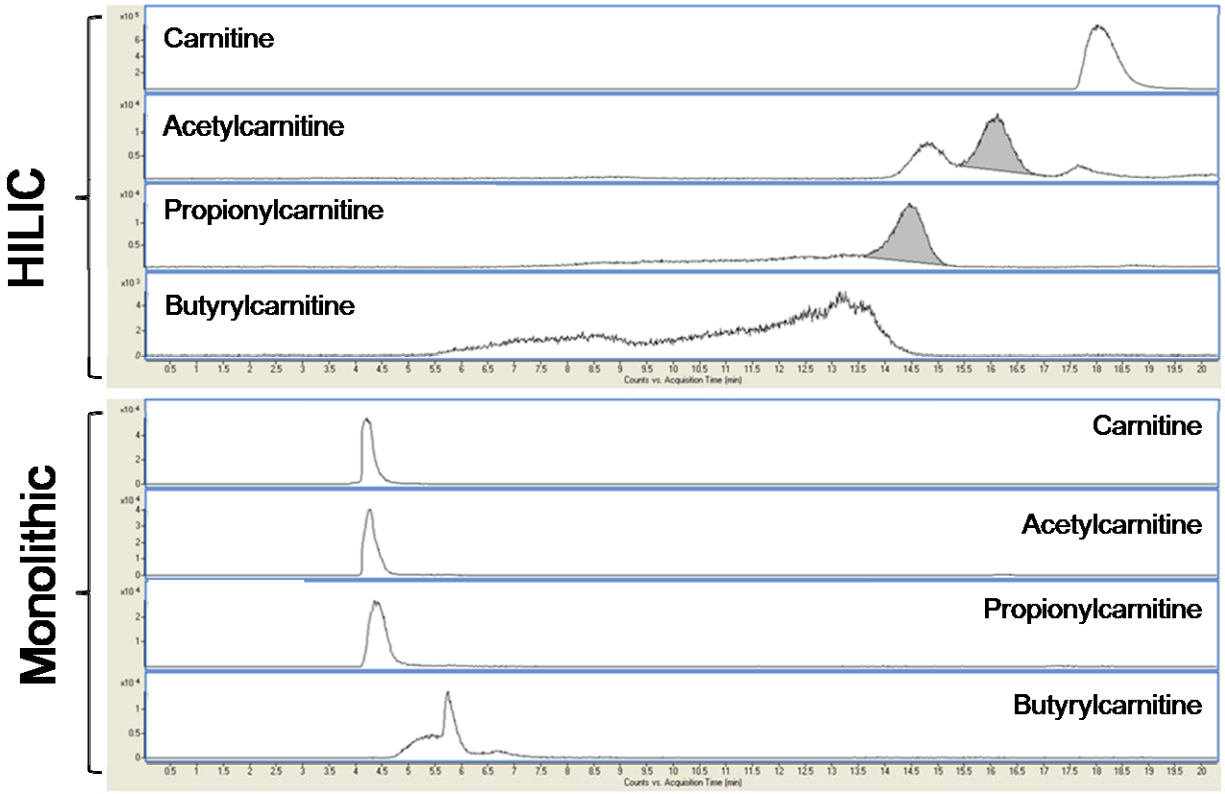


Figure 2-9. Extracted ion chromatograms of carnitine, acetylcarnitine, propionylcarnitine, and butyrylcarnitine under HILIC and C_{18}-monolithic conditions, respectively. Note: the extracted ion chromatograms for those compounds listed in Table 2-1, but not detected in the plasma samples, are not shown. The shaded peaks in each chromatogram correspond to those matching the retention times of the standards.

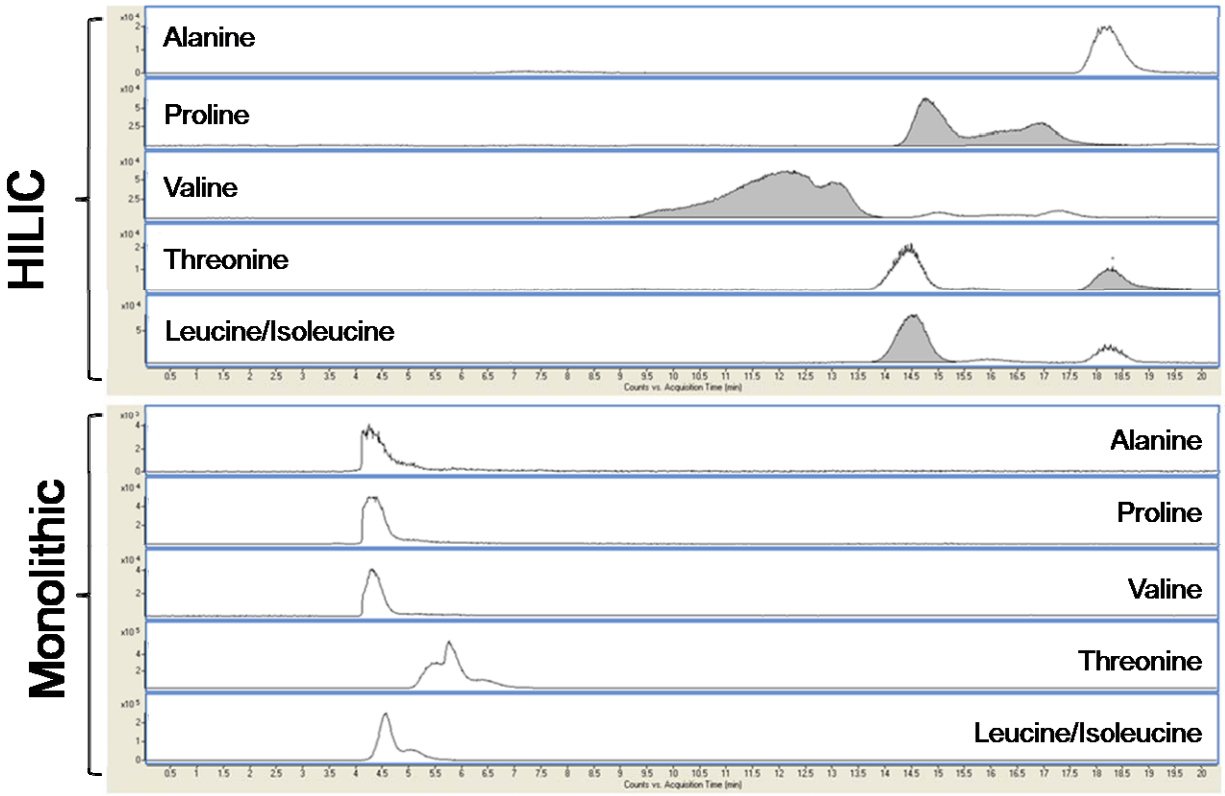


Figure 2-10. Extracted ion chromatograms of alanine, proline, valine, threonine, and leucine/isoleucine under HILIC and C_{18}-monolithic conditions, respectively. The shaded peaks in each chromatogram correspond to those matching the retention times of the standards.

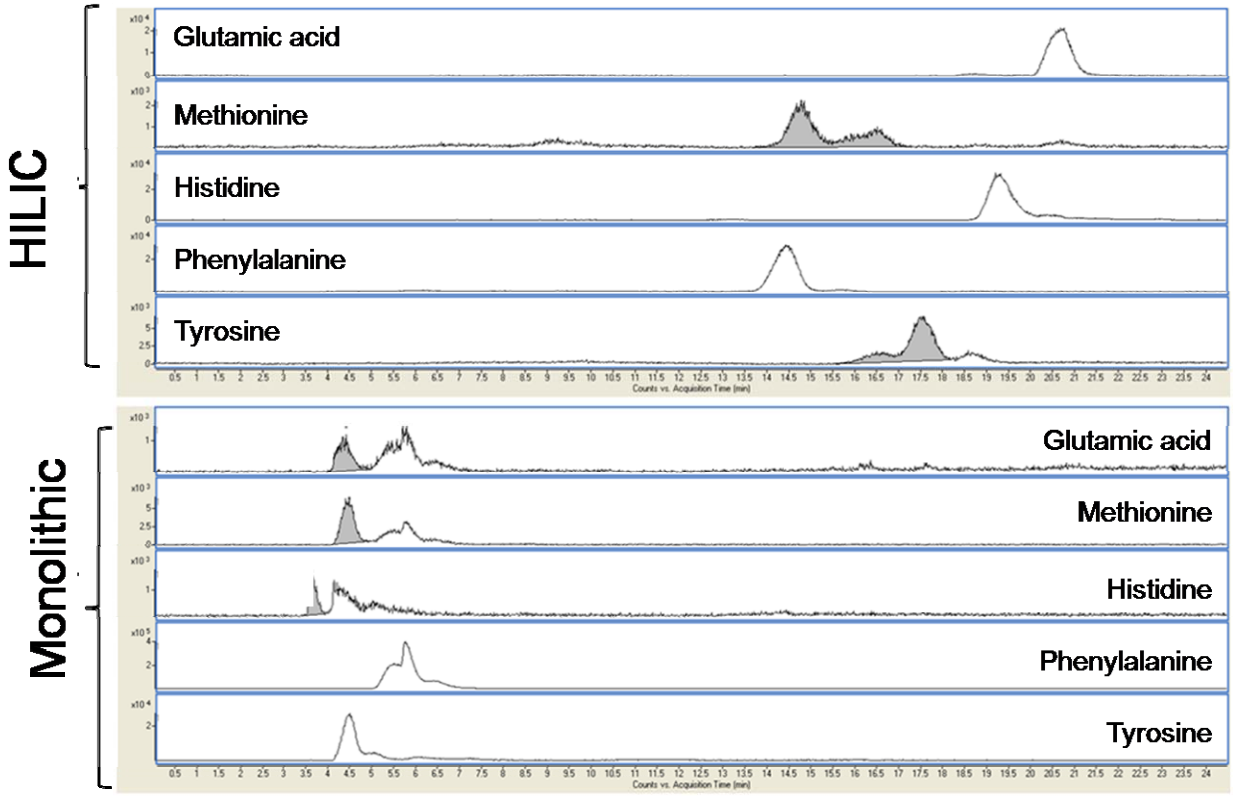


Figure 2-11. Extracted ion chromatograms of glutamic acid, methionine, histidine, phenylalanine, and tyrosine under HILIC and C_{18}-monolithic conditions, respectively. The shaded peaks in each chromatogram correspond to those matching the retention times of the standards.

CHAPTER 3
APPLICATION OF LC/MS METHODOLOGIES TO HUMAN PLASMA SAMPLES FOR GLOBAL AND TARGETED METABOLOMIC STUDIES

Metabolomics Applied to Biomarker Discovery

The metabolome is the most predictive metric of phenotype; thus, metabolomics holds the promise to extensively contribute to the understanding of phenotypic changes as an organism's response to disease, genetic changes, and nutritional, toxicological, environmental, and pharmacological influences [1]. With respect to human health, multiple benefits of metabolomics investigations can be envisioned. This field can deliver new tools to diagnose disease or monitor the success of nutritional and pharmacological interventions, provide new biomarkers to assess human health, and over time a powerful list of diagnostic markers will evolve, which can be measured using high-throughput assays.

Medical diagnosis and treatment efficacy will improve significantly when a more personalized system for health assessment is implemented. This system will require diagnostics that provide sufficiently detailed information about the metabolic status of individuals such that assay results will be able to guide food, drug and lifestyle choices to maintain or improve distinct aspects of health without compromising others [106]. The technologies to accurately and quantitatively understand how the integrated metabolism affects human health are available and they are now been brought into practice for this purpose.

To make sense of metabolite data, the metabolites must first be understood in the context of their biochemical pathways [107, 108]. To explain why the levels of particular metabolites are outside a normal range, and more importantly to predict how to alter inputs –e.g. diet, drugs and lifestyle – to accomplish a selected change in metabolism, an understanding of how pathways and their respective reactions function is required [106]. This is the body of knowledge that the field of biochemistry built throughout the 20th century and continues today. Metabolomics is thus

ideally situated to integrate the existing knowledge of biochemistry into a single, comprehensive strategy to address the health challenges facing the modern world [106].

One of the important objectives of metabolomics is to measure all or a substantial fraction of the metabolites within a biological sample and to quantify each relative to an absolute index of the sample. Using modern analytical technologies, it is often as easy to monitor many analytes as it is to measure one. In this sense, two different analytical strategies can be defined: global and targeted metabolomic approaches. The global or non-targeted approach tries to comprehensively analyze all known and unknown metabolites in a given sample with the ultimate goal of identifying discriminating metabolites. On the other hand, targeted metabolomics has the more modest goal of quantitation of selected metabolites, most typically dozens to hundreds of known compounds. The use of these two complementary approaches will be presented in this chapter.

Ketogenic Diet Therapy

A ketogenic diet (KD) is a high-fat, low-carbohydrate/adequate-protein diet designed to increase the body's dependence on fat rather than glucose for energy and to treat disorders of the brain [109]. The two disorders more commonly treated are epilepsy and certain inborn errors of metabolism involving glucose utilization. Despite its long history of clinical use, it is still not entirely clear what mechanism(s) underlie its epileptic seizure-suppressive action and how the KD affects the brain's general metabolism.

Anti-epileptic medications (AEDs) are commonly the first choice of treatment for epilepsy due to their efficacy and convenience in administration. But despite the large array of (and combinations of) AEDs available, there remain 25% of patients that continue to have seizures, and are deemed to have intractable epilepsy [109]. Many of these intractable epileptic patients are on a heavily concentrated multiple AED regimen and their quality of life has been diminished due to large list of detrimental side effects of AEDs such as sedation, insomnia,

hyperactivity, hepatotoxicity, nausea, vomiting, decreased bone density, metabolic acidosis and behavioral issues [109-112]. Thus the goal of the KD is to improve the quality of life by reducing or eliminating seizures and by reducing the adverse side effects of AEDs by decreasing their dosage.

In the US, ketogenic diet therapy for seizures has been used clinically for more than seven decades with an efficacy similar to that of the best AEDs [109-112]. During those years numerous proposed mechanisms of action have been evaluated, but the mechanism remains unknown [113, 114]. The ketogenic diet clearly increases blood ketones which have the advantage over most other intermediates in energy generating pathways in that they are transportable from the mitochondria of the organ of synthesis to blood and then to the brain to regenerate acetyl-coenzyme A, which can enter the tricarboxylic acid cycle and be used for energy [115]. Acetylcarnitine is also synthesized in energy generating pathways and shares all the transportable advantages of ketones [116]. Whereas expenditure of adenosine triphosphate (ATP) is required for conversion of ketones to acetyl-coenzyme A; acetylcarnitine does not require ATP to produce acetyl-coenzyme A, which is critically important in a pathological condition characterized by a low metabolic energy conditions. In addition to acetylcarnitine possibly being an important source of energy for the brain in KD, many of its other newly discovered metabolic roles may position it in an important role in the mechanism of ketogenic diet therapy [115]. To our knowledge, this is a novel hypothesis concerning the mechanism of ketogenic diet therapy.

To test this hypothesis, plasma samples were collected. The carnitinome and untargeted metabolome of fasting plasma collected from healthy female and male adults (n=10) before beginning the ketogenic diet were compared to the data of fasting plasma obtained in the same

people, after a week of ketogenic therapy. Liquid chromatography interfaced with mass spectrometric analyses was conducted to permit measurement of the targeted carnitinome (free carnitine and the common acylcarnitines from acetylcarnitine to stearoylcarnitine). The objective was to test the hypothesis that ketogenic diet alters the plasma carnitinome and other metabolites. In this sense, nontargeted global metabolomic analyses were used to identify any metabolites that increased or decreased when a KD was followed. The experimental design utilized for this study is presented in Figure 3-1.

Experimental Workflow

The workflow utilized for this study consisted of the following steps:

Sample analysis

The high-throughput screening with LC/MS techniques generates large volumes of analytical data that require advanced software for data mining. MassHunter and GeneSpring MS software (Agilent Technologies, Wilmington, DE) were used to perform high-throughput processing and automated quality analysis of hundreds of gigabytes of MS-based metabolomics data. These platforms also include sophisticated statistical analysis algorithms that allow for data comparison.

Data processing. There is no universal metabolomic software due to the variety of commercial and custom-built instruments and the number of data formats. Therefore, instrument manufacturers (or independent software developers) create their own software [1]. Software is able to perform data extraction to allow data comparison and discovery of differentially expressed metabolites. In our study, data processing was performed by the instrument software. It consisted in the following steps: feature extraction, peak alignment, normalization, filtering and quality control of the data (Figure 3-2).

Feature extraction consists in the detection and extraction of three-dimensional signals, so-called features, which are caused by chemical entities [117]. From an LC/MS perspective, each chemical entity is characterized by a mass-to-charge ratio (m/z ratio), its chromatographic retention time and integrated area.

Peak alignment is the process of finding significant peaks across samples in multiple files. A significant peak is one that is prevalent across samples and is the most intense peak in a certain m/z and retention time range [117]. In the software used for this study, peak alignment was carried out on the compounds read from the data files. The file that has the maximum number of compounds (most populated sample) is picked as a reference and is used for finding significant peaks from the remaining samples. The peaks/compounds of the most populated sample are placed in a two dimensional grid (array), then the compounds from the remainder of the samples are placed at the place of best match. The result is a set of buckets containing compounds found from all the samples. In addition, internal standards can be used to correct for retention time drift. When a feature is found in a particular file whose RT and mass match the specifications of a standard, the standard is considered found.

Normalization is the process of applying mathematical modifications to the values of a variable. The data values are recalculated and used in subsequent analysis. Normalization is applied to avoid neglecting significant, but low-abundance peaks. In other words, normalization is used to adjust the intensities within each sample by reducing the systematic error [117]. Different normalization steps can be applied on the data. Per-run normalizations control for run-wide variations in intensity. Per mass normalization accounts for the differences in detection efficiency between runs and compare the relative change in mass abundance levels.

Filtering and quality control allows the analyst to sort through the large amounts of abundance data, to evaluate the quality of sample data before performing data analysis, and to identify interesting masses for further study after analysis [117]. Filters can be applied to single or multiple data objects. Masses can be filtered on specific abundance criteria and the masses that pass a filter are made into a mass list. The filters can be based on any data associated with the masses including raw or normalized intensity values, fold change comparisons, flag values, statistical information, or raw data from the scanning software. Filtering can be applied to identify masses that fall below a given intensity value threshold, data that exceed recommended signal-to-noise or signal-to-backgrounds measurements, outliers that fall outside the range of standard deviations from the mean, random quantification errors, masses that do not show any changes in abundance during the experiment, interesting masses suitable for additional analysis, etc. The ability to restrict a mass list based on the behavior of its masses in experiments or in individual samples is an important quality control tool.

Data analysis. The statistical tools for the metabolomic data analysis should be selected according to the aim of the study. Commonly used methods include unsupervised and supervised learning techniques. In unsupervised learning problems, the object is often to identify previously unknown structure in the data. In supervised learning, the 'structure' (e.g., classes or groups in the data) is assumed to be known at the outset and this 'knowledge' is used in the statistical analysis [118]. A brief description of the statistical tools applied to this study is given in the following paragraphs.

Hierarchical clustering analysis (HCA) is an iterative agglomerative clustering method that can be used to produce condition trees. These resulting trees group samples or conditions together based on the similarity of their abundance data (across a specified mass list) and allow

picking out groups of similar samples/conditions [117]. In these structures, individual samples are grouped together at the base with larger clusters forming as the tree is "grown" toward the top. Samples that are more similar to one another are shown closer together in the tree and the branching structure shows where masses and group of masses were joined during the clustering steps. At the bottom of the tree all samples exist as the most independent grouping of the objects. Moving upward shows the samples joining into larger branches until all the samples are joined. Some of the advantages of clustering trees are: clustering structure is determined completely by abundance data and if confidence levels are calculated, the significance of clusters can be assessed. On the other hand, some of the disadvantages of this analysis tool are: it is computationally intensive and can be slow for a large starting number of samples; there can be multiple ways to display the branching structure for a given clustering. The generated display is not unique, and there is no easy way to generate a set of discrete clusters from a condition tree.

Principal component analysis (PCA) is a covariance analysis of different factors [117]. Covariance is always measured between two factors. So with three factors, covariance is measured between factor x and y, y and z, and x and z. When more than two factors are involved, covariance values can be placed into a matrix. This is where PCA becomes useful.

PCA finds eigenvectors and eigenvalues relevant to the data using the covariance matrix. Eigenvectors can be thought as of as "preferential directions" of a data set, or in other words, main patterns in the data. For PCA on masses, an eigenvector would be represented as a mass profile that is most representative of the data. For PCA on conditions, an eigenvector is a condition profile. For either PCA, there cannot be more eigenvectors than there are masses or conditions in the data.

Eigenvalues can be thought of as quantitative assessment of how close an eigenvector represents the data. The higher the eigenvalue of an eigenvector, the more representative it is of the data. Eigenvalues can also indicate the level of explained variance as a percentage of total variance and the percent of variance explained is dependent on how well all the components summarize the data.

In a typical MS experiment, the abundance of thousands of m/z values is measured across many conditions or time points. Therefore, it becomes impossible to make a visual inspection of the relationship between m/z values or conditions in such a multi-dimensional matrix. One way to make sense to this data is to reduce its dimensionality. Several data decomposition techniques are available for this purpose: principal component analysis is among these techniques, it reduces the data to just a few dimensions.

PCA reduces data dimensionality by performing a covariance analysis between factors. If each of the thousands of integrated m/z intensities measured in an LC/MS experiment is plotted in a multi-dimensional scatter plot, the result is a cloud (grouping) of values in a multi-dimensional space.

To characterize the trends exhibited by this data, PCA extracts directions where the grouping, which has the appearance of a cloud, is more extended; this is called the first principal component. PCA then looks for the next direction, orthogonal to the first one, reducing the multidimensional cloud grouping into a two-dimensional space. Two components may be able to explain most of the cloud's grouping's trends. However, in a more complex data set, the third, fourth, and higher components might reveal additional information about interesting trends in the data.

PCA is recommended as an exploratory tool to uncover unknown trends in the data. In the case of LC/MS experiments, PCA on integrated mass intensities masses provides a way to identify predominant mass abundance patterns. When applied on conditions, PCA explores correlations between samples or conditions. PCA is not a clustering tool and does not attempt to group masses by user-specified criteria as do the clustering tools.

Analysis of Variance (ANOVA) is based on a probability value (p-value), and allows the analyst to determine if one given factor, such as drug treatment, has a significant effect on mass abundance behavior across any of the groups under study [117]. The p-value indicates the probability of getting a mean difference between the groups as high as or higher than what is observed. The lower the p-value, the more significant the difference between groups. In other words, a significant p-value resulting from a 1-way ANOVA test would indicate that a mass is differentially expressed in at least one of the groups analyzed. However, if there are more than two groups being analyzed, the one-way ANOVA does not specifically indicate which pair of groups is statistically differentially expressed. *Post hoc* tests can be applied in this specific situation to determine which specific pair/pairs are differentially expressed [117]. If the Tukey *post hoc* test is applied, all means for each condition are ranked in order of magnitude, from the lowest to the largest. The group with the lowest mean gets a ranking of 1. The pairwise difference between means, starting with the largest mean compared with the smallest mean are tabulated between each group pair and divided by the standard error. This value, q, is compared to a t-test range critical value. If q is larger than the critical value, the pair of groups is considered to be statistically differentially expressed.

Volcano plots are useful tools for visualizing differential abundance between two different conditions [117]. They are constructed using fold-change values and p-values, and thus allow

visualizing the relationship between fold (magnitude of change) and statistical significance (which takes both magnitude of change and variability into consideration). They also allow subsets of masses to be isolated, based on those values. A volcano plot is basically a scatter plot, where each mass is plotted on two axes corresponding to p-value (p-value for a t-test of differences between samples) and fold-change (magnitude o change). Both axes are typically on a log scale, a negative log scale for the p-value, so smaller p-values appear higher up; and a positive and negative log scale for the fold-change, so the changes appear symmetric. The fold-change indicates biological impact of the change; the p-value axis indicates the statistical evidence, or reliability of the change.

Identification

Numerous tools are normally used to identify metabolites: mass accuracy, retention times, UV signals, fragmentation patterns, elemental composition, structural information, and library and database searches. Searching a database of metabolite information can help to narrow the list of possible candidates. Additionally, accurate mass data make this database searching more effective by narrowing the mass window and thus reducing the number of possible identities that need to be searched. In these experiments, a database search was conducted on the significant masses for identification of putative biomarkers and it was constrained by the two ppm mass tolerance specification for the Agilent 6210 ESI-TOF. Two different databases were utilized and compared: METLIN (METaboliteLINk) and HMDB (Human Metabolome DataBase).

Biological interpretation

Biological interpretation is where technology, identification, and knowledge need to be integrated. However, no metabolite ontology has been developed yet for metabolites and the interpretation of correlations is still a challenge [10]. Figure 3-2 lists all the steps that form part of the experimental workflow used for this study.

The above-mentioned workflow has been applied to the evaluation of the hypothesis that the metabolome will change as a consequence of a ketogenic diet therapy. It is well-known that the mechanism behind KD is connected with the radically altered macronutrient metabolism [115]. This shift in metabolism is expected to cause major alterations in the plasma metabolites including acylcarnitines, which have been noted to change during such macronutrient changes [119]. The changes in metabolites and acylcarnitines can be measured and evaluated using the LC/MS, data processing, and data analysis global and targeted metabolomic approach. Since the study is designed to assess the effects of a ketogenic diet on healthy adults, the results found will reflect only changes due to KD, unlike the metabolome to be examined from patients with the added complications of seizures and AEDs. The ultimate goal is to understand the effect of the KD on the metabolome/carnitinome of intractable epileptic patients, which will be addressed in subsequent studies.

Experimental Methods

Sample Collection

The study was conducted at the General Clinical Research Center (GCRC) of Shands Hospital at the University of Florida and was a part of a larger study evaluating the KD's effects on dyslipidemia and insulin sensitivity [120]. The protocol was approved by the Scientific Advisory Committee of the GCRC and the Institutional Review Board (IRB) of College of Medicine. All subjects signed the informed consent form. Twenty young (10 males and 10 females), healthy first year medical students volunteered to participate in the study. The subjects were screened and excluded if they were pregnant, vegetarian or an endurance athlete, or if they had cardiovascular disease, diabetes mellitus, neurological diseases, kidney disease, any food allergies, or if on any medications. Subjects were matched for gender and body mass index (BMI) and randomized by parallel design into two groups, polyunsaturated fatty acids (PUFA) or

saturated fatty acids (SFA). There were no significant differences between age, height, and BMI between the two groups. On the first day of study the subjects were admitted to the GCRC and fasted overnight. The following morning two vials of venous blood were obtained for later carnitinomics and metabolomics analysis. The subjects then received their respected diets of 60% PUFA, 15% SFA, and 25% MUFA (monounsaturated fatty acids, the PUFA diet) or 60% SFA, 15% PUFA, and 25% MUFA (the SFA diet). Both diets consisted of 70% of fatty acids, 15% of carbohydrates and 15% of proteins, which is the equivalent of a 2.33:1 ratio. The diets were designed to maintain the subject's weight. This was done through the collection of a 4 day non-consecutive diet recall before initiating the diet and careful monitoring of weight during the diet. The subjects only ate food provided by the GCRC and on the 5th day the subjects' venous blood was again collected for acylcarnitine and metabolomic analysis.

Sample Preparation

The volunteers came in fasting and their blood was on ice or in a freezer at all times during preparation. Once blood was received, it was immediately centrifuged at 4,750 rpm and 4-6 °C to separate the red blood cells and plasma compartments. It was then brought back to the Food Science and Human Nutrition Department (FSHN) lab on ice and stored at -20 °C until further preparations took place. The experimental section as detailed in Chapter 2 is also applicable to Chapter 3.

Results and Discussion

LC/MS Experiments

In this study, the hypothesis that the plasma metabolic profile of healthy female and male adults before and after ketogenic diet therapy will differ was tested. Since ten female and ten male participated in this study and each sample was run in triplicate, 30 files were obtained for the experiments related to the female samples before and 30 files after the KD using the

monolithic column approach. In addition, 30 files were obtained for the female samples using the HILIC column. On the other hand, 30 files per chromatographic method were acquired for the male samples before and after the ketogenic diet. Thus, 120 files per chromatographic method were used for further data analysis and processing. Total ion chromatograms (TICs) corresponding to the overlapping of all files for the female and male samples before and after the ketogenic diet using monolithic and HILIC chromatography are shown in Figures 3-3 – 3-6. Changes in the chromatographic profiles can be observed from these figures. In this sense, reproducibility was noticeable better for the TICs obtained with the C_{18}-approach. The variation in the TIC's shape for the HILIC methodology can be a consequence of run-to-run changes in the interaction between samples and stationary phase.

Global Metabolomics Strategy

This approach focused on the comprehensive analysis of all known and unknown compounds present in the samples under study. For data processing and analysis, the above-mentioned statistical tools were applied to the LC/MS files.

Hierarchical clustering analysis

As mentioned previously, in a hierarchical clustering analysis samples that are more similar to one another are shown closer together in the tree structure. In the case of MS data, the tree structures are determined by abundance data. Cluster analysis was applied to the raw data (163,309 masses corresponding to C_{18}-monolithic- and 94,070 masses to HILIC-experiments) in order to evaluate the female and male samples before and after ketogenic diet using C_{18} - monolithic and HILIC strategies. The dendrograms generated revealed two major groups: male and female samples; however, no clear differentiation based on diet-before and after- was observed. These results are shown in Figures 3-7 and 3-8, in which the female samples before

and after the diet are highlighted in yellow and red, respectively; and the male samples before and after the diet are highlighted in bright and dark blue.

Principal component analysis

PCA is able to reveal relationships between different experiment interpretations, e.g. gender, diet type, etc. Unlike cluster analysis, PCA does not have the constraint of generating a hierarchic classification. Therefore, 120 files were subjected to PCA. Prior to performing PCA, the raw masses were filtered based on relative frequency tool, which allowed only the most frequent masses across all the samples to remain part of the analysis. PCA plots were obtained and they are shown in Figures 3-9 and 3-10. The PCA plots revealed the presence of gender differences, which was visualized as two well-defined groups and these results agreed for both C_{18}-monolithic and HILIC chromatography. However, PCA was not able to find diet correlations across the samples.

Analysis of variance

ANOVA is applied to compare between multiple sample classes. The so-called post-hoc tests are used to further locate the differences between any of these classes. Both methods are dependent on the number of replications and the distribution properties of measurements. The significance of differences is typically expressed by a probability value (p-value). ANOVA applied to the above-mentioned data sets (masses from C_{18}-monolithic and HILIC approaches after being filtered on relative frequency) demonstrated that the minor variances of a data set may contain valuable discriminatory information. In this sense, not only differences between genders were observed, but also between samples before and after diet. Figure 3-11 summarizes the results for the ANOVA test for both types of column approaches. In the case of the C_{18}-approach, a total of 7,385 masses were found statistically significant by the ANOVA analysis, 593 masses were different between females before and after diet and 853 masses between males.

On the other hand, when the results for the HILIC samples were obtained, a total of 1571 masses were found differentially expressed, 468 masses were different between female samples before and after diet and 67 between male samples.

Volcano analysis

In order to evaluate the magnitude of change in those statistically significant masses obtained by the ANOVA tests, Volcano analysis was used. As mentioned earlier, the Volcano analysis uses the p-value (probability value) and the magnitude of change, or fold-change, to determine the significantly expressed masses. In this study, a p-value cutoff of 0.05 (0.05 false discovery rate) and a fold change equal or greater than two were utilized to perform the analysis. The ANOVA masses obtained from the comparison between females and males before and after diet were submitted to Volcano analysis. Thus, for the female samples analysis using C_{18}-monolithic columns, 41 (7%) masses out of 593 passed the filter; whereas 112 (13%) masses out of 853 passed the filter for the male samples (Figure 3-12). On the other hand, when the dataset from HILIC column separation was used, 64 (14%) masses out of 468 passed the filter for the female samples and 25 masses out of 67 (37%) for the males (Figure 3-13).

Database search for identification

The masses that passed the Volcano test were submitted for identification to two databases: METLIN and HMDB. Using a mass tolerance equal to 2 ppm, the Venn diagrams in Figures 3-14 and 3-15 show the number of masses identified in each of the databases. In all the cases, the HMDB database contributed a higher percentage of identities for the putative biomarkers. The tentative biomarkers were sorted in different classes of compounds based on the biochemical information provided by the HMDB database (Figures 3-16 and 3-17). In this sense, no classes of compounds were common to males and females results when using the C_{18}-separation

strategy. However, three classes of compounds-acyl glycines, amino acid conjugates, and acylcarnitines- were common for both males and females results when HILIC was applied.

Targeted Metabolomic Strategy

Carnitinomics: Carnitine's intimate role in fatty acid metabolism makes it and other acylcarnitines potential biomarkers for understanding changes in macronutrient metabolism. Thus, a subset of compounds composed of carnitine and some important acylcarnitines was used for the targeted approach. These compounds have already been introduced in Chapter 1 and their structures shown in Figure 1-3. In this targeted study, the attention was focused on evaluation of the effects of the ketogenic diet on 11 acylcarnitines: carnitine (C_0), acetylcarnitine (C_2), propionylcarnitine (C_3), butyrylcarnitine (C_4), hexanoylcarnitine (C_6), octanoylcarnitine (C_8), decanoylcarnitine (C_{10}), lauroylcarnitine (C_{12}), myristoylcarnitine (C_{14}), palmitoylcarnitine (C_{16}), and stearoylcarnitine (C_{18}); respectively. The instrument response, expressed as peak area counts, was extracted for each compound in each of the data files. Then, these peak areas for each compound in each of the data files were added to obtain the total area. Thus, the areas of each compound were compared to the total area and the corresponding percentages of the total area were obtained. A ratio of the percentages of each person after (A) and before (B) the ketogenic diet was calculated and the A/B ratio was plotted versus carnitine and acylcarnitines; thus, profiles for all of the females and males were obtained (Figures 3-18 and 3-19). From these plots, it is possible to observe that the ketogenic therapy affected carnitine and acylcarnitines profiles and these changes were more pronounced for some of the male samples. On the other hand, a higher variability can be observed in the case of the female samples. This is the first time that a metabolomic approach is applied to evaluate the changes in the carnitinome of healthy adults before and after a ketogenic diet and there is no evidence or explanation in the literature about the observed profiles. Our research related to the biochemical significance of the results is

underway. In addition to the mentioned plots, averaged values of the previously calculated A/B ratios versus targeted carnitine and acylcarnitines profiles and averaged values of the A/B ratios versus total carnitine and acylcarnitines profiles were plotted and they are shown in Figures 3-20 and 3-21 for the female samples, and in Figures 3-22 and 3-23 for the male samples. In summary, it is possible to conclude that carnitine and almost all of the acylcarnitines changed due to the effects of the ketogenic diet.

Conclusions

Two complementary HPLC column technologies, C_{18}-monolithic and HILIC, were applied to underivatized human plasma samples successfully. The extensive amount of information generated was processed and analyzed with sophisticated software tools. Clustering and principal component analyses showed a distinct grouping of the males and females, but no diet effect was observed. After applying analysis of variance, statistically significant masses for females and males before and after ketogenic diet therapy were identified. As shown from the ANOVA results, five percent of the total number of masses distinguished the diet for the female samples, and a similar number for the males. Volcano analysis allowed determining the masses that increased or decreased by a factor equal or greater than two. Those masses were submitted for identification, based on mass accuracy, to METLIN and HMDB databases, different percentages of the compounds were found as illustrated with the Venn diagrams and lists of classes of compounds were generated.

In addition to the global strategy, a targeted metabolomic approach was applied and carnitine and acylcarnitines profiles were generated. These profiles demonstrated the significant effect of the ketogenic diet on the targeted compounds.

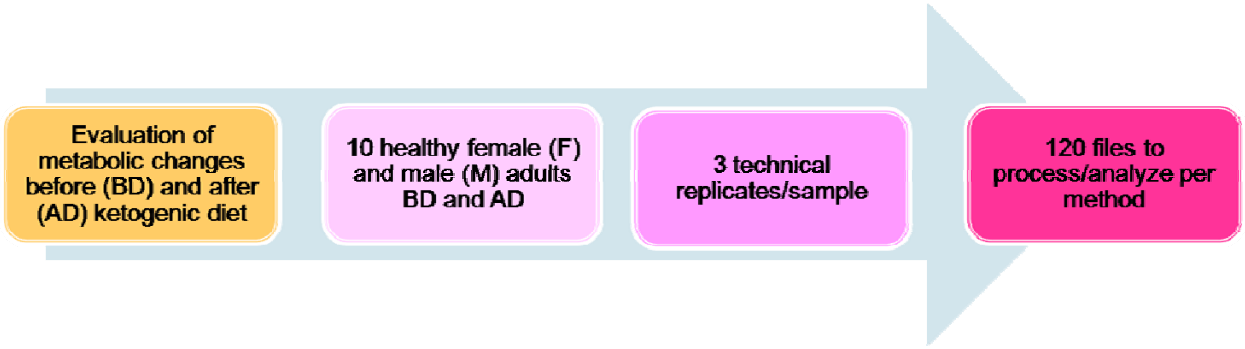


Figure 3-1. Study design for the evaluation of the metabolomic effects of a ketogenic diet on healthy adults. Plasma samples from ten healthy female and ten healthy male adults were collected before and after a ketogenic diet therapy. LC/MS approaches were applied for the analysis and three technical replicates per sample were analyzed. Both monolithic C_{18}-RPLC and HILIC column were employed followed by positive ion electrospray MS.

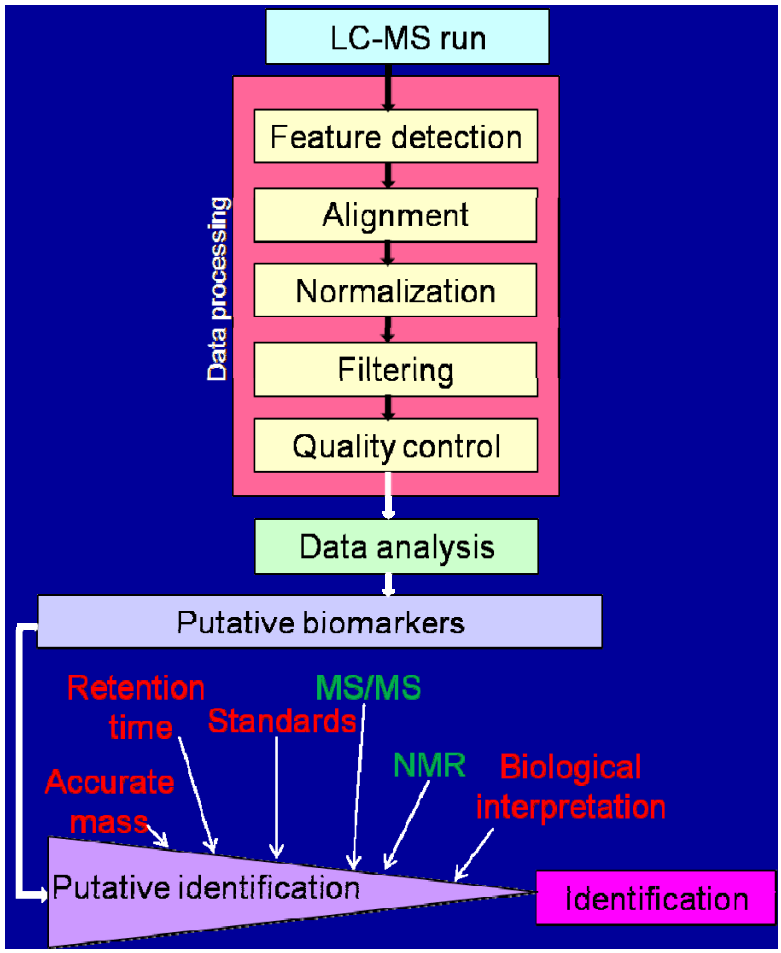


Figure 3-2. Experimental workflow. The raw instrument data were processed in several steps by the instrument software. From the obtained chromatograms, features (integrated mass-intensity pairs) were first extracted, aligned, normalized, filtered and visualized for quality control evaluation (data processing sequence), and then analyzed by different statistical methods, including hierarchical clustering analysis (HCA), principal components analysis (PCA), analysis of variance (ANOVA), and Volcano plots. Finally, using the accurate mass criterion of 2 ppm mass error, putative biomarkers were identified in the databases METLIN and HMDB.

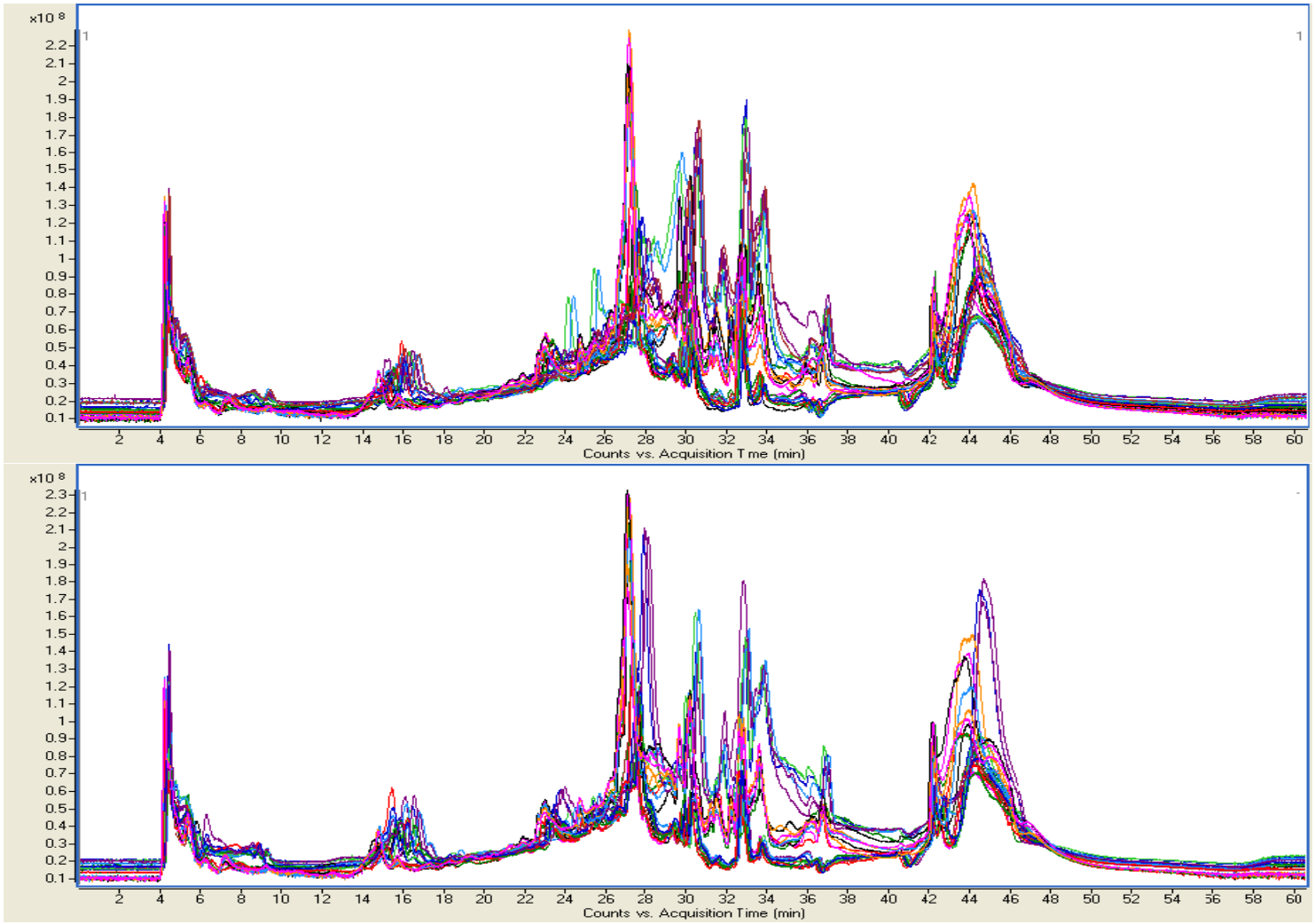


Figure 3-3. Total ion chromatograms (TICs) for the overlapping of all female samples (30 files) before (top) and after ketogenic (bottom) diet using C_{18}-monolithic chromatography. Solvent A: 1% (v/v) acetic acid in water; Solvent B: 1% (v/v) acetic acid in acetonitrile; Flow rate (mL min^{-1}): 1.0. Injection volume (µL): 15.0; Gradient: A:B (min): 95:5 (0.0-6.5); 0:100 (25.5-35.0); 95:5 (55.0-65.0).

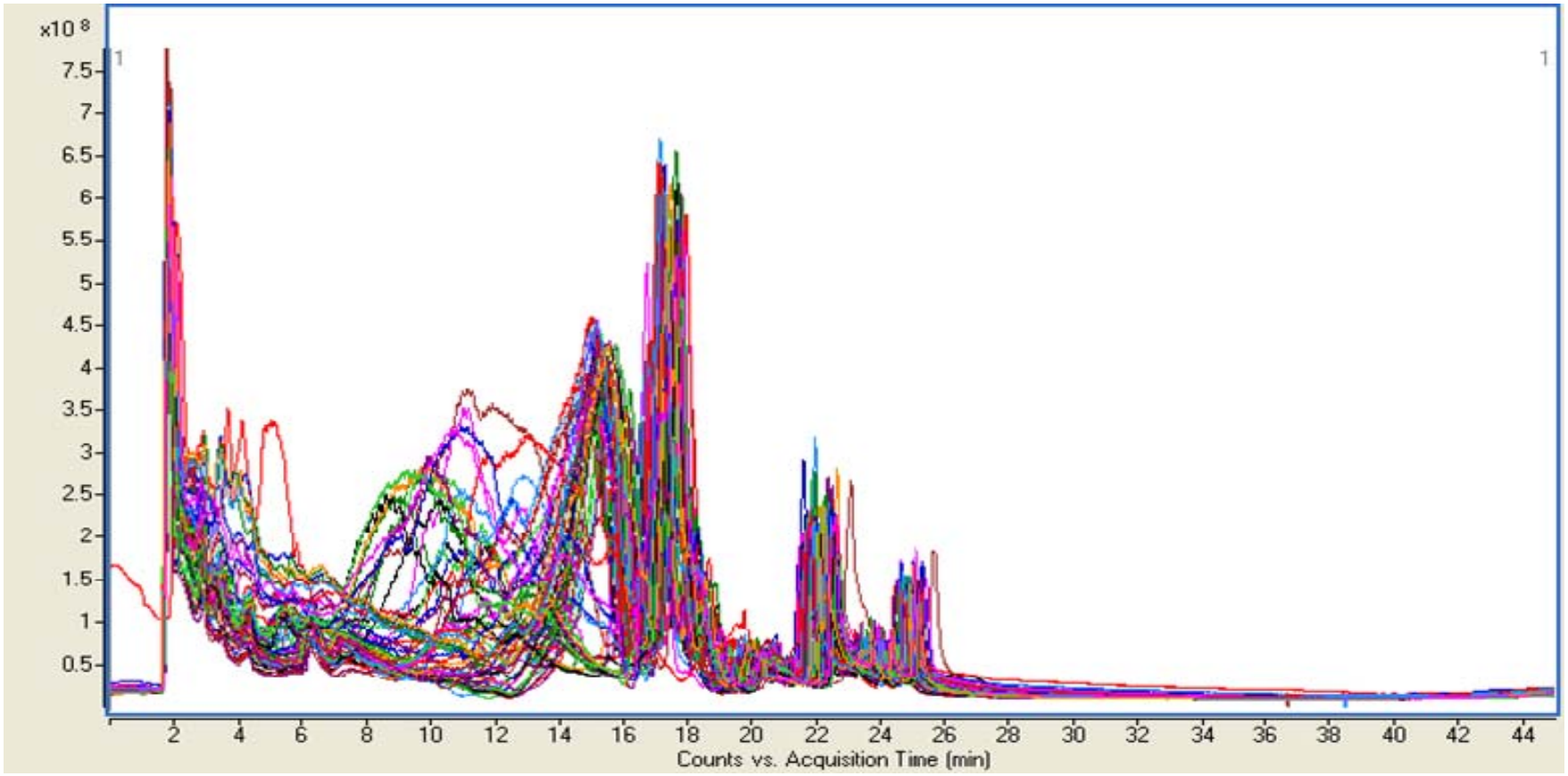


Figure 3-4. Total ion chromatograms (TICs) for the overlapping of all female samples (30 files) before and after ketogenic diet using HILIC chromatography. Solvent A: 7.5 mM ammonium formate in water; Solvent B: 7.5 mM ammonium formate in acetonitrile; Flow rate (mL min-1): 0.3. Injection volume (µL): 10.0; Gradient: A:B (min): 10:90 (0.0-5.0); 50:50 (25.0-30.0); 10:90 (40-45).

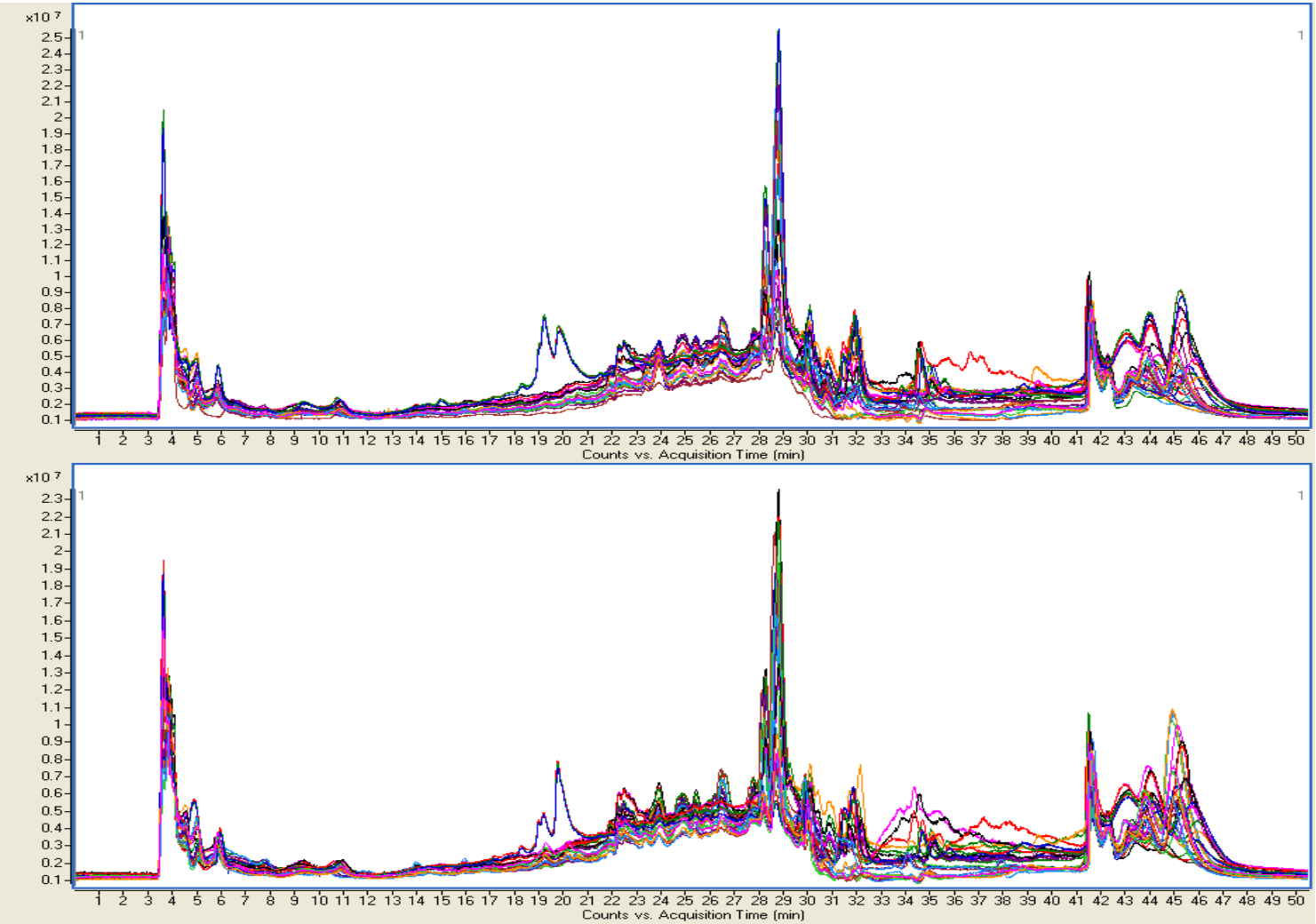


Figure 3-5. Total ion chromatograms (TICs) for the overlapping of all male samples (30 files) before (top) and after (bottom) ketogenic diet using C_{18}-monolithic chromatography. Solvent A: 1% (v/v) acetic acid in water; Solvent B: 1% (v/v) acetic acid in acetonitrile; Flow rate (mL min-1): 1.0. Injection volume (µL): 15.0; Gradient: A:B (min): 95:5 (0.0-6.5); 0:100 (25.5-35.0); 95:5 (55.0-65.0)

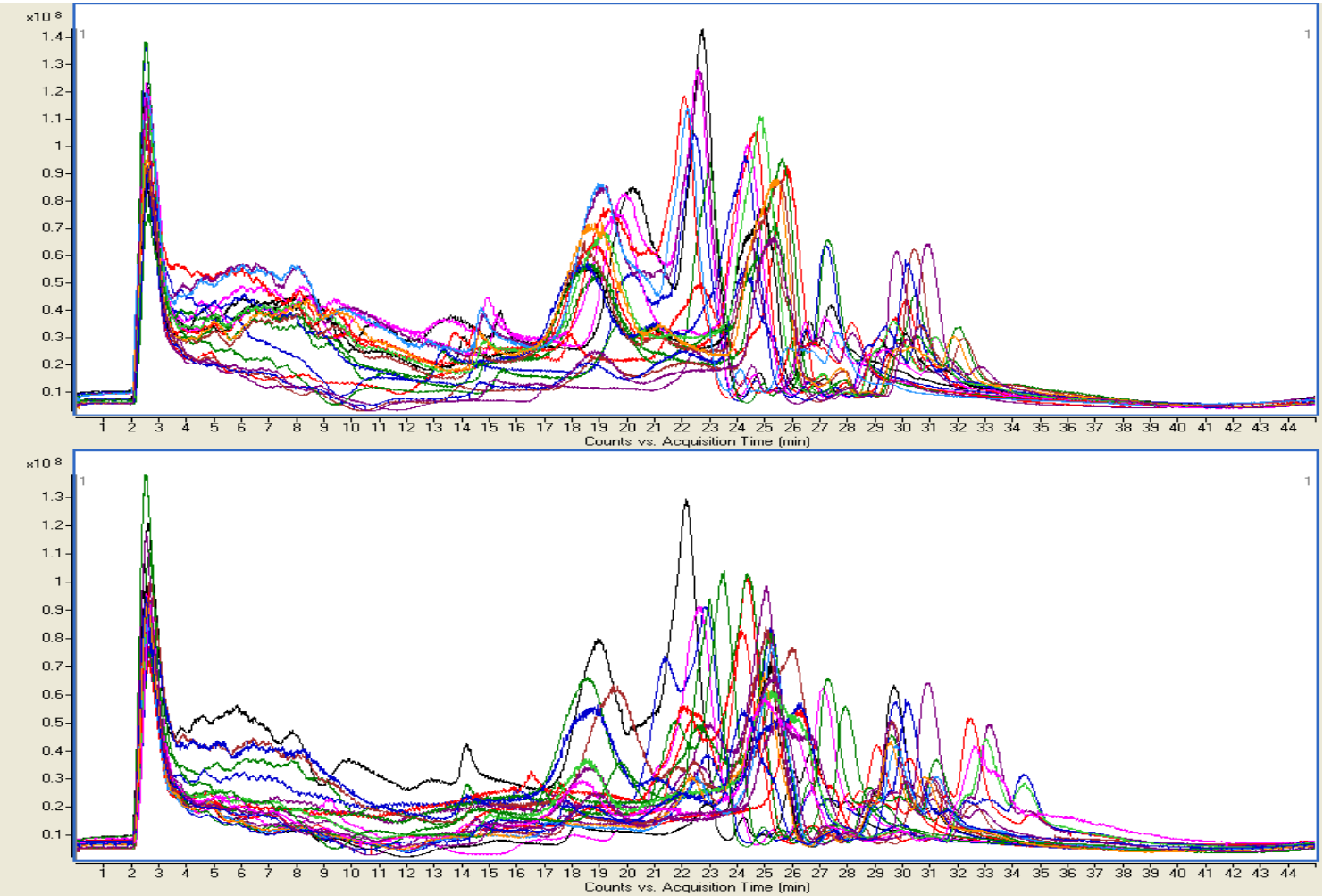


Figure 3-6. Total ion chromatograms (TICs) for the overlapping of all male samples (30 files) before (top) and after (bottom) ketogenic diet using HILIC chromatography. Solvent A: 7.5 mM ammonium formate in water; Solvent B: 7.5 mM ammonium formate in acetonitrile; Flow rate (mL min^{-1}): 0.3. Injection volume (µL): 10.0; Gradient: A:B (min): 10:90 (0.0-5.0); 50:50 (25.0-30.0); 10:90 (40-45).

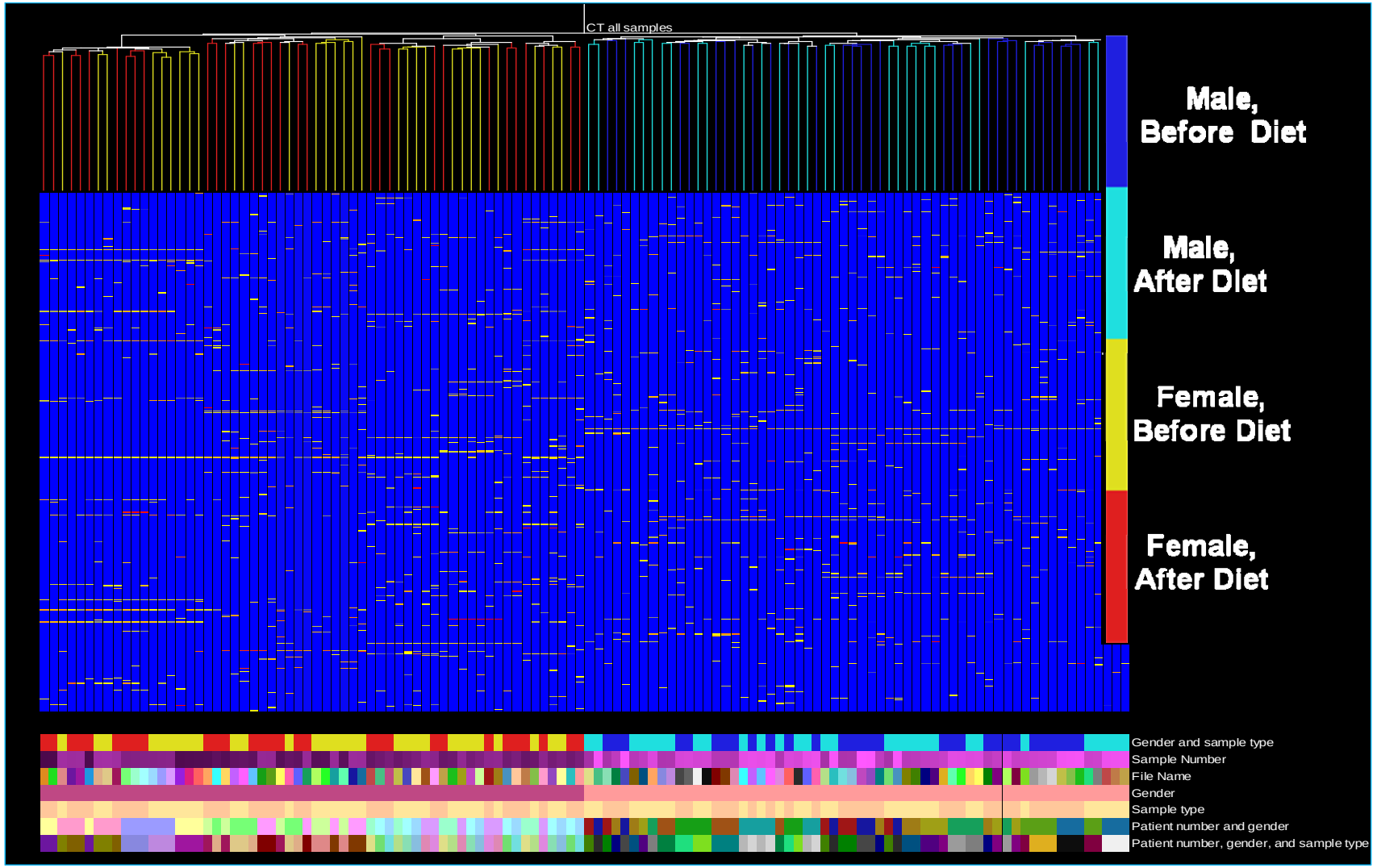


Figure 3-7. Hierarchical clustering analysis (C_{18}-monolithic column, #masses: 163,309) resulting in grouping by different experiment interpretations (gender, sample type, sample number, patient number, and file name). The tree is clearly divided into two groups: female and male samples (see color key on the right), but no differentiation based on diet was obtained.

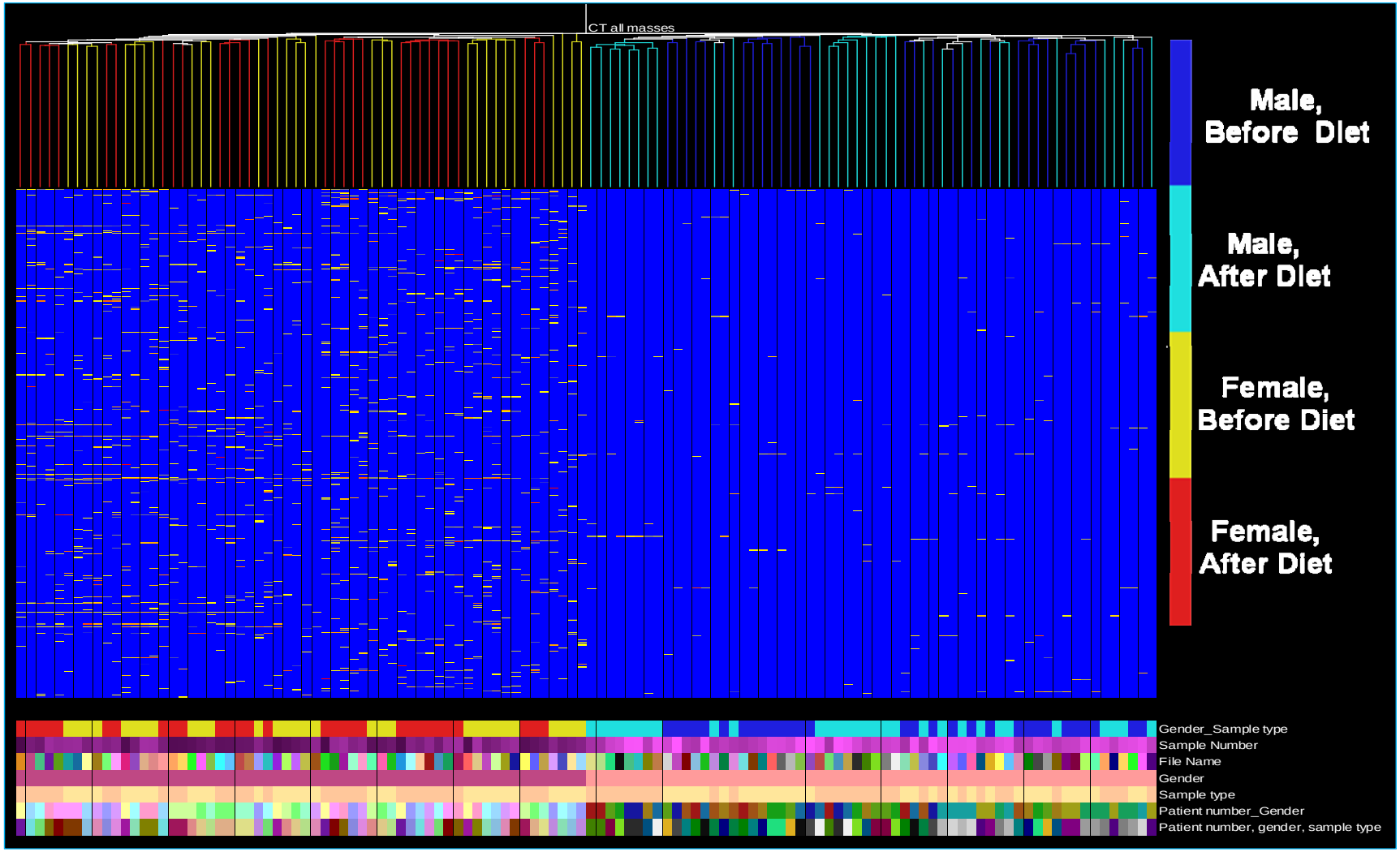


Figure 3-8. Hierarchical clustering analysis (HILIC column, #masses: 94,070) resulting in grouping by different experiment interpretations (gender, sample type, sample number, patient number, and file name). The tree is also divided into two groups: female and male samples (see color key on the right), but once again no diet differentiation was observed.

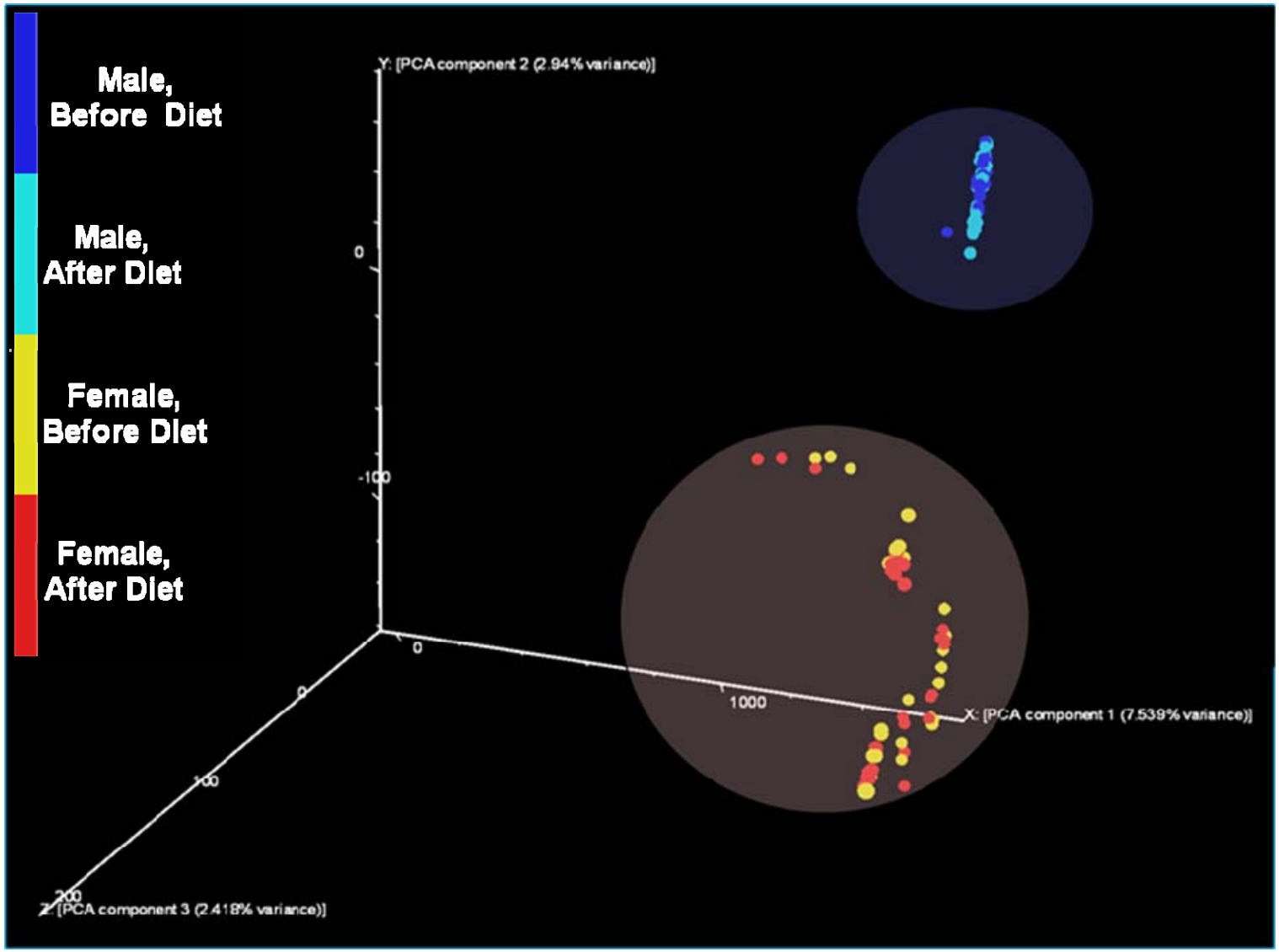


Figure 3-9. Principal component analysis on C_{18}-monolithic column data files. The PCA plot shows correlations across gender.

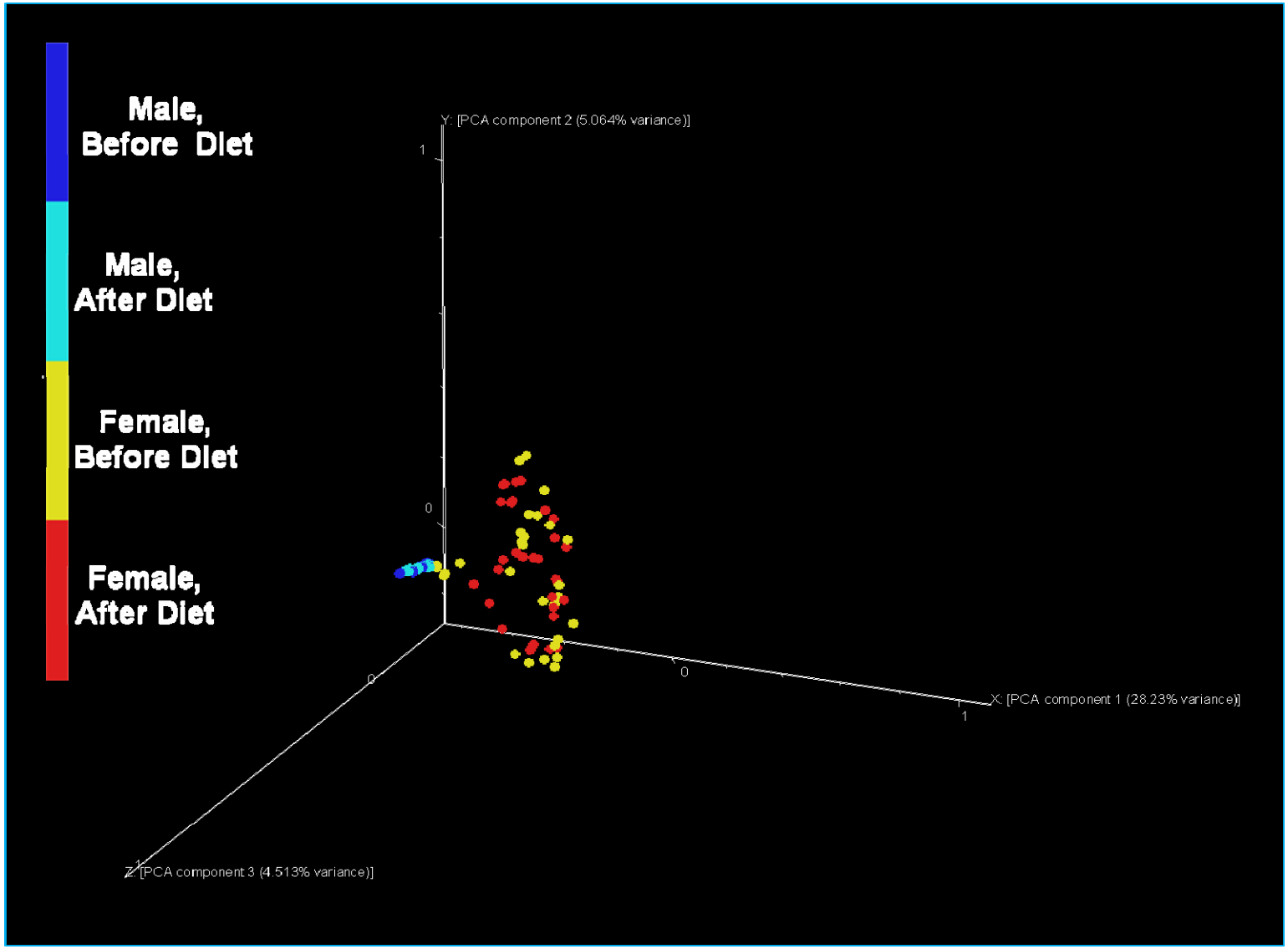


Figure 3-10. Principal component analysis on HILIC data files. The PCA plot shows correlations across gender.

C_{18}-monolithic	Female, BD	Female, AD	Male, AD	Male, BD
Female, BD	7385	593	7134	7138
Female, AD	6792	7385	7092	7087
Male, AD	251	293	7385	853
Male, BD	247	298	6532	7385
HILIC	Female, AD	Female, BD	Male, AD	Male, BD
Female, AD	1571	468	1367	1392
Female, BD	1103	1571	1386	1413
Male, AD	204	185	1571	67
Male, BD	179	158	1504	1571

Figure 3-11. One-way analysis of variance (ANOVA) for C_{18}-monolithic and HILIC approaches. Test type: parametric, do not assume variances equal; False discovery rate: 0.05; Multiple testing correction: Benjamini and Hochberg false discovery rate; Post Hoc test: Tukey; BD: before diet; AD: after diet. The red boxes represent the statistically different masses, the grey boxes the total number of masses, and the blue boxes the statistically similar masses.

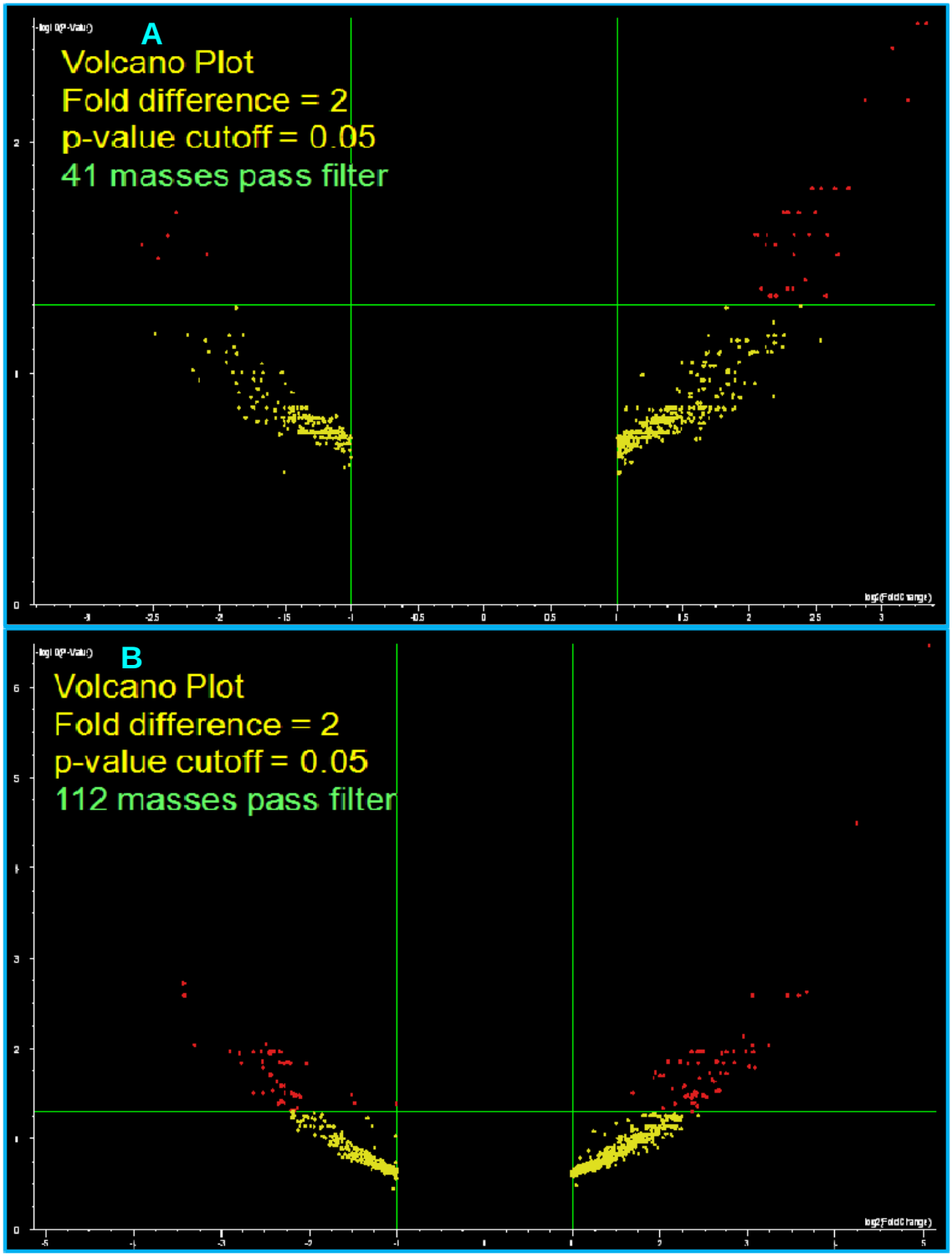


Figure 3-12. C_{18}-chromatography - Volcano plots on differential abundance (fold change-magnitude of change- and p-value-magnitude of change and variability-) between females A) and males B) before and after diet. Red dots correspond to masses that pass the test.

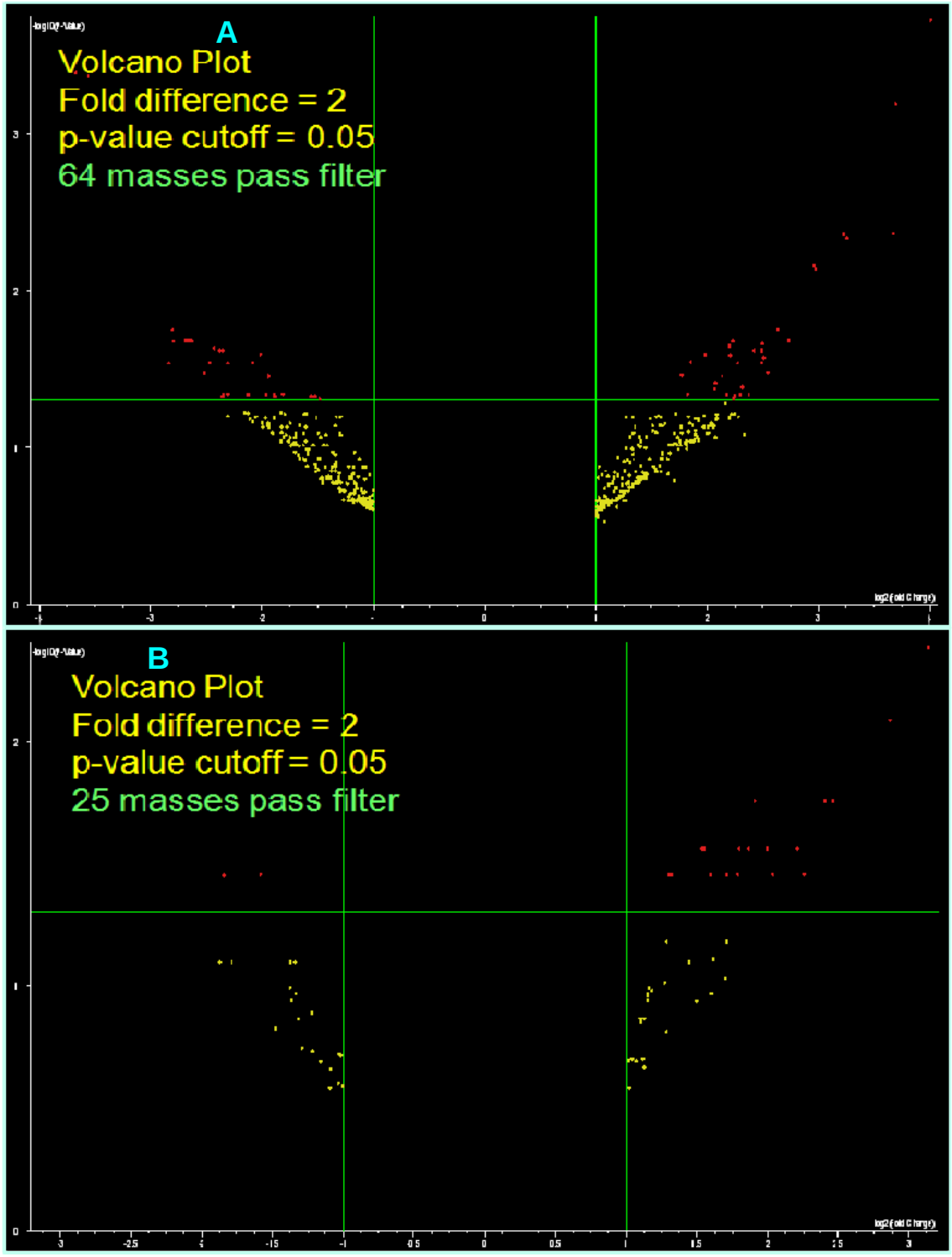


Figure 3-13. HILIC-chromatography - Volcano plots on differential abundance (fold change-magnitude of change- and p-value-magnitude of change and variability-) between females A) and males B) before and after diet. Red dots correspond to masses that pass the test.

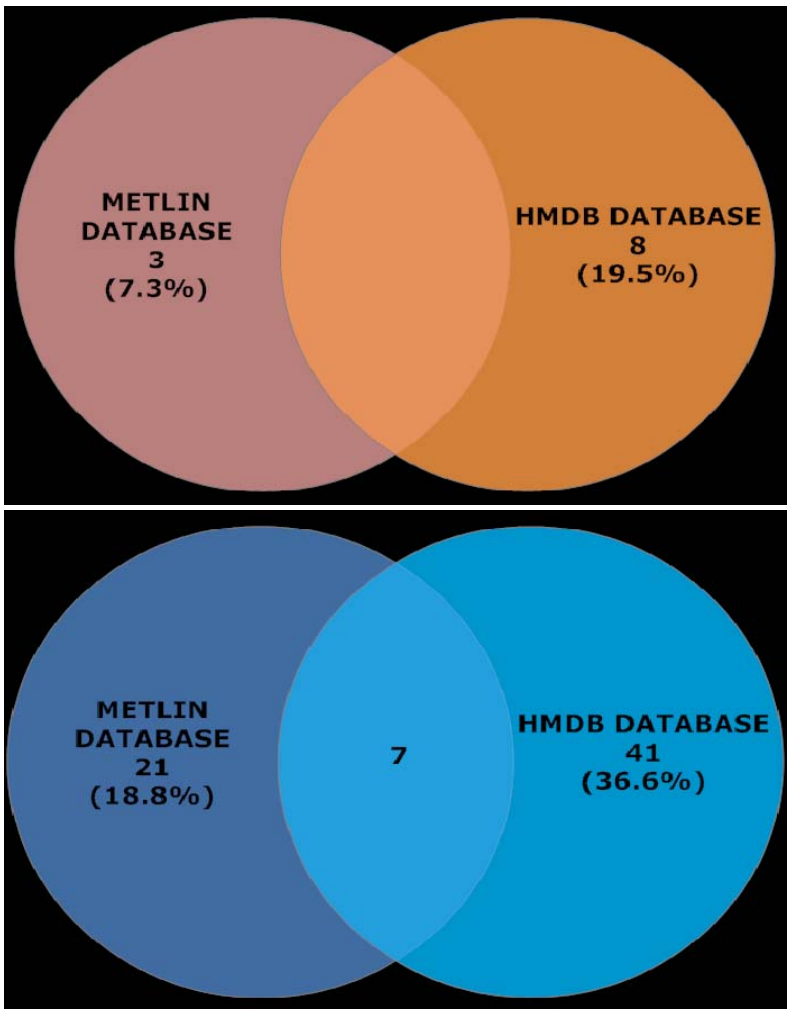


Figure 3-14. C_{18}-chromatography – Venn diagrams. The masses that passed the Volcano test were submitted for identification to several databases. The diagrams show the number of masses identified in METLIN and HMDB. The pink-orange diagrams correspond to the group of masses obtained after comparing female samples and the light-dark blue ones to the masses from the male analyses.

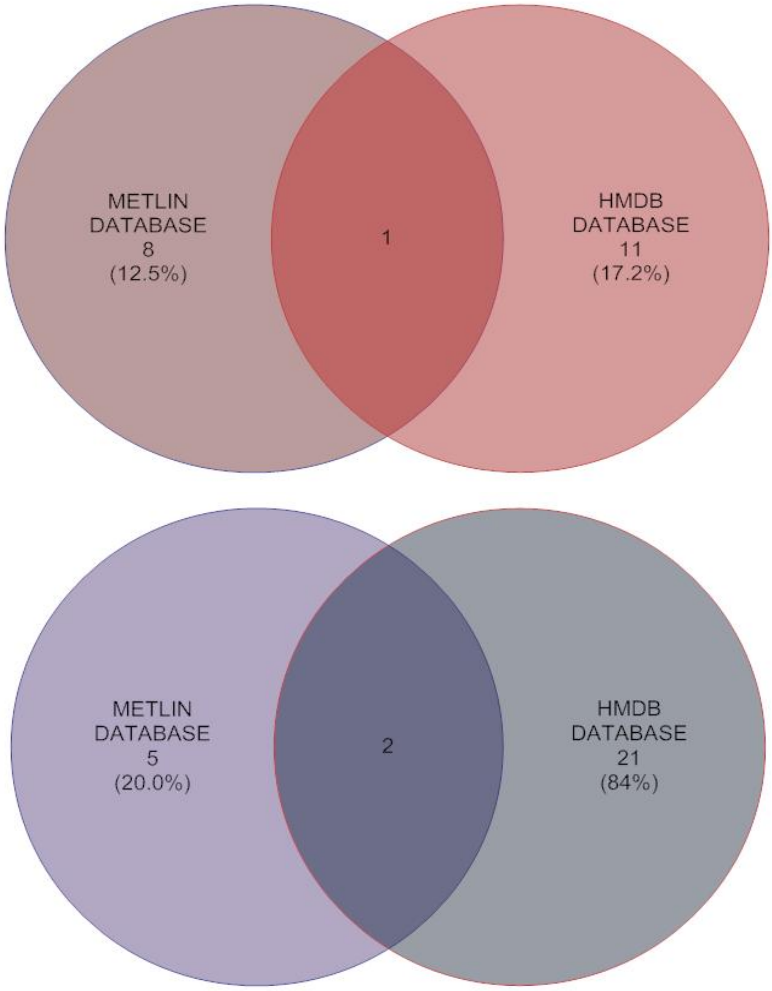


Figure 3-15. HILIC chromatography – Venn diagrams. The masses that passed the Volcano test were submitted for identification to several databases. The diagrams show the number of masses identified in METLIN and HMDB. The pink-orange diagrams correspond to the group of masses obtained after comparing female samples and the light-dark blue ones to the masses from the male analyses.

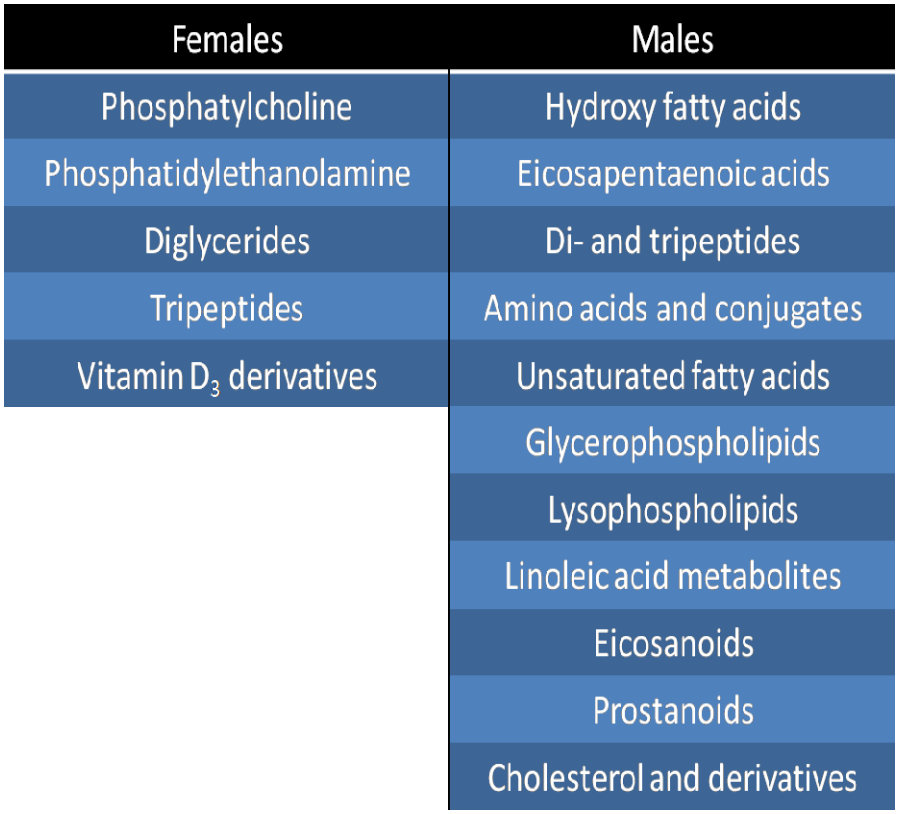


Figure 3-16. C_{18}-chromatography – Classes of compounds identified as changing as the result of the ketogenic diet.

Females	Males
Acyl glycines	Acyl glycines
Glycerolipids	Hydroxy fatty acids
Amino acid conjugates	Tripeptides
Acylcarnitines	Amino acids and conjugates
	Acylcarnitines
	Diacylglycerols
	Lysophospholipids
	Phospho ether lipids

Figure 3-17. HILIC chromatography – Classes of compounds identified as changing as the result of the ketogenic diet.

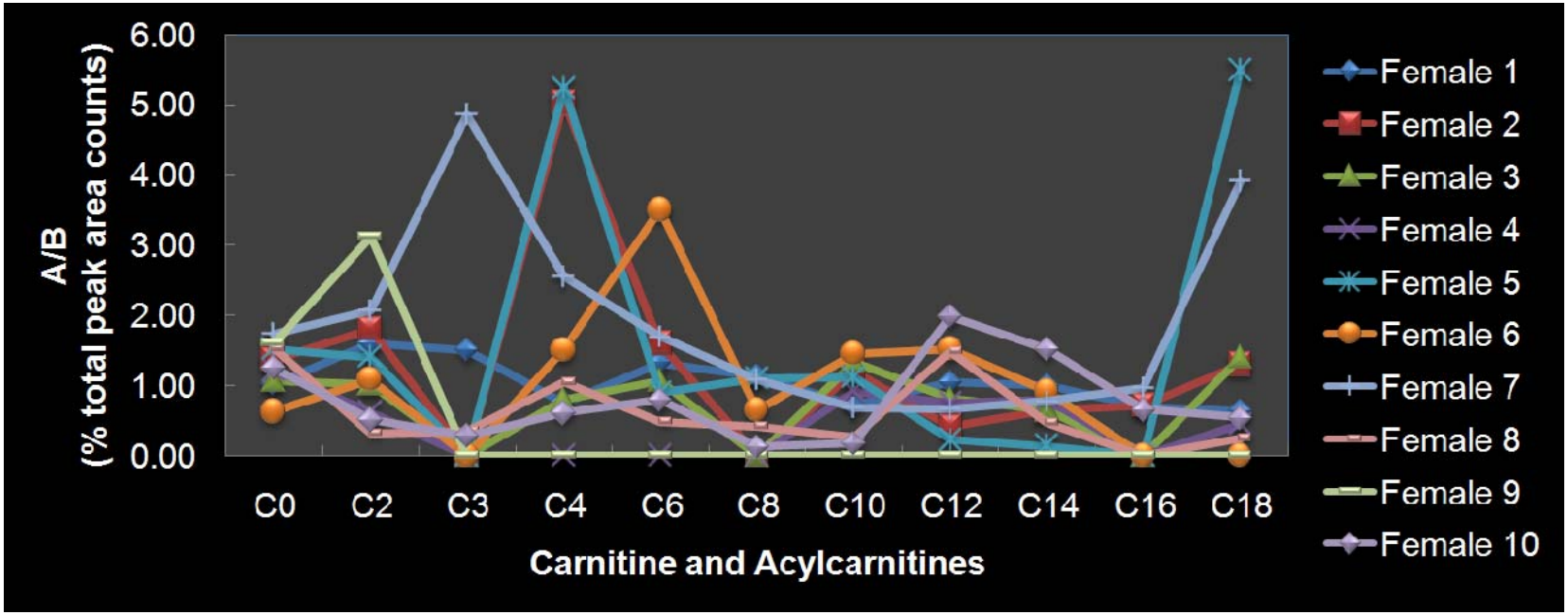


Figure 3-18. A/B (after diet/before diet) ratio response *vs.* carnitine and acylcarnitines in each of the ten female samples. C_0: carnitine, C_2: acetylcarnitine; C_3: propionylcarnitine; C_4: butyrylcarnitine; C_6: hexanoylcarnitine; C_8: octanoylcarnitine; C_{10}: decanoylcarnitine; C_{12}: lauroylcarnitine; C_{14}: myristoylcarnitine; C_{16}: palmitoylcarnitine; C_{18}: stearoylcarnitine.

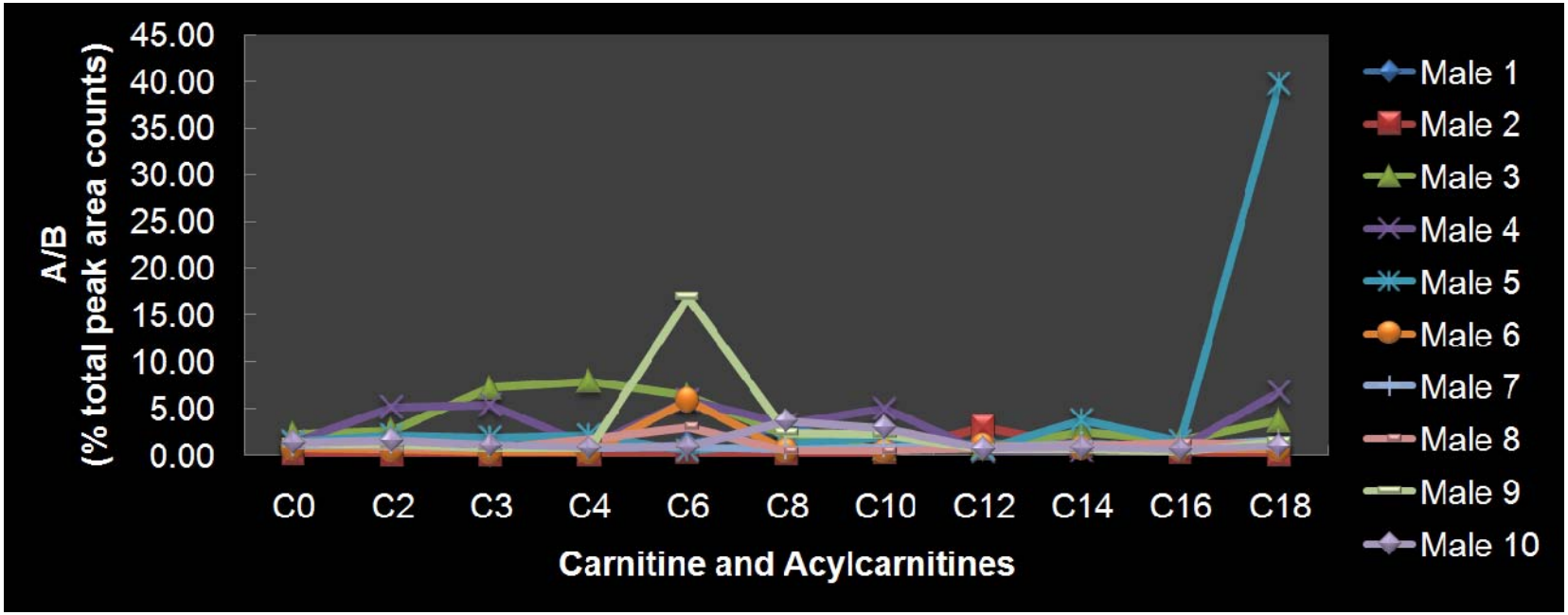


Figure 3-19. A/B (after diet/before diet) ratio response *vs.* carnitine and acylcarnitines in each of the ten male samples. C_0: carnitine, C_2: acetylcarnitine; C_3: propionylcarnitine; C_4: butyrylcarnitine; C_6: hexanoylcarnitine; C_8: octanoylcarnitine; C_{10}: decanoylcarnitine; C_{12}: lauroylcarnitine; C_{14}: myristoylcarnitine; C_{16}: palmitoylcarnitine; C_{18}: stearoylcarnitine.

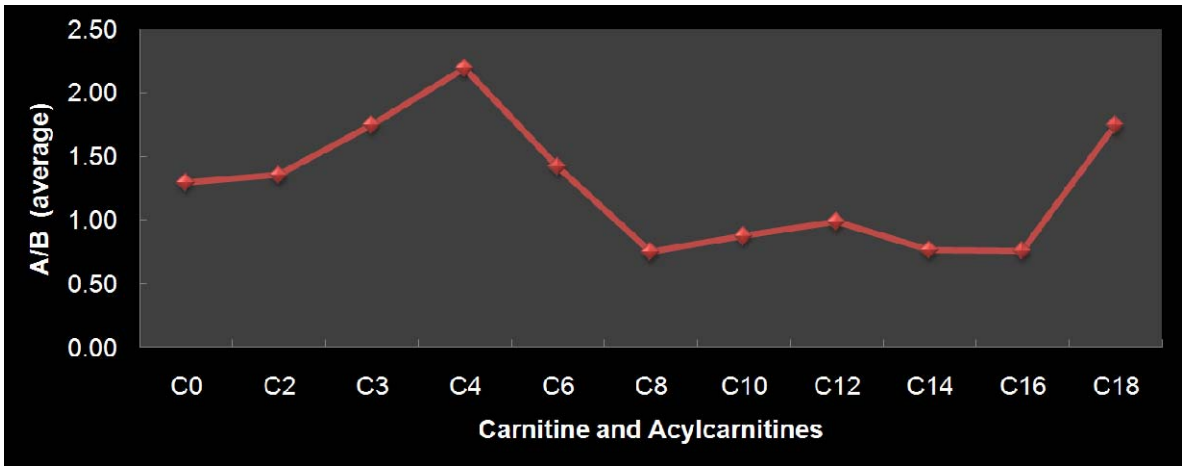


Figure 3-20. A/B (after diet/before diet) ratio versus carnitine and acylcarnitines profiles for all the female samples.

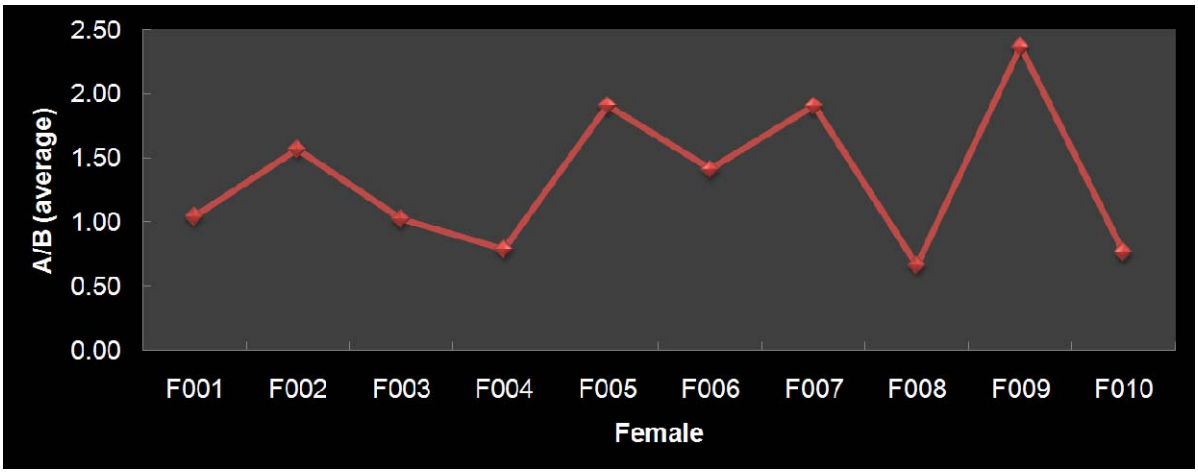


Figure 3-21. A/B (after diet/before diet) ratio versus female profiles for all the carnitine and acylcarnitines under study.

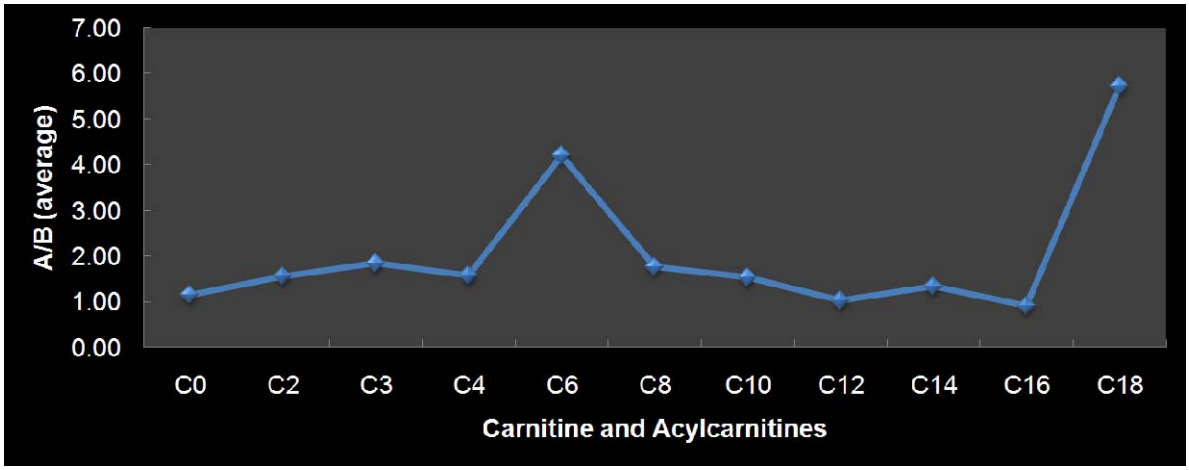


Figure 3-22. A/B (after diet/before diet) ratio versus carnitine and acylcarnitines profiles for all the male samples.

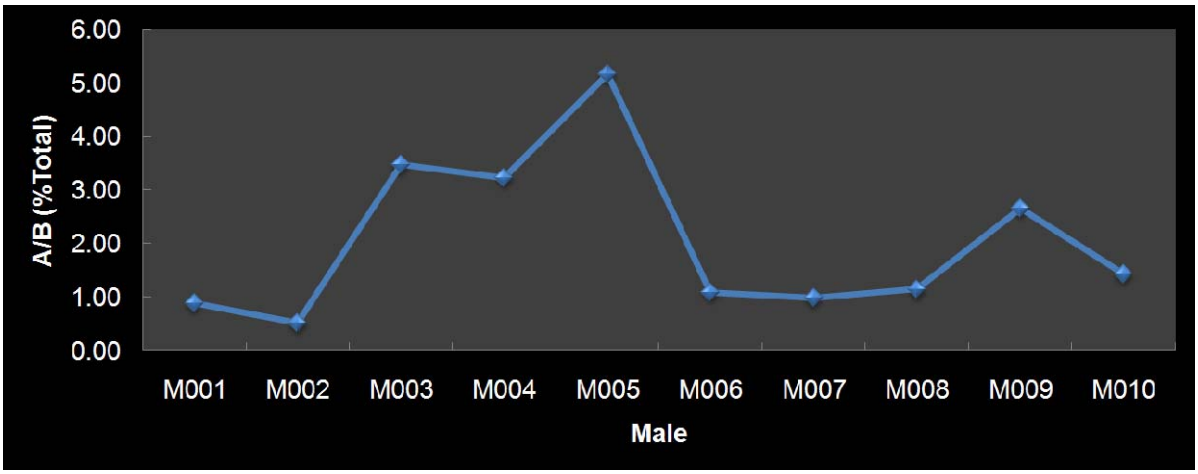


Figure 3-23. A/B (after diet/before diet) ratio versus males for all the carnitine and acylcarnitines.

CHAPTER 4
APPLICATION OF LC/MS METHODOLOGIES TO PIGLET PLASMA SAMPLES FOR GLOBAL AND TARGETED METABOLOMIC STUDIES

The Plasma Metabolome of Piglets from Days 2-8 of Life

Metabolomics is the study of all small molecules and carnitinomics is the targeted metabolomic analysis of carnitine and acylcarnitines. The importance of studying the metabolome has been described in previous chapters. The main objective of this study was the application of a global and targeted metabolomic workflow (refer to Chapter 3 and Figure 3-2) to an animal model for neonatal metabolomic fluctuations over time. The experimental design consisted of twenty-one plasma samples obtained from three piglets on each day of life from day 2 to day 8 which were analyzed using the LC/MS methodologies described previously (Chapter 2).

Targeted acylcarnitine analysis (carnitinomics) is often used in diagnosis of inborn errors of metabolism (IEM) such as medium chain acyl-CoA dehydrogenase deficiency (MCAD), one of the most common inborn errors of fatty acid metabolism [121, 122]. Esters of fatty acyl-CoA of 4-8 carbon chain length are broken down by medium chain acyl-CoA dehydrogenase, and a problem with this enzyme leads to an increase in medium chain fatty acyl groups available for binding to carnitine [121]. Newborn screening for MCAD deficiency has been implemented in many regions. The metabolic marker for MCAD deficiency, octanoylcarnitine (C_8), can be detected with a high degree of sensitivity in newborns by tandem mass spectrometry [123]. However, there is no clear evidence of metabolic information associated to changes in other acylcarnitine levels. It is here where the application of carnitinomics (targeted and untargeted) becomes clinically useful. It is important to search for not only those acylcarnitines found previously, but also possible acylcarnitines that may be specific for the patient population that is being studied.

Obtaining a sample of plasma from an adult is a relatively simple and minimally invasive procedure. However, analysis of metabolic profiles in human tissues or in infant plasma is much more difficult and impractical and may require the use of a model. Questions involving compartmentalization of compounds (e.g. acylcarnitines) in tissues, such as how increases in levels of a specific metabolite in one organ may affect another, or how this metabolite levels change over the first few weeks of life may be best suited to a non-human model in experimental design.

A number of animal species have been used to study aspects of postnatal development, particularly rodents, sheep and primates. The pig *(Sus scrofa domestica)* is a popular model for many human conditions and its value as a laboratory animal has already been assessed [124]. Some useful comparisons can be made between the newborn pigs and the human infants, e.g. their comparable level of maturity at birth, some of their anatomical similarities, their susceptibility to hypothermia, their ability to shiver, their increase in metabolic rate in the first few days after birth, and their limited thermal insulation [124, 125]. Thus, these animals are ideal for study in experiments that do not require rapid reproductive capabilities [126].

On the other hand, human neonates have an increased need for dietary carnitine due to the reduced stores and impaired synthesis of the compound [127], and at birth both the neonatal human and piglet must switch from carbohydrate to lipid metabolism involving the use of carnitine. The piglet can also be a good model for neonatal carnitine analysis because concentrations of plasma and tissue carnitine have been shown to be similar during early development [128]. Despite the fact that one day in the life of a pig does not correspond to one day in the life of a human, the life span and other developmental parameters correspond more with that of a human than almost any other animal used for experimental purposes [128].

In this study, the workflow of global and targeted metabolomics was applied to a piglet model to assess changes in neonatal metabolite levels over the first week of life (days 2 through 8). The ultimate goal was to find biomarkers for metabolic and mitochondrial function for their clinical application to assess human neonates' metabolic changes over time.

Experimental Workflow

The workflow utilized for this study consisted in the same steps described in Chapter 3.

Experimental Methods

Sample Collection

Twenty-one male piglets born between March 20, 2001 and January 29, 2003 were sacrificed in Dr. Peggy Borum's lab from day 2 to day 8 of life (Table 4-1). The animals were fed by the sow until a short time before euthanization. Three piglets were euthanized on each day of life using isoflurane and oxygen as anesthesia, and blood was extracted from the heart and quickly placed on ice for a short time before centrifugation. Plasma and red blood cells were separated using centrifugation for 15 minutes at 8,450 rpm and 4-6 °C then stored at -20°C. Piglets were also weighed and hematocrit was analyzed from the mean of the data from a hematocrit reader using three capillary tubes centrifuged for 5 minutes at 12,400 rpm.

Sample Preparation

Piglet plasma sample preparation assay has been described in Chapter 2. In addition, the experimental section as detailed in Chapter 2 is also applicable to Chapter 4.

Results and Discussion

LC/MS Experiments

In this study, metabolic changes between days two and eight in the life of piglets were evaluated. The experimental design is presented in Figure 4-1. A total of twenty-one piglets were part of this study, three biological replicates (*i.e.* three plasma samples from the same piglet) and

three technical replicates (each sample was run three times) were analyzed; thus a total of 189 files were acquired per LC/MS methodology. Total ion chromatograms corresponding to the overlapping of all files using C_{18}-monolithic and HILIC chromatography are shown in Figures 4-2 and 4-3, respectively.

Global Metabolomics Approach

This study focused on the comprehensive analysis of known and unknown compounds present in the piglet samples. Once the LC/ESI-MS data files were acquired, all the steps related to data processing and analysis were performed and plots for each of the statistical tools were generated.

Hierarchical clustering analysis

In Chapter 3, a brief description of the principles behind each of the statistical algorithms was given. Results obtained for the piglet experiments are described here.

Clustering analysis enables visualization of hierarchical relations between different groups. Ideally, samples that are similar are shown closer together in the dendograms. HCA was applied to the raw data: 15,260 mass-intensity pairs for the C_{18}-monolithic- and 34,736 mass-intensity pairs for the HILIC-experiments were obtained. For the C_{18}-approach, the generated clusters revealed single groups for the data files from days of life 3, 4, 7, and 8. However, data files related to days 2, 5, and 6 showed a greater variability, thus causing these groups to split (Figure 4-4).

On the other hand, the data files corresponding to days 3-7 for the HILIC strategy grouped in individual clusters, but the files corresponding to days 2 and 8 divided into two groups each (Figure 4-5). It was found that biological and technical variability affected the clustering of the data files.

Principal component analysis

The first few principal components represent a relevant part of the total data variance. Thus, when plotting principal component scores, the data structure can be visually inspected in two or three dimensions in order to identify groups of objects (e.g. data files, conditions, etc.) Prior to performing PCA, the raw mass-intensity pairs were filtered out using the tool "filter on relative frequency", which allowed only the most frequent mass-intensity pairs across all the samples to remain part of the analysis. After filtering, the mass-intensity pairs from all the data files were subjected to PCA and the corresponding plots were generated (Figures 4-6 and 4-7). In the case of the C_{18}-monolithic data, the PCA plot revealed the presence of four groups of samples; group 1 corresponding to days 2, 3, 5, and 6 in the life of the piglets, group 2 to day 4, group 3 to day 3, and group 4 to day 8. However, if a diagonal line is traced across this two dimensional plot, the four groups can be reduced into two: one for those data files corresponding to days 2-6 in the life of the piglets and the other for the files related to the days 7 and 8. Therefore, PCA was able to identify important changes in abundance across the first seven days of life. Similar results were obtained when PCA was applied to the HILIC data, which confirmed the above-mentioned trends.

Analysis of variance

ANOVA was applied to make comparison between multiple sample classes and extract the statistically significant mass-intensity pairs. A Tukey post-hoc test was used to further locate the differences between any of these classes and the significance of differences was expressed by the p-value. ANOVA was applied to the C_{18}-monolithic and HILIC filtered mass-intensity pairs. Statistically significant masses between the different days in the life of the piglets were found (Figures 4-8 and 4-9). In the case of the C_{18}-chromatography, a total of 1,672 mass-intensity pairs were found statistically significant, these mass-intensity pairs distributed across the

different days. On the other hand, for the HILIC mass-intensity pairs, a total of 800 mass-intensity pairs were found differentially expressed.

For both groups of statistically significant mass-intensity pairs, a list called "ANOVA masses" was generated and submitted to database identification using mass accuracy-2 ppm mass tolerance- as the criteria for searching. The METLIN database search generated 313 (18.7%) hits for the C_{18}-monolithic ANOVA masses and 198 (24.8%) hits for the HILIC ANOVA masses.

Targeted Metabolomics

Carnitinomics: One of the aims of this study was to develop and apply a metabolomic workflow to targeted carnitinomics analysis using piglets as a model for neonatal acylcarnitine metabolism. The hypothesis to test was that the plasma carnitinome would change between day 2 and day 8 of life. As mentioned earlier (Chapter 3), the interest was focused on the evaluation of the changes occurred on 11 acylcarnitines: carnitine (C_0), acetylcarnitine (C_2), propionylcarnitine (C_3), butyrylcarnitine (C_4), hexanoylcarnitine (C_6), octanoylcarnitine (C_8), decanoylcarnitine (C_{10}), lauroylcarnitine (C_{12}), myristoylcarnitine (C_{14}), palmitoylcarnitine (C_{16}), and stearoylcarnitine (C_{18}); respectively. The instrument response, expressed as peak area counts, was extracted for each of the compounds in each of the data files. Using this information, the peak area counts (averaged) versus the day of life for each carnitine were plotted; thus profiles for carnitine and the different acylcarnitines were generated (Figures 4-10 – 4-20); error bars correspond to ± 1 standard deviation and the number of areas (n) extracted per analyte per day was 27. The day of life markedly affected the relative levels of carnitines and acylcarnitines and different profiles were obtained; some of them showed a progressive increase of acylcarnitines (C_0, C_8, and C_{14}) from day 2 to 8, whereas others showed some interesting fluctuations (C_2-C_6, C_{10}, C_{12}, C_{16} and C_{18}). In addition, the total area (%) versus the day of life was plotted for all the carnitines under study (Figure 4-21). In summary, carnitine and acylcarnitines changed

significantly across the first seven days of life and the biological interpretation of this information is underway.

Conclusions

In this study, LC/MS-based global and targeted metabolomic workflows were successfully applied to assessing changes in the piglet plasma metabolome during the first seven days of life. Results obtained after data processing and compound identification revealed that there are differences in metabolite levels from day 2 to day 8. These results confirmed the initial hypothesis that the plasma metabolome would change within this time period.

In the case of the global strategy, although statistical differences were observed day after day, the most important changes happened for days 7 and 8 in life. When using different statistical tools, the data files corresponding to these two days discriminated very well against the others.

After birth both piglets and humans must make the transition from using primarily carbohydrate as an energy source to the metabolism of lipids from milk [126]. However, neonatal piglets differ from humans in the fact that they have very small amounts of adipose tissue at birth and experience an extremely rapid accumulation of fat before weaning. Also in contrast to humans, piglets do not experience hyperketonaemia as a result of dietary fat absorption. Studies have suggested the tendency of fatty acyl-CoA towards re-esterification for fat accumulation directly after birth instead of producing energy by beta-oxidation requiring the use of carnitine. This suggests that changes in acylcarnitine levels may begin later in life as the piglet begins to accumulate fat and move towards increased lipid metabolism for energy. However, this trend was not directly observed in this study. Differences in acylcarnitines were observed over time, but the most differences were seen on days 4 and 7 with a decrease in some acylcarnitine differences on days 5, 6 and 8. These unexpected decreases suggest that while the

differences in acylcarnitines may be attributed to age there seem to be other factors influencing carnitinome changes after day 2. Also while these changes are not linear and differ among acylcarnitines they may still be important considerations in the analysis and diagnosis of metabolic diseases that result in small fluctuations in metabolites according to disease state.

Table 4-1. Piglet ID, day of life, and MS ID.

Piglet ID #	Day of life	MS ID #
362	2	7
370	2	7
371	2	7
357	3	8
361	3	8
369	3	8
358	4	9
363	4	9
368	4	9
356	5	10
359	5	10
376	5	10
360	6	11
364	6	11
365	6	11
373	7	12
374	7	12
375	7	12
366	8	13
367	8	13
372	8	13

Figure 4-1. Study design for the evaluation of the metabolomic changes from day 2 to 8 in the life of piglets. Plasma samples from twenty one piglets were obtained. LC/MS approaches were applied for the analysis of three biological and three technical replicates per piglet, the dataset consisted in 189 files per chromatographic approach.

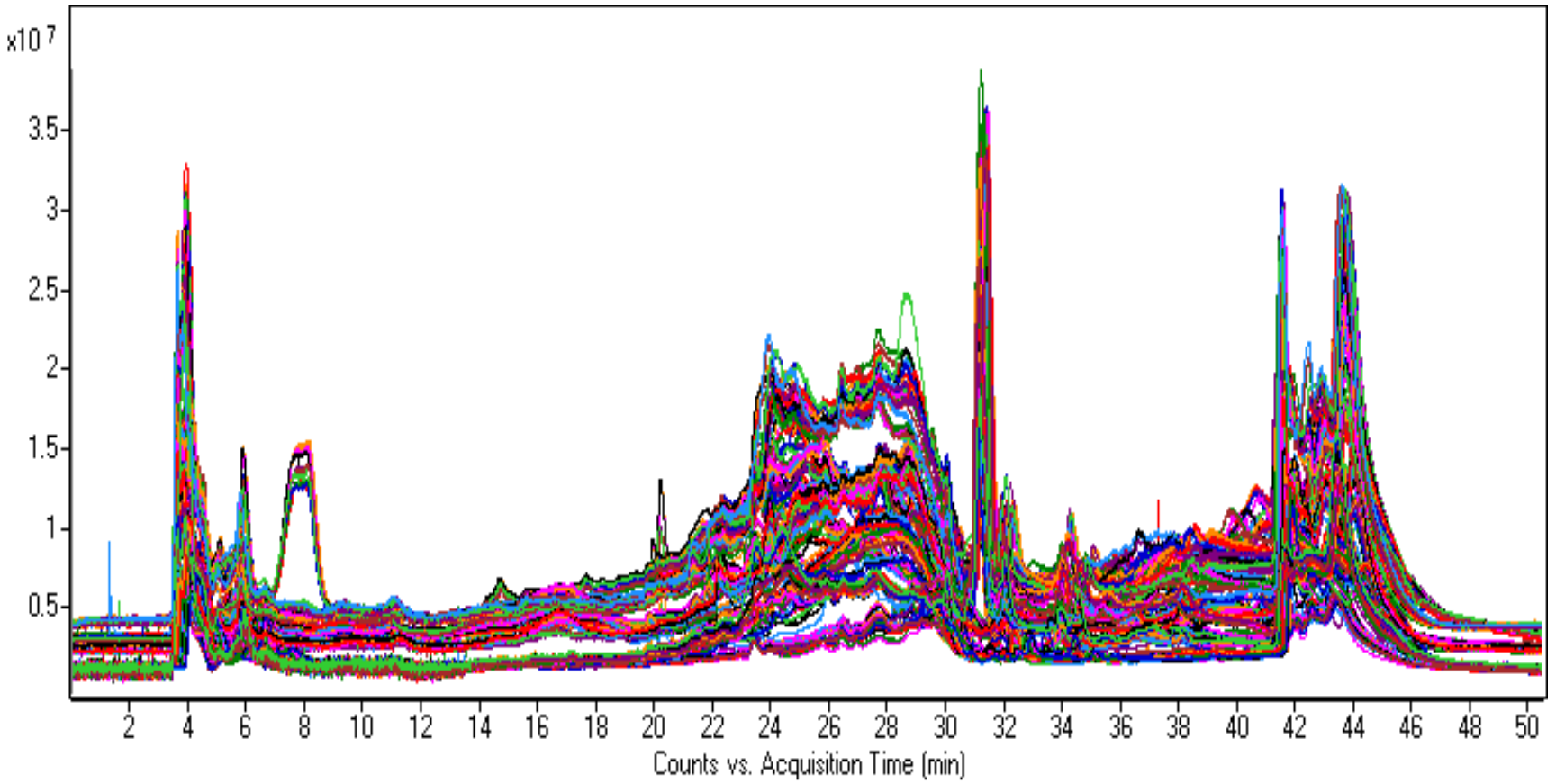


Figure 4-2. Total ion chromatograms (TICs) for the overlapping of all piglet samples (189 files) using C_{18}-monolithic chromatography. Solvent A: 1% (v/v) acetic acid in water; Solvent B: 1% (v/v) acetic acid in acetonitrile; Flow rate (mL min^{-1}): 1.0. Injection volume (µL): 15.0; Gradient: A:B (min): 95:5 (0.0-6.5); 0:100 (25.5-35.0); 95:5 (55.0-65.0)

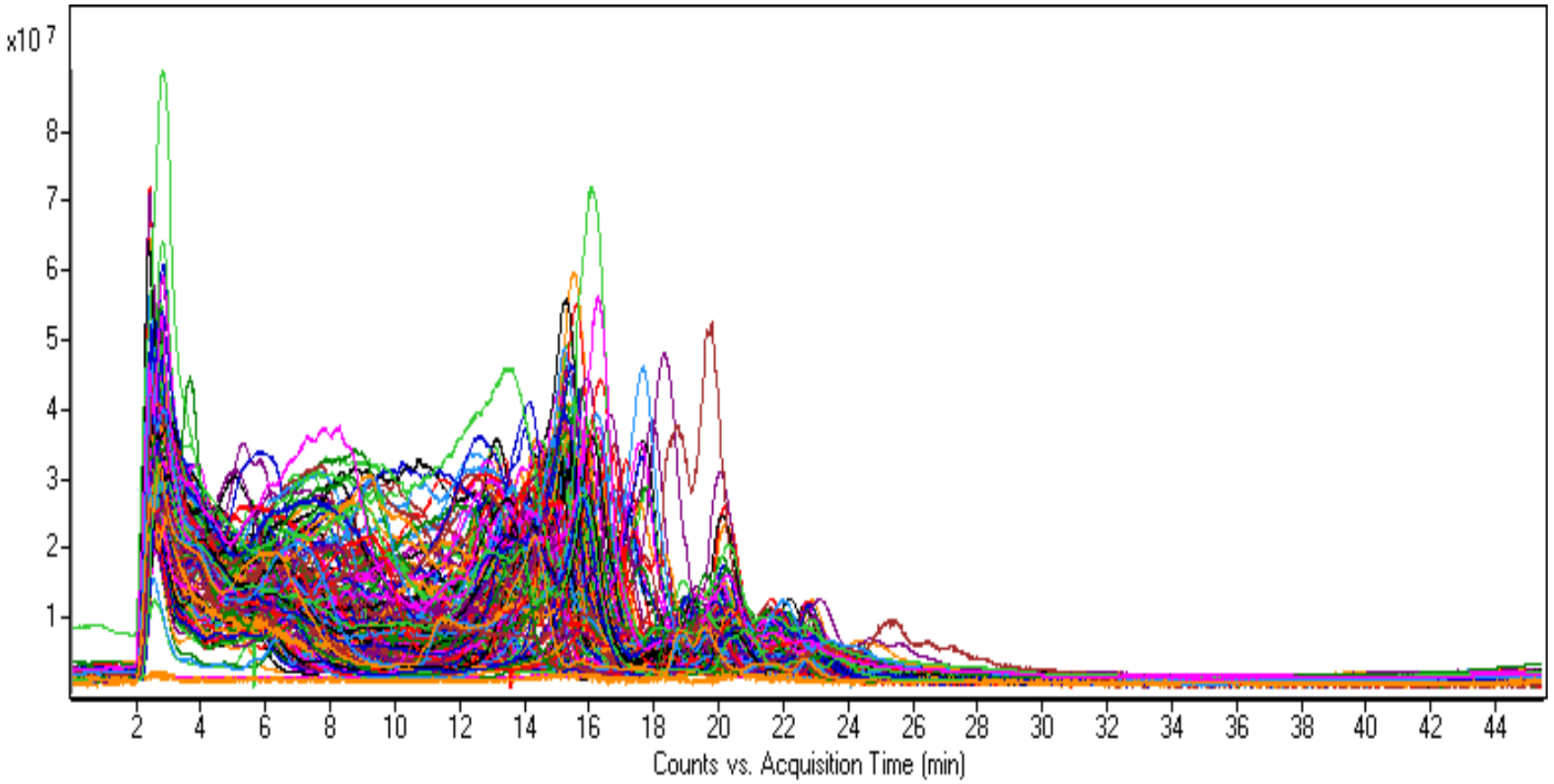


Figure 4-3. Total ion chromatograms (TICs) for the overlapping of all piglet samples (189 files) using HILIC chromatography. Solvent A: 7.5 mM ammonium formate in water; Solvent B: 7.5 mM ammonium formate in acetonitrile; Flow rate (mL min^{-1}): 0.3. Injection volume (µL): 10.0; Gradient: A:B (min): 10:90 (0.0-5.0); 50:50 (25.0-30.0); 10:90 (40-45).

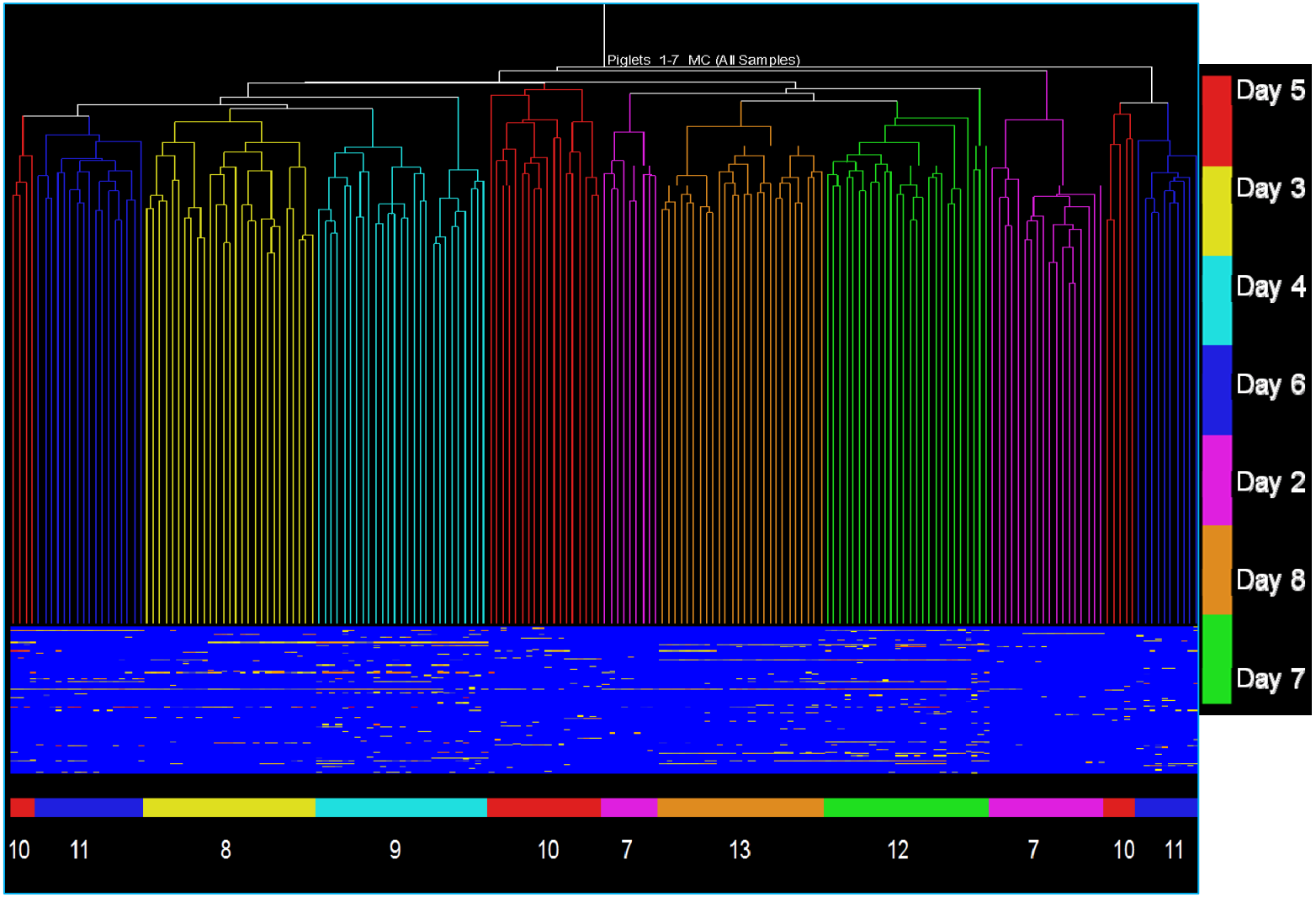


Figure 4-4. Hierarchical clustering analysis (monolithic column, #masses: 15,260) resulting in grouping by MS ID# (color key below clustering tree) or day of life (color key on the right). The tree is clearly divided into several groups, some of them clustered all the data files for a specific day of life (days: 3, 4, 7, and 8) and some of them were divided into different subgroups (days 2, 5, and 6).

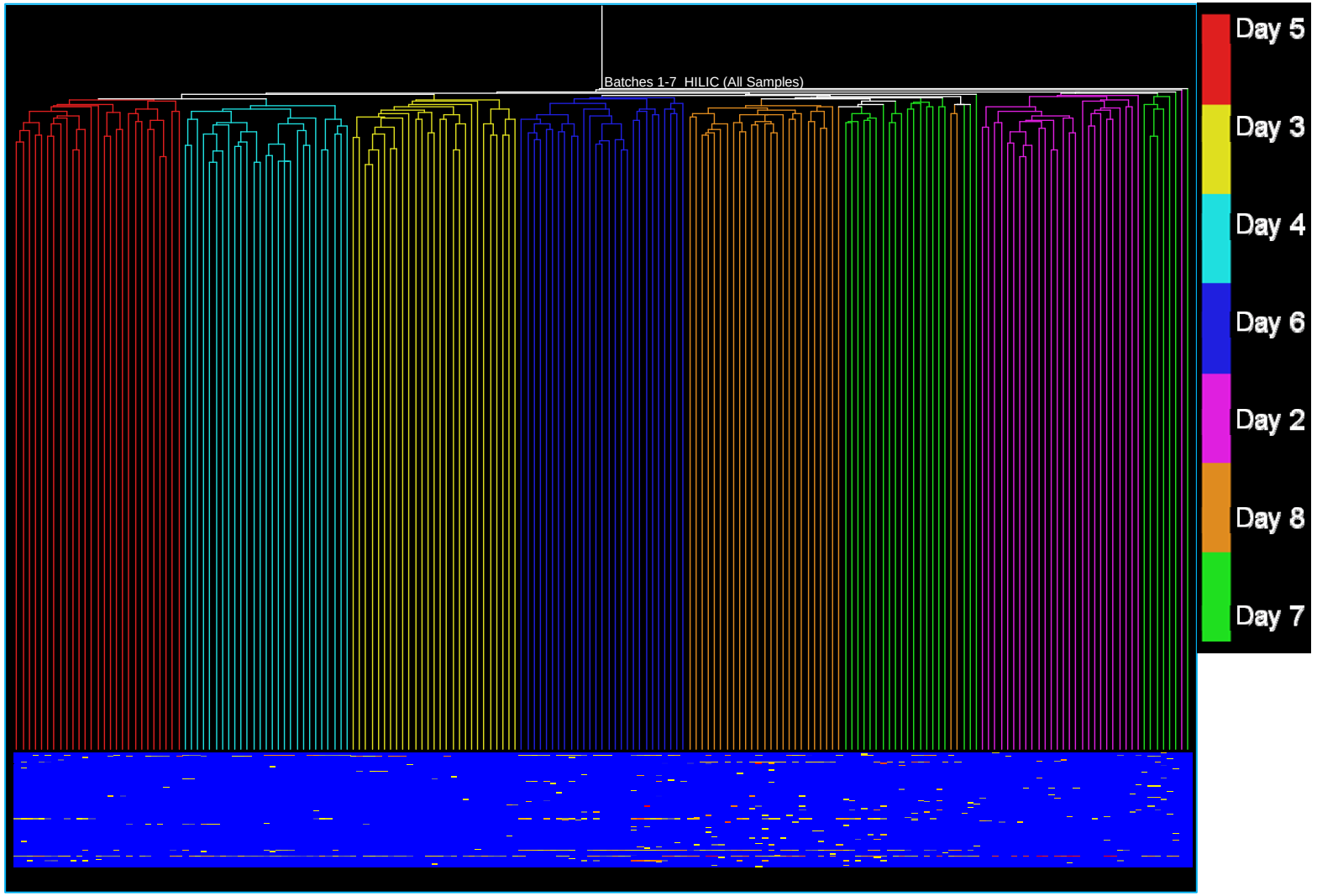


Figure 4-5. Hierarchical clustering analysis (HILIC column, #masses: 34,736) resulting in grouping by day of life (color key on the right). The tree is divided into several groups, data files corresponding to all of the days, except for days 2 and 8, were clustered together.

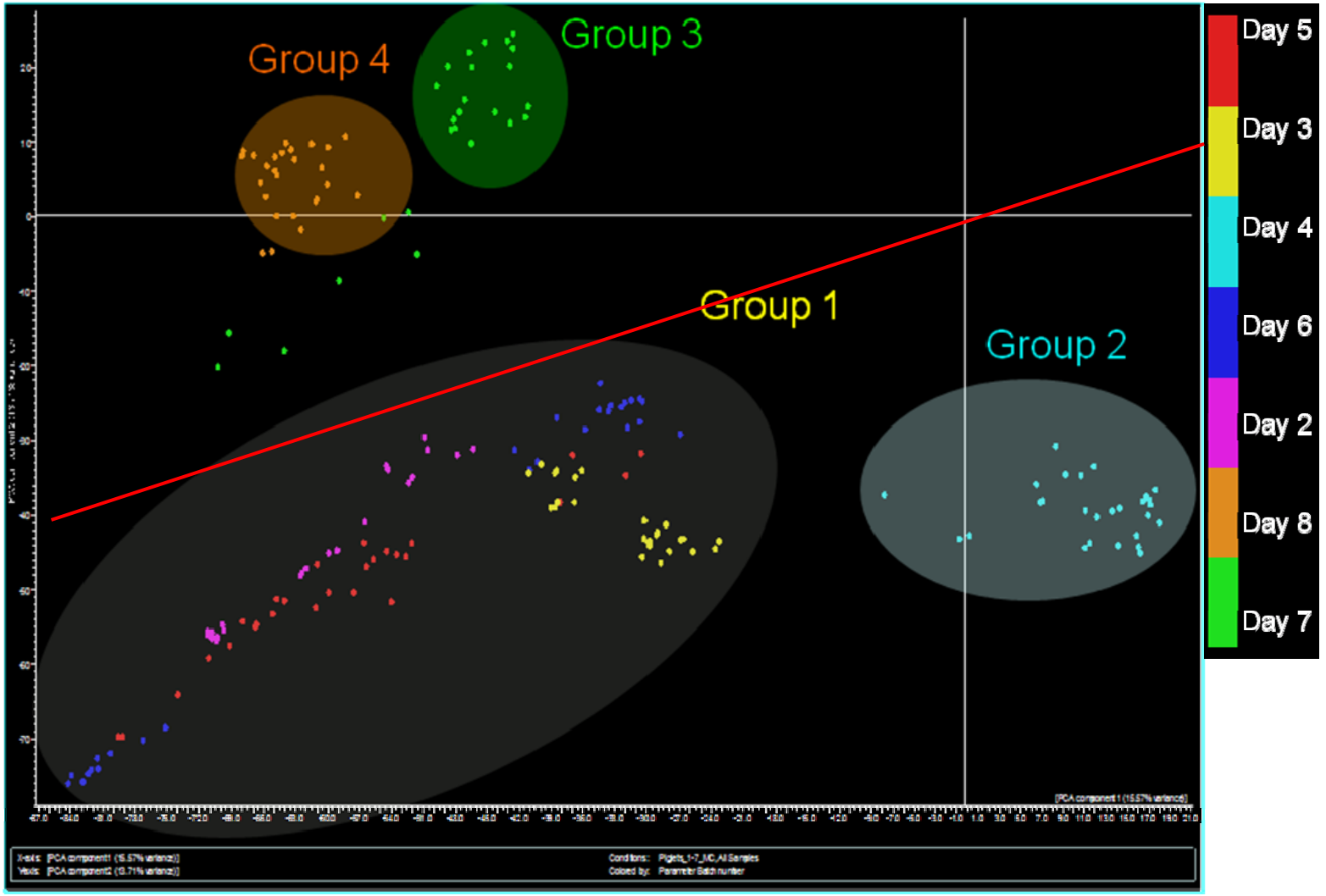


Figure 4-6. Principal component analysis (monolithic column). The two dimensional PCA plot shows four different group of samples, which corresponds to days 2, 3, 4, and 5 for group 1, day 4 for group 2, day 7 for group 3, and day 8 for group 4. If a red diagonal is traced, two major groups can be visualized: days 2-6 in the life of the piglets below the diagonal and days 7 and 8 above the diagonal.

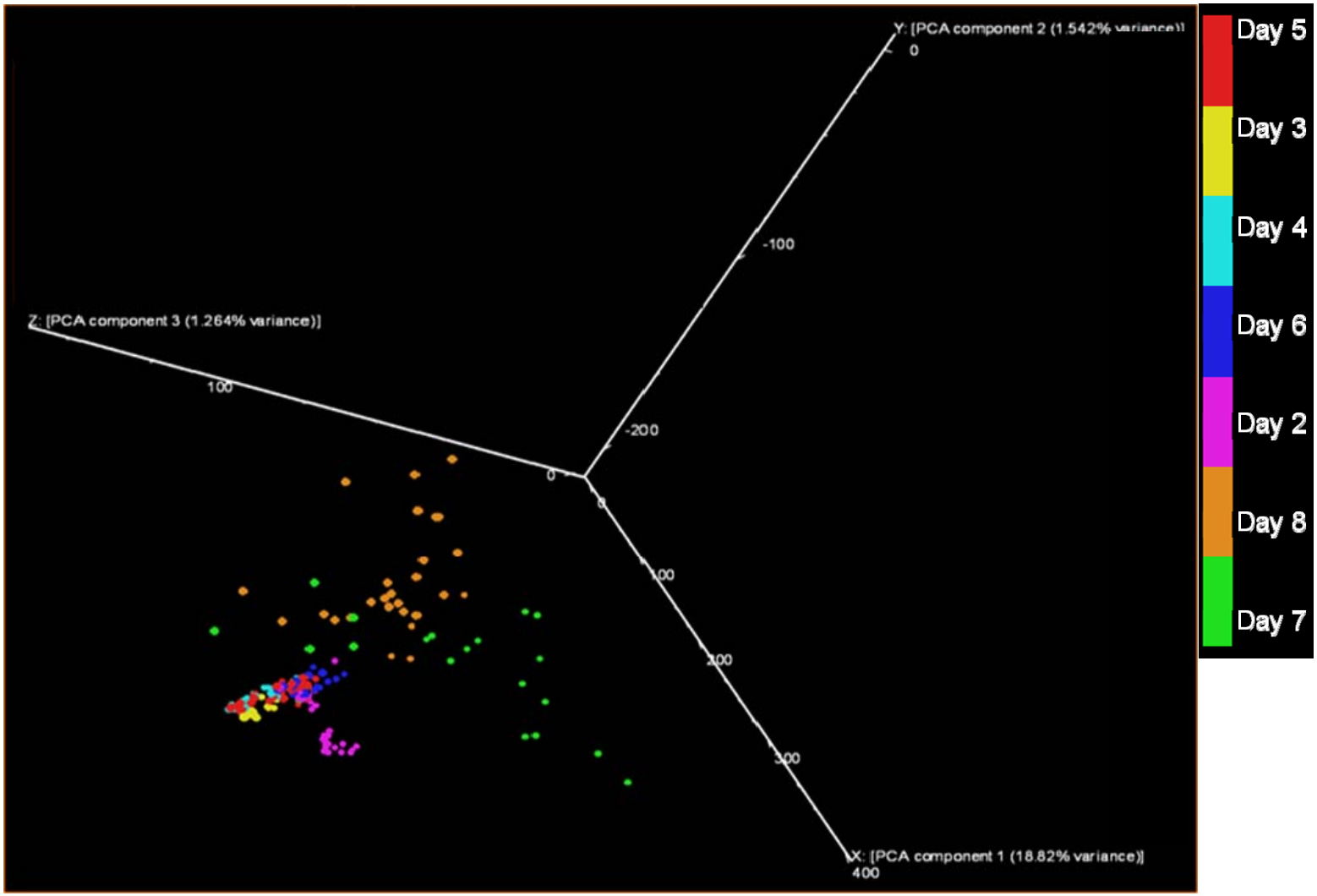


Figure 4-7. Principal component analysis (HILIC column). The rotated three dimensional PCA plot shows correlations across samples from days 2-6 and between days 7 and 8. These results are consistent with those obtained for the C_{18}-monolithic approach.

Batch/Day	10/5	8/3	9/4	11/6	7/2	13/8	12/7
10/5	1672	550	949	618	411	775	832
8/3	1122	1672	781	739	510	768	805
9/4	723	891	1672	1038	921	1076	1047
11/6	1054	933	634	1672	729	910	905
7/2	1261	1162	751	943	1672	653	777
13/8	897	904	596	762	1019	1672	646
12/7	840	867	625	767	895	1026	1672

Figure 4-8. One-way analysis of variance (ANOVA) for C_{18}-monolithic. Test type: parametric, do not assume variances equal; False discovery rate: 0.05; Multiple testing correction: Benjamini and Hochberg false discovery rate; Post Hoc test: Tukey. The red boxes represent the statistically different masses, the grey boxes the total number of masses, and the blue boxes the statistically similar masses.

Batch/Day	10/5	8/3	9/4	11/6	7/2	13/8	12/7
10/5	800	243	212	319	277	395	273
8/3	557	800	229	334	226	418	326
9/4	588	571	800	348	321	461	353
11/6	481	466	452	800	380	400	261
7/2	523	574	479	420	800	404	323
13/8	405	382	339	400	396	800	259
12/7	527	474	447	539	477	541	800

Figure 4-9. One-way analysis of variance (ANOVA) for HILIC. Test type: parametric, do not assume variances equal; False discovery rate: 0.05; Multiple testing correction: Benjamini and Hochberg false discovery rate; Post Hoc test: Tukey. The red boxes represent the statistically different masses, the grey boxes the total number of masses, and the blue boxes the statistically similar masses.

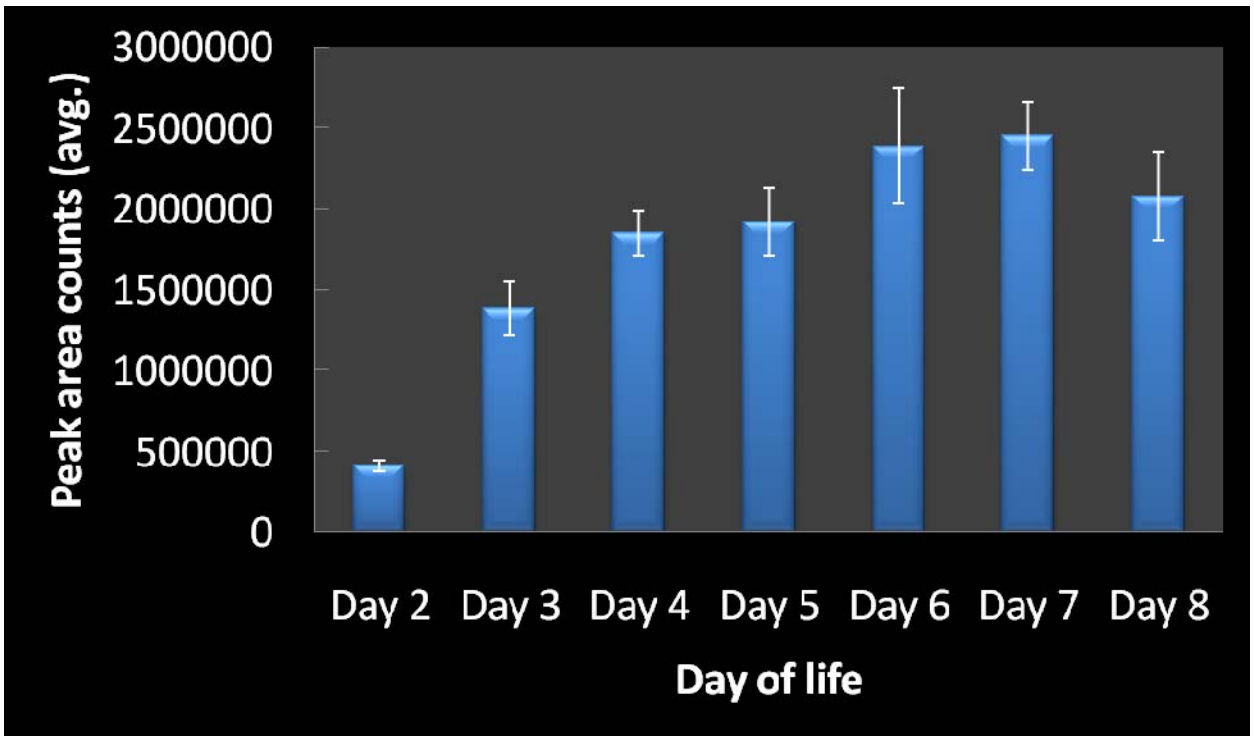


Figure 4-10. Targeted metabolomics analysis. Carnitine profile (m/z = 162.1125).

Figure 4-11. Targeted metabolomics analysis. Acetylcarnitine profile (m/z = 204.1230).

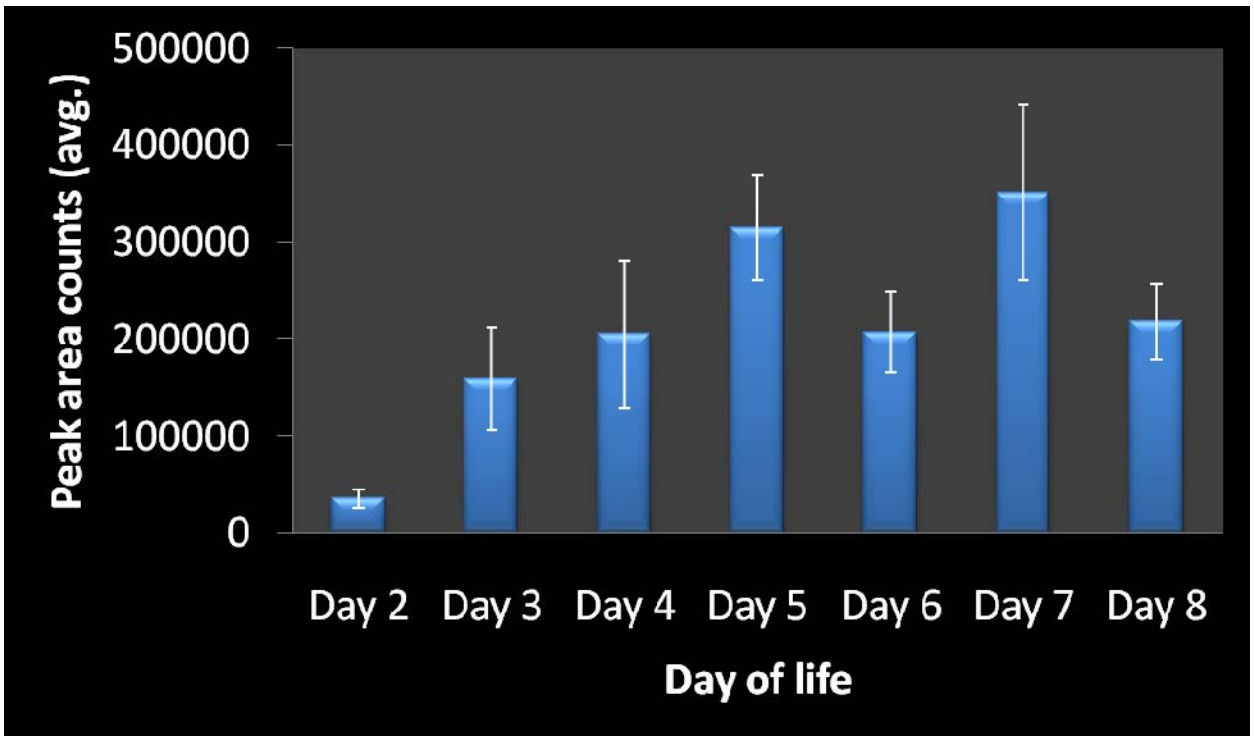


Figure 4-12. Targeted metabolomics analysis. Propionylcarnitine profile (m/z = 218.1387).

Figure 4-13. Targeted metabolomics analysis. Butyrylcarnitine profile (m/z = 232.1543).

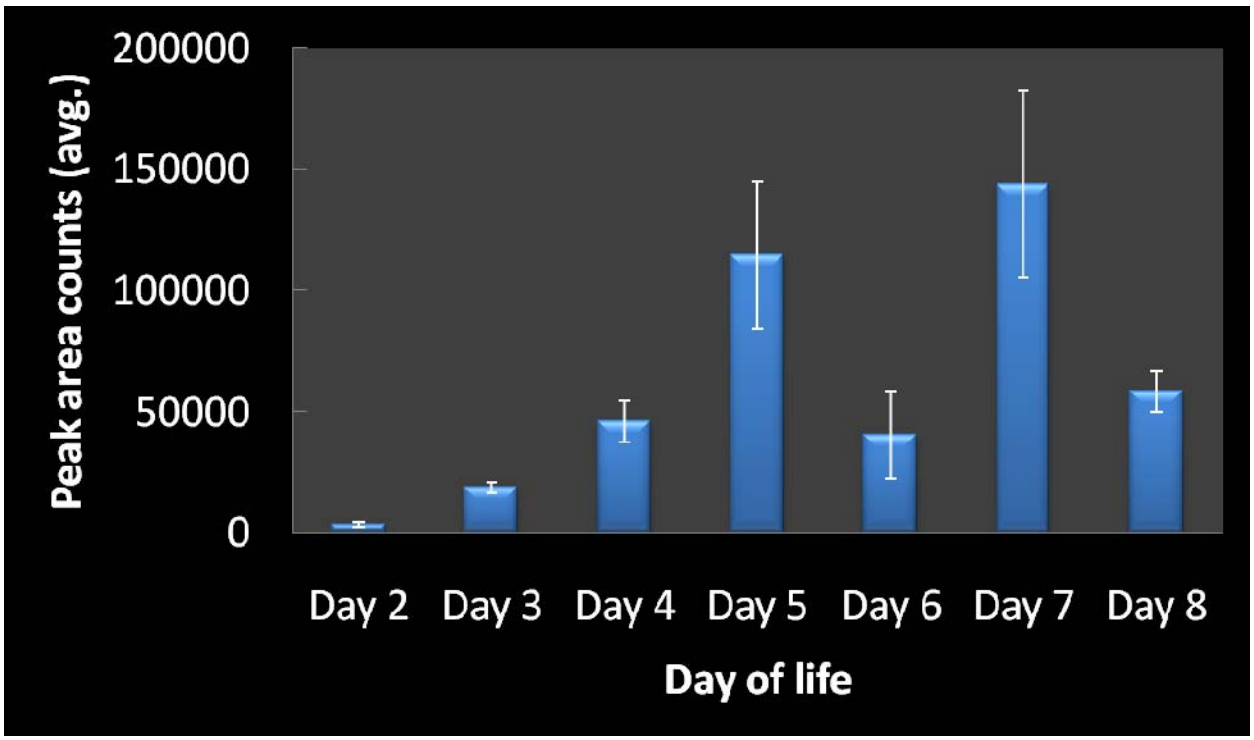


Figure 4-14. Targeted metabolomics analysis. Hexanoylcarnitine profile (m/z = 260.1856).

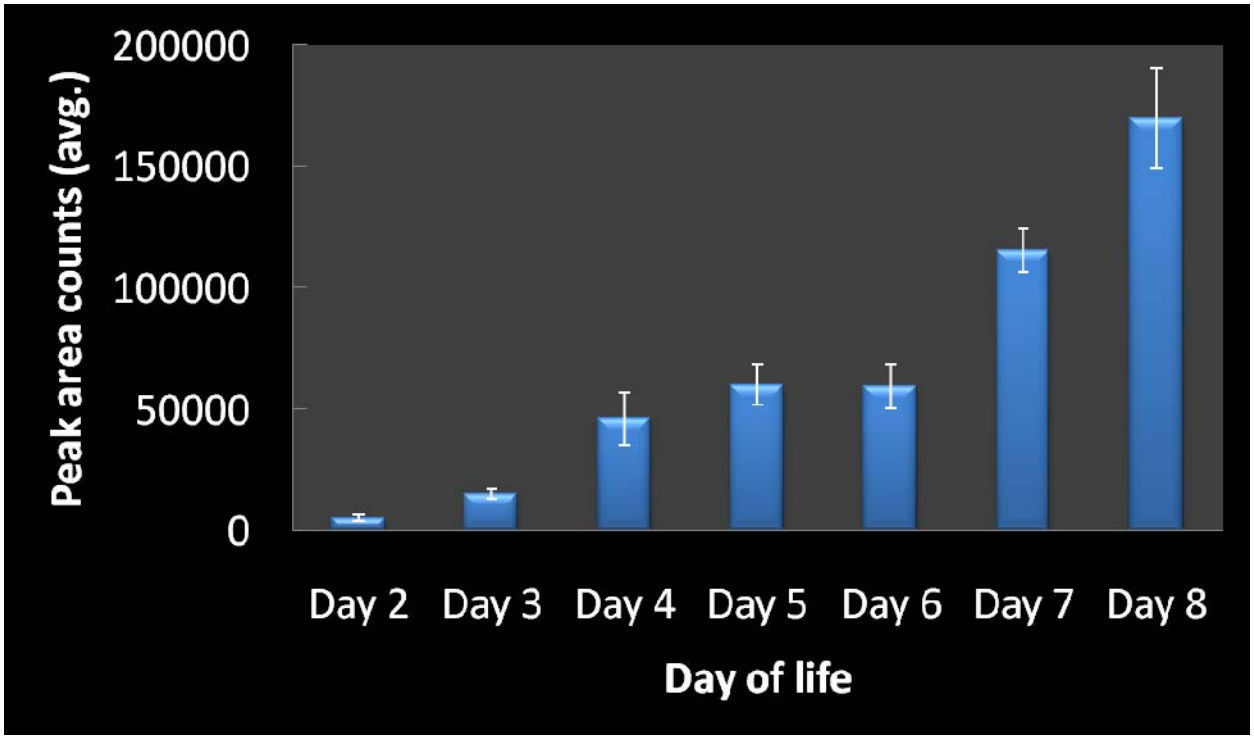


Figure 4-15. Targeted metabolomics analysis. Octanoylcarnitine profile (m/z = 288.2169).

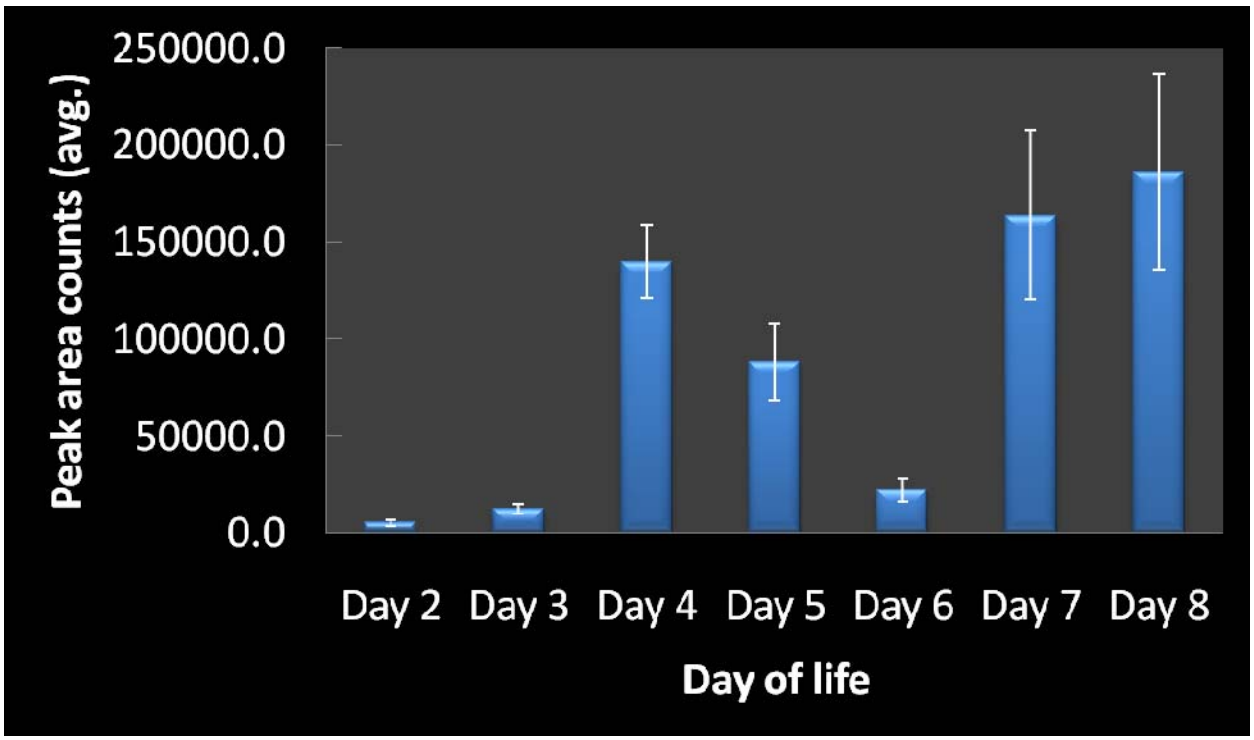


Figure 4-16. Targeted metabolomics analysis. Decanoylcarnitine profile (m/z = 316.2482).

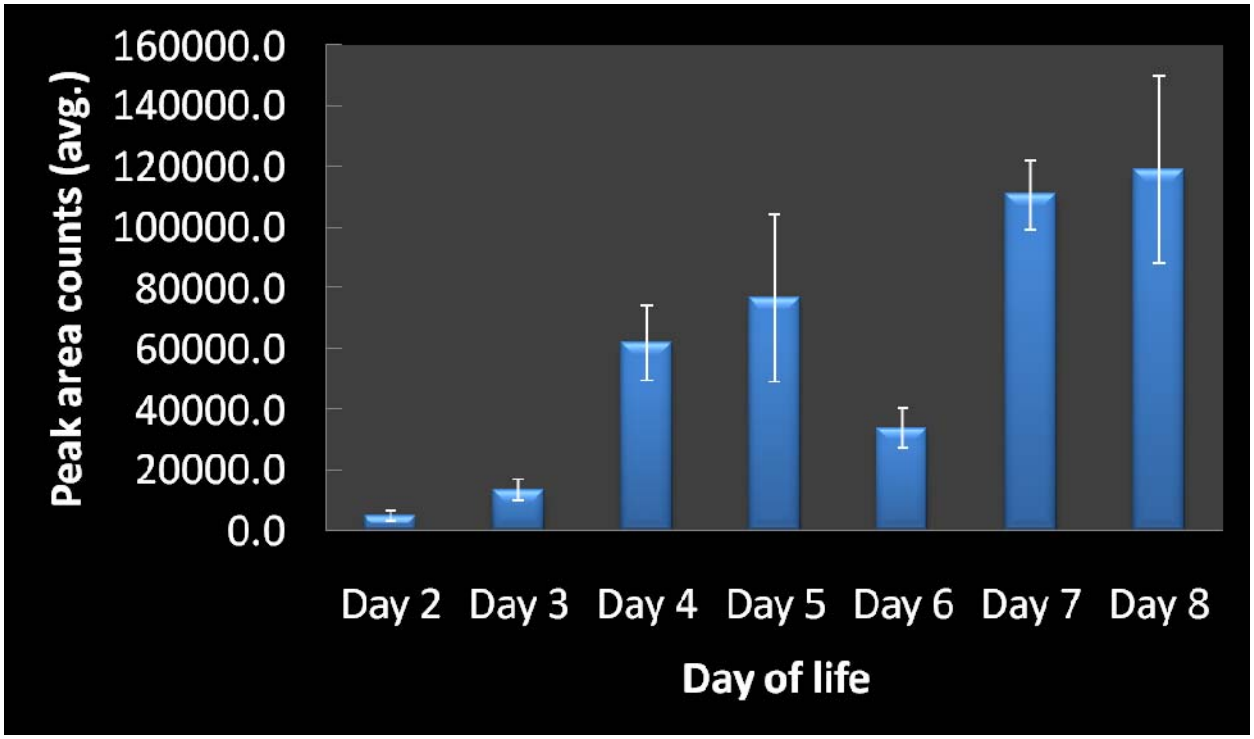


Figure 4-17. Targeted metabolomics analysis. Lauroylcarnitine profile (m/z = 344.2795).

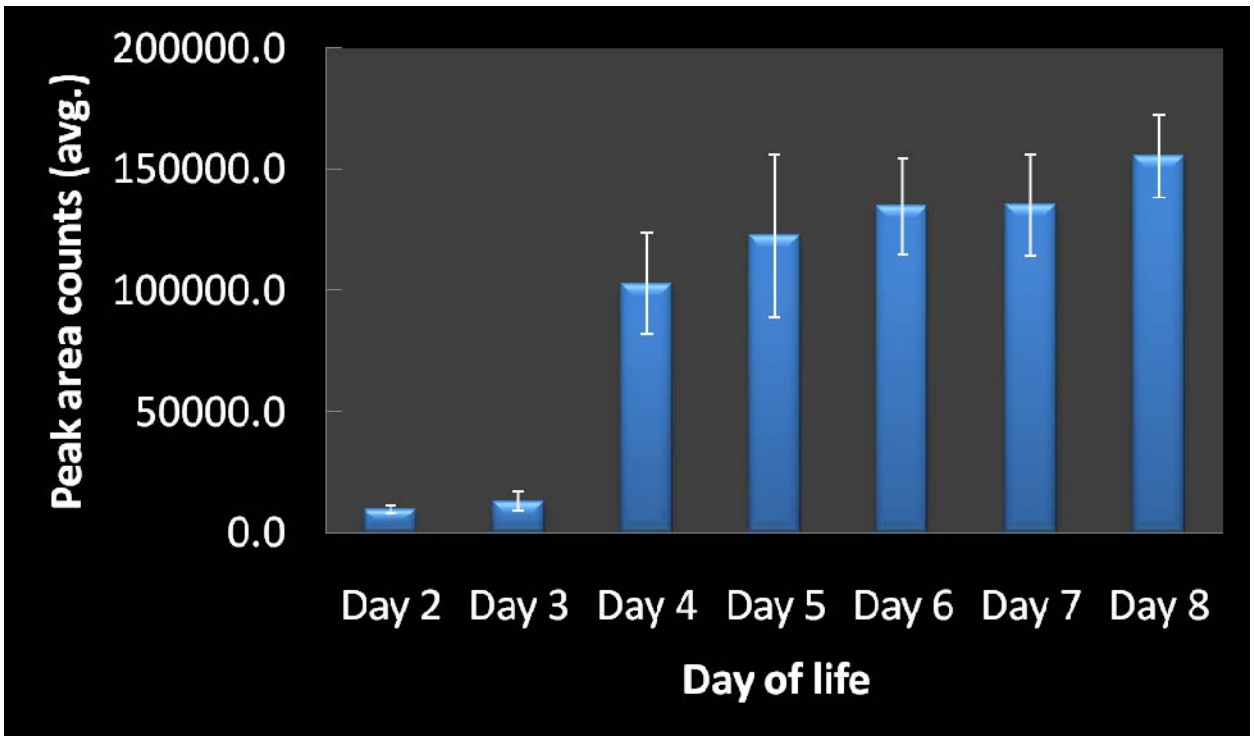


Figure 4-18. Targeted metabolomics analysis. Myristoylcarnitine profile (m/z = 372.3108).

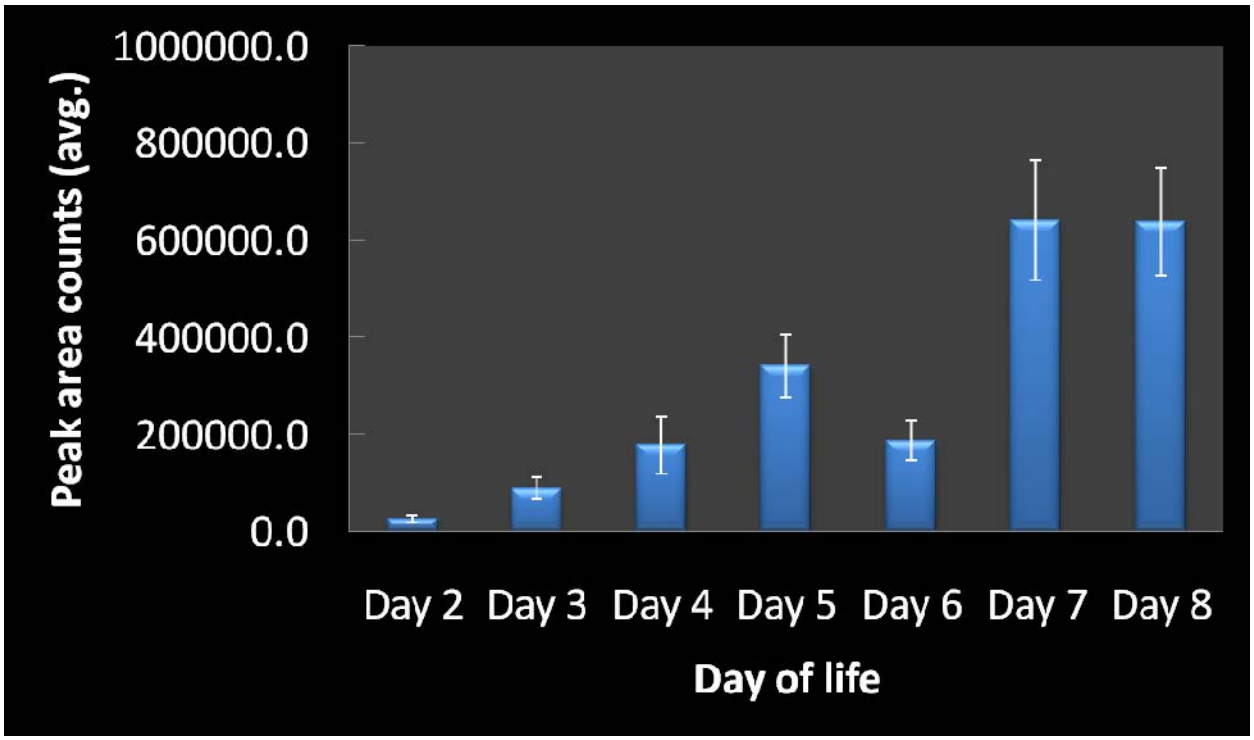


Figure 4-19. Targeted metabolomics analysis. Palmitoylcarnitine profile (m/z = 400.3421).

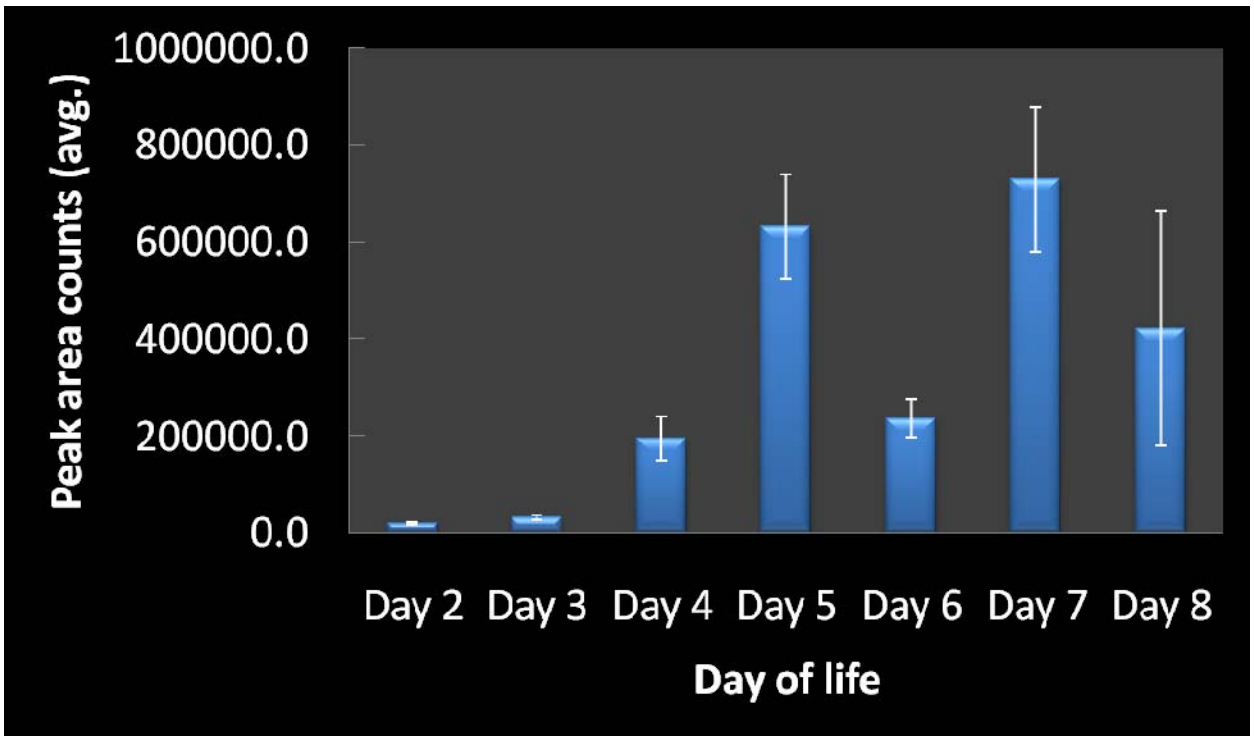


Figure 4-20. Targeted metabolomics analysis. Stearoylcarnitine profile (m/z = 428.3734).

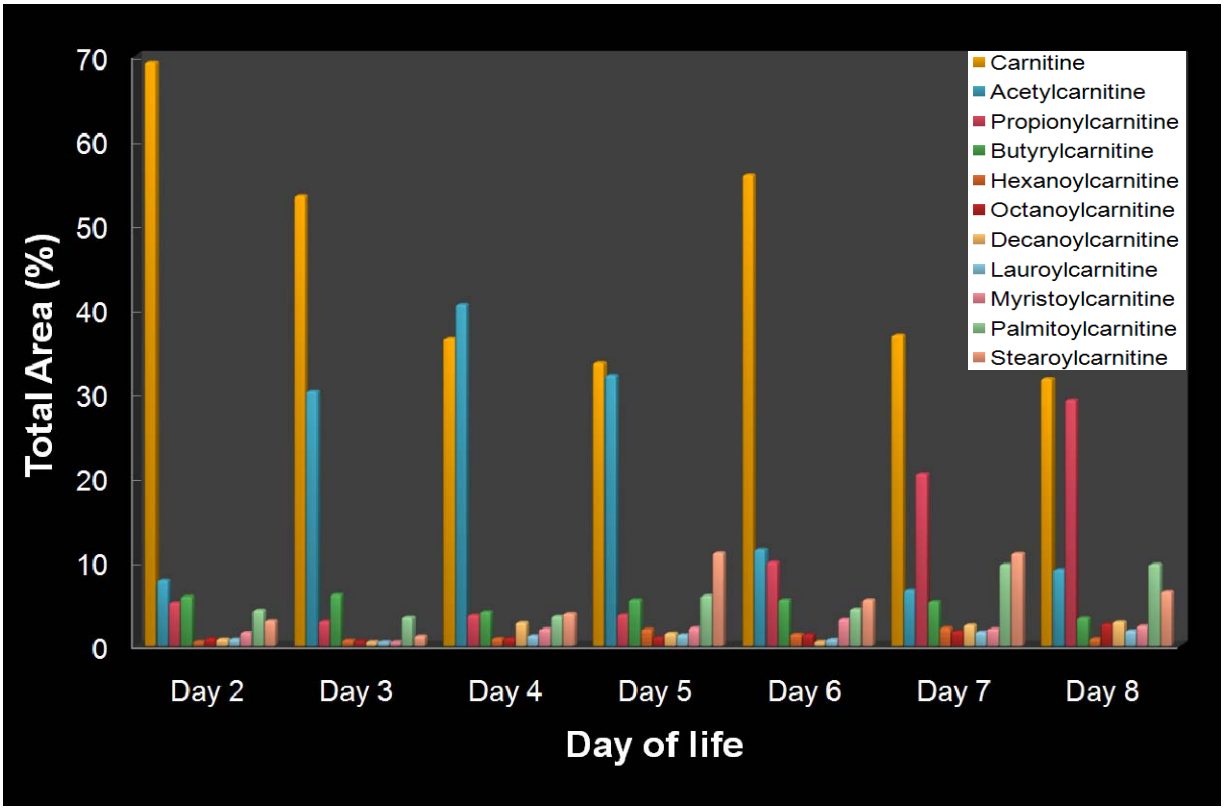


Figure 4-21. Targeted metabolomics analysis. Carnitine and acylcarnitines profiles between days 2 through 8 in the life of piglets.

CHAPTER 5
CONCLUDING REMARKS AND FUTURE WORK

LC-MS methods were developed for improving the efficiency of separation of nonpolar and polar compounds in a metabolic context. High precision and accuracy of the proposed chromatographic approaches were demonstrated. The results clearly suggest that the use of both C_{18}-monolithic and HILIC columns in HPLC are complementary and can serve as an advantageous complement to conventional methods for separation.

The chromatographic methodologies described in the present manuscript dealt with samples without any kind of derivatization (neither prior, nor after column separation), thus achieving a less complex, less expensive, and less time-consuming way to carry out the analysis of the metabolites. In addition, this fact differentiates this study and its procedures from other approaches normally reported in the literature for the separation and/or determination of carnitine and carnitine-based compounds in clinical samples.

Using the features provided by the proposed chromatographic techniques plus the mass-accuracy and sensitivity characteristics of the *oa*-TOF mass spectrometer, targeted and untargeted metabolites were efficiently separated and accurately detected. The aim of this study was to achieve a comprehensive understanding of the metabolic role of known and unknown metabolites. The development of a reliable analytical workflow based on the utilization of cutting-edge technology together with sophisticated data processing and analysis software tools enabled the accomplishment of this goal.

Putative biomarkers for the different applications were generated and their unambiguous identification is still underway. In fact, the unequivocal identification of metabolites is still one of the bottlenecks of any metabolomics study. In this sense, it is well-known that tandem mass spectrometry provides the ability to positively identify the analytes with the correct compound

assignments using the information provided by characteristics fragmentation patterns. In addition, NMR can generate valuable structural information, but much less sensitively than MS. Therefore, implementation of these analytical methodologies to improve the identification of the tentative biomarkers is part of the future work. In addition, different ionization methodologies have been demonstrated to contribute to the detection of unique metabolites. The application of ionization techniques other than ESI needs to be carried out.

This study also generated many biological questions to be answered through further experimentation. Ideas have previously been proposed for reasons why there may be more or less difference in levels of certain compounds. These compounds may have important diagnostic and biological implications previously unknown, and the hypotheses generated from these analyses must be tested through other studies.

While the changes in metabolites in plasma are important due to their clinical relevance, compartmentalization of metabolites in other tissues may also play a role in the interpretation of the data. Plasma fluctuations may be influenced by the uptake of compounds into other biological media. Further studies on differences in putative metabolites in other tissues need to be conducted to apply fully the proposed analytical models to clinical care. Understanding the changes in metabolites and acylcarnitines as a result of a ketogenic diet may lead to further understanding of the mechanism of ketogenic therapy.

The fluctuations in differences among targeted metabolites seen in this study may indicate a need for a more controlled experimental design when using humans or piglets as a model for patient populations. The findings described would allow further biological implications to be made and investigated. These data are preliminary and must be investigated further in order to accurately analyze the results.

REFERENCES

[1] K. Dettmer, P. Aronov, B.D. Hammock, Mass Spectrom. Rev. 26 (2007) 51.

[2] S.H. Gary Khoo, M. Al-Rubeai, Cell Engineering 5, Systems Biology, Springer, Netherlands (2007).

[3] C.W. Schmidt, Environ. Health Perspect. 112 (2004) A410-A415.

[4] S.G. Olivier, M.K. Winson, D.B. Kell, F. Banganz, Trends Biotechnol. 16 (1998) 373.

[5] D. Ryan, K. Robards, Anal. Chem. 78 (2006) 7954.

[6] Fiehn, Comp. Funct. Genom. 2 (2001) 155.

[7] M. Mayr, Circ. Cardiovasc. Genet. 1 (2008) 58-65.

[8] W.B. Dunn, D.I. Ellis, Trends Anal. Chem. 24 (2005) 285.

[9] R. Ramautar, G.W. Somsen, G.J. de Jong, Electrophoresis 30 (2009) 276.

[10] R. Goodacre, S. Vaidyanathan, W.B. Dunn, G.H. Harrigan, D.B. Kell, Trends Biotechnol. 22 (2004) 245.

[11] S. G. Villas-Boas, J. Nielsen, J. Smedsgaard, M.A.E. Hansen, U. Roessner-Tunali, Hoboken, New Jersey, John Wiley & Sons, Inc. (2007).

[12] K. Dettmer, B.D. Hammock, Environ. Health Perspect. 112 (2004) A396.

[13] I.M. Keseler, J. Collado-Vides, S. Gama-Castro, J. Ingraham, S. Paley, I.T. Paulsen, M. Peralta-Gil, P.D. Karp, Nucl. Acids Res. 33 (2005) D334.

[14] K-M. Oksman-Caldentey, D. Inze, M. Oresic, Proc. Natl. Acad. Sci. 101 (2004) 9949.

[15] I.F. Duarte, J. Marques, A.F. Ladeirinha, C. Rocha, I. Lamego, R. Calheiros, T.M. Silva, M. P.M. Marques, J.B. Melo, I.M. Carreira, A.M. Gil, Anal. Chem. 81 (2009) 5023.

[16] G.G. Harrigan, R.H. LaPlante, G.N. Cosma, G. Cockerell, R. Goodacre, J.F. Maddox, J.P.Luyendyx, P.E. Ganey, R.A. Roth, Toxicol. Lett. 146 (2004) 197.

[17] H.E. Johnson, D. Broadhurst, D.B. Kell, M.K. Theodorou, R.J. Merry, G.W. Griffith, Environ. Microbiol. 70 (2004) 1583.

[18] W. Lu, B.D. Bennett, J.D. Rabinowitz, J. Chromatogr. B 871 (2008) 236.

[19] A. Koulman, G. Woffedin, V.K. Narayana, H. Welchman, C. Crone, D. Volmer, Rapid Commun. Mass Spectrom. 23 (2009) 1411.

[20] J. A, J. Trygg, J. Gullberg, A.I. Johansson, P. Jonsson, H.Antti, S.L. Marklund, T. Moritz, Anal. Chem. 77 (2005) 8086.

[21] W. Wikoff, A.T. Anfora, J. Liu, P.G. Schultz, S.A. Lesley, E.C. Peters, G. Siuzdak, Proc. Natl. Acad. Sci. 106 (2009) 3698.

[22] D.T.-T Nguyen, D. Guillarme, S. Heinisch, M-P. Barrioulet, J.-L. Rocca, S. Roudaz, J.-L. Veuthey, J. Chromatogr. A 1167 (2007) 76.

[23] L.W. Summer, P. Mendes, R.A. Dixon, Phytochem. 62 (2003) 817.

[24] A. Koulman, G.A. Lane, S.J. Harrison, D.A. Volmer, Anal. Bioanal. Chem. 394 (2009) 663.

[25] R.J. Wenzel, U. Matter, L. Schultheis, R. Zenobi, Anal. Chem. 77 (2005) 4329.

[26] W.E. Stephens, Phys. Rev. 69 (1946) 691.

[27] A.E. Cameron, D.F. Eggers, Rev. Sci. Instrum. 19 (1948) 60.

[28] W.C. Wiley, I.H. McLaren, Rev. Sci. Instrum. 26 (1955) 1150.

[29] W.C. Wiley, Science 124 (1956) 817

[30] B.A. Mamyrin, V.I. Karataev, D.V. Shmikk, V.A. Aagulin, Sov. Phys. JETP 37 (1973), 45.

[31] J.H.J. Dawson, M. Guilhaus, Rapid Commun. Mass Spectrom. 3 (1989) 155.

[32] A.F. Dodonov, I.V. Chernushevich, V.V. Laiko, Proceeding of 12th International Mass Spectrometry Conference, Amsterdam, Netherlands, (1991) 153.

[33] M. Guilhaus, D. Selby, V. Mlynski, Mass Spectrom. Rev. 19 (2000) 65.

[34] A.P. Bruins, J. Chromatogr. A 794 (1998) 345.

[35] M. Dole, L.L. Mack, R. Hines, M.B. Alice, L.D. Ferguson, R.C. Mobley, J. Chem. Phys. 49 (1968) 2240.

[36] J.B. Fenn, M. Yamashita, J. Chem. Phys. 88 (1984) 4451.

[37] J.B. Fenn, C.K. Meng, Org. Mass Spectrom. 26 (1991) 542.

[38] J.B. Fenn, T. Nohmi, J. Am. Chem. Soc. 114 (1992) 3241.

[39] S. Nguyen, J.B. Fenn, Proc. Natl. Acad. Sci. 23 (2007) 1111.

[40] N.J. Haskins, J. Pharm. Biomed. Anal. 25 (2001) 767.

[41] C.E.C.A. Hop, R. Bakhtiar, Biospectroscopy 34 (1998) 259.

[42] Chhabil Dass, Hoboken, Fundamentals of Contemporary Mass Spectrometry, John Wiley & Sons, Inc., New Jersey, (2007).

[43] S. Sanglier, H. Ramstrom, J. Haiech, E. Leize, A. Van Dorsselaer, Int. J. Mass Spectrom. 12208 (2002) 1.

[44] W.E. Haskins, Z. Wang, C.J. Watson, R.R. Rostand, S.R. Witowski, D.H. Powell, R.T. Kennedy, Anal. Chem. 73 (2001) 5005.

[45] J.H. Gross, Mass Spectrometry. A Text Book, Springer, Heidelberg, Germany, (2004).

[46] J.J. Kirkland, J. Chromatogr. Sci. 38 (2000) 535.

[47] Y. Shen, R.D. Smith, Electrophoresis 23 (2002) 3106.

[48] K.K. Unger, R. Skudas, M.M. Schulte, J. Chromatogr. A 1184 (2008) 393.

[49] S. Svec, C.G. Huber, Anal. Chem. 78 (2006) 2101.

[50] E. Klodzinska, D. Moravcova, P. Jandera, B. Buszeewski, J. Cromatogr. A 1109 (2006) 51.

[51] F. Svec, L. Geiser, R.E. Majors, LC GC N. Am. 24 (2006) 22.

[52] L. Rieux, D. Lubda, H.A.G. Niederländer, E. Verpoorte, Rainer Bischoff, J. Chromatogr. A 1120 (2006) 165.

[53] A.M. Siouffi, J. Chromatogr. A 1126 (2006) 86.

[54] N. Wu, J. Dempsey, P.M. Yehl, A. Dovletoglou, D. Ellison, J. Wyvratt, Anal. Chim. Acta 523 (2004) 149.

[55] J. Billen, P. Gzil, G.V. Baron, G. De Smet, J. Chromatogr. A 1077 (2005) 28.

[56] K. Cabrera, J. Sep. Sci. 27 (2004) 843.

[57] M. Al-Bokari, D. Cherrak, G. Guiochon, J. Chromatogr. A 975 (2002) 275.

[58] A.M. van Nederkassel, A. Aerts, A. Dierick, D.L. Massart, Y.Van der Heyden, J. Pharm. Biomed. Anal. 32 (2003) 233.

[59] S. El Deeb, H. Watzig, Turk. J. Chem. 30 (2006) 543.

[60] V.V. Tolstikov, A. Lommen, K. Nakanishi, N. Tanaka, O. Fiehn, Anal. Chem. 75 (2003) 23.

[61] J.K. Nicholson, P.J. Foxall, M. Spraul, R.D. Farrant, J.C. Lindon, Anal. Chem. 67 (1995) 793.

[62] E. Holmes, P.J. Foxall, M. Spraul, R.D. Farrant, J.K. Nicholson, J.C. Lindon, J. Pharm. Biomed. Anal. 15 (1997) 1647.

[63] M. Godejohann, J. Chromatogr. A 1156 (2007) 87.

[64] T. Ikegami, H. Fujita, K. Horie, K. Hosoya, N. Tanaka, Anal. Bioanal. Chem. 386 (2006) 578.

[65] F. Beaudry, P. Vachon, Biomed. Chromatogr. 20 (2006) 200.

[66] B.A. Olsen, J. Chromatogr. A 913 (2001) 113.

[67] J. Layne, J. Chromatogr. A 957 (2002) 149.

[68] E.S. Grumbach, D.M. Wagrowski-Diehl, J.R. Mazzeo, B. Alden, P.C. Iraneta, LC GC N. Am. 22 (2004) 1010.

[69] A.J. Alpert, J. Chromatogr. 499 (2007) 46.

[70] D.V. McCalley, J. Chromatogr. A 1171 (1996) 378.

[71] A.R Oyler, B.L. Armstrong, J.Y. Cha, M.X. Zhou, Q. Yang, R.I. Robinson, R. Dunphy, D.J. Burinsky, J. Chromatogr. A 724 (1996) 378.

[72] T. Yoshida, J. Biochem. Biophys. Methods 60 (2004) 265.

[73] B.Y. Zhu, C.T. Mant, R.S. Hodges, J. Chromatogr. 594 (1992) 75.

[74] S.C. Churms, J. Chromatogr. A 720 (1996) 75.

[75] H. Lindner, B. Sarg, C. Meraner, W. Helliger, J. Chromatogr. A 743 (1996) 137.

[76] V.V. Tolstikov, O. Fiehn, Anal. Biochem. 301 (2002) 298.

[77] J.K. Troyer, K.K. Stephenson, J.W. Fahey, J. Chromatogr. A 919 (2001) 299.

[78] P. Hemstrom, K. Irgum, J. Sep. Sci. 29 (2006) 1784.

[79] Y. Lin, D. Si, Z. Zhang, C. Liu, Toxicol. 256 (2009) 191.

[80] J.J. Pesek, M.T. Matyska, S.M. Fischer, T.R. Sana, J. Chromatogr. A 1204 (2008) 48.

[81] F.B. Collings, V.S. Vaidya, Toxicol. 245 (2008) 167.

[82] C. Hoppel, Am. J. Kidney Dis. 41 (2003) S4–S12.

[83] C.R. Roe, C.L. Hoppel, T.E. Stacey, R.A. Chalmers, B.M. Tracey, D.S. Millington, Arch. Dis. Child 58 (1983) 916.

[84] L. Vernez, M. Dickenmann, J. Steiger, M. Wenk, S. Krahenbuhl, Nephrol. Dial. Transplant. 21 (2006) 450.

[85] C.L. Hoppel, Techniques in Diagnostic Human Biochemical Genetics: A Laboratory Manual, Wiley-Liss, New York, (1991).

[86] L. Vernez, M. Wenk, S. Krahenbuhl, Rapid Commun. Mass Spectrom. 18 (2004) 1233.

[87] S. Ahmad, Semin. Dial. 14 (2001) 209.

[88] P.E. Minkler, J. Kerner, K. North, C.L. Hoppel, Clin. Chim. Acta 352 (2005) 81.

[89] M. Yoshino,Y. Tokunaga,Y. Watanabe,I. Yoshida, M. Sakaguchi, I. Hata,Y. Shigematsu, S. Yamaguchi, J. Chromatogr. B 792 (2003) 73.

[90] S. Vijay, A. Patterson, S. Olpin, M.J. Henderson, S. Clark, C. Day, G. Savill, J.H. Walter, J. Inher. Metab. Dis. 29 (2006) 627.

[91] T. Werner, I. Treiss, D. Kohlmueller, P. Mehlem, M. Teich, E. Longin, T. Gerstner, A. Schulze, Epilepsia 48 (2007) 72.

[92] K. Imura, A. Okada, Nutrition 14 (1998) 143.

[93] R. Lourenco, M.E. Camilo, Nutr. Hosp. 17 (2002) 262.

[94] T.E. Creighton, Protein Folding, W.H. Freeman & Co. New York, (1992).

[95] R.P.W. Scott, Techniques and Practices of Chromatography, Marcel Dekker, New York, (1995).

[96] V.R. Meyer, Practical High-performance Liquid Chromatography, Wiley, New York, (1994).

[97] M.E. Swartz, LC GC N. Am. (2005) 8.

[98] Department of Health and Human Services, Food and Drug Administration. “Good laboratory practice regulations. Final Rule” Federal Register 52 (1987) 33768.

[99] P.E. Minkler, S.T. Ingalls, C.L. Hoppel, Anal. Chem. 77 (2005) 1448.

[100] S. Cerutti, Masters Thesis, University of Florida, Florida (2007).

[101] P. Hemstrom, K. Irgum, J. Sep. Sci. 29 (2006) 1784.

[102] Y. Guo, J. Liq. Chromatogr. & Rel. Tech. 28 (2005) 497.

[103] X. Wang, W. Li, H.T. Rasmussen, J. Chromatogr. A 1083 (2005) 58.

[104] Y. Guo, S. Gaiki, J. Chromatogr. A 1074 (2005) 71.

[105] P. Molander, R. Trones, K. Haugland, T. Greibrokk, Analyst 124 (1999) 1137.

[106] J.B. German, B.D. Hammock, S.M. Watkins, Metabolomics 1 (2005) 3.

[107] S.M. Watkins, P.R. Reisfsnyder, H.J. Pan, J.B. Leiter, J. Lipid Res. 43 (2002) 1809.

[108] S.M. Watkins, X. Zhu, S.H. Zeisel, J. Nutr. 133 (2003) 3386.

[109] J.M. Freeman, E.H. Kossoff, A.L. Hartman, Pediatrics 119 (2007) 535-543.

[110] J.M. Freema, E.P. Vining, D.J. Pillas, J.C. Casey, L.M. Kelly, Pediatrics 102 (1998) 1358.

[111] C.B. Henderson, F.M. Filloux, S.C. Alder, J.L. Lyon, D.A. Caplin, J. Child Neurol. 21 (2006) 193.

[112] D.L. Keene, Pediatr. Neurol. 35 (2006) 1.

[113] K.J. Bough, J.M. Rho, Epilepsia 48 (2007) 43.

[114] A.L. Hartman, M. Gasior, E.P. Vining, M.A. Rogawski, Pediatr. Neurol. 36 (2007) 281.

[115] A.A. Morris, J. Inher. Met. Dis. 28 (2005) 109.

[116] A. Virmani, Z. Binienda, Mol. Aspects Med. 25 (2004) 533.

[117] GeneSpring MS Manual, Agilent Technologies, California, (2006).

[118] M.J. Baxter, Archaeometry 4 (2006) 671.

[119] X. Zhao, A. Peter, J. Fritsche, M. Elcnerova, A. Fritsche, H.U. Haring, E.D. Schleicher, G. Xu, R. Lehmann, Am. J. Physiol. Endocrinol. Metab. 296 (2009) 384.

[120] B.S. Fuehrlein, M.S. Rutenberg, J.N. Silver, M.W. Warren, D.W. Theriaque, G.E. Duncan, P.W. Stacpoole, M.L. Brantly. J. Clin. Endocrinol. Metab. 89 (2004) 89 1641.

[121] Bodamer, A. Muhl, Monatshefte fur Chemie 136 (2005) 1293.

[122] A. Schulze, M. Lindner, D. Kohlmuller, K. Olgemoller, E. Mayatepek, G.F. Hoffmann, Pediatrics 111 (2003) 1399.

[123] H-W. Hsu, T.H. Zytkovicz, A.M. Comeau, A.W. Strauss, D. Marsden, V.E. Shih, G.F. Grady, R.B. Eaton, Pediatrics 121 (2008) e1108.

[124] J.E. Cooper, Lab. Anim. 9 (1975) 329.

[125] S.A. Book, L.K. Bristad, J. Anim. Sci. 38 (1974) 997.

[126] E.M. Glauser, Exp. Med. Surg. 24 (1966) 181.

[127] P.R. Borum, Nutr. Rev. 39 (1981) 385.

[128] J.K. Baltzell, F.W. Bazer, S.G. Miguel, P.R. Borum, J. Nutr. 117 (1987) 754.

BIOGRAPHICAL SKETCH

Estela Soledad Cerutti was born in the beautiful province of San Luis, Argentina. She was a curious child and discovered her fascination for chemistry at an early age. The search for explanations, the "magic" colors behind the chemical reactions, and the understanding that "we all are a chemistry factory" led Soledad to choose chemistry as her major.

Soledad attended Universidad Nacional de San Luis, graduating with a Bachelor's degree in chemistry in March, 2002. After graduation, she immediately started working toward a Ph.D. and completed her studies with a doctoral degree in chemistry on December 15, 2006.

Soledad realized the importance of travel to United States to explore the field of mass spectrometry. After applying and being awarded with a Fulbright scholarship, she moved to Gainesville and started with her Master's program at University of Florida and completed these studies in August, 2007. Positive personal and scientific experiences convinced Soledad to continue with her research toward her Ph.D. degree in mass spectrometry under the supervision of Dr. Richard Yost and Dr. David Powell at University of Florida.