THE INVESTIGATION, DEVELOPMENT AND APPLICATION OF NON-TARGETED METABOLOMIC METHODS APPLYING DRIED BLOOD SPOT COLLECTION

by

ELLIOTT ANDREW PALMER

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences

College of Life and Environmental Sciences

University of Birmingham

September 2019

Abstract

The use of dried blood spots (DBS) as an alternative sampling method to venous blood draws is a well established technique within clinical and pharmaceutical research settings, with routine assays applied to quantify a small number of analytes applying targeted analytical assays. However the use of DBS within non-targeted metabolomics research is limited. Concerns about DBS have limited its use for non-targeted work and therefore its applicability in large scale studies where DBS is a very attractive alternative. With lower costs associated with collecting, storing and transporting samples, DBS offer a great alternative to whole biofluid samples.

This thesis aims at assessing the applicability of DBS collection in non-targeted metabolomics scenarios. Initially, the stability of DBS was studied over the course of 12 months, at various temperatures and guidelines surrounding the best practices of storing DBS identified with suitable stability at room temperature in the short term however for longer term storage of samples should be at -20°C or -80°C.

Secondly a surface analysis method was developed with the aim of increasing the throughput of DBS data collection. The method developed applied a simple 5 second extraction from the surface of a DBS with parameters optimised to be able to extract multiple thousands of features and similar analytical performance to a comparable direct infusion mass spectrometry method.

Additionally, a new method for normalisation of whole blood volume spotted on

to the DBS card was developed for non-targeted analysis with the goal of reducing non-biological variation within the sample. Results demonstrated applying the Bradford assay to the measurement of total protein content with a DBS reduced the non-biological variation present within the sample.

Finally, a surrogate of DBS in dried serum and plasma spots were applied in a medium scale biological study in developing countries, proving that whether or not you are using the whole biofluid or whether it is dried, a comparable biological conclusion is reached with similar lipid classes significantly changing in relation to hypervitaminosis A in both plasma or serum and their associated dried biofluid spots.

The research presented within this thesis has demonstrated the applicability of DBS for non-targeted metabolomics in mammalian research. With new methods to normalise data, to reduce technical variation and analyse DBS using a surface analysis technique. Further assessment in real world biological research applications are required to expand the application of DBS.

Acknowledgements

First and foremost I would like to extend my sincere gratitude to my supervisor Prof. Warwick Dunn for all the help and guidance he has provided throughout the period of my studies at Birmingham.

Secondly, my thanks go to all the members of the Fourth floor for being such a great support network over the past four years. In particular. special thanks go to:

- Will; for the laughs and motivation whilst we were preparing many hundreds of samples in the middle of a lab in San Jose.
- Cate; for providing much needed personal advice at key points throughout the last four years whilst also putting up with my incessant interruptions that came with being my desk neighbour.
- Martin; for his constant willingness to listen and provide helpful insight into problems I may have been having.
- Andy; for being a constant source of cynicism and making me realise there is a bright side to everything.
- Rosie (Rose, Rosemary); for providing me with much needed coffee breaks, 'dirty' pizzas and a new perspective on many issues that the last four years have bought, both personally and academically.

Thanks are also given to the Birmingham Metabolomics Training Centre, Phenome Centre Birmingham and the Advance Mass Spectrometry Facility for access to instrumentation and training.

Additionally I would also like to thank Helen Cooper and the members of the Fifth floor for their ability to provide a different perspective on my work and ask questions that allow me to think about my research from a different angle.

I would also like to extend my thanks to the Biotechnology and Biological Sciences Research Council (BBSRC) and Thermo Fisher Scientific for funding for this project.

My family also contributed in helping me get to this point and I would very much like to thank them all for the love and support that they have given throughout my education.

Finally, my thanks go to Laura for all the emotional help and support that has been given over these past 4 years in Birmingham, if not for you I know i would not be in this position right now and for that I will forever be grateful.

Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA	**An**alysis **o**f **Va**riance
APCI	**A**tmospheric **P**ressure **C**hemical **I**onisation
ATP	**A**denosine **Tri**p**h**osphate
CBB-G	**C**oomassie **B**rilliant **B**lue - **G**
CI	**C**hemical **I**onisation
d.c.	**D**irect **C**urrent
d.H_2O	**D**eionised **Water**
DBS	**D**ried **B**lood **S**pot
DIMS	**D**irect **I**nfusion **M**ass **S**pectrometry
DPS	**D**ried **P**lasma **S**pot
DSS	**D**ried **S**erum **S**pot
DUS	**D**ried **U**rine **S**pot
EI	**E**lectron **I**onisation
ESI	**E**lectro**s**pray **I**onisation
FDR	**F**alse **D**iscovery **R**ate
FT-ICR	**F**ourier **T**ransform - **I**on **C**yclotron **R**esonance
FWHM	**F**ull **W**idth **H**alf **M**aximum
GC	**G**as **C**hromatography
HCD	**H**igher Energy **C**ollision **D**issociation

HILIC	**H**ydrophilic **I**nteraction **Li**quid **C**hromatography
HPLC	**H**igh **P**erformance **L**iquid **C**hromatography
IPA	**I**so**p**ropyl **A**lcohol
LC	**L**iquid **C**hromatography
LESA	**L**iquid **E**xtraction **S**urface **A**nalysis
LLE	**L**iquid **L**iquid **E**xtraction
LMJ	**L**iquid **M**icro**J**unction
MS	**M**ass **S**pectrometry
nESI	**n**ano-**E**lectro**s**pray **I**onisation
NIR	**N**ear **I**nfra**r**ed Spectroscopy
nLC	**n**ano-**L**iquid **C**hromatography
NMR	**N**uclear **M**agnetic **R**esonance
PCA	**P**rincipal **C**omponent **A**nalysis
PKU	**P**heyl**k**eton**u**ria
PLS-DA	**P**artial **L**east **S**quares-**D**iscriminant **A**nalysis
Q-ToF	**Q**uadrupole-**T**ime **o**f **F**light
QC	**Q**uality **C**ontrol
QqQ	Triple Quadrupole
r.f.	**R**adio **F**requency
RBC	**R**ed **B**lood **C**ells
RSD	**R**elative **S**tandard **D**eviation
SPE	**S**olid **P**hase **E**xtraction
ToF	**T**ime **o**f **F**light
UHPLC	**U**ltra **H**igh **P**erformance **L**iquid **C**hromatography
WHO	**W**orld **H**ealth **O**rganisation

Dedicated to my grandfather

(1947-2017)

Chapter 1

Introduction

1.1 Metabolomics

Metabolomics as a recognised field has been around since 1998 [1], but the theory behind metabolomics has been around since the 1950s [2]. Metabolomics can be described as the study of the metabolome which is the total representation of metabolites present within a specific biological sample. Metabolomics compared to other 'omic' fields can be considered a dynamic measurement of all of the metabolic processes occurring within the biological system being interrogated [3, 4]. Due to metabolomics being thought of as 'downstream' from the other 'omics' processes such as genomics, transcriptomics and proteomics [5], often changes occurring at these levels, present themselves as changes within the biochemical pathways that make up the metabolome.

The range of compounds measured in the metabolome are generally less than 1500 Da [6]. This range covers a large number of classes of biologically relevant compounds from sugars to phospholipids which can be used to investigate the physiological state of the subject from which the biological sample originated. Many of the compounds analysed within metabolomics have key biological roles and most are involved in the different biochemical pathways within the biological subject which take part within

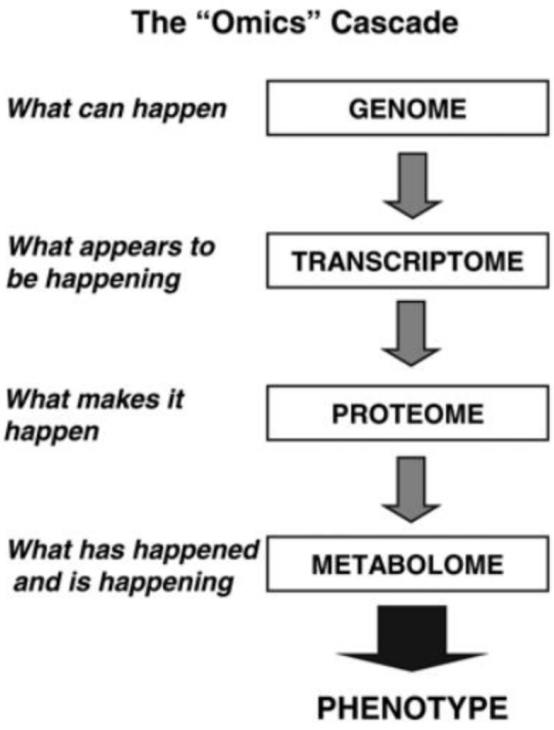


FIGURE 1.1: Representation of the 'omics' cascade with all areas contributing to the overall phenotype of the biological subject. Figure reproduced from [10].

many different functions in order for it to survive. Within humans for example the up-regulation and down-regulation of many metabolites can have profound effects on the physiological health of the subject due to the many metabolic pathways that metabolites may be a component of [7–9]. Therefore, identification of any anomalous levels which fall outside of typical metabolite ranges can be key in predicting physiological health or disease status. Consequently, it is important that the interactions and functions between the many thousands of metabolites are understood and how the different metabolic pathways which rely on these metabolites function and relate to physiological health.

In addition to the key biological functions that metabolites have on physiological

health, the metabolic phenotype is another highly dynamic product that has a large effect on the metabolome. The metabolic phenotype can be defined as the interaction of different environmental and lifestyle factors as well as genotype influences on the metabolome [11]. Due to the regularity in which these interactions occur, the metabolic phenotype can change very quickly. Thus it can be a good predictor of health and disease [12, 13].

The role in which metabolites have on the general health in biological organisms is wide ranging, from the regulation of insulin to control blood glucose levels [14], to the break down of acetylcholine by acetylcholinesterase [15] to allow muscles to relax. Due to the large number and complex nature of the compounds which makeup the metabolome, there has been a huge effort within the research community in providing databases of known and predicted compounds that are expected. These included databases such as The Human Metabolome Database [16], LIPID MAPS [17] and KEGG[18]. Currently HMDB reports above 114,008 compounds (September 2019), both measured and expected, but the volume of which are actually real metabolites is heavily debated. However, whilst these databases are useful resources in the identification and annotation of the metabolome, none of these databases provide a definitive and complete view of the metabolome. Whilst the use of these databases has allowed putative annotations of many compounds within non-targeted studies, the actual identification of metabolites requires further analysis than just matching to an *m/z* value. Fragmentation and measurement against chemical standards further help and increases the confidence that the compound being measured is the one expected and not just an isomer or isobar.

As mentioned towards the beginning, although metabolomics and it's theory has been around since the 1950s, the technical ability to measure the large number of

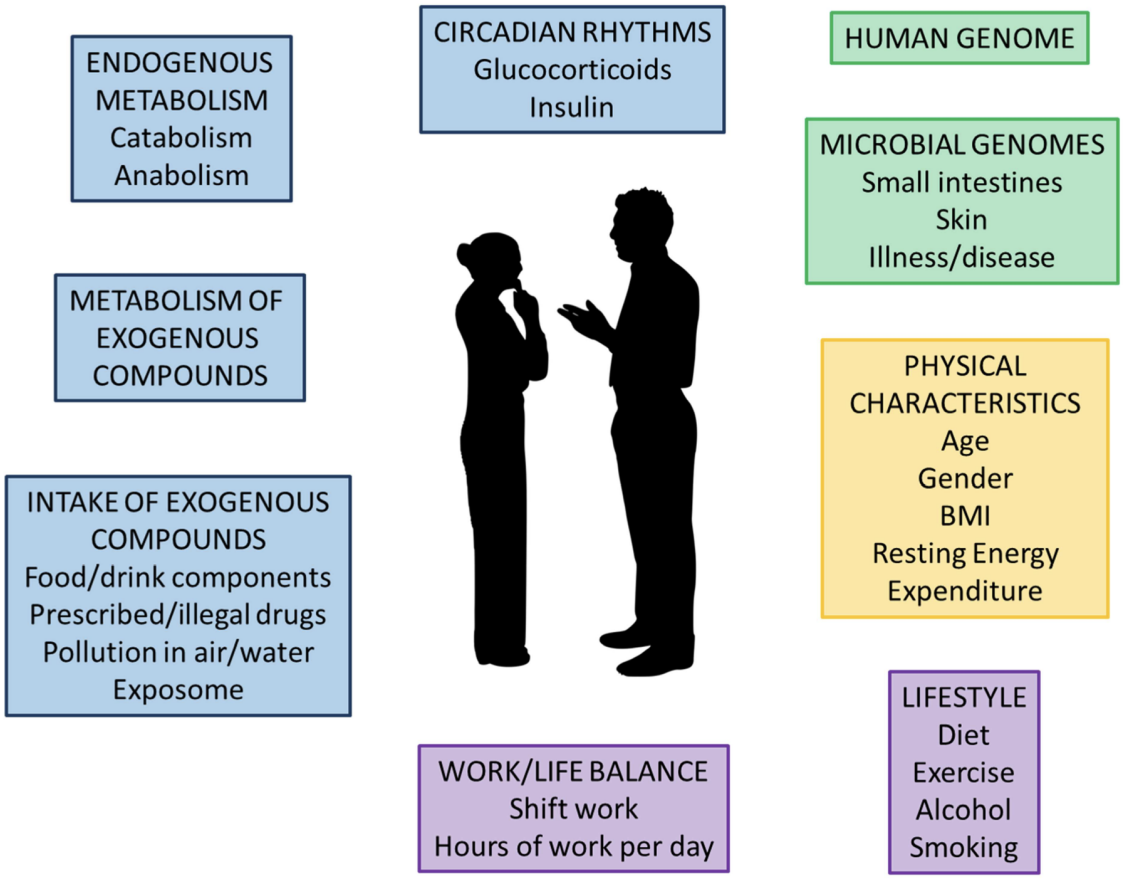


FIGURE 1.2: Environmental and physiological factors which influence the metabolism and subsequently alter the phenotype of an individual. Figure reproduced from [19]

metabolites present in a non-targeted manner has only recently been developed within the last 20 years. Initially, nuclear magnetic resonance spectroscopy was used to investigate perturbations within the metabolome before the advent of GC-MS and high resolution mass spectrometry systems which vastly increased the total number of metabolites able to be detected within a single sample. Combined with the advancement in separation sciences (both in liquid chromatography and gas chromatography), there are a large number of metabolites which can now be measured and subsequently identified through mass spectrometric measurement of various biofluids (serum and plasma [20–22], urine [23–25], cerebrospinal fluid[26–28], tears[29, 30], tissue [31–33] and many more). Furthermore, the use of different LC column chemistries within chromatography allow a wider range of compounds to be separated and subsequently detected, thus leading to better and more confident identification. Major examples include the complementary use of reversed-phase and hydrophilic liquid interaction column chemistries to ensure as many metabolites as possible [34], from the polar to the non-polar can be separated from each other and be detected as a individual feature rather than an annotation of one of a possible many.

Alongside the term metabolomics also comes the term 'metabonomics' [35]. Whilst the two are often used interchangeably, there is often a defined difference between the two. While metabolomics looks at the analytical approach as to how systems work and aims to quantify the presence of metabolites within a biological sample, metabonomics looks at the response of a system to external changes and identifies what changes occur within the metabolome and is defined as 'the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification' [4].

1.1.1 Types of Metabolomic Study

Within metabolomics there are different areas of experimental design which apply fundamentally different approaches to the measurement of metabolites within the metabolome. These different areas can broadly be defined as targeted [36, 37], semi-targeted [38] and non-targeted [39–41]. Firstly, targeted studies are aimed towards quantifying the absolute concentrations of a select few metabolites that are of interest. This may be compounds involved in a specific metabolic pathway or biomarkers of a particular disease being investigated. Conversely, non-targeted strategies aim at detecting as many metabolites as possible with minimal loss of biological compounds. With regard to data analysis, due to the large number of metabolites being present in non-targeted studies instead of absolute quantification, relative responses between metabolic features is the main parameter used in extracting significant changes occurring within the metabolome, with preliminary data analysis often applying multivariate statistical techniques such as principal component analysis (PCA) to identify and select features of potential interest. Finally, a third area of metabolomics study design is semi-targeted. Whereas with targeted studies the number of compounds detected is typically a maximum of 20 which are fully quantified, and in non-targeted studies low thousands are detected but with quantification limited to relative peak intensities. In semi-targeted, the aim is to quantify a few hundred metabolites which may make up a series of metabolic pathways of interest. However quantification within semi-targeted studies is usually limited to 1-point calibrations or performing a calibration against a compound of the same class of compound such as a particular lipid species. Within recent years there have been commercial suppliers offering semi-targeted kits with the materials required to perform these studies. One of the leading companies in this field is Biocrates who offer kits that claim to quantify hundreds of metabolites

applying their p180 [42] and p400 HR [43] kits.

With targeted studies, the requirement to analyse a select number of metabolites usually involves more extensive sample preparation such as solid phase extraction [44, 45] to remove unwanted compounds whilst increasing the sensitivity and specificity of any analysis performed. Instrumentation used to perform targeted studies is also different with triple quadrupole based mass spectrometers dominating studies applying targeted metabolomics [46]. Non-targeted studies on the other hand aim to have sample preparation procedures which involve as fewer steps as possible to prevent any and as little loss of metabolites [47, 48] from the sample as possible, with preparation limited to the removal of protein content from the sample typically [49]. Non-targeted studies also tend to use instruments [50, 51] with high mass resolving power to ensure metabolites of similar mass are able to be distinguished, many of which are present in complex samples such as biofluids. Targeted and non-targeted studies can also be categorised as either hypothesis testing or hypothesis generating respectively [52]. Targeted metabolomics is generally the last step in the validation of a hypothesis which has been generated by a non-targeted study. A panel of potential biomarkers is then selected for further investigation to ensure that the hypothesis generated in a non-targeted scenario, is still valid in a targeted scenario. Subsequently, absolute concentrations of metabolites can be then be measured between different biological classes such as a treatment and control group.

1.1.2 Metabolomics studies involving blood and urine

For metabolomics studies involving whole blood, it is required that soon after collection it is processed into either serum or plasma to ensure the long term integrity of the sample [53, 54]. This is due to the inability to freeze whole blood which would lead

to cell lysis [55, 56]. Therefore many studies involving the collection of whole blood through venipuncture, are often performed using serum or plasma.

Serum and plasma are prepared through the low speed centrifugation of whole blood. In the case of serum preparation, whole blood is collected into tubes without an anticoagulant present and the blood is allowed to clot. After clotting the sample can then be centrifuged and the supernatant remaining is serum [22]. With plasma preparation, whole blood is collected into collection tubes containing an anticoagulant such as lithium heparin or sodium EDTA [57]. The sample is then centrifuged with the plasma being left as the supernatant layer. Although the two processed biofluids are metabolically similar they do differ in part due to the lack of clotting factors that have been removed in the serum compared to plasma [22, 58]. Additionally, because of the presence of anticoagulant, adverse effects on MS sensitivity can appear when analysing plasma due to ionisation suppression [59]. However having the flexibility to collect and prepare blood in either scenario is a benefit to metabolomics and can afford a wide range of experimental design approaches.

Another major biofluid that has been widely studied in non-targeted metabolomics is urine [25, 60]. Urine undergoes a variety of changes depending on the physiological condition of the subject (pH [61], ionic strength [62, 63] and osmolality [64]), thus often complicating experimental analysis when compared to serum or plasma. To negate this effect normalisation methods are used to ensure that the physiological changes that occur readily within urine samples such as the dilution due to hydration or concentration of various compounds that are excreted via the kidneys due to toxicity have no effect on the resulting data that is collected [64, 65].

1.1.3 Extraction methods for non-targeted studies

There are a variety of methods used for the extraction of metabolites from the biological matrix into LC and MS compatible solvents. The two main methods used within the literature today apply liquid-liquid extraction (LLE) and solid phase extraction (SPE) [66] (including solid phase micro-extraction (SPME) [67]. LLE involves the addition of organic solvents into to a sample and thorough mixing to partition metabolites from their native environment within the biofluid into the organic solvent. Depending on the choice of extraction solvent, major influences can occur on the classes of compounds which are extracted from the biofluid, with different solvent mixtures being more suited towards the extraction of specific compounds over others. Thus depending on the assay that is to be performed can cause major decisions to be made within experimental design to ensure the correct classes of molecules can be extracted.

One of the most common methods of metabolite extraction methods is applying a monophasic solvent mixture typically based off a polar organic solvent such as methanol, isopropanol or acetonitrile mixed with water [32]. Extraction applying a monophasic solvent system is simpler and requires less steps, than biphasic and solid phase extraction methods, however optimal extraction of the many classes of compounds present within the sample is not possible with a monophasic system. Therefore biphasic solvent extraction systems are available and are frequently used within non-targeted metabolomics research. Solvent systems developed by Bligh & Dyer [68] and Matyash [69, 70] are common throughout non-targeted metabolomics in the optimal extraction of multiple classes of compounds from a single sample. Bligh & Dyer's biphasic system is based on a combination of chloroform/methanol/water, with lipid

classes preferentially extracting into the chloroform phase, with polar metabolites extracted into the methanol/water phase. Matyash's biphasic system employs methyl-tert-butyl ether as the non-polar phase and can be viewed as a more preferred alternative due to the carcinogenic nature of chloroform. Additionally, with Matyash's method, the protein pellet formed forms at the bottom of the extraction vial instead of in between the two phases as with the Bligh & Dyer method. This allows it to be a more suitable extraction method for automated systems which lack the fine motor skills of manual handling to avoid the protein pellet [70].

When applying SPE, specific chemistries of stationary phase are loaded with sample upon which differing strengths of mobile phase are then washed through. This removes components which are unwanted leaving behind metabolites of interest on the stationary phase which can then subsequently be eluted for analysis. This procedure can also be performed the other way around where components of interest are eluted first off the stationary phase. SPE is often more suited towards targeted analysis due to the removal of much more of the metabolome leading to a more specific group of compounds ready to be analysed [71]. Due to the presence of proteins within serum and plasma, it is required that these are removed from the sample prior to analysis. This is achieved alongside metabolite extraction through the addition of organic solvents such as methanol and acetonitrile [72]. With the addition of organic solvents, proteins are crashed out of solution leaving a purely metabolite based extract present within the organic solvent.

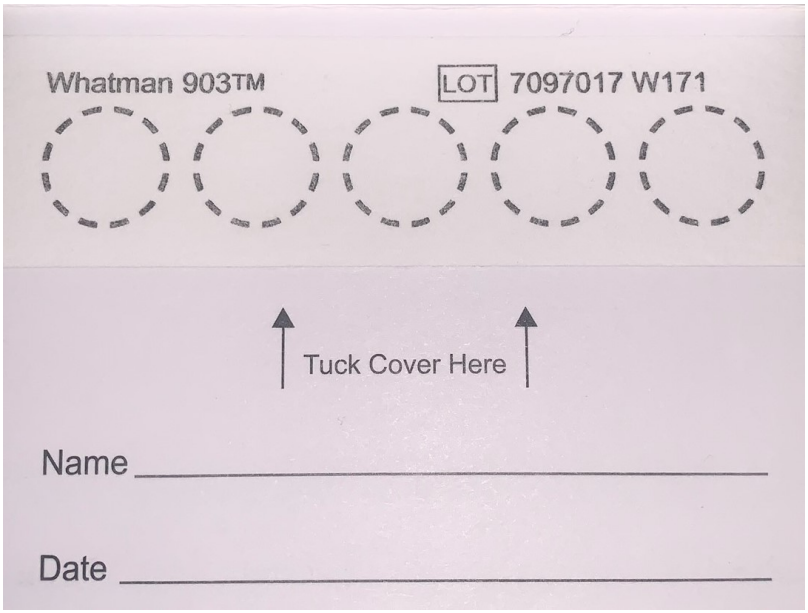


FIGURE 1.3: Example of a Whatman 903 Collection card

1.2 Dried Blood and Urine Spots

Due to the challenges in the analysis of whole blood and the processing requirements to be able to store and transport samples confidently, there has been a large increase in interest in using filter paper based collection cards not just for metabolomics but in other research areas. Through the use of collection cards, samples can be collected directly onto filter paper based mediums before allowing to dry and then storage. This is in direct comparison to the processing requirements for either plasma or serum where samples are required to be collected by trained personnel and then subsequent physical processing through centrifugation to prepare samples that can be stored appropriately with minimal changes to the biofluid.

The use of dried blood spots as an alternative to whole blood collection dates back to the 1960s where Guthrie and Susi [73] spotted blood from the heel of neonates onto filter paper. Subsequent assays were then performed to test for phenylketonuria

(PKU). PKU is a genetic condition that can lead to impeded learning ability, behavioural problems and epilepsy due to the inability to metabolise phenylalanine [74]. Thus early detection of the disease is key to prevent any long term effects for the neonate. Since this novel assay was developed, there has been a large increase in the number of assays performed using DBS. Most of this increase in use of DBS can be explained through the large number of inborn errors of metabolism which are now screened for in developed countries around the world. With inborn errors of metabolism it is key that any of these conditions are caught early on, and with the use of DBS, discomfort to the neonate is kept to a minimum whilst still being able to collect enough blood upon which is then sent for analysis to detect any inborn error of metabolism which occurs in 1 in every 2500 births [75]. Therefore targeted analysis of DBS is commonplace within clinical settings.

1.2.1 Advantages

Compared to typical blood sampling procedures, where trained clinical staff are required to ensure the safe collection of blood, DBS can be collected by the patient themselves [76, 77] if required using only a spot of capillary blood, usually collected from the tip of a finger or in the case of neonatal patients, the heel. Blood is collected through a small puncture in the skin created using a lancet device.

Due to the low complexity required to collect samples, training materials can be prepared in advance before either sending to patients who have agreed to collect their own samples, or to researchers who are not trained in phlebotomy but are still interested in the use of DBS for their own research. Therefore samples can be collected away from the lab in environments where specialist laboratories are not availabel for

FIGURE 1.4: Image describing how a DBS is created through the pressure of a spot of blood at the end of a finger to a collection card. Image reproduced from [78].

the processing of whole blood directly. Another key advantage that is of huge benefit to patients is the volume required to prepare a single DBS compared to that of a serum or plasma sample. With DBS, sample volumes range from 10-20 μL [79] and thus discomfort is kept to a minimum whilst also maintaining sample integrity. This is compared to whole blood draws through venipuncture where volumes can range from 0.1 mL up to 10 mL [80]. Additionally, due to the low volumes required, samples can be collected in challenging environments [81–83] where collecting a venous sample would either require any potential dynamic research to be halted such as collecting samples whilst an athlete was exercising, or from a patient who has a severe phobia of needles. Due to the low volumes of blood, costs of collection, transport and storage of samples is drastically reduced when compared to whole biofluids [83, 84]. If the stability of analytes has been determined to be stable whilst present as a DBS, then collection cards can be sent via typical commercial mail routes [85, 86] rather than

employing specialist couriers to transfer plasma and serum samples using cold storage techniques such as dry ice or ice packs. Therefore, opportunities arise where samples can be collected in remote locations before being sent back to the analysis laboratory, again, widening the range of research that can be conducted. Also, for researchers that have already prepared plasma or serum, yet are either volume limited to reluctant to send large volumes of whole biofluid, the volumes required to prepare replicates of these samples as dried biofluids and then shipped to collaborators is much cheaper.

1.2.2 Identified Drawbacks

Although DBS analysis has been around now for over half a century, using DBS for high throughput, non-targeted metabolomics has only become technically possible in the last 10 years. With the advent of advanced high resolution, high sensitivity mass spectrometric techniques, the routine analysis of DBS is now possible. However, validating DBS for use in non-targeted metabolomic research has also become a wide research areas in itself.

Due to the nature of DBS, there can be a wide variety of ways in which samples can be collected. This could be down to the nature of the filter paper used, to the amount of training given to the person collecting the samples, as well as physical properties of DBS which could have effects on any analysis completed. One of the major concerns with DBS is the effect of haematocrit [87–90] and chromatographic effects arising from the deposition of whole blood onto filter paper collection cards [91]. Haematocrit levels can be defined as the amount of red blood cells (RBC) within a sample of blood and is expressed as a percentage of the total volume [92]:

$$\text{Haematocrit} = \frac{\text{Volume of Red Blood Cells}}{\text{Total Volume of Blood}} * 100 \tag{1.1}$$

Typically haematocrit levels vary on a variety of external factors such as sex, diet and race, however accepted levels for men are in the range of 40-53% and for women are 36-48% [93]. Higher haematocrit levels lead to a more viscous blood sample which when deposited onto a collection card will lead to a smaller DBS in radius [94]. This is in comparison to low haematocrit samples which when deposited will spread further even if the same volume of blood has been deposited onto the collection card.

In many studies, standard procedures indicate that when a DBS is to be extracted for analysis, a small sample of the DBS, typically taken using a hole punch is removed from the larger spot. Thus, if separate samples are prepared with differing haematocrit levels, when a fixed size is removed from the original DBS, two different volumes of blood will be represented, consequently leading to potential for quantitative bias in any analysis performed.

In addition to haematocrit issues surrounding DBS, chromatographic effects can also arise in DBS analysis. Due to the multicomponent nature of whole blood compared to serum or plasma, there is a potential that whenever a DBS is prepared, the separation of these various components within the filter paper matrix can occur [95, 96]. This therefore can lead to variation in response of analytes, depending on where a subsection of the DBS was removed.

An obvious solution to the above issues of haematocrit and chromatography effects is the removal and extraction of any metabolites from the whole spot instead of a subsection of the sample. However due to the way in which DBS are collected (typically through applying the collection card directly to the finger or heel), ensuring the same volume of blood is collected onto the collection cards is incredibly difficult and instead there ends up being a wide variation in the volume of whole blood that ends up on the collection card. Combine this with the issue of haematocrit, it is therefore not possible

to deduce an accurate estimate for the amount of blood deposited onto the collection card purely from measuring the radius of the DBS itself.

1.2.3 DUS compared to DBS

In addition to DBS, collection cards can also be used for a variety of other biofluids such as plasma and serum, but also for urine [97–99], saliva [100–102] and cerebrospinal fluid [103, 104]. From these it is most likely that any further development of non-targeted metabolomics research using collection cards, would have more applicability for using dried urine spots that any other. This is due to its already widespread use within the metabolomics community as a way to look at biological breakdown products of many pathways within the human body as well as an endpoint for many bi-products of food [105] and drug metabolism [106]. It is also easy to obtain in larger volumes compared to blood (>50 mL) and due to the method in which it is collected, would allow it to be prepared in a dried format more easily than most other biofluids. With little in the way of protein or lipids present within urine, it is also a favourable biofluid to analyse, with fewer steps required to ensure the comprehensive analysis of the many small molecules present.

1.2.4 Collection devices for dried biofluid collection

There are many different collection devices that can be used to prepare DBS & DUS alongside various other dried biofluid spots. Most are based on filter paper which is a cellulose based material [107, 108], which allows the biofluid to disperse within the matrix of the material. These collection cards are also prepared in different variations depending on the application upon which is required. Some collection cards

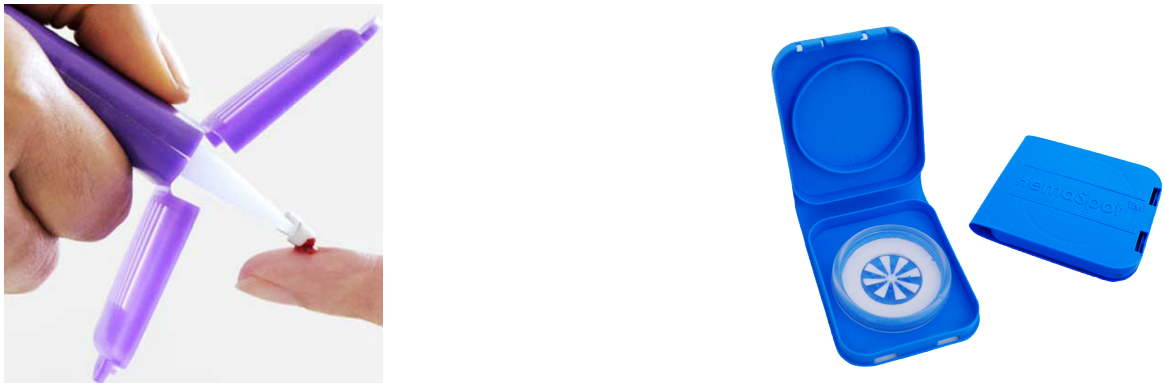
are treated with specific materials such as FTA® cards which have been shown to deactivate certain pathogenic viruses such as Avian Influenza Virus after just 1 hour of being deposited [109]. Most filter cards used however in DBS collection are untreated such as the Whatman 903 and Perkin Elmer 226 brand of cards which are commercially available from many suppliers.

In addition to the untreated cards being most common, the Whatman 903 and Perkin Elmer 226 cards are the only devices approved by the FDA in the USA for medical use [91, 107]. Consequently they are highly monitored for performance and variation in batch quality.

Alongside the filter paper based collection cards however, is a developing industry of alternatives which aim at ensuring the downsides associated with spot based collection such as haematocrit and chromatography effects are negated.

First is a company called Neoteryx which produce a 'stick' like device with an absorbent material on the end [110]. The device can be purchased in different forms which allow either 10 μL or 20 μL of blood to be collected within the absorbent material. The company claims that the reproducibility of volumes that are collected fall within an acceptable range for non-targeted metabolomic studies.

In addition to the above there is also a company called Spot On Sciences which produce an all in one device called Hemaspot [113] which acts to collect three drops of blood which then disperse onto a fan shaped material where the whole blood dries in a cartridge. Once the blood has dried it can then be shipped to the laboratory where each 'segment' of the fan can be removed before extraction and analysis. However unlike the Neoteryx kit which aims to remove any volume related issues in collection, with Spot On Sciences kit, there is still the issue of unknown volumes of blood being deposited onto the material itself. So whilst it is a convenient kit to use, issues are still



(A) (B)

FIGURE 1.5: Examples of (A) Neoteryx Mitra and (B) Spot On Science Hemaspot Devices. Images reproduced from [111, 112]

present which may prevent its use within non-targeted metabolomics studies.

Whilst manufactured kits are possible alternatives for dried biofluid collection, the cost of using these kits still remains greater than that of filter paper based collection cards therefore specific applications would have to be required which takes advantage of one of the unique selling points to be justified within studies.

1.2.5 Extraction of Dried Blood Spots

As with the extraction of whole biofluids such as plasma or serum, dried biofluids can be extracted in same manner with the addition of organic solvents. Many organic solvents have been utilised for extracting DBS such as acetonitrile [114], methanol [115] and ethyl acetate [116]. Because whole blood cannot be directly infused into a mass spectrometer, whole blood in its liquid form is typically prepared as plasma or serum before extraction. This is not required for DBS analysis and therefore removes an additional preparation step. Due to this, any extraction of DBS, also contains the intracellular components that are present within the cells in blood and additional metabolic

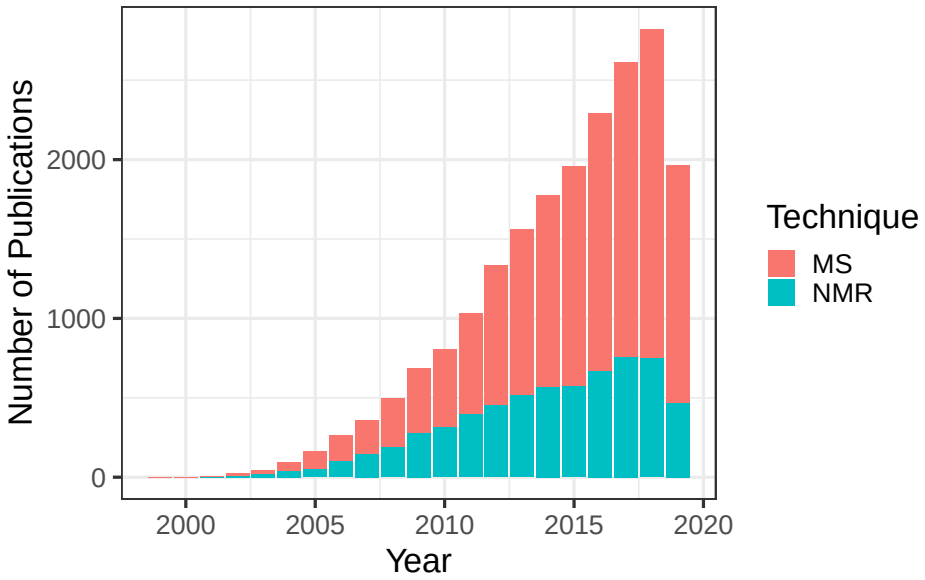


FIGURE 1.6: Total number of publications as indexed by Scopus on 5th September 2019 applying the search terms "Metabolomics" AND "NMR" or "Metabolomics" AND "Mass Spectrometry"

information may be gathered. Organic solvents however are still required to remove protein content and clotting factors, which can then be removed through centrifugation. The resultant supernatant can then be removed and transferred to separate tubes for drying down under vacuum or nitrogen gas.

1.3 Analytical Techniques

The total number of analytical techniques available for metabolomics researchers has vastly increased in recent decades. However the two main areas of focus in which developments have occurred are within nuclear magnetic resonance (NMR) spectroscopy [117] and mass spectrometry (MS) [10] instrumentation. Whilst NMR was the most prevalent towards the start of metabolomics research, since the turn of the twenty-first century [4], MS has become the dominant technique in use.

The primary reason for MS becoming the most prevalent technique can be attributed to a variety of reasons. The skill level required to operate instrumentation and interpret MS data has a lower level barrier to entry so the training of researchers in the operation of MS instrumentation can be easier and quicker than training someone to operate NMR spectroscopy instruments and interpret data. Furthermore, with the recent advances in high resolution mass analysers and the ease in coupling liquid chromatography devices prior to MS detection, the total number of compounds which can be detected is vastly higher than what can be hoped for NMR spectroscopy experiments. However where NMR spectroscopy excels is its ability to elucidate structures of compounds within the sample [50] in addition to the excellent technical reproducibility and ability to integrate data collected in different batches across multiple months and years. Structural confirmation and therefore metabolite identification is still an area in which MS is inferior to NMR spectroscopy. MS often relies on accurate mass annotation of compounds and/or MS/MS fragmentation to elucidate structures [118]. Both of these methods are susceptible to false annotations which may direct the researcher interpreting the data down the wrong pathways or hypotheses [19].

Because this thesis solely uses MS instrumentation, the theory behind NMR analyses will not be discussed and instead the next section will solely focus on the detection and applicability of MS instrumentation in non-targeted metabolomics.

1.3.1 Mass Spectrometry

As briefly mentioned in the previous section, the use of MS in metabolomics research is now much larger than NMR. The versatility of MS has allowed a wide range of applications to be developed, mainly in part due to the large number of ion sources and separation modes that can be placed before the MS detector.

Today a wide variety of mass spectrometric methods are available for the metabolomics analysis of different biofluids. These include hyphenated techniques comprising of a separation technology followed by MS detection such as:

- Ultra-high performance liquid chromatography – mass spectrometry (UHPLC-MS). - Note: UPLC is often used within the literature however is not used within this thesis because UPLC is a trademark of Waters Corporation. Therefore UHPLC refers to LC systems by other manufacturers.
- Nano-liquid chromatography – mass spectrometry (nLC-MS)
- Gas chromatography – mass spectrometry (GC-MS)

In addition to separation based techniques, direct analysis techniques are also very commonly employed for non-targeted research [119–121]. Whilst lacking separation removes a vital component in separating complex mixtures such as biofluids, the increase in throughput can allow large scale studies to be performed in a much shorter space of time. This presents itself well if the technique is used purely as a discovery based tool with further analysis performed later on a much more select group of potential markers that have been identified. However with no separation, correct annotation of metabolites becomes much more difficult without the addition of MS/MS data collection, which can still prove problematic with complex spectra where the isolation of narrow *m/z* ranges for fragmentation may be difficult.

1.3.2 Chromatography

The use of chromatographic techniques allows the separation of a complex mixture based on the interaction of the different compounds within a sample with a stationary

phase before finally eluting from the column before detection. Compounds present can either be separated based on their interaction with a stationary phase either in a gaseous or liquid form. Whilst GC-MS is more suited towards volatile compounds or compounds which can be derivatised to ensure volatility [50], LC-MS has come to the forefront of hyphenated metabolomics techniques due to its ability to handle a wide range of molecular classes through separation on an equally wide range of stationary phases that are commercially available from many manufacturers [122]. In recent years, both reversed-phase [123, 124] stationary phases and hydrophilic interaction [125–127] phases have been the primary columns used in LC-MS metabolomics . Further technical advances within the LC area have also led to the development of columns with narrower internal diameters [128, 129] and smaller particle sizes [130, 131], increasing the sensitivity and resolution of any separation performed [132]. In addition to smaller internal diameters and particle sizes, the advancement of instrumentation which can handle the increased back pressures have also allowed for faster LC run times whilst maintaining separation performance [133].

1.3.2.1 Column Parameters and Effect on LC Performance

As previously mentioned, there are a few key parameters that can affect the performance of any LC separation. These are:

- Particle size
- Length of column
- Internal diameter
- Stationary phase

- Mobile phase selection
- Column temperature

Many columns used in metabolomics research are now based on particle sizes of 2 μm, with a length of between 5 and 15 cm using an internal diameter of 2.1 mm. Especially for high throughput and large scale studies, these parameters can have a major influence on the time taken to separate the complex mixtures and the overall run time for the study.

The main parameter that describes the efficiency of any separation in LC and GC is the total number of theoretical plates. The Van Deemter Equation as described by Equation 1.2 is used to calculate the efficiency of a column where HETP is the height equivalent to the theoretical plates, A is the Eddy diffusion parameter, B is the random molecular diffusion term and C is the mass transfer term.

A larger number of theoretical plates for a column will lead to an increased chromatographic resolution and lower full width half maximum (FWHM) for any particular peak, indicating better separation of a complex mixture. Whilst decreasing particle size increases the total number of theoretical plates, as described by the Eddy diffusion term in the Van Deemter equation, decreasing particle size also leads to an increase in back pressure as described by the Kozeny-Carman Equation (Equation 1.3).

$$\text{HETP} = A + \frac{B}{u} + Cu \tag{1.2}$$

In the Kozerny-Carman Equation ΔP relates to the change in back pressure, η is the viscocity, F is the flow rate, L is the length of the column, K^0 is the specific permeability, r is the radius and d_p is the particle size.

$$\Delta \mathrm{P} = \frac{\eta F L}{K^0 \pi r^2 d_p^2} \tag{1.3}$$

From Equation 1.3 any decrease in particle size, has an inverse square effect on back pressure. Therefore instrumentation has to be able to handle the increased back pressures that come with columns of smaller particle sizes. Additionally, by decreasing the particle size, an increase in resolution is possible which can be explained by the efficiency term of the Purnell equation.

$$R_s = \frac{\sqrt{N}}{4}\left(\frac{\alpha - 1}{\alpha}\right)\left(\frac{k}{1+k}\right) \tag{1.4}$$

With α denoting separation factor of two peaks and k defining the retention factor, N is the efficiency term and can be calculated by the following equation:

$$\mathrm{N} \approx \frac{L}{2d_p} \tag{1.5}$$

Upon which L is column length and d_p denotes particle size

Therefore any decrease in particle size, has an inverse effect on the efficiency term which when increased, increases the resolution of any peak separation. Therefore higher column separation performance is possible through the decrease of particle size, providing the instrumentation and method developed can cope with the increased back pressures.

From Equation 1.5 we can also see that an increase in column length will lead to an increase in the efficiency term of the Purnell equation which defines resolution. Therefore, by using longer columns, resolution can be increased however this comes at the cost of decreased throughput and higher costs if analyses time is increased for fewer samples.

Finally, the effect of reducing the column diameter also increases the resulting back pressure as explained by Equation 1.3 again.

Therefore any choice of column to be used for any separation has to be a careful choice of optimising the performance of the separation whilst maintaining compatibility with any LC system that is to be used and which can handle either the optimum flow rates and pressure required to use a column efficiently.

1.3.2.2 Different Stationary Phases used in Metabolomics

Within metabolomics, there are a few main types of stationary phases that are common and presently used in many applications. Firstly is the use of RPLC columns which involve the use of a non-polar stationary phase and the use of a polar mobile phase such as a methanol:water mixture. Thus RPLC excels in separation of semi-polar and non-polar compounds. RPLC stationary phases are often made of particles consisting of C_{18} carbon chains bonded to silica beads, consequently, non-polar compounds interact with these chains leading to better separation and polar molecules pass through along with the mobile phase and are poorly retained. Other carbon chain lengths have also been used for analysis of larger molecules such as proteins, or the use of C_{30} columns for the separation of lipid classes.

Whilst reversed-phase chromatography has poor retention for highly polar compounds, normal phase chromatography can retain these compounds on a polar (commonly silica) stationary phase and using a non-polar mobile phase such as hexane. However due to the organic nature of the mobile phase, this technique becomes incompatible when coupling with MS systems and the ionisation source required such as electrospray ionisation. This incompatibility and requirement to separate polar compounds which dominate a huge part of the metabolome, has led to the development

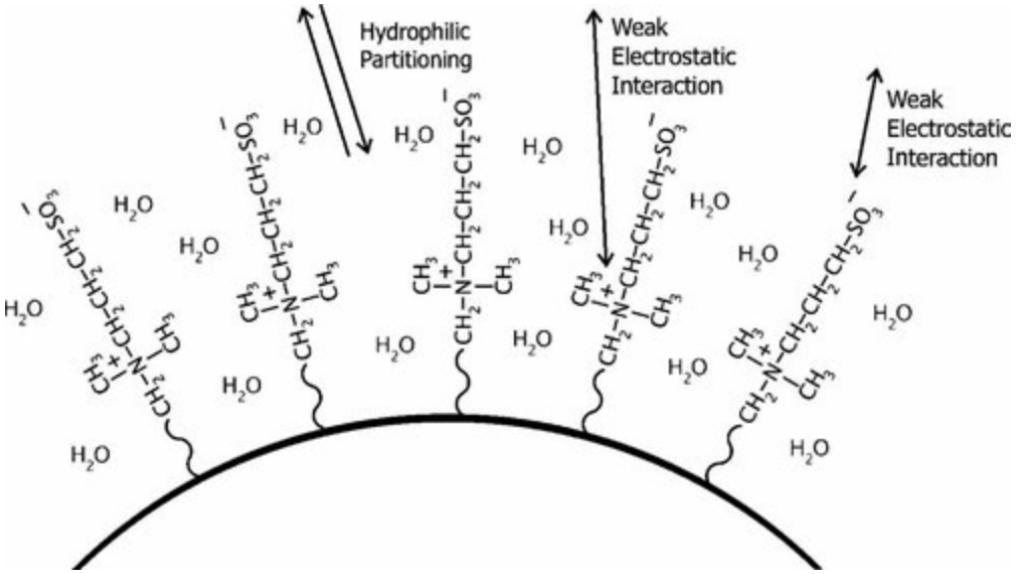


FIGURE 1.7: Retention mechanism of a sulfoalkylbetaine zwitterionic based HILIC stationary phase. Compounds first partition into the water rich layer before further electrostatic interactions with the stationary phase. Figure reproduced from [126].

of hydrophilic interaction liquid chromatography columns (HILIC). The mechanism by which HILIC works is based upon two steps. Firstly a water rich layer is formed between the mobile phase and the stationary phase upon which polar metabolites partition into. Secondly whilst in the water rich layer, further retention on the column is then based on the electrostatic interaction that occurs between the analytes and the stationary phases at specific pH.

Polar and thus more hydrophilic compounds partition into this water phase and interact with the stationary phase and are retained longer than any non-polar compounds. Typically solvents applied for HILIC separation are based on a binary mixture of acetonitrile and water, which are both compatible with MS systems in direct comparison to normal phase chromatography.

Therefore there are a wide range of separation techniques available for almost any

type of compound that may be of interest to the researcher. However, where discovery based studies are not focused on a specific area of the metabolome then a combination of both RP and HILIC columns can be used to ensure as many metabolites as possible is gathered and any biochemical pathway perturbations can be detected.

1.3.2.3 Nano Liquid Chromatography

As with the move from HPLC to UHPLC columns, where the internal diameter and particle sizes have decreased to improve resolution and flow rates can be increased to achieve the same performance in quicker run times, there has also been further miniaturisation of internal diameters to sub-milimeter diameters to further increase the performance of separation. Where column diameters are less than 1 mm, the term micro liquid chromatography is often applied and internal diameters of less than 300 microns, nano liquid chromatography used. Making matters complicated however is that the boundaries for these terms are blurred, and terminology used is often different depending on the author of any corresponding research.

However as mentioned earlier, any decrease in particle size has an inverse square effect on the total back pressure. To be able to deal with the increase in column back-pressures, a reduction in flow rate is often applied, however this leads to increased run times and costs per sample. To date, several companies have developed commercial systems to perform nano-LC separations, including all the major MS instrument manufacturers. The reduction in internal diameters leads to stronger chromatographic bands eluting from the column from the same volume of sample, this increase in sensitivity allows smaller volumes of samples to be used and due to the reduced flow rate to ensure optimum performance, as well as a massive decrease in the amount of solvent used.

Many studies using nano-LC have been applied to proteomics [134–136], with very few studies relating to the analysis of small molecules [137, 138]. The advantages offered by applying nano-LC therefore offer a promising potential in the future of non-targeted metabolomic applications. However due to the complex nature of sample preparation to ensure high column performance and the inability to run large scale studies in 'reasonable' run times, this leads to a requirement in the improvement in performance of the instrumentation required to cope with the flow rates and back pressures required to reduce run times to allow large scale studies.

1.3.3 Ionisation Methods

After sample separation, compounds are required to be ionised before detection by mass spectrometry. Furthermore, by imparting a charge on an analyte, it can then be manipulated based on electrical, magnetic and RF potentials for movement throughout the mass spectrometer before mass analysis and detection.

To be able to be detected, a molecule must be present in the gaseous phase. Unlike for GC where analytes are already in the gaseous phase either through their natural volatility or through derivatisation, analytes in LC must first be converted into the gas phase. Additionally, all mass spectrometry systems rely on the measurement of an analytes mass-to-charge ratio (m/z, thus analytes need to be charged prior to detection.

To achieve this there are two main techniques used which can either be classed as ionisation under high vacuum or ambient pressures. Each technique has a variety of characteristics which can be utilised either for the identification of analytes or the flexibility to modify and operate at atmospheric conditions.

1.3.3.1 High Vacuum Ionisation for Gases

The two major techniques which ionise under high vacuum are electron ionisation (EI) and chemical ionisation (CI). Both methods produce significantly different types of ions with EI being classed as a high energy technique which produces many fragments, whilst CI is much 'softer' ionisation technique leading to more intact and solely protonated molecules.

1.3.3.1.1 Electron ionisation

Initially developed in 1929 by Bleakney [139], electron ionisation as a technique is mainly used for GC-MS analysis and involves the bombardment of analytes with electrons. Due to this, many fragments are produced [140] leading to easier structural elucidation. Most EI source operate at 70 eV which produces electrons with a corresponding wavelength similar to that of a covalent bond. When the molecule interacts with specific wavelengths, complex interactions occur which promote the loss of an electron [140]. The molecule remaining then undergoes fragmentation or rearrangement to lose the energy imparted. GC-MS EI fragmentation libraries have vast wealth of knowledge and make the annotation of analytes more straightforward than other methods, and when all EI sources operate at the same energy of 70 eV, direct comparisons between fragmentation libraries are possible [141]. Electrons for the fragmentation are typically produced from a heated filament and then accelerated towards the gaseous analytes. EI is a useful ionisation tool due to the similar wavelength specific energies of electrons and bond lengths of analytes leading to the formation of fragment ions.

1.3.3.1.2 Chemical Ionisation

Chemical ionisation can be viewed as a complementary technique to EI due to the much lower energy of ionisation that is achieved through the collision of analytes with primary ions already present in the source [140]. First developed in 1966 by Muson et al [142], CI, using reagent ions such as NH_3^+ which have already been ionised by EI are collided with analyte ions entering the source. The protonated CI gas then provides the charge to the analyte molecule through a series of gas phase reactions. This charge transfer is at much lower energy than EI, leading to less chance of fragmentation.

1.3.3.2 Atmospheric Pressure Ionisation

Due to the high vacuum nature of EI and CI, coupling systems such as UHPLC instrumentation can become problematic. Therefore since the early 1990s there have been a huge research focus on the development and application of on-line coupling of techniques which operate at atmospheric pressures such as UHPLC to the high vacuum status of detectors within a mass spectrometer.

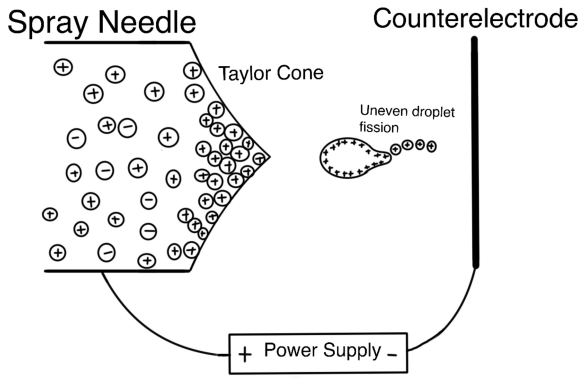


FIGURE 1.8: Representation of the formation of a Taylor Cone at the end of the ESI needle and subsequent formation of ions from larger droplets. Figure redrawn from [144].

1.3.3.2.1 Electrospray Ionisation

In 1989 Fenn et al [143] published research which demonstrated that multiply charged ions could be obtained from proteins applying an electrospray ionisation (ESI) technique. Because the ionisation of analytes can be described as 'soft' and is much more comparable to CI than EI, ESI provides the ability to convert analytes present in a liquid to a gaseous form whilst maintaining the chemical structure of the analyte itself.

ESI works through the application of a high electric field to the analyte mixture passing through a capillary under high temperature. When the mixture arrives at the end of the capillary, due to the large electric field applied to the mixture and the potential difference between the capillary and counter electrode, a charge accumulation on droplets occurs at the end of the capillary. Along with the continuous flow of mixture through the capillary, this droplet (commonly referred to as a Taylor cone) breaks off. The droplets are then pulled towards the MS orifice due to the negative pressure differential between the atmospheric pressure and high vacuum operation of the instrument

Whilst moving towards the MS orifice, temperatures within the source are often elevated much higher than the boiling point of solvents used in the mobile phase. Thus any solvent still present within the droplet evaporates off, and as the volume of droplet decreases the amount of charge per unit volume is increasing. Once droplets have reached a specific size and the charge present per droplet reaches the Rayleigh limit, further and smaller droplets of analytes are ejected eventually leading to a repeating cycle of ejection of smaller and smaller droplets until individual ions are infused directly into the MS.

Typically, efficiency of the ionisation process can be improved through the reduction in size of initial droplets [145]. Therefore the time taken to reach individual analytes in the gaseous phase is reduced. Predictably, the use of what is termed nano-electrospray ionisation (nESI) is often coupled with nano-LC due to the low flow rates involved. This partnership makes nESI and nLC one of the most sensitive MS techniques commercially available from suppliers.

1.3.3.2.2 Atmospheric Pressure Chemical Ionisation

Much like with CI, atmospheric pressure chemical ionisation (APCI) can be viewed as an analogous technique to the use of CI in GC-MS. However as with ESI, APCI can work at atmospheric pressures without the requirement of high vacuum leading to the ability to couple to separation techniques such as LC.

Developed in 1975 by Carroll et al. [146], APCI works through the nebulisation of the analyte mixture alongside a nebulisation gas, typically N_2. The nebulised gas then passes through an environment of increased temperature where mobile phase solvents are allowed to evaporate leaving purely the analytes of interest. After this the analytes passes through an area of reagent molecules produced from a corona discharge which

when at a specific voltage and current produces reagent ions. These reagent ions then react with the analyte molecules, passing their charge based on their proton affinities. After reaction with the reagent ions, the protonated or deprotonated ions in the case of negative ion mode analysis, pass into the MS for analysis.

Both ESI and APCI have broad applicability for use in non-targeted metabolomics workflows, with each technique being more favourable to the ionisation of specific compounds over the other [50]. Therefore the use of a particular technology is often at the researcher's discretion and what instrumentation ultimately is available for their use.

1.3.4 Direct Infusion Mass Spectrometry

Direct infusion mass spectrometry (DIMS) is another technique commonly used in the metabolomic community. The advantages of removing the chromatographic separation step prior to MS analysis allows the high throughput of samples with run times of less than 5 minutes per sample commonly seen within the literature [147, 148]. This reduction in analysis time allows the high throughput of samples with many hundreds able to be analysed each day. For global profiling studies where large numbers of samples are available, then DIMS analysis can direct further research appropriately whilst reducing cost and time taken to perform the research.

Removal of chromatography also introduces a dependant axis that is highly reproducible and based solely upon the mass to charge ratio of ions instead of retention time that is used in LC-MS studies. The retention time of LC studies can be highly variable and is dependant on many properties, many of them physical factors of the column separation itself. Whilst the retention time parameter in LC studies can be reproducible, it can vary widely on depending on conditions. One of the major factors

which alter the reproducibility of retention time is the different mobile phase conditions that are often present between studies spread over multiple batches of mobile phase and different procedures in the preparation of mobile phases between laboratories. With LC, age, solvent composition, dead volume and pump motions all can have effects on the retention time of the resulting separation. This can lead to issues in large scale or multi-batch experiments where the retention time for the same compound can vary leading to annotation or grouping issues for the creation of data matrices.

Two major disadvantages however surround the usage of DIMS. Due to the lack of separation, ion suppression caused by matrix effects can significantly lower the sensitivity of any analyses performed. With all components of the sample being delivered to the ionisation source at the same time, only the compounds which have the highest ionisation efficiency are going to dominate any resultant spectrum. Additionally with instruments which rely on combining 'packets' of ion together before detection (e.g., Orbitrap based systems), this 'packet' of ions once again will be dominated by the most abundant ions and any ions which either struggle to ionise or are low in abundance will not appear in the data.

To reduce the effect of ionisation suppression, reduction of flow rates and the use of nanoESI sources have vastly improved the performance of DIMS techniques in recent years [149]. As mentioned in Section 1.3.2.3, the reduction in flow rate and use of nESI reduces the ionisation suppression that is a major hurdle in non-targeted metabolomic direct analysis techniques. When the flow rate of the infusion into the MS is sufficiently reduced, the ionisation efficiency is increased due to the smaller droplets being formed in the nESI process, thus more analytes are being ionised at any one time.

Currently the commercial leader in DIMS nESI technology is Advion with the Triversa NanoMate system [149]. As a nESI source it is versatile allowing both coupling

FIGURE 1.9: Image of an Advion Triversa NanoMate source coupled to an Orbitrap Elite mass spectrometer

to LC systems and a wide variety of instrument manufacturer's whilst also allowing DIMS experiments using an automated robotic platform.

1.3.4.1 Liquid Extraction Surface Analysis

As part of the Advion instrument, a surface analysis technique is also available. Marketed under the Liquid Extraction Surface Analysis (LESA) brand, LESA has the ability to perform direct surface sampling as part of the DIMS workflow [150, 151]. In recent years LESA has also been coupled to LC to provide additional separation.

The operation behind the LESA technique was first developed in 2005 by Kertesz et al [151] and involves several automated steps to complete a surface extraction. Firstly, the Nanomate robot picks up a conductive pipette tip and aspirates extraction solvent from a pre-selected reservoir. This pipette tip is then moved to a user selected x and y coordinate above the surface of the sample before dispensing of a specific volume of

solvent directly onto the surface of the sample. This dispensation of solvent forms a liquid micro-junction (LMJ) between the sample and the conductive pipette tip which is then allowed to remain on the sample for a selected amount of time (usually a few seconds however this can be user selected based on the application). Following the dwell of solvent on the surface of the sample, the Nanomate instrument respirates the solvent forming the LMJ into the conductive pipette tip and then moves to the nESI chip present in front of the orifice of the MS. An application of voltage is then made to the conductive pipette tip to start the ESI process and a back-pressure applied to the solvent within the tip to start the flow of extracted analytes through the nESI nozzle and into the MS.

The use of LESA within non-targeted metabolomics is limited and most studies involving small molecules are targeted or based on mediums other than collection cards such as glass slides [152, 153]. Additionally, the main application focus of LESA has been heavily towards proteomics with a wide range of research applying LESA to DBS [154–158].

1.3.5 Mass Analysers

The ability to have mass analyser resolutions of greater than 50,000 and beyond at FWHM for *m/z* 200 has been one of the major driving forces behind the expansion of metabolomics, particularly non-targeted research. Compared to LC where resolution defines the ability to separate two peaks eluting at similar times, MS resolution refers to the mass measurement made and its precision to the actual theoretical mass of the ion measured. When resolution is low, such as for single quadrupole mass analysers, typically only nominal mass measurement can be made, limiting its use for the analysis of complex mixtures such as in non-targeted metabolomics.

However at high resolutions as with FT-ICR and Orbitrap based instruments, accurate measurement of molecules which are separated by sub-ppm mass differences is possible, this leads to increased confidence in any subsequent annotations as well as improved sensitivity and number of metabolic features ultimately detected.

After either the separation of metabolites via LC or the direct infusion followed by ionisation, metabolites can then be analysed via a variety of different methods to obtain a mass to charge (m/z) ratio. This m/z ratio is not the actual mass of the metabolite but instead the ionised form which bears a charge due to either it's protonation/deprotonation, or it's combination with a variety of adducts that commonly combine with analytes alongside the charge due to the complex gas phase chemistry that occurs at the temperatures and voltages present in the ionisation process. Additionally, due to the presence of electrical fields, in-source fragmentation can occur based on the desolvation and activation energies that are present throughout the ESI process or additionally in the ion optics present within the instrument. This once more can produce complex fragments which are then subsequently analysed by the MS producing complex spectra.

1.3.5.1 Quadrupole

As mentioned in the previous section, quadrupole mass analysers are typically the lowest resolution of those found in modern instrumentation. Quadrupole mass analysers separate molecules based on their interaction with oscillating electric fields applied to four rods in a circular geometry, where the electric fields are made up of both radio frequency (r.f.) and direct current (d.c.). When an ion enters the quadrupole, through careful selection and oscillation of both the r.f. and d.c. fields, an ion will pass through

towards the detector. Whether or not an ion reaches the detector is based upon a series of equations of motion, which can be described by the Mathieu stability diagram which describes the regions of stability for ions of different masses.

Due to lack of precision in mass measurement, applying single quadrupole mass analysers to metabolomics, particularly non-targeted is deemed as a poor tool for the measurement of the complex mixtures that biofluids are considered to be.

In addition to being used as a mass analyser, quadrupoles can also be used as a mass filter device when operated in r.f. only modes. This mode focuses the trajectory of ions as they pass through the instrument and can be enhanced through the addition of further parallel sets of rods to create hexapole and octapoles.

1.3.5.2 Ion Trap

Like quadrupole mass analysers, ion trap technologies apply an r.f. and d.c. electric field to manipulate and 'trap' ions within a confined region. Ion traps have been around since the 1950s and has since produced two types of analyser, firstly was the 3D ion trap which is commonly referred to as the Paul trap [159] and secondly is the 2D ion trap [160].

The 3D ion trap was developed by Paul and Steinwedel and consisted of a circular electrode with two caps on either end. Ions enter through one end of the caps as one before being subject to an electric field. As this electric field is altered, ions will leave the trap based on their molecular mass before being detected.

A 2D ion trap is much more similar to a quadrupole in geometry and is often referred to as a linear ion trap. Ions will enter a 2D ion trap and will travel towards one end of the linear configuration of 4 rods, before being repelled back towards another end cap. As with 3D ion traps, ions are then ejected from the trap based on their

interaction with a changing electric field.

Due to configuration and method of analysis, both traps experience a phenomenon called 'space charge' effects. 'Space charge' effects occur when a high concentration of charged ions are confined within a region. The ions towards the edge of the 'packet' present within the ion trap then act as a shield to the electric field present, thus changing the voltages at which specific ions are then ejected from the trap [161]. This ultimately has an effect on the total accuracy of the mass measurement made. One of the major benefits of 2D ion traps is the higher capacity of ions that can be present before 'space charge' effects start to have an effect on mass accuracy and therefore are much more sensitive and have a greater dynamic range.

1.3.5.3 Fourier Transform Ion Cyclotron Resonance

Fourier transform ion cyclotron resonance (FT-ICR) mass analysers achieve by far, the highest mass resolving power of any analyser on the market [162]. However the cost of installing and maintaining these instruments often make then prohibitive to many research groups. The main cost behind the installation and maintenance of FT-ICR instrumentation is the superconducting magnets that are required to operate in the region of around 3 to 9.4 Tesla.

FT-ICR works through the trapping of ions in a circular, magnetic trap. Ions are then excited by a fixed magnetic field. For a uniform magnetic field, ions of differing masses will be bent into unique circular paths base on their charged molecular mass which will be perpendicular to the magnetic field. Subsequent measurement of the ion cyclotron frequency is then able to be measured.

After the measurement of the ion cyclotron frequency, a Fourier transform can be applied to convert the intensity measured as a function of frequency, to a measurement

of intensity as a function of time which can then be converted into a mass spectrum.

1.3.5.4 Orbitrap

The Orbitrap mass analyser is the most recently commercially developed mass analyser available on the market. First developed in 2000 by Makarov [163, 164], Orbitrap analysers have been one of the most influential factors in the development and growth of non-targeted metabolomics research. Due to the ability to reach sub parts per million mass accuracy and resolutions of greater than 500,000 [165], Orbitrap mass analysers are now common pieces of instrumentation in most metabolomics research groups.

Because the Orbitrap analyser can be manufactured within a 'bench top' size instrument and the lack of requirement for a superconducting magnet, the requirements for installation within a laboratory are not as burdensome as a FT-ICR, whilst still being able to achieve incredibly high mass resolving power.

Orbitrap analysers work through the injection of packets of ions in-between two electrodes (an inner spindle electrode and an outer electrode) which form to make a spindle shaped device. An electrostatic field is then applied between the two electrodes and kinetic motion is imparted to the ions present within the Orbitrap analyser.

Ions then travel in an oscillating path up and down the length of the inner electrode. The frequency of these oscillations can be described by Equation 1.6.

$$\omega = \sqrt{\frac{z}{m} \times k} \tag{1.6}$$

Where ω is the frequency of harmonic oscillations along the central electrode and k is the instrumental constant. Ions of different masses then take their own trajectories along the inner electrode oscillating at different frequencies depending on their mass.

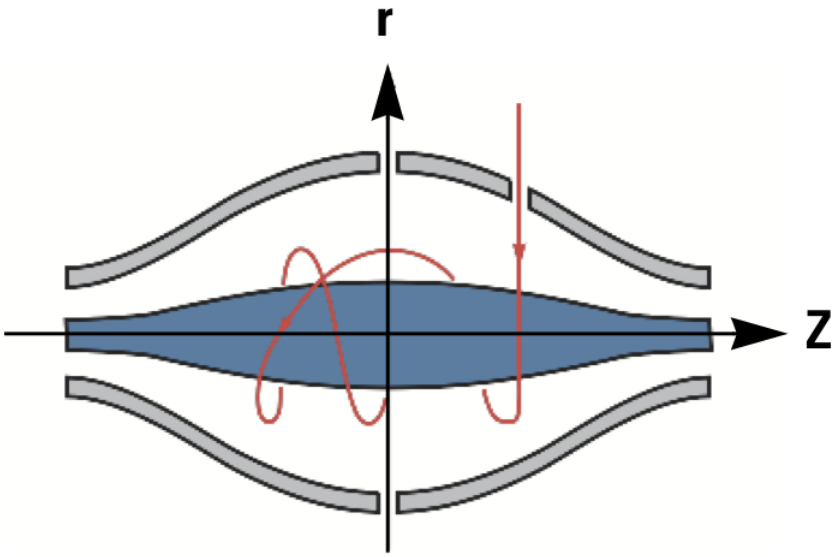
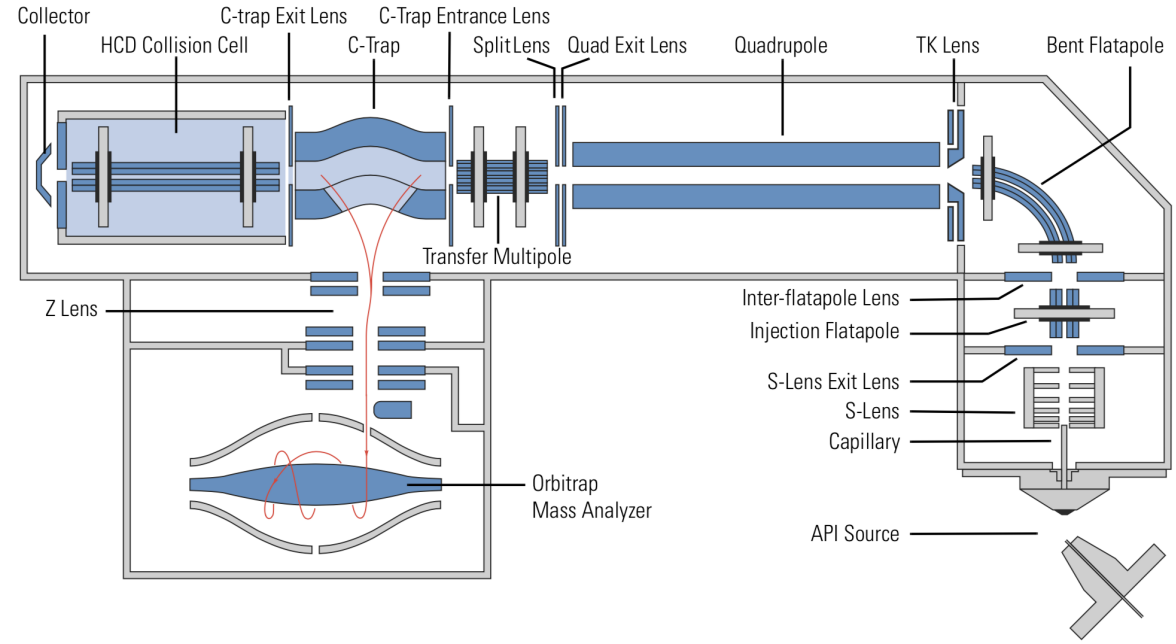


FIGURE 1.10: Schematic diagram of an Orbitrap mass analyser. Reproduced from [166]

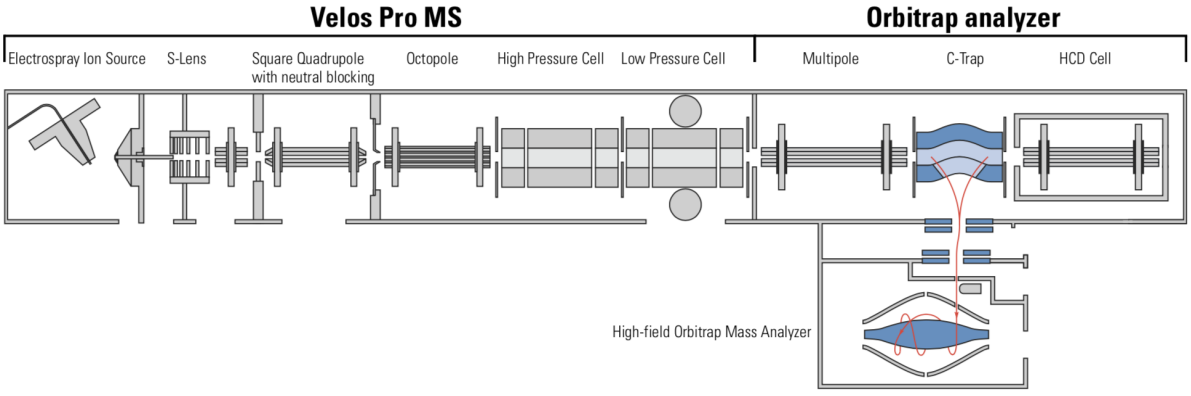
Lighter ions will travel further and oscillate at an increased frequency compared to heavier ions.

Following excitation and oscillation, the current produced by the oscillating ions is then measured and converted via a Fourier transform from the frequency domain to the time domain to produce a mass spectrum.

Another key element of Orbitrap based MS is the 'C-trap'. Ions entering the C-trap collide with nitrogen gas to dissipate any energy before injection into the Orbitrap for mass analysis. The use of a C-trap allows ions to be focused and then introduced into the Orbitrap analyser as a single packet of ions instead of a continuous flow. Additionally, depending on the configuration of the Orbitrap MS system used, ions can then undergo further fragmentation by being passed into a higher-energy collisional dissociation (HCD) cell, passed back into the C-trap again and then the fragmented ions injected into the Orbitrap analyser.



(A)



(B)

FIGURE 1.11: Schematic diagrams of (A) Q-Exactive Plus and (B) Orbitrap Elite mass spectrometers. Figures reproduced from [166]

One of the major advantages that Orbitrap analysers provide is the wide range of configurations which can be manufactured and the ability to couple with other mass analysers such as ion traps to perform simultaneous detection and fragmentation of ions. This flexibility in configurations has allowed Orbitrap based systems to be installed in proteomic based research labs in addition to metabolomics.

Throughout this thesis, all work was conducted on Orbitrap based systems to take advantage of the high mass resolving power whilst maintaining flexibility in the MS experiments and method development that was to be performed.

1.4 Informatics and Data Processing

Performing non-targeted experiments often leads to the measurement of >100,000 individual variables. Therefore the computational need for resources to be able to interpret large amounts of non-targeted metabolomics data is key in ensuring viable conclusions are made.

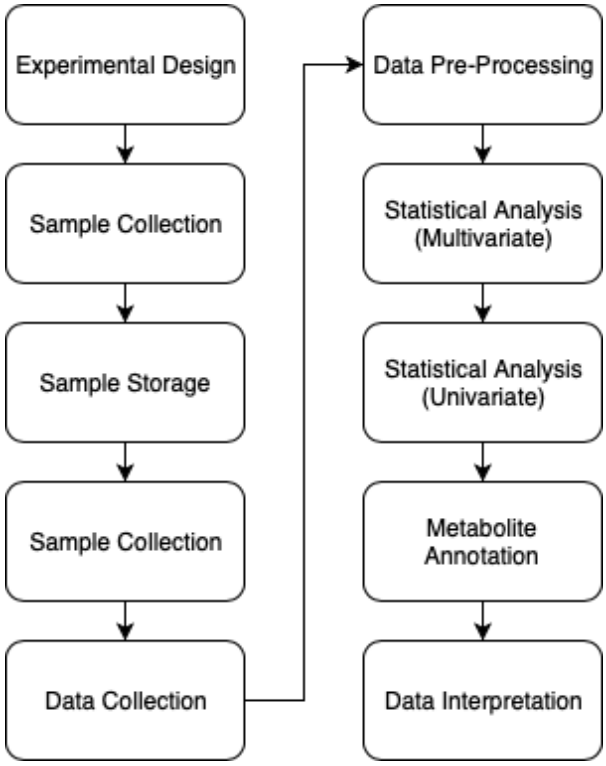


FIGURE 1.12: Typical non-targeted metabolomics experimental workflow

Figure 1.12 describes a typical non-targeted metabolomics workflow. Before data can be statistically analysed however, a data matrix containing the response of each variable (metabolite feature) per sample has to be produced from the raw data that has been collected by the instrument.

1.4.1 Data-processing

To convert raw data into a data matrix there are several open source and closed source programs available to use. Typically the closed source packages are developed directly by the instrument manufacturer and only handle the specific format of data that is output from that specific instrument. Examples of this included Thermo Scientific's Compound Discoverer which only handles Thermo .raw data format. Therefore, the

computational tools and methodologies are generally regarded as commercial secrets and are not publicly available.

On the other side is open source software which generally handles a much wider range of data formats. Currently the most popular open source pre-processing tools are XCMS [167], MZmine [168], with many other smaller tools available such as OpenMS [169], MS-Dial [170] and MAVEN [171].

1.4.1.1 Pre-Processing of UHPLC-MS data

Typically the major function of these open source tools is to convert raw data into a human readable peak matrix upon which statistical analysis can be conducted. This is achieved through the peak picking and grouping of features within samples applying algorithms specialised in the detection of true peaks against background noise. The data matrix output of these programs contains *m/z*-RT pairs corresponding to a singular metabolite feature (or in the case of DIMS data only *m/z*) followed by intensities of each metabolite feature in each sample processed.

This data matrix can then undergo statistical analysis to attempt to pull out features of interest such as those which are changing significantly between biologically different classes.

1.4.1.2 Normalisation and Transformation of Data for Statistical Analysis

Prior to the univariate and multivariate analysis of non-targeted metabolomics data, it is often suitable to apply various different mathematical procedures to ensure the validity of any statistical technique undertaken. Examples of this include the normalisation of data. Due to non-biological variance that often enters the workflow through

sample preparation and instrument analysis, correction of this variate is often necessary to ensure variance within the dataset is stable [172]. Additionally, for multivariate analysis it is preferential that missing value imputation is performed to ensure multivariate statistical methods are valid due to their improved performance with no missing values within the dataset [173]. Finally, data is often transformed and scaled to remove any present heteroscedasticity and to adjust for any significant fold changes between metabolites [174].

1.4.1.3 Pre-Processing of DIMS data

Due to the high degree of ionisation suppression that occurs with DIMS data, typical full-scan analysis modes often have low sensitivity and dynamic range. With hybrid Orbitrap based instruments such as the Orbitrap Elite and Velos Pro, the configuration and hybridisation of an ion-trap and Orbitrap mass analyser, affords the ability to minimise the effect of any ionisation suppression. This is through the utilisation of a method called SIM-stitching.

Developed by Southam et al. [175] SIM-stitching, as the name suggests, stitches multiple SIM-windows together computationally, instead of the analysis of one large mass range. Through the filtering of ions in the ion-trap, only a select mass range enters the Orbitrap for mass analysis. By removing any dominant features that may be present within a specific mass range, both the dynamic range and sensitivity of the instrument can be improved leading to higher quality data. After the collection of data, the multiple SIM 'windows' are then computationally stitched together to produce a single, comprehensive spectrum and subsequent data matrix.

1.4.2 Statistical Analysis

1.4.2.1 Mulitvariate

The large number of variables present within non-targeted metabolomics data leads to the requirement of all observations to be viewed at simultaneously and with relation to one another. Therefore multivariate statistical techniques are commonplace within any non-targeted analysis. One of the most common is the use of principal component analysis (PCA). Due to the large number of variables, there can be viewed to be as many dimensions to the data as there are variables. PCA reduces the dimensionality of the data producing a plot where all variables are condensed into a single representation of the sample it belongs too. A PCA scores plot will have two axes which represent the direction of the data which explains the most variance and a loadings plot which explains how each variable relates to each principal component. [176, 177] Initial conclusions can be made from PCA scores plot due to biologically similar sample grouping closer together than those which are biologically different. Overlaid on top of sample metadata and class information, conclusions can be made as to whether or not there are initial, obvious biological changes happening between two classes of samples.

Whereas PCA can be described as an unsupervised method of multivariate analysis due to the lack of class information present in calculating variance, supervised methods of multivariate analysis are also useful in determining important features which are different between classes. In supervised analysis, class information is present when calculating the model. One example of a method where the class information is present, is partial least squares discriminant analysis (PLS-DA), with the objective to maximise the variance between groups based on the class information present within

each sample [178].

1.4.2.2 Univariate

Applying statistical analyses to each individual feature independent of every other feature present within the dataset can be classed as univariate analyses. Common univariate analyses performed are the t-test and analysis of variance (ANOVA) statistical tests, however these tests are parametric and can only be implemented when certain assumptions about the data are met. In the case of metabolomics data, these assumptions are often broken so non-parametric methods are often applied such as Mann-Whitney U-Test and Kruskall-Wallis tests. [179]

Applying both multivariate and univariate statistical tests allow robust interpretation of the data to be made and any discriminant features that are determined as significant in a multivariate model, can further be tested for statistical significance through univariate tests.

1.4.3 Metabolite Annotation

Following the pre-processing of data to obtain a single data matrix, describing the response of each metabolite feature with each sample, the annotation and identification of data is key in the correct and valid interpretation of data.

Depending on the level of information present within the data and experimental design, correct annotation and identification can be difficult [180, 181]. For example, with DIMS data, there is only mass information present, so it is not possible to separate isobaric compounds that would potentially elute from an LC column at different times to represent two separate compounds and would instead be reported as a single annotation. Additionally complicating procedures is the presence of adducts and

fragments [19], with individual metabolites able to be represented as multiple metabolite features. Therefore attempts should be made to filter out these extraneous features through the grouping of potential mono-isotopic masses and fragments through retention time similarity and response correlation. Additional correlation calculations can be made to check for intensity correlations between two ions. This is because if the two features are not correlated then they may not be two features of the same compound but instead two features of different compounds and should therefore be retained within the data matrix.

Confident annotation of metabolites though can be made through the use of MS/MS and RT data and comparison against chemical standards and with an increasing number of parameters that can be matched against then high confidence in annotations can be made. For example, matching against *m/z* can only provide a list of potential metabolites related to a molecular formula. Through the acquisition of MS/MS data, additional structural information about the metabolite can be gathered and with further RT data, the physiochemical properties of the compound and its interaction with a specific stationary phase can be gathered on an independent orthogonal parameter. Finally, through the acquisition of an authentic chemical standard and comparison against the data described above, high confidence in a specific metabolite annotation can then be made.

Due to the complexity and wide range of annotations that are reported within the literature, the Metabolomics Standards Initiative has defined guidelines and levels which metabolite features can be reported as [182] to help both the researcher and the readers in determining the level of confidence in the data that has been presented. These are the following:

- Level 1 – Identified Compounds

- Level 2 – Putatitvely annotated compounds
- Level 3 – Putatitvely characterised compound classes
- Level 4 – Unknown compounds

Thus annotations that have been made based purely on *m/z* along can be defined as having met Level 3, whereas annotations that have been made based on *m/z*, RT and potentially a spectral match with a public MS/MS library can be Level 2.

1.5 Conclusion

Overall, the impact that high resolution mass spectrometry has had for metabolomics cannot be understated. The ability to detect hundreds to thousands of compounds in a single sample extract has allowed a huge variety of non-targeted applications to be performed that would not have been possible without the recent developments both in separation science and mass analyser technologies.

Applying HRMS to the analysis of DBS opens up the ability to collect samples away from clinical environments to then be sent back to laboratories for processing and analysis. This widens applicable research, but could also allow patients to collect their own samples, allowing greater opportunity for recruitment to large scale population based studies at reduced cost.

However, because of the lack of applications surrounding DBS and non-targeted metabolomics, there are optimisations and recommendations to be made, methods to be developed and applications of which can be used to validate DBS use within non-targeted metabolomics.

1.5.1 Research Objectives

This thesis aims to demonstrate the applicability of DBS in non-targeted metabolomics research as well as define the limitations that may be present when applying DBS collection. Therefore the subsequent four research chapters aim to investigate four separate areas of metabolomics research where the knowledge gaps present currently limit DBS from being applied. These are as follows:

Chapter 2 aims at defining the stability of DBS & DUS over an extended period of time at different temperatures across the whole metabolome applying multiple UHPLC-MS assays. Current research investigating the stability of the metabolome when applying DBS collection is currently limited, with many studies focused on specific areas of the metabolome, or only investigating stability over a short period of time. Additionally, there is very little in the way of published research surrounding the use of DUS within the application of non-targeted metabolomics research.

Therefore this chapter aims to quantify the proportion of the metabolome which could be expected to be stable when performing a long term study which would require knowledge about the behaviour of the metabolome over an extended period of time.

Chapter 3 aims at developing a novel method of normalising for blood volumes that are present when collecting blood spots without any accurate measurement of volume at the time of sample collection. The requirement to normalise spots to collected blood volume is important for the accurate relative quantification of metabolites present within any DBS. Errors in measurement of these metabolites could introduce greater technical variance between individual DBS, thus masking any genuine biological variance between DBS.

Current methods used for the normalisation of blood volume in DBS either employ

expensive technical equipment which may not be present in every laboratory, or destroy the DBS during the measurement of any factor used for normalisation. The assay applied, whilst not novel to the field of quantitative biology, has not been applied for the normalisation of DBS before the publication of this thesis. The method, applying the Bradford assay, is quick to perform and can be performed on left over sample after the initial extraction of metabolites from the sample, therefore gaining information from a sample that is destined for waste.

Consequently, this chapter aims at defining the performance of employing such a method and its ability to normalise samples both prior to the acquisition of data and subsequently after data has been acquired.

Chapter 4 explores the development of a LESA-MS method to directly analyse a DBS without prior need for sample extraction. Whilst LESA-MS has been applied for the analysis of DBS before, it has only been used for the proteomic analysis of DBS. Through the development of a fast and reliable LESA-MS method, the ability to screen samples regularly could become commonplace in scenarios where looking for anomalous levels of metabolites is required such as in areas involving inborn errors of metabolism. In this chapter, different parameters which can be altered on the LESA platform used to observe their effect on the data collected with recommendations given based on their individual performance, with a final method recommendation made for those that wish to apply LESA-MS to their samples.

Finally, Chapter 5 involves the study of two separate cohorts in developing world countries where the excess of Vitamin A in young children's diet is potentially causing harmful side effects. To investigate the effect that excess Vitamin A has on a subjects metabolome, the use of both dried serum spots and dried plasma spots are compared against the current gold standard of metabolomics samples such as serum and plasma.

Through comparing metabolomes which have been exposed to excess Vitamin A levels in two independent populations and comparing the dried filter paper results to the whole biofluid results, similarities and differences between the two samples types can be reported and thus identifying whether there is potential for the application of dried substrates in non-targeted metabolomic workflows.

Chapter 2

Stability of Dried Blood and Urine Spots

Note: Many parts of this chapter have been submitted for publication as part of a peer-reviewed article in Analytical Chemistry. Palmer, E. A; Cooper, H. J.; Dunn, W. B.; Investigation of the 12-month stability and detectable metabolomes of dried blood and urine spots applying untargeted UHPLC-MS metabolomic assays; Analytical Chemistry. Article has been re-submitted following reviewer comments.

2.1 Introduction

The use of dried blood spots (DBS) in non-targeted metabolomics research has advantages over the typical whole blood sample collection and associated sample preparation and processing [183]. Samples can be collected quickly without the need for qualified phlebotomists and the volumes collected are much smaller than those through venipuncture draws.

Potential scenarios for collection of DBS away from clinical settings are therefore possible whether that be in patient homes or in foreign countries where the presence

of laboratory facilities to collect and process whole blood samples may not be present. Additionally, the transportation of DBS samples is considered to be at much lower cost due to the decreased volume of samples and can therefore be sent in an envelope rather than require the use of specialist couriers and packaging to ensure sample integrity [184]. Furthermore, where cost is considered, applying DBS to large scale population based studies represents significant cost savings, in both the collection and storage of many thousands of potential samples required. Therefore DBS collection can be viewed as a future alternative to whole biofluid collection.

However, whilst the application of DBS collection cards within clinical settings has been commonplace for many years with neonates and detecting inborn errors of metabolism, their use for non-targeted metabolomics research is limited however with issues surrounding normalisation, chromatography effects and stability.

The disadvantages associated with varying hematocrit level (correlated to red blood cell count) and chromatographic effects of spotting biofluids onto collection cards (with different rates of diffusion of different metabolites) can cause quantitative issues. Importantly, hematocrit has been normalized through various factors such as potassium content [94], using near infrared (NIR) spectroscopy [185] and hemoglobin measurement using non-contact diffuse reflectance spectroscopy [186].

Chromatography effects [187] can be adjusted for through accurate volume collection[188, 189] (for example, Neoteryx kits) and extraction of the whole biofluid dried spot, though different blood volumes are introduced on to the cards during sampling and standardization of this volume is not possible.

Currently literature surrounding the application of non-targeted metabolomics for the analysis of DBS is limited with papers published limited to investigating the issues preventing the widespread adoption of metabolic phenotyping of DBS. Whilst

the hematocrit variation between individuals and chromatography effect of spotting samples has been widely discussed[190–192], the stability of the metabolites present in a biofluid once dried on filter paper card has not been investigated in detail with most studies focused on a single or a small set of metabolites applying targeted assays [193–195] than a global view of metabolites applying non-targeted metabolomics assays [196–198]. Therefore, current literature that has published which demonstrates the use of metabolic phenotyping when applied to DBS is limited and is mainly focussed on areas where DBS are already used such as inborn errors of metabolism [199], rather than new and novel areas where DBS have yet to be applied. As such, current literature is limited to either targeted metabolomic [200], semi-targeted [198] (low hundreds of metabolites) or shorter term[196] (<1 month) approaches to defining stability.

The study described here presents a 12-month study of metabolite stability in DBS and DUS samples stored at different temperatures (-20°C, +4°C and room temperature (defined from now as +21°C)) compared to plasma and urine samples stored under the same conditions. Two complementary non-targeted ultra high performance liquid chromatography-mass spectrometry (UHPLC-MS) assays were applied to maximize metabolite detectable coverage. The influence of storage temperature on stability at 5 different time points (days 1, 14, 28 and months 6 and 12) were investigated.

2.2 Methods and Materials

2.2.1 Materials

All extraction (HPLC grade Methanol and Water) and UHPLC solvents (Optima grade Water, Isopropanol and Acetonitrile) were purchased from Fisher Scientific (Loughborough, UK). Formic acid ($\geq$98.0% purity) was purchased from VWR International (Lutterworth, UK) and ammonium formate ($\geq$98.0% purity) was purchased from Sigma Aldrich (Poole, UK).

2.2.2 Sample Preparation

A single rat whole blood sample (collected in to Lithium Heparin plasma tubes) and a single rat urine sample (pooled from multiple Wistar Hannover rats) were purchased from Seralab (West Sussex, UK). Both pooled samples were delivered on ice and processed within 36-48 hours of sample collection. Upon delivery, 20 μL of whole blood or urine was pipetted onto Whatman 903 Protein Saver cards (Fisher Scientific, Loughborough, UK) and allowed to dry at room temperature for 3 hours before storage in foil "Ziploc" bags alongside silica-based desiccant (Fisher Scientific, Loughborough, UK).

Plasma and urine samples were prepared from the same pool of whole blood and urine, respectively, by centrifugation (2500g, 4°C, 15 minutes) followed by aliquoting of 20 μL into 2.0mL Eppendorf SafeLock tubes (Scientific Laboratory Supplies Ltd, Hessle, UK).

All samples were stored in triplicates at three different temperatures (temperature controlled at +4°C and -20°C and at room temperature (variable temperature, approximately +21°C). All samples were kept away from light at all times throughout storage to prevent any UV degradation of metabolites.

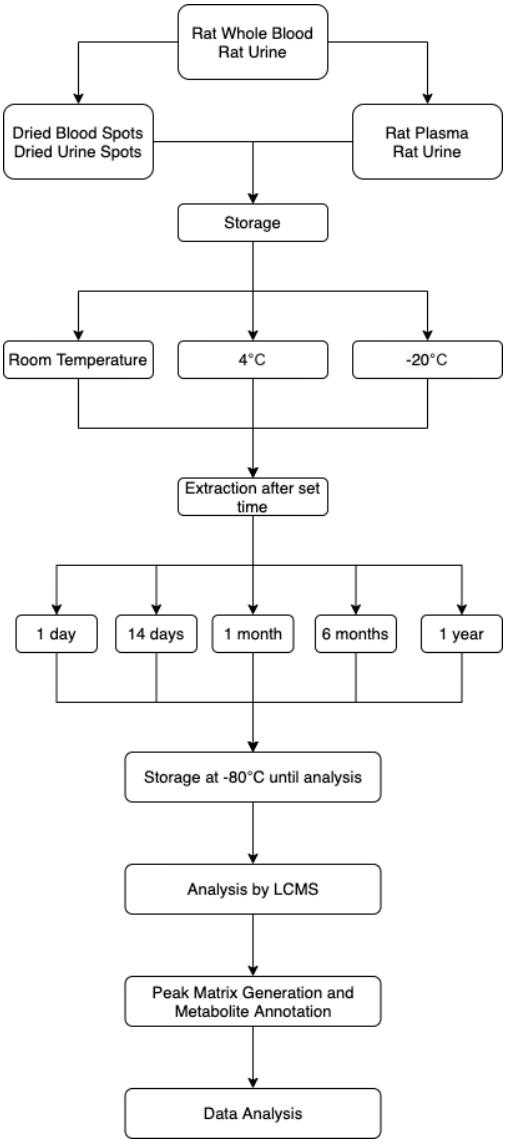


FIGURE 2.1: Flow diagram demonstrating the different conditions which samples were kept in and for how long, including the general experimental design and workflow leading up to the analysis of data.

Figure 2.1 demonstrates the conditions and length of time that samples were stored in. Additional consideration had to be made when deciding whether or not to analyse samples directly after extraction and perform multiple analyses across the year compared to one large batch at the end. However the expected technical variation between instrument condition was expected to be higher than the biological variation, thus potentially introducing difficulties when integrating multiple MS data sets collected across an extended time span. Therefore, the degradation that could potentially occur to metabolites, even when stored at -80°C was considered to be preferable that the expected alternative of large technical variation between experimental batches.

2.2.3 Sample Extraction

At set time points (days 1, 14 and 28 and months 6 and 12), samples stored at -20°C were removed from storage and allowed to defrost on ice for 30 minutes. Samples stored at 4°C and 21°C were removed from storage and placed on ice for 30 minutes. For both DBS and DUS samples, whole spots were removed from the collection card using a sterile scalpel and cut into multiple pieces to aid the extraction process.

DBS samples were extracted in 100 μL of methanol/water (4:1 (v/v)), DUS were extracted in 180 μL of methanol/water (8:1 (v/v)), 20 μL plasma was extracted in 80 μL methanol (assuming 20 μL water volume in plasma) and 20 μL urine was extracted in 160 μL methanol (assuming 20 μL water volume in urine). All DBS and DUS samples were then extracted under gentle agitation using a vortex mixer for 20 minutes before centrifugation (21,000g, 4°C, 15 minutes).

All plasma and urine samples were extracted under agitation for 30 seconds before centrifugation (21,000g, 4°C, 15 minutes). The supernatant for each sample was then separated into two equal volume aliquots before being dried down under vacuum

for 2 hours at 35°C (Savant SPD111V SpeedVac Concentrator connected to a RVT5105 Refridgerated Vapout Trap, Thermo Scientific, UK). After drying, samples were stored at -80°C until analysis by UHPLC-MS after the 12-month storage period.

2.2.4 Ultra High Performance Liquid Chromatography - Mass Spectrometry (UHPLC-MS) Analysis

When ready for analysis, samples were reconstituted in suitable starting solvents (DBS - 80 μL, DUS – 89.6 μL, Plasma - 112 μL, Urine – 142.4 μL) before vortexing and centrifugation (21,000g, 4°C, 15 minutes) and transferring into low recovery LC vials (Chromatography Direct Ltd., Runcorn, UK). A pooled QC sample was then prepared to monitor instrument variability during analytical batches. QC samples were produced from reconstituted samples by aliquoting 10 μL from each sample and pooling together into a 2mL Eppendorf SafeLock tubes (Scientific Laboratory Supplies Ltd, Hessle, UK) and vortexing thoroughly for 5 minutes.

Blank samples were additionally created through the extraction of blank filter paper discs using the extraction methods outlined above. 40 μL of each reconstituted sample was then subsequently aliquoted into low recovery LC vials.

Samples were analysed separately in positive and negative ion modes using a Vanquish UHPLC coupled to an electrospray Q Exactive Plus mass spectrometer (Thermo Scientific, San Jose, CA, US). Data were acquired at a mass resolution of 70,000 (FWHM at *m/z* 200), with the AGC target set at 3e6.

Three separate assays were applied (assay 1, assay 2 and assay 3). Assays 1 and 2 were applied to study DBS and plasma samples and assays 1 and 3 were applied to study DUS and urine samples. Samples to be analysed by assay 1 or 2 were reconstituted in 95% acetonitrile/water (v/v) and 60% acetonitrile/water (v/v) respectively,

and samples to be analysed by assay 3 were reconstituted in 100% water. All assays were 14 minutes in length.

Assay 1 was a hydrophilic interaction chromatography (HILIC) method which used a Thermo Accucore 150 Amide column (2.1 x 100mm, 2.6 μm) with a flow rate of 500 μL/min and with two mobile phases, mobile phase A (10mM ammonium formate in 95% acetonitrile/water (v/v) + 0.1% formic acid) and mobile phase B (10mM ammonium formate in 50% Acetonitrile/water (v/v) + 0.1% formic acid). A 14 minute gradient elution was applied as follows: 0 min, 1% B; 1 min, 1% B; 3 min, 15% B; 6 min, 50% B; 9 min, 95% B; 10 min, 95% B; 10.5min, 1% B. A 4 μL injection volume was applied.

Assay 2 was a lipidomics method which used a Thermo Hypersil GOLD column (2.1 x 100mm, 1.9 μm), with a flow rate of 400 μL/min flow rate and with two mobile phases, mobile phase A (10mM ammonium formate in 60% acetonitrile/water (v/v) + 0.1% formic acid) and mobile phase B (10mm ammonium formate in 90% propan-2-ol/acetonitrile (v/v) + 0.1% formic acid). A 14 minute gradient elution was applied as follows: 0 min, 20% B; 0.5 min, 20%B; 8.5 min, 100% B; 9.5 min, 100% B; 11.5 min, 20% B. A 4 μL injection volume was applied.

Assay 3 was an aqueous C_{18} reversed-phase method which used a Thermo Hypersil GOLD aQ column (2.1 x 100mm, 1.9 μm) with a flow rate of 300 μL/min and with two mobile phases, mobile phase A (0.1% formic acid in water) and mobile phase B (0.1% formic acid in acetonitrile). A 14 minute gradient elution was applied as follows: 0 min, 1% B; 0.5 min, 1% B; 2 min, 50% B; 9 min, 99% B; 10 min, 99% B; 10.5 min, 1% B. A 4 μL injection volume was applied.

Assay 1 applied a *m/z* range of 150-2000 Da, Assay 2 applied a *m/z* range of 70-1050 Da and Assay 3 applied a *m/z* range of 100-1500 Da. In total 8 batches of samples were analysed with each batch representing one biofluid comparison using one assay in a

single polarity mode. DBS and plasma samples were analysed together in the same batch. DUS and urine samples were analysed in the same batch. All batches were randomized across groups as such that 3 replicates of the same biofluid, time point and temperature were measured sequentially within a run to ensure instrument drift did not affect any subsequent precision calculations. Each batch was preceded by 5 leading QC samples followed by an extraction blank sample and an additional 5 QC samples. Batches were then terminated by 2 QC samples followed by a second extraction blank sample. QC samples were analysed after every 6th injection.

2.2.5 Data Processing

Data were analysed applying Compound Discoverer 3.0 with processing parameters shown in Appendix A. (Thermo Fisher Scientific, San Jose, USA). QC filtering of data was performed and compounds with a relative standard deviation >20% for all QC samples (from QC sample nine) were removed from the dataset prior to data analysis. Compounds were retained when showing a RSD of less than 20% in QC samples. Data matrices were exported as .csv files. Noise (peak response <10,000) were reported as Not a Number (NaN). Metabolite features that were detected in <2 of 3 triplicates of the identical sample type, storage temperature and storage time were replaced with a missing value (NaN) for all three replicates.

Univariate statistical analysis (two-way repeated measures ANOVA and t-test) were applied after normalization by sum and glog transformation and were performed using MetaboAnalyst (version 4) [200].

Prior to multivariate analysis, data were normalised by sum and underwent glog transformation and pareto scaling [201]. Missing values were imputed using the k-nearest neighbour method for multivariate analysis. Multivariate statistical analysis

was performed using Metaboanalyst [202]. To assess the blank contribution from the filter cards we also processed the raw data with XCMS [203]. The parameters applied are included in Appendix B.

2.2.6 Metabolite Identification

Metabolite identification was performed by applying the mzCloud & Chemspider Search modules in Compound Discoverer applied MS1 matching to molecular formula and MS2 spectral matching to compounds present in the mzCloud mass spectral database. Additionally, in-house mass spectral annotation software applying MS1 data only were applied to provide further annotations. Accurate mass annotations were made against HMDB, KEGG and LIPIDMAPS databases using a 5ppm mass error window. The workflows produced annotations consistent with reporting level 2 of the Metabolomics Standards Initiative recommendations [182] for annotated metabolites.

2.3 Results and Discussion

UHPLC-MS data were collected and processed to generate four data matrices constructed of metabolites (rows) vs. samples (columns). Metabolites were reported after grouping of all metabolite features in to a single metabolite (or compound as defined in Compound Discoverer). All metabolites demonstrated a relative standard deviation (RSD) of less than 20% and a detection rate >70% for all QC samples following data filtering in Compound Discoverer 3.0.

Principal Components Analysis scores plots for all four combined biofluid-dried biofluid datasets showed compact clustering of replicate injections of QC samples compared to the same single biological sample stored at different temperatures and periods of time and therefore demonstrates suitable data quality to proceed with further analysis of the data (Appendix C).

2.3.1 Metabolite information content of DBS and DUS samples

DBS filter paper cards are expected to contain non-biological chemicals released during sample extraction. The software applied (Compound Discoverer) performed removal of blank-related peaks from the dataset. To investigate whether chemicals were present and extracted from the DBS cards, data were processed with XCMS to define the number of blank-related peaks. These data showed that across all assays applied, 14.7-19.4% of all detected signals were extracted chemicals from the filter paper used in DBS cards. These results highlight the importance of analyzing an extracted blank DBS card to remove detected peaks which are not of biological origin.

Base peak chromatograms for each assay and pair of samples are shown in Appendix D and demonstrate that the biofluid and associated dried sample are visually comparable. Further investigations showed that the majority of metabolites detected in the biofluid (plasma or urine) were also detected in the dried sample (Table 2.1).

One limitation of dried blood and urine spot sample collection is that lower biofluid volumes are typically collected compared to plasma, serum and urine samples and therefore analytical sensitivity can be decreased. Comparisons were made to data acquired from the analysis of plasma applying the same analytical methods but using 60 μL rather than 20 μL biofluid volumes and the results are shown in Table 2.1; 60 μL

TABLE 2.1: Comparison of metabolites detected in dried blood spots, plasma, dried urine spots and urine following extraction of either 20 μL or 60 μL sample volumes. Two assays were applied in positive and negative ion modes to analyse each sample type to maximize metabolite coverage. Reported detected metabolites is before filtering of data related to stability.

Sample Type	Assay	Polarity	Sample Volume Collected (μL)	Number of Metabolites Detected
Dried Blood Spot	HILIC	Positive	20	537
		Negative	20	417
	Lipidomics	Positive	20	2162
		Negative	20	1110
Plasma	HILIC	Positive	20	364
		Negative	20	337
	Lipidomics	Positive	20	1554
		Negative	20	981
Dried Urine Spot	HILIC	Positive	20	2449
		Negative	20	1029
	C_{18} Reversed-phase	Positive	20	4558
		Negative	20	3019
Urine	HILIC	Positive	20	2327
		Negative	20	1022
	C_{18} Reversed-phase	Positive	20	4900
		Negative	20	3043
Plasma	HILIC	Positive	60	1000
		Negative	60	628
	Lipidomics	Positive	60	1889
		Negative	60	1221
Urine	HILIC	Positive	60	2638
		Negative	60	2327
	C_{18} Reversed-phase	Positive	60	3684
		Negative	60	4785

volumes are the standard volume applied in the author's laboratory. For plasma, increased numbers of metabolites are detected when applying the larger sample volume. Increased analytical sensitivity may provide greater metabolite coverage (for example, the use of 1.0mm ID instead of 2.1mm ID analytical columns [182]).

Comparison of data collected for urine showed a mixed set of conclusions; negative ion mode assays showed higher numbers of detected metabolites when the higher sample volume was applied but this was not observed for positive ion mode data in either assay.

TABLE 2.2: The number of metabolites identified as unstable in dried blood spots and plasma derived from rats following storage for 14, 28, 180 and 365 days at three different storage temperatures (-20°C, +4°C and +21°C))

Sample Type	Assay	Polarity	Time Comparison (Days)	Storage Temperature (-20°C)	Storage Temperature (+4°C)	Storage Temperature (+21°C)
DBS	HILIC	Positive	1 vs. 14	0 (0.0%)	0 (0.0%)	2 (<0.1%)
			1 vs. 28	5 (0.1%)	64 (6.0%)	260 (24.4%)
			1 vs. 180	0 (0.0%)	237 (22.3%)	394 (37.0%)
			1 vs. 365	67 (6.3%)	249 (23.4%)	503 (47.3%)
Plasma			1 vs. 14	112 (10.5%)	16 (1.5%)	385 (36.2%)
			1 vs. 28	25 (2.3%)	155 (14.6%)	475 (44.6%)
			1 vs. 180	553 (52.0%)	588 (55.3%)	698 (65.6%)
			1 vs. 365	690 (64.8%)	688 (64.7%)	806 (75.8%)
DBS		Negative	1 vs. 14	0 (0.0%)	0 (0.0%)	0 (0.0%)
			1 vs. 28	0 (0.0%)	0 (0.0%)	4 (0.5%)
			1 vs. 180	12 (1.6%)	6 (0.8%)	215 (29.4%)
			1 vs. 365	5 (0.7%)	79 (10.8%)	506 (69.4%)
Plasma			1 vs. 14	0 (0.0%)	37 (5.1%)	168 (23.0%)
			1 vs. 28	4 (0.5%)	0 (0.0%)	227 (31.1%)
			1 vs. 180	170 (23.3%)	177 (24.3%)	233 (32.0%)
			1 vs. 365	391 (53.6%)	176 (24.1%)	520 (71.3%)
DBS	Lipidomics	Positive	1 vs. 14	0 (0.0%)	75 (1.9%)	0 (0.0%)
			1 vs. 28	0 (0.0%)	2 (0.1%)	24 (0.3%)
			1 vs. 180	0 (0.0%)	216 (5.5%)	549 (13.9%)
			1 vs. 365	-	-	-
Plasma			1 vs. 14	0 (0.0%)	9 (0.2%)	2979 (75.6%)
			1 vs. 28	0 (0.0%)	147 (3.7%)	2948 (74.8%)
			1 vs. 180	483 (12.3%)	2443 (62.0%)	3264 (82.9%)
			1 vs. 365	1196 (30.4%)	2586 (65.7%)	3276 (83.2%)
DBS		Negative	1 vs. 14	0 (0.0%)	2 (0.2%)	0 (0.0%)
			1 vs. 28	0 (0.0%)	5 (0.4%)	12 (0.9%)
			1 vs. 180	2 (0.2%)	30 (2.4%)	86 (6.8%)
			-	-	-	-
Plasma			1 vs. 14	0 (0.0%)	22 (1.7%)	61 (4.8%)
			1 vs. 28	0 (0.0%)	69 (5.4%)	962 (75.6%)
			1 vs. 180	329 (25.8%)	735 (57.7%)	1027 (80.7%)
			1 vs. 365	508 (39.9%)	667 (52.4%)	1079 (84.8%)

TABLE 2.3: The number of metabolites identified as unstable in dried urine spots and urine derived from rats following stor-age for 14, 28, 180 and 365 days at three different storage temperatures (-20°C, +4°C and +21°C)

Sample Type	Assay	Polarity	Time Comparison (Days)	Storage Temperature (-20AC)	Storage Temperature (+4AC)	Storage Temperature (+21AC)
DUS	HILIC	Positive	1 vs. 14	0 (0.0%)	0 (0.0%)	5 (0.2%)
			1 vs. 28	0 (0.0%)	5 (0.2%)	312 (9.7%)
			1 vs. 180	0 (0.0%)	127 (4.0%)	1807 (56.4%)
			1 vs. 365	2 (0.1%)	486 (15.2%)	1272 (39.7%)
Urine			1 vs. 14	0 (0.0%)	8 (0.2%)	1162 (36.3%)
			1 vs. 28	0 (0.0%)	65 (2.0%)	1544 (48.2%)
			1 vs. 180	97 (3.0%)	796 (24.8%)	1737 (54.2%)
			1 vs. 365	193 (6.0%)	1037 (32.4%)	2096 (65.4%)
DUS		Negative	1 vs. 14	0 (0.0%)	0 (0.0%)	0 (0.0%)
			1 vs. 28	0 (0.0%)	0 (0.0%)	67 (6.0%)
			1 vs. 180	0 (0.0%)	0 (0.0%)	232 (20.9%)
			1 vs. 365	5 (0.5%)	286 (25.7%)	419 (37.7%)
Urine			1 vs. 14	0 (0.0%)	1 (0.1%)	920 (82.8%)
			1 vs. 28	0 (0.0%)	0 (0.0%)	550 (49.5%)
			1 vs. 180	10 (0.9%)	227 (20.4%)	877 (78.9%)
			1 vs. 365	86 (7.7%)	370 (33.3%)	889 (80.0%)
DUS	Reversed-phase	Positive	1 vs. 14	3 (<0.1%)	0 (0.0%)	0 (0.0%)
			1 vs. 28	0 (0.0%)	0 (0.0%)	48 (0.8%)
			1 vs. 180	0 (0.0%)	0 (0.0%)	2335 (37.8%)
			1 vs. 365	0 (0.0%)	609 (9.9%)	1488 (24.1%)
Urine			1 vs. 14	0 (0.0%)	15 (0.2%)	1500 (24.3%)
			1 vs. 28	1 (<0.1%)	1 (<0.1%)	3643 (59.0%)
			1 vs. 180	1 (<0.1%)	1250 (20.2%)	3680 (59.6%)
			1 vs. 365	0 (0.0%)	1402 (22.7%)	4080 (66.1%)
DUS		Negative	1 vs. 14	1 (0.0%)	2 (0.0%)	2 (0.0%)
			1 vs. 28	0 (0.0%)	0 (0.0%)	16 (0.4%)
			1 vs. 180	1 (0.0%)	58 (1.6%)	801 (21.7%)
			1 vs. 365	4 (0.1%)	381 (10.3%)	981 (26.5%)
Urine			1 vs. 14	0 (0.0%)	0 (0.0%)	2143 (58.0%)
			1 vs. 28	0 (0.0%)	24 (0.6%)	2973 (80.5%)
			1 vs. 180	10 (0.3%)	722 (19.5%)	2794 (78.6%)
			1 vs. 365	6 (0.2%)	860 (23.3%)	2856 (77.3%)

For urine, no significant number of additional metabolites were detected in the biofluid compared to the dried sample and vice versa (Table 2.1). However, and as expected, additional metabolites were detected in the DBS samples with the additional metabolites suggested to originate from the red and white blood cells present in whole blood and which are removed using centrifugation in the preparation of plasma (Table 2.1). 23.7 and 47.5% of additional metabolites were detected in the HILIC negative and positive ion modes, respectively. 13.1 and 39.1% of additional metabolites were detected in the lipidomics negative and positive ion modes, respectively.

The annotated metabolites detected in DBS and not in plasma are comprised of five major enriched classes of acyl carnitines (21 metabolites), aromatic metabolites (18 metabolites), fatty acids/esters/oxidized fatty acids (34 metabolites), glycerophospholipids (42 metabolites) and lysoglycerophosphoserines (12 metabolites). The fatty acids and (lyso)glycerophospholipids suggest detection of lipid components of cell membranes and potentially lipoproteins. Acyl carnitines shuttle long chain fatty acids across the inner mitochondrial membrane and so can be expected to be cell-derived. The exclusion of glycolytic intermediates suggest that either (i) the whole blood volume used does not allow detection of many metabolites expected to be present only within the cell, (ii) that cells in DBS samples are not lysed during storage or extraction and therefore metabolites are not released from the cell for analysis or (iii) the metabolites are not present at any concentration or are present but at a concentration below the assay's limit of detection.

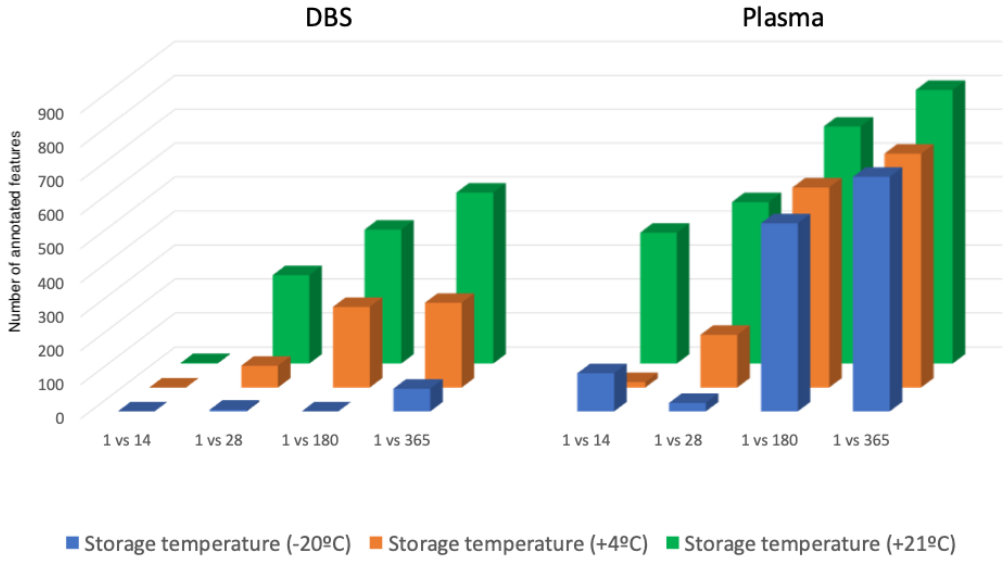


FIGURE 2.2: Plot demonstrating the number of statistically significant annotated features at each time point and temperature. Plot represents the HILIC positive results for the DBS and plasma comparison. All other plots can be found in Appendix E

2.3.2 Stability of dried blood and dried urine spots stored at different temperatures

We performed t-test analyses to compare day 1 data to each other time point separately for each storage temperature and UHPLC-MS assay. The results are shown in Table 2.2 for DBS/plasma and Table 2.3 for DUS/urine. A critical p-value of 0.1 (following correction for multiple testing using Benjamin Hochberg method) and a relative change of greater than 20% was applied to define a metabolite which was identified as being unstable or whose relative concentration changed for another reason. Visual comparison of the data described in Tables 2.2 and 2.3 are included in Appendix E with an example shown in Figure 2.3.

A lowest storage temperature of -20°C was chosen to study as previous studies

have assessed storage at -80°C and demonstrated this as an appropriate storage temperature [198, 204]. The results showed that DBS samples were stable at -20°C over the first six months (less than 1.6% of detected metabolites showed a change in relative concentration). Some additional instability is observed between months 6 and 12 (in total less than 6.3% of detected metabolites showed a change in relative concentration). We therefore conclude that collection and immediate storage of dried blood spots at -20°C provides a stable and representative sample for metabolomics studies for a minimum of 12 months and for the majority (but not all) metabolites. DBS samples showed a greater level of instability if stored at 4°C or room temperature for 6 or 12 months, with 2.4-23.4% of metabolites showing a relative concentration change depending on the UHPLC-MS assay. These results validate other short-term studies of up to 1 month applying non-targeted and targeted assays and long-term studies of up to 2 years applying a targeted assay for detection of amino acids and acyl carnitines.

These studies demonstrated short-term stability at room temperature but increased instability with storage at room temperature over time and recommendations for storage at -20°C [195, 198, 204, 205]. Importantly, storage at -20°C and -80°C across 1 month has shown similar levels of stability and so samples do not necessarily have to be stored at -80°C [205]. In the study reported here, more metabolites were unstable in the HILIC assay data compared to the lipidomic dataset suggesting changes in polar metabolites. We therefore recommend that DBS samples can be collected and transported at room temperature for up to 28 days but recommend storage at -20°C or -80°C for longer term storage. DBS samples were more stable than plasma samples at all storage temperature and time comparisons investigated in this study. Metabolites

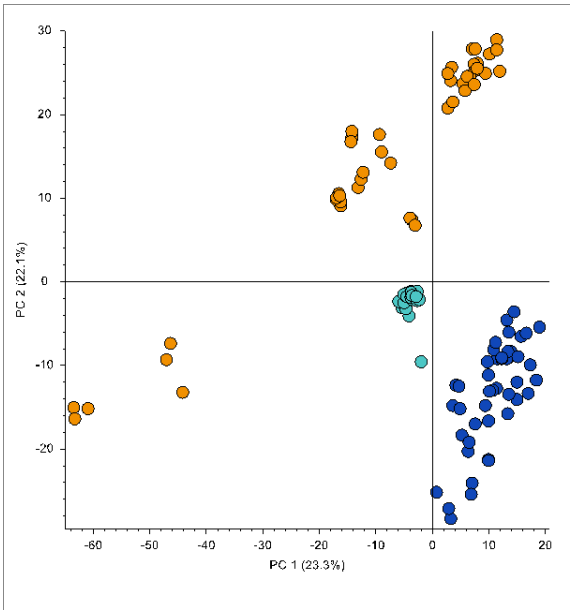


FIGURE 2.3: PCA plot demonstrating the separation and clustering of plasma and dried blood spot samples respectively. Plot represents the HILIC positive results for the DBS and plasma comparison. All other plots can be found in Appendix E

which showed instability within the first 28 days after dried blood spot collection included sphingosine, vitamins A and E and tryptophan/phenylalanine-related metabolites with lower relative concentrations at 28 days compared to 1 day storage at room temperature. (This data and following conclusions related to this data can be found in Appendix F.) Increased proteolysis was observed also with many peptides detected at higher concentrations at day 28 compared to day 1 storage at room temperature.

Longer term instability at 6 and 12 months was observed in multiple metabolite classes including primarily glycerophospholipids, lysoglycerophospholipids and acyl carnitines which all showed large decreases in their concentrations at day 365 compared to day 1 storage at room temperature. Decreased concentrations of (lyso)glycerophospholipids and acyl carnitines does not appear to be caused by their enzymatic degradation and release of fatty acids because increased concentrations of

fatty acids were not observed. The decreased concentrations may be caused by non-enzymatic degradation which releases other chemicals and products which are not fatty acids or may be caused by continued drying and absorption on to the filter paper over 12 months with the use of methanol as an extraction solvent not providing the release of glycerophospholipids from the filter paper. Further studies are required to confirm this though a similar number of glycerophospholipids were detected in this study which showed no relative change in concentration which suggests that poor extraction efficiency is not the cause.

Increased proteolysis was observed also with many peptides detected at higher concentrations at day 28 compared to day 1 storage at room temperature. We expected to see significant increases in lipid oxidation but only a small percentage of oxidized lipids were detected applying our UHPLC-MS assays and only a small number of those detected showed changes in relative concentration. Oxidised lipids may be present but at concentrations lower than were detectable applying the volumes and assays applied in this study. The storage of samples in a sealed dark bag with a desiccant may also help to protect lipids from oxidation but further studies would be required to test this hypothesis.

A similar conclusion can be derived for stability of DUS samples as was discussed for DBS samples above. Less than 0.5% of detected metabolites were unstable if stored at -20°C for up to 365 days compared to urine which demonstrated up to 80% of metabolites to be unstable. Storage of DUS at 4°C for up to 180 days showed less than 4% of metabolites to be unstable and if extended to 365 days showed less than 23.4% of metabolites to be unstable. If stored at room temperature then up to 10% of metabolites are unstable after 28 days and up to 40% of metabolites are unstable up to 365 days. We therefore recommend that DUS samples can be collected and transported

at room temperature for up to 28 days but recommend storage at -20°C or -80°C for longer term storage.

DUS samples were more stable than urine samples at all storage temperatures and time comparisons investigated in this study. Dopamine showed a 500-fold decrease in relative concentration after 12 months compared to day 1 DUS samples. Vitamin C (ascorbate metabolites showed lower relative concentrations at 365 days compared to day 1 also. Of interest was that a large number of metabolites were stable when present in plasma and urine at +21°C for 12 months. Both of the biofluids studied were acellular and therefore no cellular metabolism can be expected to be occurring as both biofluids do not contain enzymes at a high concentration and therefore lack of metabolism will help maintain the stability in metabolite concentrations across the 12 months. Instability of metabolites can be proposed to occur through thermal degradation.

This study has provided a broad panel of detected metabolites but shows limitations. Firstly, we applied whole blood and urine samples which had been stored on ice for 36-48 hours before being processed and spotted on to DBS cards. This storage period could allow for degradation of highly labile metabolites (e.g. glutathione and ATP) and bias the results though the cellular metabolic flux would be much lower than observed at a body temperature of 37°C. Secondly, we applied one pooled sample for each biofluid and did not assess across multiple rats and other mammalian species. Thirdly, the precision measurement used to define whether a metabolite was unstable was +/-20% which is very conservative because biological variability between different subjects in serum and urine is typically much higher (see [206, 207] for examples). Finally, we used a single extraction solution for both UHPLC-MS assays applied to each sample type and therefore detection of all lipids was not possible (for example,

only a small number of triacylglycerides and cardiolipins were detected).

2.4 Conclusion

The study presented has shown that DBS samples and DUS samples can provide broad metabolome coverage and therefore offer an alternative to collection of whole blood, plasma, serum and urine away from a clinic and where specialized staff and laboratories are not available. Broad metabolome coverage was observed with 20 μL of biofluid which was subsequently dried compared to a typically used volume of 60 μL which is applied in the author's and others laboratories. However, when applying standard UHPLC-MS assays, fewer metabolites were detected when compared to the analysis of standard volumes of biofluid, thus development of an assay providing increased analytical sensitivity would be appropriate (for example, using 1.0mm ID or narrower UHPLC columns or through sample enrichment).

Collection, transport and short term storage of dried blood and urine spot samples at room temperature is appropriate and provides only a small number of metabolites where relative concentration changes were observed. However, instability increased across time for the two higher storage temperatures and long term storage at room temperature is not recommended, instead storage at -20°C or -80°C being recommended.

The study has provided a general recommendation for temperatures for collection, transport and short-term and long-term storage. However careful design of experiments and thought should be undertaken to ensure that using cellulose based collection cards are suitable for the experiment to be considered. This is due to the still high proportion of metabolites that have unacceptably high RSD's, when just a reduction in storage temperature for the whole biofluid can outweigh the benefit of using

DBS and DUS. This could especially be the case for small scale local studies where the logistical effort of collecting DBS/DUS samples will not necessary be financially or scientifically beneficial. However, benefits are clear for larger scale studies where samples are required to be sent through local postal mail services and the logistical challenge of recruiting subjects to travel to clinics for whole blood sample collection can be challenging. The use of filter paper collection cards within this scenario could provide a large cost saving, and the benefits of being able to recruit larger cohort sizes can potentially increase the statistical power of studies which can be a major issue in metabolomic studies.

Although a large number of blank features were present within the datasets, due to the way in which DBS and DUS are collected, direct comparison against blank filter paper from the same card can also be made, thus removing any potential interfering features that will inevitably be extracted directly from the collection card. This is not necessarily possible with biofluid samples where samples may be sent between collaborators where samples have been processed using multiple tubes and unused tubes are not normally available to determine their contribution to the metabolic profile acquired.

Whilst DBS samples have a much greater use within clinical research, DUS also have been demonstrated to provide an alternative means of sample collection, with appropriate stability demonstrated at room temperatures, once again providing suitable processes are in place for samples to be placed into long term storage at reduced temperatures at a later date. Once again, with smaller volumes required and the ease of transportation and storage, the use of DUS afford the same advantages as DBS whilst also removing the minimal invasive collection procedure that DBS require. Additionally, due to the large volumes of urine that are collected in a single sample, many more

replicates can be made, whilst also allowing accurate volume samples to be prepared, thus removing volume related variance issues that are present in with the collection of DBS samples.

Further work that could be conducted would be looking at the effect that different collection devices (such as Neoteryx's Mitra and Spot On Science's HemaSpot collection devices) also has on the stability of samples. Furthermore, additional semi-targeted research could be conducted applying Biocrates p180 and p400 kits to identify stability of a smaller, more select number of metabolites compared to the number measured in a non-targeted study.

Studies also looking at the short term stability of metabolites throughout the drying process would also be of interest, with potential implications on the storage of DBS and DUS whilst samples are being prepared. With typical procedures relying on the drying of samples at ambient temperatures for multiple hours, there will no doubt be biological changes occurring much faster than the time scale that this research observes. Therefore it would be of interest to identify these changes to further understand the limitations of DBS collection.

Overall DBS and DUS will remain popular choices for the collection of biofluid samples in targeted studies for the foreseeable future, and with the development of techniques aimed at reducing the quantitative effects of storing blood and urine as dried specimens, they will become of greater use in non-targeted settings and in the development of personal health monitoring scenarios.

Chapter 3

Applying the use of a total protein content assay to normalise for variation in dried blood spot volume

3.1 Introduction

Non-targeted metabolomics data contains both non-biological and biological sources of variation. Ensuring that these sources of variation have minimal effects on the data is important in the valid interpretation of any analysis performed. Typically, normalisation to correct for this variance can occur prior to the analysis of samples, through alteration of reconstitution volumes, or computationally afterwards applying calculations to every measurement made.

Sources of biological variation can arise from many factors which can be difficult to account for, even in the most controlled experimental settings. The most common sources of biological variation within mammalian systems include the age and sex of the subject in which the sample has been taken from. Due to the different biochemical processes that occur between different sexes and subjects of different ages, these

ultimately effect the levels of metabolites which are subsequently analysed in any non-targeted study. Additionally, the fed and fasted states of subjects can have large influences on specific metabolites. Therefore careful design of experiments are required in ensuring that large sources of biological variation that are unrelated to potential class information are not introduced.

Non-biological variation occurs after the sample has been collected. This can be introduced through poor liquid handling techniques or instrument setup procedures which subsequently effect any measurement made. Careful sample handling and preparation is therefore required in addition to proper standard operating procedures in setting up and calibrating instruments. Any failure in the above can drastically effect results obtained from incorrectly calibrated instruments to large drops in sensitivity based on poor sample preparation procedures.

A common technique in correcting for biological variation is normalising data based on the measurement of a physiological paramter. Normalisation of metabolomics data based on a physiological measurement within the sample itself has been commonplace within many non-targeted and targeted metabolomics applications for many years [208, 209]. One of the most common biofluids that is corrected for is urine. Due to the large variation in volume of urine that is collected day to day, it is necessary to normalise data to ensure integrity. This variation can be due to many factors including simply the hydration status of the subject donating the sample to the kidney function and its ability to excrete urine [210]. Therefore there has been extensive research performed in the area of methods to normalise urine after the donation of the sample but before acquisition itself.

More straight forward is the normalisation of biofluid samples which can be seen as having low biological variability within the donor subject itself such as serum or

plasma. Due to the low variation in samples collected simply based on basic physiological health such as weight or hydration status, these types of samples are not normally normalised prior to data acquisition, and instead normalisation typically occurs post acquisition of data [211–214]. This is to correct for any inherent biases in the analytical collection of the data instead of the biologically variation within the sample. This could either be arising from instrument drift or from biases occurring throughout sample preparation. As with normalising urine prior to the collection of data, normalising post acquisition has been the focus of many researchers and can be computationally achieved through multiple techniques such as through probabilistic quotient normalisation [211] or normalisation by sum [**Xia2011xx**], mean [**martuccixx**] or median [214] of metabolite responses across a sample as a normalisation factor [201].

In addition to serum/plasma and urine based metabolomics experiments, studies involving tissues also require the use of normalisation to ensure that the mass differences that will arise due to the collection of different samples, has no overall effect on the outcome of the experiment. Typically, samples are normalised via their weight and then reconstituted in varying amounts of solvent depending on the overall weight of the sample itself [215].

Both methods of normalisation, whether it is normalising prior to data acquisition, or computationally, after the data has been collected, are key aspects of the experimental design to ensure that any external non biological variability within the samples is controlled and any genuine biological variation between sample classes is not masked and can be interpreted correctly to ensure either a valid hypothesis being generated or validated.

Compared to the techniques presented above, if a normalisation procedure were to be applied to DBS, it would easily be thought of as most convenient to use a method

which corrects for the total volume of blood that has been spotted onto the card, as per procedures undertaken when normalising for tissues. However, if DBS are collected directly onto a collection card with no prior measurement of blood volume as is very typical in clinical and research settings, then this is not possible, and the volume of blood actually deposited onto the card is unknown. An additional complication with DBS, is the heamatocrit level of each individual patient. Due to varying haematocrit levels, the volume of blood within a DBS is not proportional to its size on the collection card. Ultimately this means that spots of higher haematocrit demonstrate a smaller radius of DBS compared to those with lower haematocrit levels. Additionally, this also means if fixed size hole punches are taken from spots, then different volumes of blood are being extracted from sample to sample.

There are volume specific collection devices within the field of collecting whole blood [216–220], however in terms of the popularity of filter paper based collection cards both in regard to the minimal cost in being able to allow patients to collect their own samples, to the many hundreds of thousands of samples that have already been collected and stored in biobanks around the world, these devices are not going to solve the issue of volume variation for studies which wish to look at these samples directly.

Methods have been published which demonstrate the ability to measure several characteristics of DBS of which can then be normalised against [191, 221, 222], however many of these methods either require spare samples due to the destructive nature of the test or expensive instruments in order to characterise various aspects.

This chapter aims at demonstrating a cheap, easy, quick and non-destructive method for the determination of total protein content within a DBS in order to normalise against varying blood volumes that will be present from sample to sample. To measure the total protein content of a sample, a Bradford assay was used which fulfils

all the criteria for a suitable test.

In order to test the validity and effectiveness of any normalisation that has been performed, various performance metrics will be used to identify any improvement in the quality of the data. This includes monitoring the range of RSD's seen across all detected metabolites, the total number of features which demonstrate an RSD of less than 30%, the sum of all intensities for all features as well as the total number of significant features between different volumes.

Through the monitoring of these features a recommendation can be made regarding the usefulness and effectiveness of performing a Bradford Assay on DBS as a means to normalise for differing total blood volumes on each spot.

3.2 Materials and Methods

3.2.1 Materials

HPLC Grade methanol, water, acetonitrile and isopropanol were purchased from VWR Chemicals. Rat whole blood (Pooled gender, Wistar Hannover strain) was purchased from BioIVT. Whole blood was collected in lithium heparin collection tubes to prevent coagulation before preparation of DBS. Time between collection of whole blood and preparation of DBS was less than 48 hours. Whatman 903 filter paper collection cards, dessicant and foil ziploc bags were purchased from Fisher Scientific (Loughborough, UK). 2M hydrochloric acid, bovine serum albumin and Coomassie Brilliant Blue G (CBB-G) were purhcased from Merck (Poole, UK). Deionised water (d.H_2O) was prepared in laboratory using an Elix Essential 10 UV water purification system (Merck, UK). .

3.2.2 Sample Preparation

DBS samples were prepared by pipetting accurate volumes of rat whole blood onto Whatman 903 Filter Paper based collection Cards. Cards were then allowed to dry for three hours at ambient temperature before storage at -80°C until analysis. In total five different volumes of blood were deposited onto DBS collection cards (10 μL, 15 μL, 20 μL, 25 μL, 30 μL).

3.2.3 Metabolite Extraction

Samples were removed from storage and allowed to defrost on ice for 30 minutes prior to extraction. Samples were then incised from the collection card using a sterile scalpel and cut into smaller sizes to aid in extraction by increasing surface area of the spot. Samples were then extracted in 500 μL of methanol under gentle agitation using a vortex mixer for 20 minutes before centrifugation (21000 g, 4 °C, 20 minutes). Supernatant from each sample was then transferred into two separate aliquots (150 μL each) before being dried down under vacuum for 40 minutes at 35 °C. After drying samples were stored at -80 °C until analysis by UHPLC-MS.

3.2.4 Extraction of DBS protein

After aliquots had been taken for the metabolomics analysis, the remainder of the sample (including the DBS as prepared in Section 3.2.3) was dried down under vacuum for 30 minutes at 35 °C to remove any excess organic solvent. Subsequently, 1 mL of deionised water was then added to the DBS sample and mixed under high agitation using a vortex mixer followed by centrifugation (21000 g, 4 °C, 5 minutes). A 125 μL

aliquot was then taken and diluted with 875 μL of deionised water and mixed thoroughly. Another 125 μL was then taken in duplicate and added to a 96 well plate ready for total protein content determination.

3.2.5 Preparation of Bradford Assay standard curve

The protocol used follows a modified version of [223]. Stock solution of 1 mg/ml of BSA was prepared and subsequently diluted into the following concentrations with deionised water, 100 μg/ml, 70 μg/ml, 60 μg/ml, 50 μg/ml, 40 μg/ml, 35 μg/ml, 25 μg/ml, 16 μg/ml, 10 μg/ml, 8 μg/ml, 6 μg/ml, 4 μg/ml, 2 μg/ml, 1 μg/ml. 125 μL of each concentration was then prepared in triplicate into a 96 well plate alongside samples.

3.2.6 Preparation of CBB-G reagent and total protein content analysis

A CBB-G stock solution was prepared by dissolving 60 mg of CBB-G in 100 mL of 2M HCl and stirring for 40 minutes before being filtered using a 0.2 μm syring filter to ensure a particulate free solution. A CBB-G working solution is then prepared by mixing together d.H_2O, CBB-G stock and 2M HCl (3:1:1, v/v/v). After preparation, 125 μL of CBB-G working solution is then added to each sample in the 96 well plate. The 96 well plate was then incubated for 10 minutes at room temperature before analysis using a TECAN Infinite F200 Pro microplate reader running TECAN iControl (v1.10) software applying an absorbance wavelength of 595 nm.

3.2.7 Reconstitution of Samples for Metabolomics Analysis

3.2.7.1 Post-Data Acquisition Normalisation

Samples to be normalised post data acquisition were all resuspended in 56 μL of appropriate solvent depending on the LC assay performed. For samples analysed applying a lipidomics assay this was a mixture of 75% IPA and 25% water. For samples analysed applying a HILIC assay this was a mixture of 95% acetonitrile and 5% water.

3.2.7.2 Pre-Data Acquisition Normalisation

Samples which were normalised pre-data acquisition, were resuspended according to the total protein content measured in Section 3.2.6. Volumes samples are resuspended in are shown in Table 3.1.

3.2.8 UHPLC-MS Analysis

Samples were analysed in both positive and negative ion modes applying both a HILIC and lipidomics method using a Thermo Vanquish UHPLC coupled to a Thermo Q-Exactive Plus MS (Thermo Scientific, San Jose, CA).

A pooled QC sample was prepared from the extracts of all samples (including samples which were normalised both pre and post data acquisition) and then prepared in the same way as all samples before being injected every 8th sample with 10 leading QC's to condition the column. Both LC methods were 14 minutes in length and data was collected at mass resolution 70,000 resolution (FWHM at m/z 200). The HILIC method used an Accucore 150 Amide column (100 mm x 2.1 mm i.d., 1.9μm pore size). Solvent A was acetonitrile/water/formic acid (95/4.9/0.1, v/v/v) with 10 mM ammonium formate. Solvent B was acetonitrile/water/formic acid (50/49.9/0.1, v/v/v)

TABLE 3.1: Reconstitution volumes (μL) for each sample based on the method of normalisation.

Volume	Replicate	Pre Data Acquistion (μL)	Post Data Acquisition (μL)
10	A	56	56
	B	62	56
	C	59	56
	D	60	56
15	A	109	56
	B	117	56
	C	106	56
	D	102	56
20	A	136	56
	B	148	56
	C	122	56
	D	120	56
25	A	157	56
	B	157	56
	C	156	56
	D	168	56
30	A	183	56
	B	162	56
	C	175	56
	D	193	56

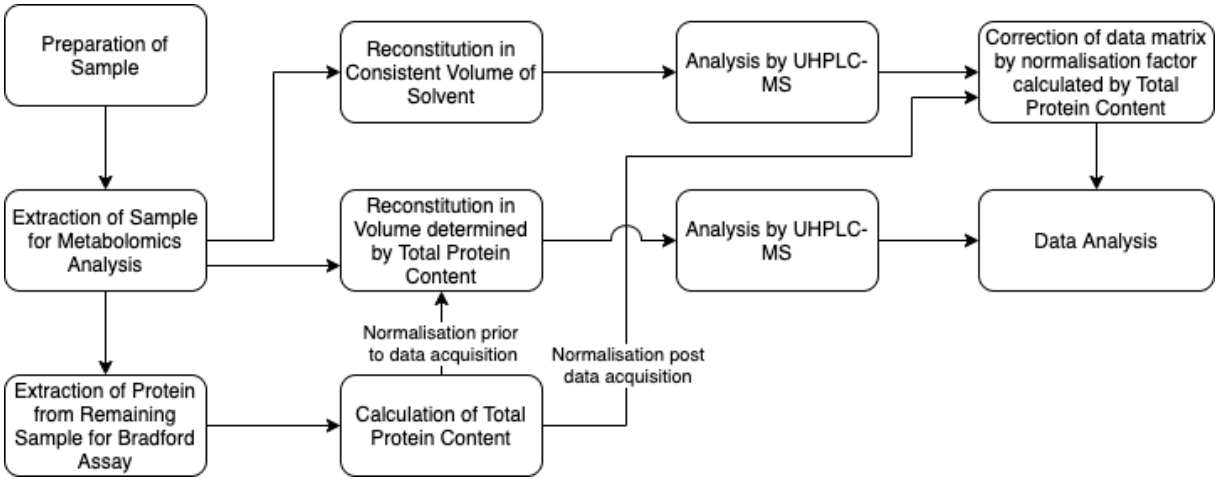


FIGURE 3.1: Experimental workflow demonstrating the two separate normalisation methods and when normalisation was applied.

with 10 mM ammonium formate. A 14 minute gradient elution was applied as follows using a 500 μL/min flow rate: 0 min, 1% B; 1 min, 1% B; 3 min, 15% B; 6 min, 50% B; 9 min, 95% B; 10 min, 95% B; 10.5min, 1% B. The lipidomics method used a Hypersil Gold C_{18} (2.1 x 100mm, 1.9 μm), with a flow rate of 400 μL/min flow rate and with two mobile phases, mobile phase A (10mM ammonium formate in 60% acetonitrile/water (v/v) + 0.1% formic acid) and mobile phase B (10mm ammonium formate in 90% IPA/acetonitrile (v/v) + 0.1% formic acid). A 14 minute gradient elution was applied as follows applying a 400 μL/min flow rate: 0 min, 20% B; 0.5 min, 20%B; 8.5 min, 100% B; 9.5 min, 100% B; 11.5 min, 20% B.

3.2.9 Data Pre-Processing

Raw data was first converted to .mzML format using ProteoWizard (version 3.0.1921) before peak picking and retention time alignment using XCMS (version 3.8.0). Parameters for processing are described in Appendix B. The resultant peak matrix was then filtered to remove low quality peaks applying the following filters: each feature had

to be present in at least 60% of all QC samples; features had to show an RSD of less than 30% across all QC samples (except the first 8 which were used for column equilibration); finally the mean QC response for each feature had to be 10 times greater than the average blank response. All other features were removed. All data was processed as one data matrix in R (version 3.6.0) applying both in-house scripts and the MetaboAnalystR package.

3.2.10 Normalisation of samples post-data acquisition

After peak picking and data filtering, a normalisation factor was calculated for each sample based on the total protein content measured for that sample against the sample with the lowest total protein content.

$$\text{Normalisation Factor} = \frac{\text{Total Protein Content of Sample}}{\text{Total Protein Content of Sample with Lowest Total Protein Content}} \tag{3.1}$$

Each feature of a specific sample was then divided by the normalisation factor calculated for that sample as determined by Equation 3.1.

$$\text{Normalised Feature Intensity} = \frac{\text{Feature Intensity}}{\text{Normalisation Factor}} \tag{3.2}$$

The exact value of each samples normalisation factor can be found in Table 3.2.

Whilst the normalising factor appears to be linear initially, after 25 μL the normalisation factor appears to decrease again. This could potentially be down to the extraction solvent becoming saturated with protein, however further investigation would

TABLE 3.2: Normalisation factor for each sample which is normalised post data acquisition.

Volume	Replicate	Normalisation Factor
10	A	1.13
	B	1.06
	C	1
	D	1.03
15	A	1.43
	B	1.47
	C	1.77
	D	1.47
20	A	2.30
	B	2.22
	C	2.43
	D	2.28
25	A	3.58
	B	3.43
	C	3.21
	D	3.22
30	A	3.18
	B	3.01
	C	2.82
	D	3.02

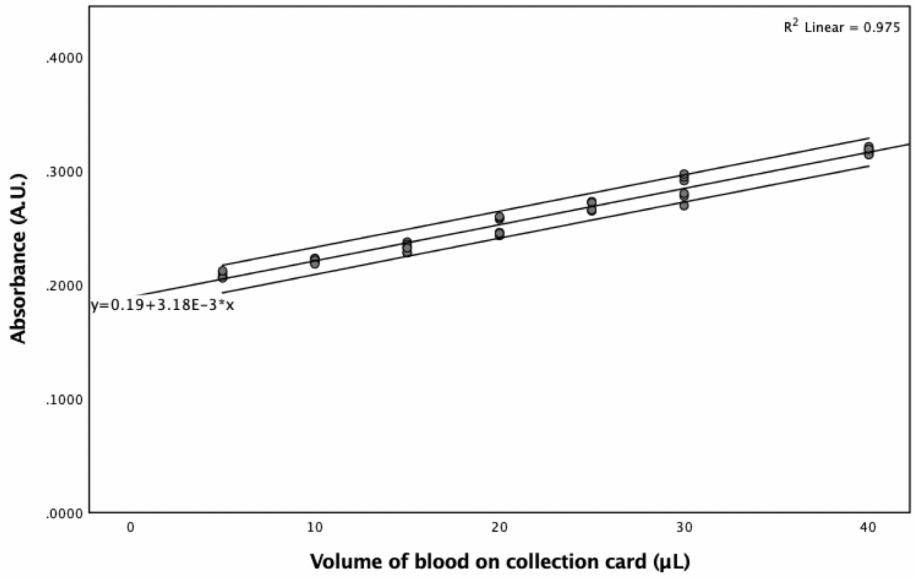


FIGURE 3.2: Scatter plot demonstrating the linear response between total volume of blood spotted onto collection card against absorbance measured at 590nm.

need to be completed to determine the cause of this effect. Initial thoughts were mislabelling, however subsequent processing and analysis of the data demonstrated that this was not the cause.

3.3 Results and Discussion

3.3.1 Linear response between protein content and volume of blood

Firstly to ensure linearity between the total protein content and the volume of blood within a DBS, controlled volume spots were collected and subsequent Bradford assays performed. Total absorbance was then plotted against volume of blood on each collection card. In total each volume was replicated 6 times.

TABLE 3.3: Number of features before and after filtering for each method of normalisation

Assay		Total Number of Features		
Separation Mode	Polarity	Before Filtering	After Filtering	
			Post Data Acquisition & TIC Normalisation	Pre Data Acquisition
HILIC	Positive	20953	2779	2759
	Negative	14587	1083	1049
Lipidomics	Positive	45171	3626	3884
	Negative	26808	3597	3746

The results from this test are shown in Figure 3.2 and a clear linear trend between whole blood volume and protein concentration can clearly be seen, with an R^2 of 0.975. Most of the variation seen between the individual volumes can most likely be explained down to errors in the preparation and extraction of the DBS themselves.

3.3.2 Quality of UHPLC-MS data

The total number of features remaining after filtering are described by Table 3.3. To ensure that the features which are normalised were of high quality and ensure confidence in the data, only peaks which were present in at least 3 of 4 replicates and also present in every single class (different volumes of whole blood present on each collection card) were kept for each different normalisation type. All figures shown in this chapter represent only one assay (Lipidomics - Positive Ion Mode). Result for all remaining assays can be found in Appendix G.

3.3.3 Comparison of data normalisation techniques

To compare the performance of both normalisation techniques (pre and post data acquisition), several performance characteristics can be compared between normalised data and data that had not had any normalisation procedures applied. This can then be used to compare and measure the improvement in the quality of the data itself.

3.3.3.1 Relative Standard Deviation (RSD)

Firstly we can look at the relative standard deviation of each feature and compare both pre and post data normalisation to data which has not been normalised. RSD is often used within non-targeted metabolomics data analysis as a quality filter within the QC samples. Typically any features that show an RSD of greater than 30% across all QC samples are removed as these features therefore show a response that means the feature is unreliable and therefore any biological variation within actual samples will be masked by analytical variation of that feature. As a point to note, due to each sample originating from a single biological source, there will be no biological variation between samples. Variation between samples will have only be due to non-biological sources such as sample preparation variation or instrument bias.

To calculate the RSD's of each feature and ensure a reliable representation of the dataset, only features which were present in at least 75% of all replicates within all groups were kept. This was to prevent any features that may not have been detected due to falling below the limit of detection of the detector having an effect on the calculation.

Figure 3.3 shows the effect that the different normalisation procedures have on the RSD of all features across the volumes of blood that have been spotted on the collection cards.

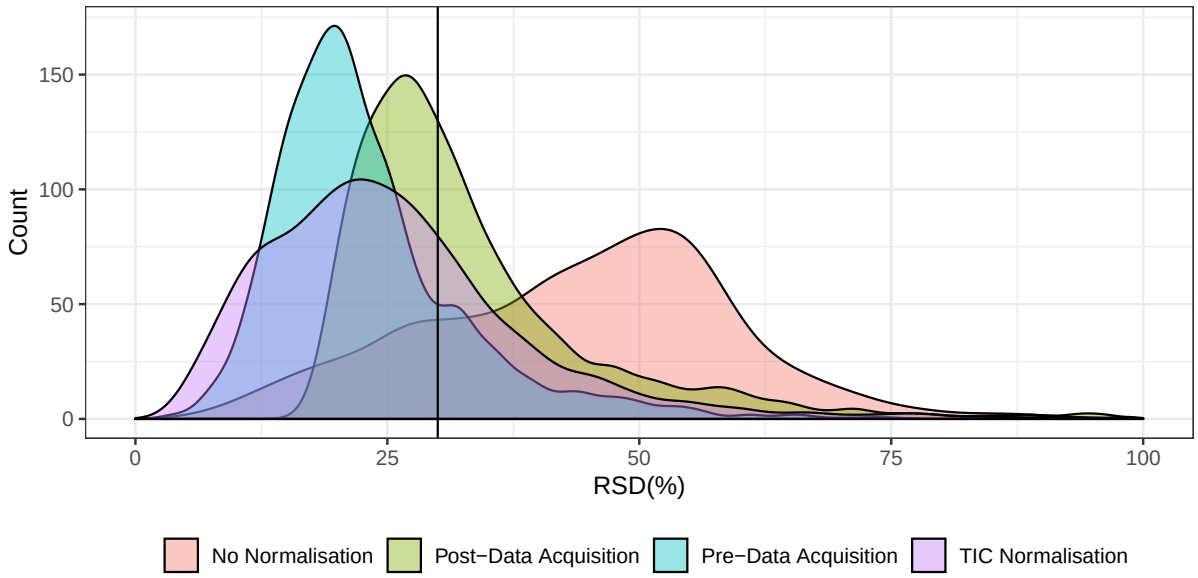


FIGURE 3.3: RSD of all high quality features in the lipidomics, positive ion mode assay before and after normalisation either pre or post data acquisition or normalising applying total ion current alone. Verticle line indiciates a 30% RSD cut off.

TABLE 3.4: Percentage of total features that show a RSD of less than 30% in each assay.

Assay		Normalisation Type			
Column	Polarity	No Normalisation	TIC Normalisation	Post-Data Acquisition	Pre-Data Acquisition
HILIC	Positive	42.1	70.8	61.1	89.0
	Negative	27.5	53.4	57.5	85.4
Lipidomics	Positive	20.0	70.2	51.1	81.8
	Negative	15.0	68.1	63.7	78.4

It can clearly be seen that there is a vast improvement in the number of features that

show an RSD of less than 30% for both normalisation methods, however the improvement when reconstituting pre data acquisition is better than computationally normalising post data acquisition. Another observation made is the similarity in RSD distribution between both datasets which have been normalised before and after data acquisition. Although the distribution for the post data acquisition set is shifted towards higher percentage RSD's by about 5%, it demonstrates that the majority of features have improved RSD's when compared to applying no normalisation procudures.

The total percentage of features that show an RSD of less than 30% depending on the normalisation method applied are shown in Table 3.4.

A comparison between the performance of normalising to total protein content alone was made to normalising against the total ion current (TIC) which is a commonly used normalisation method used within non-targeted studies. Normalising against the TIC works through each feature within a sample is divided by the sum of all intensities present within the same sample to produce a normalised value. The performance of the TIC normalisation can be seen in Figure 3.3 and Table 3.4. As with normalising to total protein content, both before and after data acquisition, an improvement in the RSD's are demonstrated in Figure 3.3, however with a wider distribution. Table 3.4 also demonstrates that the improvement to the number of features which demonstrate an RSD of less than 30% is better than normalising post data acquisition, but not as good as normalising prior to the collection of data through volume. This is the case for all assays apart from HILIC negative where TIC normalisation doesn't perform as well as both methods of normalising to total protein content.

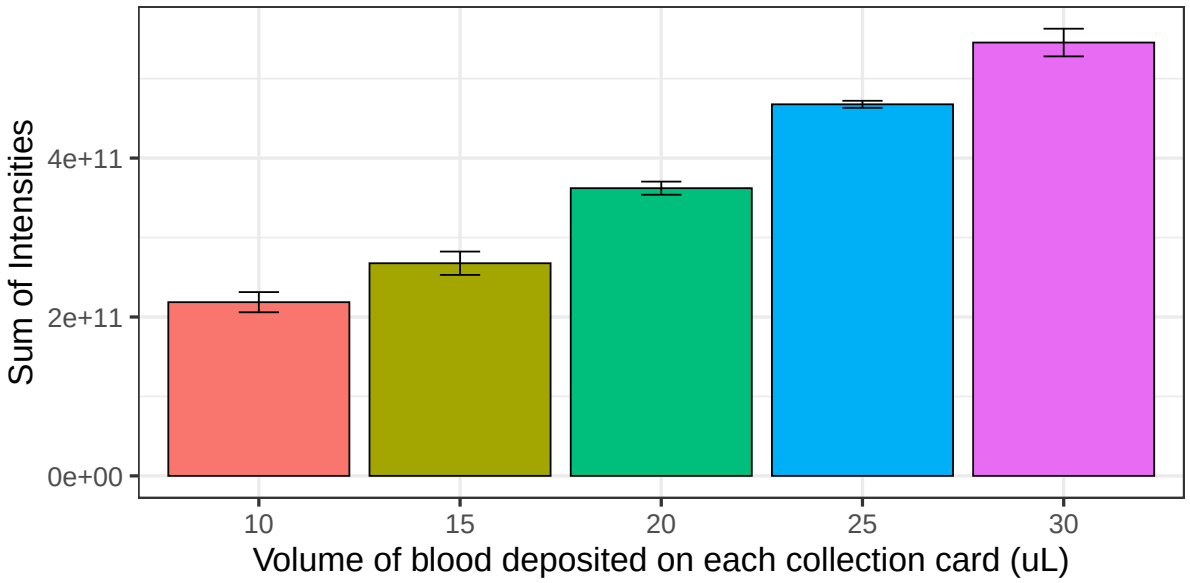
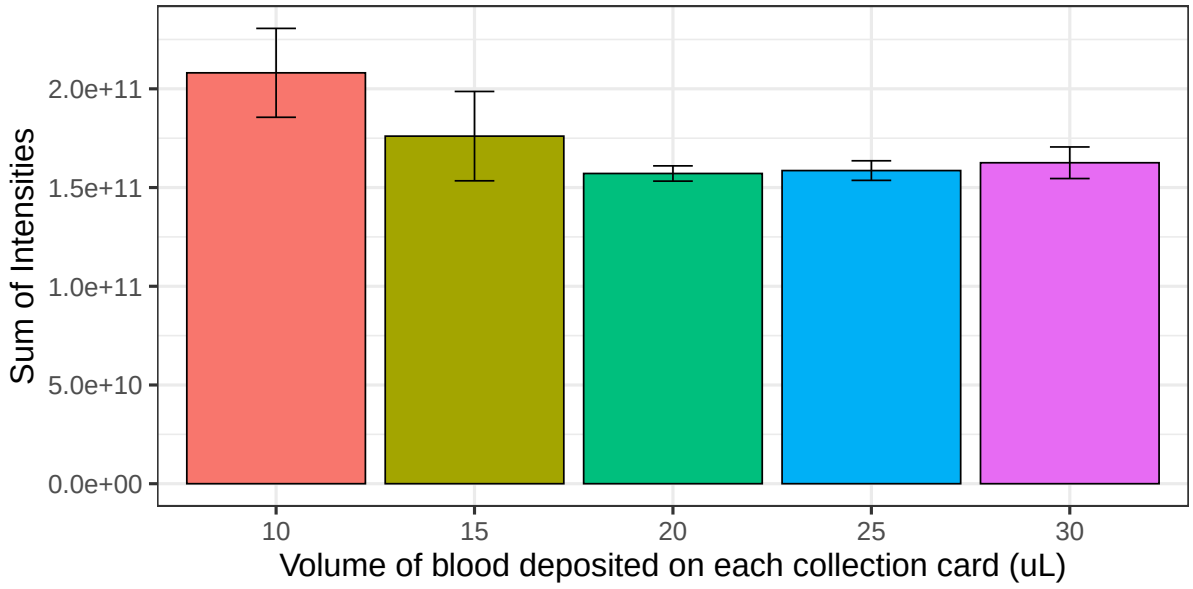


FIGURE 3.4: Sum of all intensities from features that are present in at least 75% of all replicates within every group.

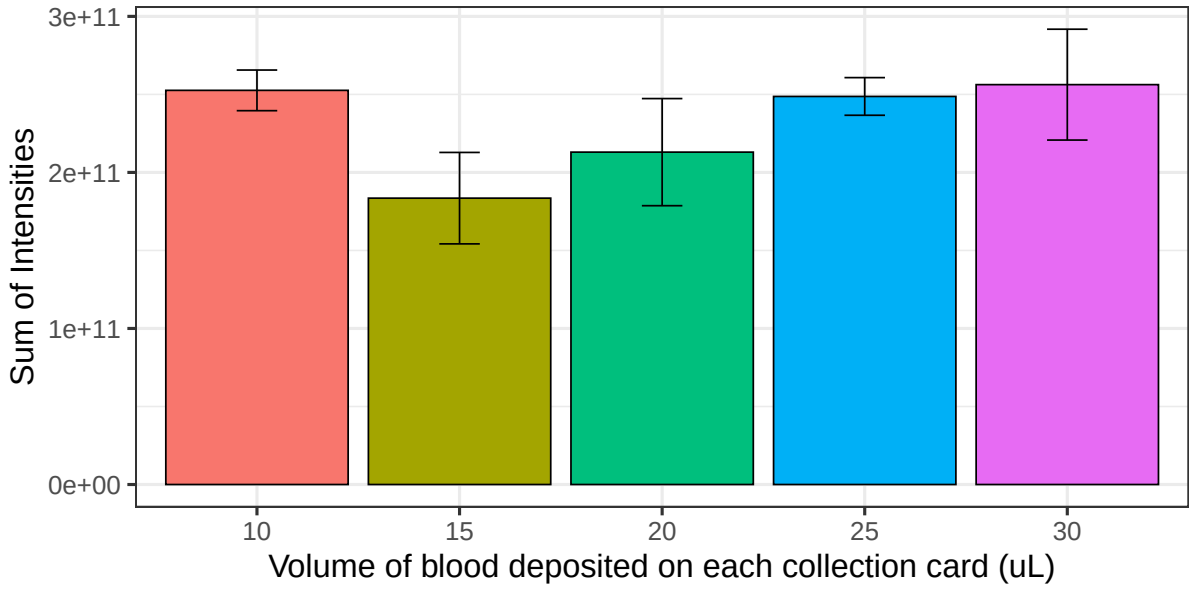
3.3.3.2 Sum of all features

Another procedure in determining whether or not normalisation has improved the quality of the data is to calculate the sum of all intensities within the data matrix. This can identify if there is a general trend within the data and can give an indication of a general increase in concentration of the sample. Once again, the sum of all intensities was only calculated on features that were present in at least 75% of all replicates that were present within each and every group.

Figure 3.4 shows non-normalised data with a clear increase in the sum of intensities of all features. This it to be expected due to the increase in concentration of all features that are present within a fixed reconstitution volume down to the volume of blood that is present on the collection card. To represent an improvement in quality of the data, the differences between each volume should be as minimal as possible to ensure that any variation that is seen is biological in nature rather than analytical.



(A) Post-data acquisition



(B) Pre-data acquisition

FIGURE 3.5: Sum of all intensities from features that are present in at least 75% of all replicates within every group. (A) Post-data acquisition normalisation. (B) Pre-data acquisition normalisation.

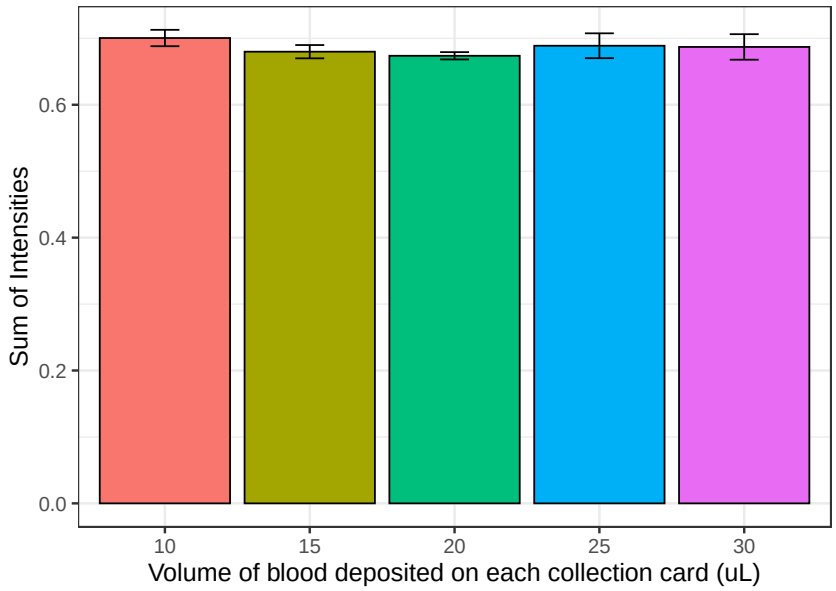


FIGURE 3.6: Sum of all intensities from TIC normalised features that are present in at least 75% of all replicates within every group.

After normalisation, both pre and post normalisation techniques show a decrease in the gradient of the total sum of intensities as shown in Figure 3.5. However whilst normalising post-data acquisition appears to bring the higher volume groups into line with each other, there is a significant decrease in the sum of all intensities after the 10 and 15 μL spots. This drop off could potentially be explained by the reconstitution volume that is used. Due to reconstituting in the same volume, where larger volumes of blood have been placed on the cards it could be the case that not all of the sample which has been dried down is going into solution and therefore the reconstitution solvent is saturated with sample.

Normalising pre-data acquisition shows good performance when normalising both the highest and lowest volume groups together. Although there does seem to be wide variation within each group, this variation is still better than not normalising at all, and in conjunction with the results seen in the improved RSD's for features, then this

TABLE 3.5: Number of significant features as defined by a one way ANOVA with a p value (FDR corrected) ≤ 0.05 for each assay and normalisation method.

Assay		No Normalisation	TIC Normailisation	Post-Acquisition Normalisation	Pre-Acquisition Normalisation
Seperatation Mode	Polarity				
HILIC	Positive	2565 (92.3%)	2285 (82.2%)	2336 (84.1%)	1315 (47.7%)
	Negative	924 (85.3%)	857 (79.1%)	854 (78.9%)	563 (53.7%)
Lipidomics	Positive	3285 (90.6%)	3061 (84.4%)	2661 (73.4%)	2075 (53.4%)
	Negative	3370 (93.6%)	3108 (86.4%)	3097 (86.0%)	2221 (59.3%)

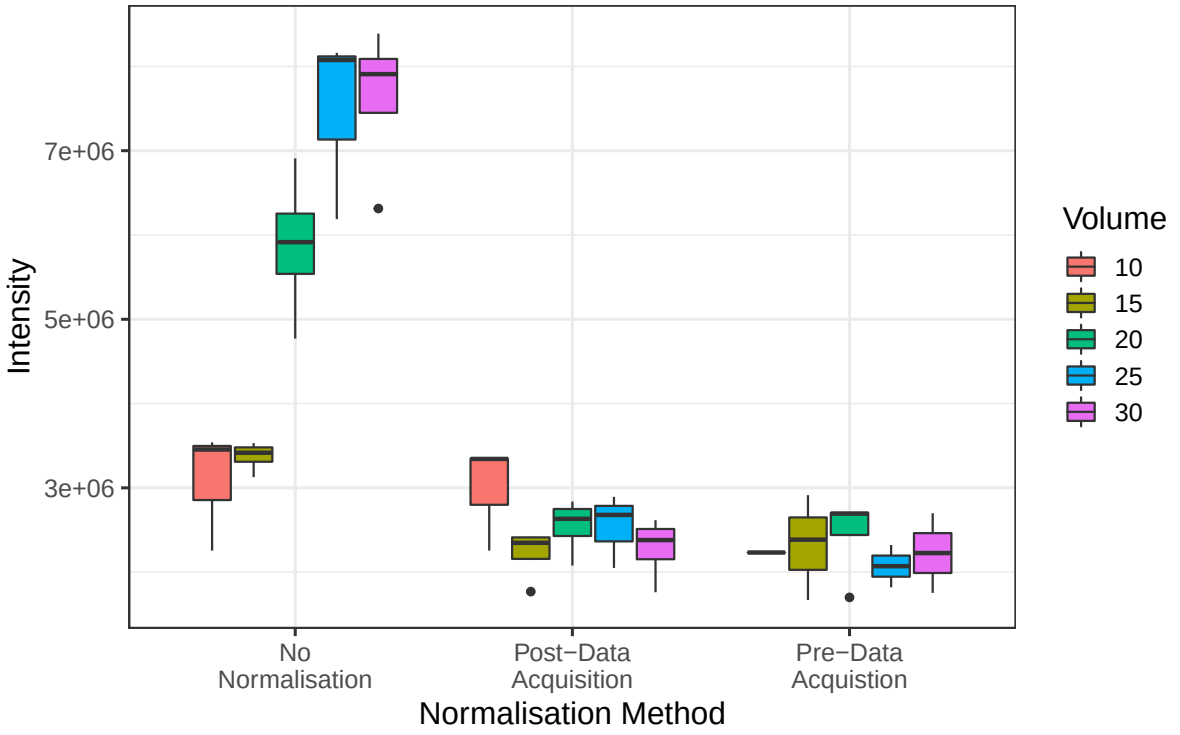
is still much preferred and the quality of the data is vastly improved.

Due to the addition of an extra step of measurement of total protein content and thus subsequent changes in reconstitution volume when preparing samples for normalisation prior to the acquisition of data, variance within samples is larger. This is demonstrated in Figure 3.5 where the errors bars are larger in the pre data acquisition dataset compared to post.

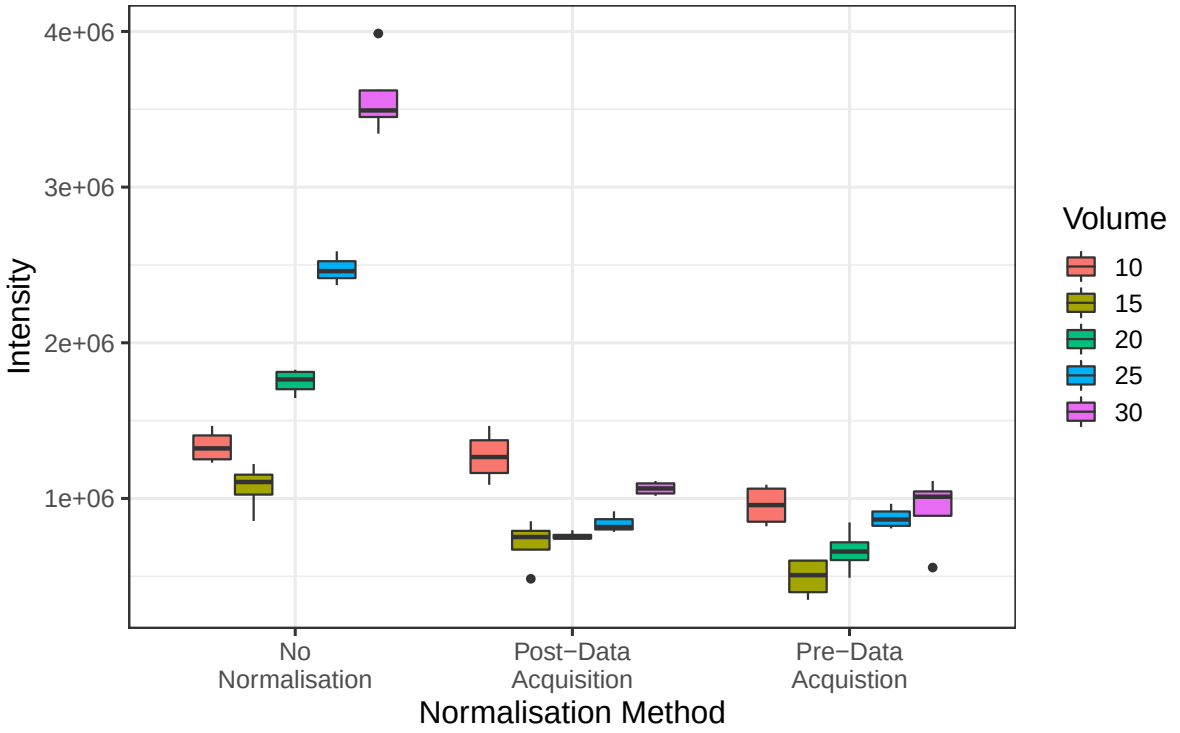
3.3.3.3 Individual Features

To demonstrate the effect that normalising both pre and post data acquisition has on individual features, as well as comparing to a purely computational metric such as TIC normalisation, an ANOVA was performed and the number of features determined to be statistically significant before normalisation then compared to the number which are not significant after normalisation. These results are described in Table 3.5 and a few select peaks are plotted in Figures 3.7a & 3.7b.

There is a clear trend where features which are statistically significant prior to normalisation are then no longer significant afterwards. This demonstrates a clear improvement in the quality of the data for individual features and could allow for potential biological variances to be seen.



(A) (-)-Carvone



(B) Stearoylcarnitine

FIGURE 3.7: Intensities for (A) (-)-Carvone and (B) Stearoylcarnitine before and after normalisation applying total protein content normalisation. Volumes in μL.

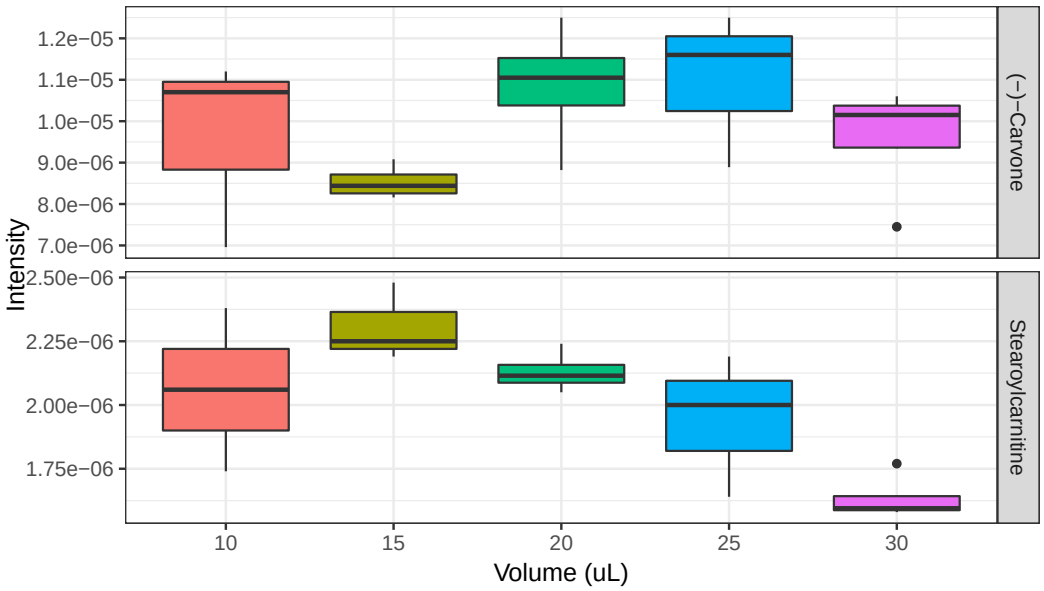


FIGURE 3.8: Intensity of (-)-Carvone and Stearoylcarnitine after TIC normalisation

Whilst there does appear to be a dilution trend as demonstrated by Figure 3.7b, this does not appear to occur in specific classes of metabolites and occurs randomly throughout the dataset. Whilst it would be preferable if this trend was not present, the quality of the data has still vastly been improved through the application of normalising against total protein content.

When comparing data which has solely been corrected for by applying TIC normalisation, very little reduction in the total number of significant features is observed, with both post and pre data acquisition normalisation strategies based on total protein content, demonstrating improved performance. This is not what was observed when comparing the improvements in RSD's in Section 3.3.3.1. In Section 3.3.3.1, the improvement in RSD ranges fell between normalising post and pre data acquisition. Therefore from that conclusion alone it could have been suggested as an alternative strategy for normalising dried blood spots of different volumes. However, from Table

3.5, with minimal reduction in the total number of features which have been determined to be statistically significant when compared to the other normalisation strategies based on total protein content, normalising against TIC cannot be recommended.

Whilst there are still a significant number of metabolites which show statistically significant differences after pre-data acquisition normalisation, this can be explained through the use of a single source of whole blood to prepare the spots of different volumes. Therefore there is no biological variability present within the spots and all variance seen is technical in nature. In a typical study where DBS will have been collected from multiple sources, biologically variability will be much greater than technical variability. Thus the technical variability within this study has been magnified due to the lack of biological variability and the number of statistically significant differences.

3.4 Conclusion

Throughout this chapter I have described a method of normalising DBS which is non-destructive to the original sample destined for metabolomics analysis. This technique represents a cheap (approximately 40 pence per sample), easy and reliable method for protein determination within the DBS and can therefore offer widespread utilisation for anyone performing experiments with DBS. The assay itself is quick to perform meaning the experiment can be performed prior to the collection of data. Additionally, the materials required for this experiment are readily found within many laboratories and the most expensive piece of equipment required is the spectrophotometer itself which is already a key part of many laboratories. However, if the assay was to be applied to large studies then proper experimental planning would be required and

samples dried and stored for longer periods of times whilst the total protein measurement is determined for large numbers of samples before reconstitution in appropriate solvents.

Being able to normalise DBS also opens up many prospects for large scale studies that utilise samples within biobanks around the world. Unless the DBS has been collected using an accurate volume collection device, then the volume on the DBS is an unknown. Coupled with the varying haematocrit that each individual patient has then normalising based on spot size can also become problematic.

The technique applied here demonstrates a large improvement in the RSD for many features in a single biological sample, thus removing as much non-biological source of variation as is possible applying the Bradford assay. Whilst there will still be features which demonstrate a large variation, if volume has not been recorded then this technique reduces the number of those features by a large degree, with a decrease in the number of significant features of between 31.6% and 44.6%.

This chapter has also demonstrated that normalisation prior to data acquisition establishes a greater improvement in the quality of the data when compared to normalising computationally after the collection of data with an increase of between 46.9 to 63.4% to the total number of features demonstrating an RSD of less than 30%. With improved RSD performance and a more consistent response relating to the total number of features detected, normalisation should occur at the reconstitution step if at all possible.

Finally, this chapter has demonstrated that whilst applying TIC normalisation alone improves the RSD of features present to levels which are comparable to normalising post acquisition of data, there was very little effect on the number of features which remained statistically significant, thus preventing its use in a non-targeted metabolomics

scenario where relative concentrations between classes is key in distinguishing important biological features.

However, whilst the technique demonstrates that normalising to total protein content reduces the number of statistically significant features and improves RSD performance, there are still improvements that could potentially be made. The variance observed in the total protein content of each spot is large, therefore improvements could potentially be made in both the method in which the Bradford assay is applied to attempt to reduce these variances. Although more sensitive protein assays are avialable for use, these were not considered for this chapter due to the Bradford assay having more than sufficient sensitivity for the assay performed, with the sample required to be diluted to ensure that measurements would fall within the linear dynamic range.

Potential further developments to the application of this method would be the automation of the assay. This would reduce the time taken to perform the assay whilst reducing any human sources of variation that are naturally introduced when performing a manual extraction. Further work could also be focused on ensuring the assay is valid for other sample types which are not whole blood. Additionally, it would be useful to understand whether this assay could be applied to the variety of collection devices that are available, both those which are as a card format, or those which are specially manufactured by commercial suppliers with the specific aim of controlling for volume based effects such as Neoteryx's Mitra device.

In summary, the method demonstrated here, whether or not it is utilised prior to the collection of data or afterwards, represents a method with the ability to improve the quality of data acquired especially with spots that are widely varying in volume. Preferentially, normalising prior to acquiring data gives best performance with regard to decreasing RSD's and ensuring the number of features remains constant.

Chapter 4

Optimisation and Assessment of Liquid Extraction Surface Analysis methods for Non-targeted Metabolomics studies

4.1 Introduction

The use of Liquid Extraction Surface Analysis (LESA) as a surface analysis technique has been around since 2005 [151, 224] and has been applied to the direct analysis of DBS in a variety of applications mainly focused around proteomics [225–228]. Whilst LESA has been applied towards the detection of small molecules, these applications have instead focused on other surfaces than paper based collection cards, using the likes of thin tissue sections [229] and oral fluid, urine and fingerprints [230].

Using LESA allows the direct analysis and infusion of a sample without prior manual extraction into solution. Due to this, the technique is fast and automated and a sample can be analysed from the moment it is collected. With DBS, because there has been no prior preparation of the blood that has been used to create the DBS, a more

complete picture of the metabolome can be detected at any one time from the surface. This is because all the constituent components that make up whole blood such as plasma, platelets and cells (red and white) are still present within the sample as they have not been removed because of prior sample preparation procedures.

LESA works through the deposition of solvent onto the surface of the sample and then waiting a pre-determined amount of time before reaspirating into the same conductive pipette tip. The extraction solution is then consistently infused directly into the MS where data can be acquired. Almost every aspect of the LESA process can be modified, making it a versatile technique with a wide range of applications that can be developed. Developing a LESA method for metabolomics could potentially open up applications for a rapid, high throughput procedure of analysing DBS where the resources and knowledge required to manually extract samples for more extensive analysis such as UHPLC-MS may not be available. This could potentially be in environments away from a research lab or in a clinical setting where the high throughput of samples is required.

LESA utilises a chip based on nano electrospray ionisation (nESI) system that allows for an increase in sensitivity compared to typical electrospray ionisation (ESI) operating in a direct infusion mode, whilst also allowing minimal volumes of solvent to be used, therefore lending itself well to the analysis of small molecules that have been extracted from a brief dwell on the surface of the spot. In order to be able to utilise LESA for a non-targeted metabolomics experiment, a method needs to be optimised to ensure enough information is extracted from the surface of the DBS to be useful.

This chapter describes the optimisation of a LESA method capable of the extraction of metabolites directly from the surface of a DBS for non-targeted metabolomics studies. The method was then compared directly against a DI-MS method as well as

UHPLC-MS to determine where LESA-MS of DBS has its benefits but also where other techniques might be more applicable depending on the application.

4.2 Materials and Methods

HPLC Grade methanol, water, acetonitrile and isopropanol were purchased from VWR Chemicals (Lutterworth, UK) along with formic acid (>98%). Rat whole blood (female, male & pooled, Wistar Hannover strain) was purchased from BioIVT (West Sussex, UK). Whole blood was collected in Lithium Heparin collection tubes to prevent coagulation before preparation of DBS. Time between collection of whole blood and preparation of DBS was less than 48 hours. Whatman 903 filter paper collection cards, desiccant and foil ziploc bags were purchased from Fisher Scientific (Loughborough, UK). Formic acid ($\geq$98.0%) was purchased from VWR International (Lutterworth, UK)

4.2.1 Sample Preparation for LESA-MS

Samples to be analysed by LESA-MS were prepared by pipetting 20 μL of rat whole blood for each gender separately onto Whatman 903 Filter Paper based collection cards. Cards are then allowed to dry for three hours at ambient temperature before storage at room temperature in foil ziploc bags alongside desiccant until analysis. All samples were analysed within 1 week of receipt of the whole blood.

4.2.2 Sample Preparation for DI-MS & UHPLC-MS

Samples to be analysed by DI-MS and UHPLC-MS were prepared in the same manner as in Section 4.2.1, except samples were stored at -80°C until analysis.

4.2.3 Extraction of DBS for UHPLC-MS and DI-MS analysis

Before extraction, samples were removed from storage and allowed to defrost on ice for 30 minutes prior. Samples were then removed from the collection card using a sterile scalpel and cut into smaller sizes to aid in extraction by increasing surface area of the spot. Samples were then extracted in 480 μL of methanol and 20 μL of water under gentle agitation for 20 minutes before centrifugation (21000 g, 4°C, 20 minutes). The supernatant from each sample was then dried down under vacuum for 40 minutes at 35°C. After drying, samples were stored at -80°C until analysis by UHPLC-MS.

4.2.4 Reconstitution of samples for UHPLC-MS analysis

Samples were resuspended in 56 μL of acetonitrile/water/methanol (3:1:1, v/v/v), before vortexing for 30 seconds and centrifugation (21000 g, 4°C, 5 minutes). Samples were then transferred to low volume LC vials for analysis.

4.2.5 Reconstitution of samples for DI-MS analysis

Samples were resuspended in 56 μL of methanol/water/formic acid (80%/19.9%/0.1%, v/v/v) before vortexing for 30 seconds and centrifugation (21000 g, 4°C, 5 minutes).

4.2.6 LESA extraction and nESI parameters

Samples were analysed applying a LESA method as developed in Section 4.3. Table 4.1 describes the parameters applied with corresponding references to sections later in the chapter where optimisations have been discussed and presented.

TABLE 4.1: Major parameters of LESA method developed. * Relative height above the surface of the sample, actual value used in software is different and is adjusted based on sample height using LESA stage

Parameter	Description	Value	Section
Solvent Volume	Volume of extraction solvent aspirated into conductive pipette tip before extraction	6 μL	4.3.7
Dispense	Volume of extraction solvent dispensed onto surface of sample	5 μL	4.3.7
Delay	Amount of time to dwell on the surface of sample	5 seconds	4.3.8
Aspirate	Volume of extraction to reaspirate into the conductive pipette tip before infusion into MS	5 μL	
Dispensation Height *	Height to dispense extraction solvent onto the surface of the sample	0.4 mm	4.3.9
Aspiration Height *	Height to aspirate extraction solvent back into conductive pipette tip	0.0 mm	
Voltage (Positive)	Voltage to apply to the conductive pipette tip for the nESI process in positive ion mode	1.7 kV	
Voltage (Negative)	Voltage to apply to the conductive pipette tip for the nESI process in negative ion mode	1.4 kV	
Gas pressure	Pressure to apply to solvent within the conductive pipette tip to infuse it directly into MS	0.3 psi	

4.2.7 DI-MS and LESA-MS data acquisition

Data was collected in both positive and negative ion mode using an Orbitrap Elite (Thermo Scientific, San Jose, CA, US) applying an mass range of 50 - 620 Da. Data for LESA was collected applying a 'looped' SIM-stitch method described in Section 4.3.3. Data for DI-MS was collected in quadruplicate and subsequently applying only a single SIM-stitch pass as demonstrated in Figure 4.5.

4.2.8 UHPLC-MS data acquisition

Data was collected in both positive and negative ion mode applying a HILIC method using a Thermo Vanquish LC coupled to a Thermo Q-Exactive Plus MS (Thermo Scientific, San Jose, CA). A pooled QC sample was prepared from the extracts of all samples

and then prepared in the same was as all samples before being injected every 5th sample with 10 leading QC's to condition the column. The LC method was 14 minutes and length and data was collected at 70000 resolution (at m/z 200) using an Accucore 150 Amide column (100 mm x 2.1 mm i.d., 1.9 μm pore size. Solvent A was acetonitrile/water/formic acid (95/4.9/0.1, v/v/v) with 10 mM ammonium formate. Solvent B was acetonitrile/water/formic acid (50/49.9/0.1, v/v/v) with 10 mM ammonium formate. A 14 minute gradient elution was applied as follows using a 500 μL/min flow rate: 0 min, 1% B; 1 min, 1% B; 3 min, 15% B; 6 min, 50% B; 9 min, 95% B; 10 min, 95% B; 10.5min, 1% B.

4.2.9 DI-MS and LESA-MS data pre-processing

Data was processed using Galaxy-M [231] using the workflow described in Figure 4.5 and the parameters described in Section 4.3.3.1. Thermo .raw files were passed into the workflow alongside a sample list describing class structures and QC samples with a data matrix consisting of samples and peak intensities as the final output.

4.2.10 UHPLC-MS data pre-processing

Raw data was first converted to .mzML format before peak picking and retention time alignment using XCMS (version 3.8.0). Parameters for processing are described in Appendix B. The resultant peak matrix was then filtered to remove low quality peaks applying the following filters: each feature had to be present in atleast 60% of all QC samples; features had to show an RSD of less than 30% across all QC samples; finally features had to beat the blank signal by a factor of 10 to remain in the matrix, all other features were removed. All data was processed in R (version 3.6.0) applying both in-house scripts and the MetaboAnalystR package.

4.2.11 Statistical Analysis

All statistical analysis was performed in R (version 3.6.0) using both the base R-stats package, MetaboanalystR in addition to in-house scripts.

4.2.12 Optimisation Workflow

To define the best performance characteristics a range of parameters will be monitored and used to define an optimal method when applying LESA to the non-targeted analysis of DBS. The main performance parameters used to define which instrument setting or method optimisation is best are the following:

- Relative standard deviation of features
- Total number of features
- Number of successful infusions
- Total peak area

Through monitoring these performance characteristics, decisions can be made to move forward with specific settings to ensure an optimal method is developed.

In addition to performance parameters, a specific order in which to complete optimisations is required. Thus the following order of optimisations was decided when conducting the method development section of this chapter.

1. Selection of solvents
2. Selection of modifier content percentage
3. Data processing procedures

4. Full scan vs SIM-Stitch
5. Repeated sampling
6. Solvent dispense volume
7. Solvent dwell time
8. Solvent depth of dispensation

4.3 Results and Discussion - Method Development

4.3.1 Selection of extraction solvents

4.3.1.1 Effect of different binary solvent systems on liquid microjunction stability

One of the key choices to be made when developing a LESA-MS method is the choice of solvent mixture that is used for the extraction of compounds from the DBS. Therefore the selection of solvents can have a major effect on the ability to form and maintain a liquid microjunction (LMJ) between the conductive pipette tip and the DBS. If a LMJ is unstable then the ability to have reproducible and reliable extractions can be compromised, thus affecting data quality.

To test for a suitable extraction solvent, a selection of binary solvent mixtures were prepared at different compositions of aqueous/organic solvent ratios. Binary solvent mixtures were chosen over single solvents due to the inability of organic solvents individually to form stable LMJ. Organic solvents selected for use were acetonitrile, isopropanol and methanol, all solvents which have been used extensively within non-targeted metabolomics extractions. Binary solvent mixtures were prepared by mixing each organic solvent with a select volume of water.

TABLE 4.2: List of different binary solvent mixtures created. Organic solvents relate to the use of either acetonitrile, isopropanol or methanol.

Mixture No.	Organic Solvent (%)	Water (%)
1	0	100
2	10	90
3	20	80
4	30	70
5	40	60
6	50	50
7	60	40
8	70	30
9	80	20
10	90	10
11	100	0

All of the binary mixtures prepared can be found in Table 4.2.

For each solvent mixture 2 μL was spotted onto the centre of a DBS. The time taken for the solvent droplet to diffuse into the DBS itself and the LMJ lost was then timed. These results for each different binary solvent composition and each different mixture can be seen in Figure 4.1

An interesting observation is made when looking at the length of time a droplet lasts on the surface of a DBS as shown in Figures 4.1 & 4.2. At high fractions of water ($\geq$90%) , a stable LMJ of 10 seconds or greater is only observed for binary solvent mixtures containing methanol and acetonitrile, with IPA mixtures showing a LMJ stability

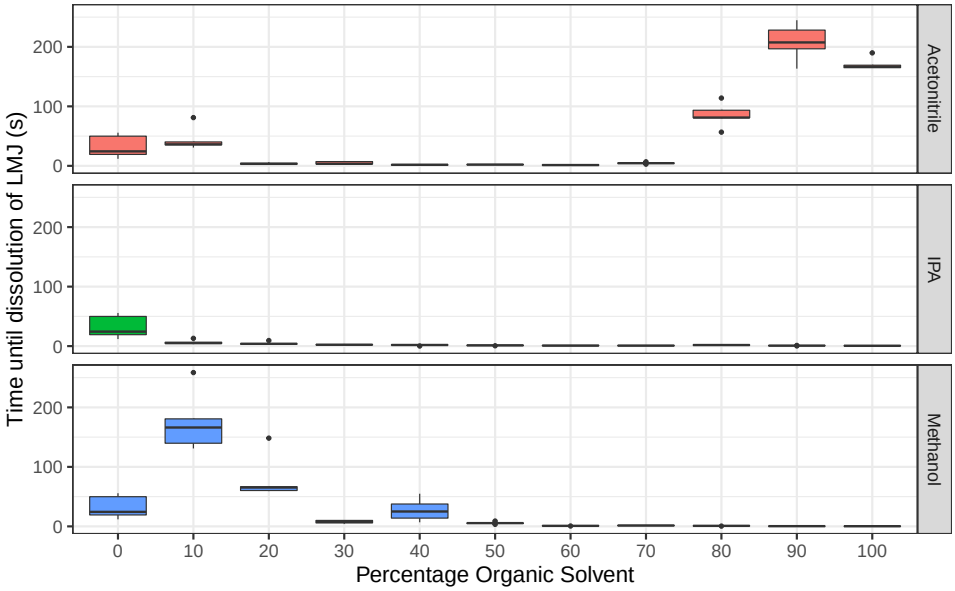


FIGURE 4.1: Time taken for 2 μL of each solvent mixture to dissolve into the surface of a DBS.

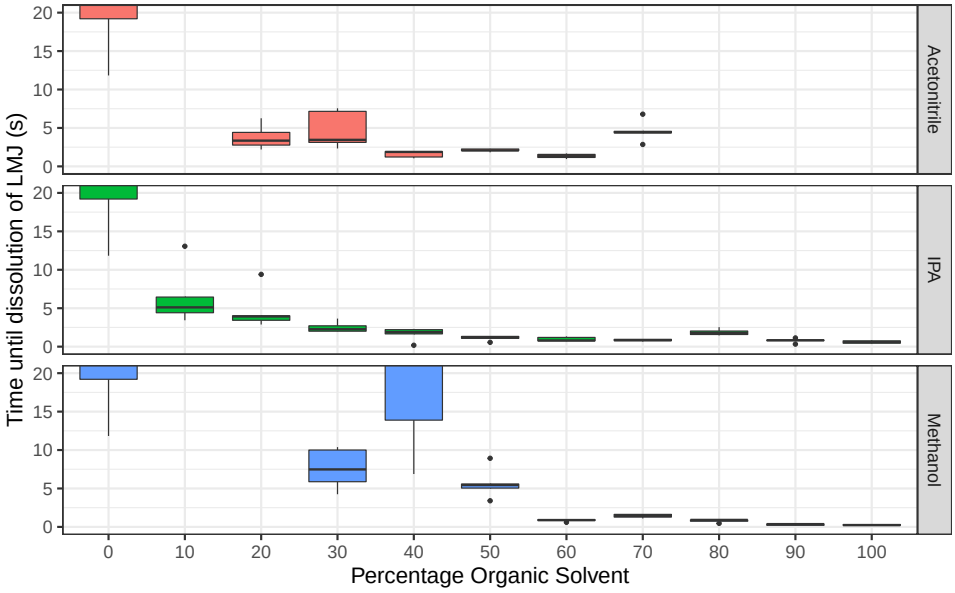


FIGURE 4.2: Time taken for 2 μL of each solvent mixture to dissolve into the surface of a DBS. Same as Figure 4.1 except y-axis magnified to between 0 and 20 seconds.

of approximately 5 seconds or less. At mixtures of greater than 10% IPA, no further stable LMJ were observed. For mixtures involving methanol, a stable LMJ of more than 5 seconds are observed for mixtures up to 50% water. Below this percentage of water, no further LMJ were observed.

In comparison to both IPA and methanol binary solvent mixtures, acetonitrile based mixtures demonstrated a different pattern. As mentioned in the previous paragraph, at high percentage water content, with acetonitrile, a LMJ is able to be formed. The stability of LMJ then gradually reduces until the percentage of organic solvent reaches approximately 80% where the stability of the LMJ then returns. This pattern is not observed for both methanol and IPA binary mixtures. Due to this return in stability when the organic content of the LMJ is high ($\geq$80%), this has the potential to be a useful extraction solvent to apply when using LESA. A common non-targeted metabolomics solvent composition consisting of 80% methanol and 20% water is used [232, 233], however as can be seen by Figure 4.1 this leads to a very unstable LMJ which lasts much less than the required 60 seconds that would be optimal for optimising many of the other LESA parameters available. Therefore, by applying a mixture of 80% acetonitrile and 20% water, LESA compatibility with DBS is maintained whilst also ensuring a commonly used organic/aqueous mixture is applied.

With the surface tension of organic solvents (acetonitrile - 28.16 mN/m, methanol 22.10 mN/m, isopropanol - 23.30 mN/m [234]) being much less than water (72.05 mN/m) [234], it is expected that as the aqueous content increases, the total surface tension of the LMJ droplet increases, thus there is better LMJ stability. This explains the results for both methanol and IPA, but does not explain the reason for the increased stability at high acetonitrile content. When looking at the literature for surface tension

of binary mixtures, there is to be no increase in the surface tension of acetonitrile/water mixtures as the organic content is increased [234], suggesting a different physical property of the mixture that is being altered at high organic percentage.

Therefore it can be assumed that surface tension is not the main factor that leads to a LMJ being formed and is instead controlled by another property alone or in addition to surface tension. The answer as to what this property may be is unknown, with no clear answer present within the literature. Potential options could be various interactions between the intermolecular forces present within the binary solvent mixture and the DBS allowing a stable LMJ to be formed, however further work would be required to either confirm this or suggest alternative reasons for this observation.

Moving forward for subsequent optimisations, it was decided that an 80%/20% mixture of acetonitrile and water was to be used.

4.3.1.2 Effect of charge modifier on acetonitrile/water LMJ stability

To ensure that any change in modifier concentration has no large impact in the stability of the LMJ, the same experiment is repeated, except with varying concentrations of formic acid. The mixtures tested are shown in Table 4.3.

TABLE 4.3: List of different concentrations of extraction solvent with varying contents of formic acid.

Mixture No.	Acetonitrile (%)	Water (%)	Formic Acid (%)
1	80	20	0
2	79.95	19.95	0.1
3	79.75	19.75	0.5
4	79.5	19.5	1
5	78.5	18.5	3

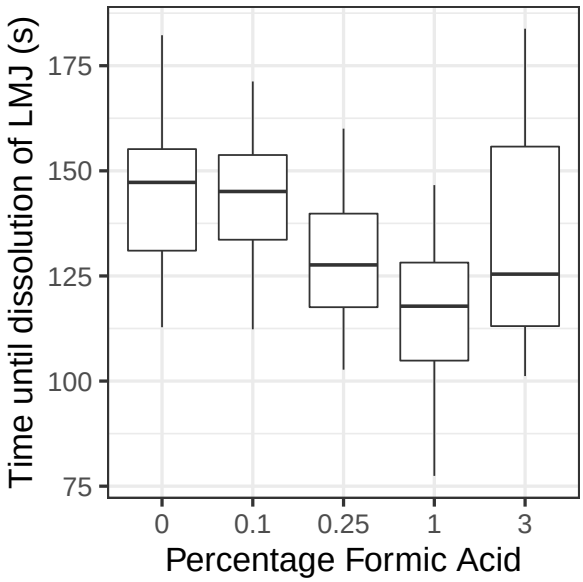


FIGURE 4.3: Time taken for 2 μL of each mixture of 80% acetonitrile/20% water with varying amounts of formic acid to dissolve into the DBS. Box represents the 25th and 75th quartile with the horizontal line equivalent to the median. Whiskers indicate the highest and lowest values.

The results from this experiment are shown in Figure 4.3 and whilst there is an effect on the ability to form a LMJ, it does not bring the instability down to below 60 seconds which is the aim for carrying on for the remainder of the optimisation. The inclusion of a modifier such as formic acid is important to provide a source of protons for the ionisation process. There is no difference in LMJ stability with the addition of 0.1% of formic acid to the binary solvent mixture, or no formic acid at all. An effect can only be seen when 0.25% of formic acid has been added to the mixture. This effect could be caused by the decrease in organic solvent content with increased formic acid concentration. As demonstrated by Figure 4.1, there is a region of LMJ stability at high percentage of acetonitrile in the binary mixture. Therefore with the addition of increased formic acid, the organic content has been reduced, thus reducing the stability of the LMJ but not to a level which would compromise any future method

development.

Due to there being no significant effect to the LMJ with the addition of 0.1% of formic acid, this was the value chosen for the extraction solvent mixture.

4.3.2 Preparation of Dried Blood Spots

All samples were prepared by accurately pipetting 20μL of whole blood onto Whatman 903 filter paper based collection cards. Cards were then allowed to dry at ambient temperature before storage in foil ziploc bags alongside dessicant. Cards were stored at room temperature and analysed within two weeks of preparation.

Samples were unable to be stored at below freezing temperatures due to the subsequent inability to form a LMJ once cards had been defrosted. This could potentially be down to the increase in the absorption of the filter paper once the collection had been frozen. This may have been caused by the remaining water in the DBS freezing and expanding disrupting the clot that initially forms on the surface of the DBS which creates a semi-impermeable barrier to prevent the LMJ from forming. When this clot is disrupted the LMJ fails almost immediately due to the solvent being able to soak into the filter paper based collection card beneath.

Therefore all cards which are intended to be analysed by LESA should be stored above freezing temperature until they have been analysed at which point they could then be stored at below freezing temperatures if further analyses is to be performed.

4.3.3 Data Processing

To increase the dynamic range and sensitivity of DI-MS analysis, a data collection technique called SIM-Stitch can be used. SIM-stitch works through the collection of multiple SIM windows within a single acquisition and then through post analysis computational processing, these multiple SIM windows can be 'stitched' together to form a single continuous peaklist and peak matrix.

Whilst the orginal implementation of the technique developed by Southam et al [121, 235], is suitable for DI-MS experiments, it is unsuitable for LESA-MS analysis. This is because the original technique employs 4 technical replicates of the same sample which is not possible with LESA-MS due to the semi-destructive nature of the extraction procedure. Due to the modification of the surface it is therefore impossible to infuse the 4 technical replicates as required by the original SIM-stitch data processing workflow.

To be able to use the SIM-stitch technique with LESA-MS it is therefore required to modify the original SIM-stitch workflow. These modifications include multiple SIM-stitch loops within a single .raw file. This is compared to the original method where only 1 SIM-stitch loop is required before moving onto the next technical replicate. The number of SIM-stitch loops set for a single LESA-MS infusion was 3.

Due to the lack of technical replicates, it is then necessary to modify the standard SIM-stitch data processing workflow as shown in 4.5 A. To process data which has been collected in the method described above where multiple SIM-stitch windows have been collected within a single .raw file, there is no need for the replicate filter step where peaks are filtered out based upon there presence of lack of presence within each of the technical replicates. Instead the filtering of replicate peaks that occur within

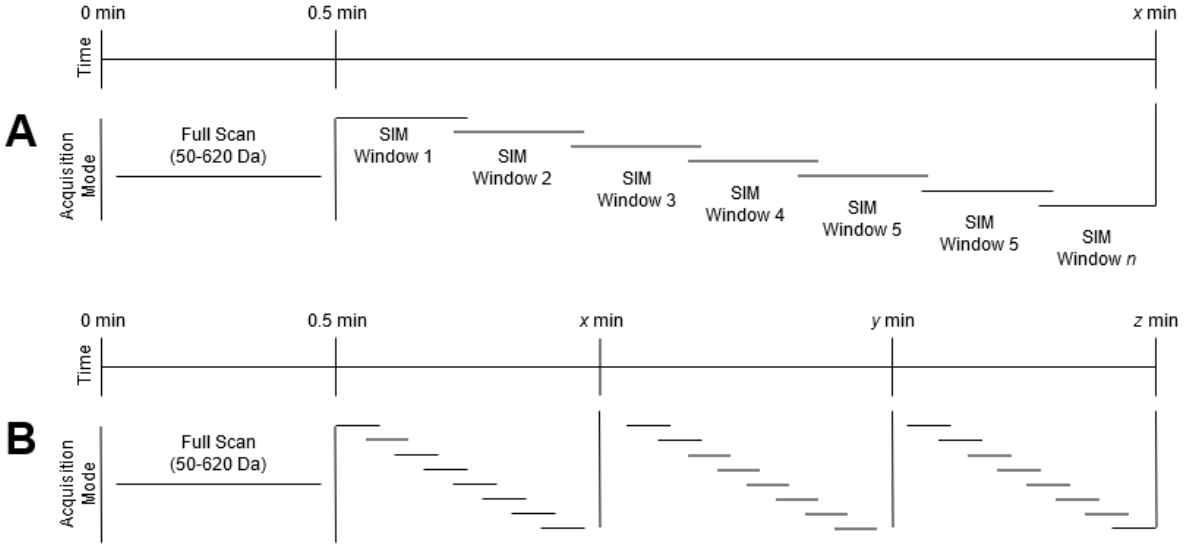


FIGURE 4.4: **A** - Method setup for a DI-MS infusion. Each DI-MS infusion is repeated multiple times. **B** - Method setup for a LESA-MS infusion. Each infusion is only repeated once and 'technical' replicates arise from each loop of the SIM-stitch process.

each 'loop' of the analysis occurs within the process scans step. Due to this, data processing occurs using the steps described in Figure 4.5 B.

4.3.3.1 LESA-MS Data Processing Parameters

To process LESA-MS data the following settings were applied to each section of the data processing workflow.

Process Scans - Used to convert data from the .raw file into a single peak list containing *m/z* and intensities of each feature detected. For peaks to be classed as real features, the peak has to beat the noise threshold by a factor of ten and also has to be present in all three SIM-stitch 'loops' within the .raw file.

Align Samples - Take each individual peak list produced from the Process Scans step and align all detected peaks within a 2 ppm mass *mz* tolerance. The output of this

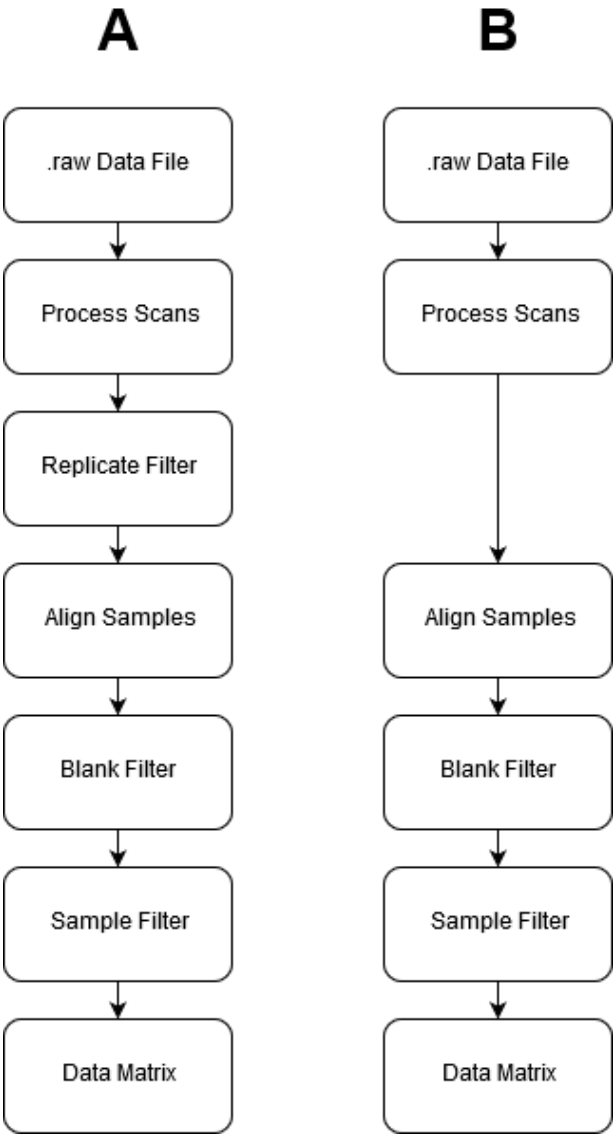


FIGURE 4.5: **A** - Typical DI-MS processing workflow including the replicate filter which is used to filter peaks present in x out of n technical replicates. **B** - LESA-MS data processing workflow. No replicate filter step due to the filter step being.

step produces a data matrix of peaks and corresponding intensities within each data file.

Blank Filter - For each feature detected, a peak has to be present in at least 75% of all samples and have an mean intensity of greater than five times the mean blank intensity to be retained.

Sample Filter - For each feature to be retained, it has to be present in at least 80% of all samples within a sample class. A peak can still be retained if not present in at least 80% of samples in one class as long as it is present in at least 80% of sample in another class.

4.3.4 3D Printing

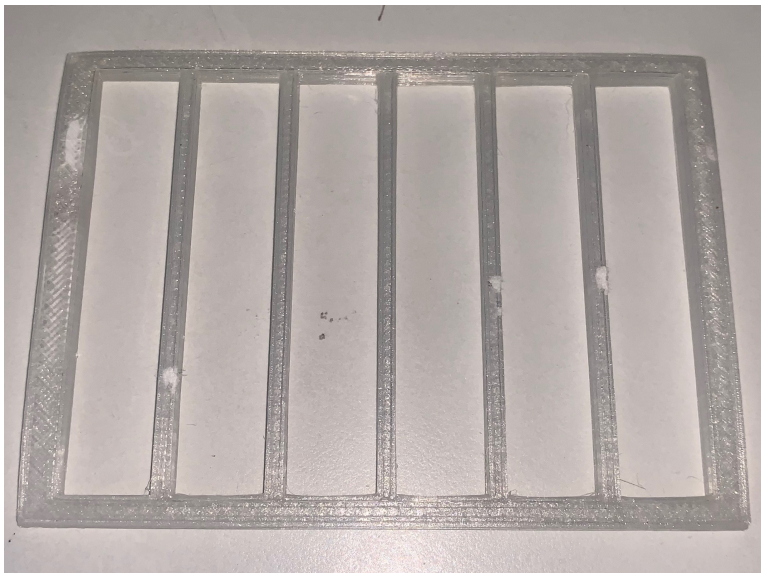


FIGURE 4.6: 3D printed template used to ensure consistent height between DBS

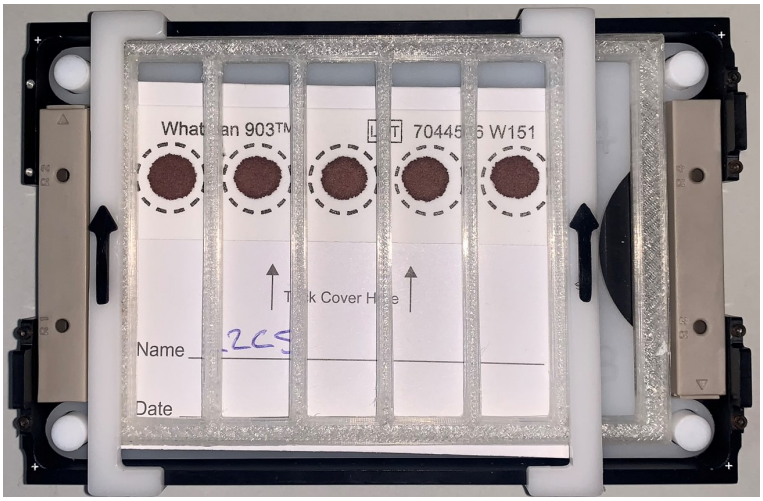


FIGURE 4.7: Picture demonstrating the use of the 3D printed template in use with a DBS collection card and LESA adapter.

Due to the way DBS cards are held in place in the LESA template on the Advion Triversa Nanomate device, there can be variations in the height of each DBS due to only having two anchor points. Therefore an additional 3D-printed template was produced to ensure a consistent height of each DBS through the addition of more anchor points which keep each spot at the same level. Photographs of the template can be seen in Figure 4.6

4.3.5 MS Data Acquisition

To increase sensitivity and dynamic range when performing DI-MS analysis, it is favourable to perform a SIM-stitch style of data acquisition. However due to the nature of surface sampling and only being able to sample a spot once due to the nature of altering the surface of a spot after a single extraction, it is necessary to perform an altered SIM-stitch workflow compared to the typical workflow of performing 4 identical replicate infusions.

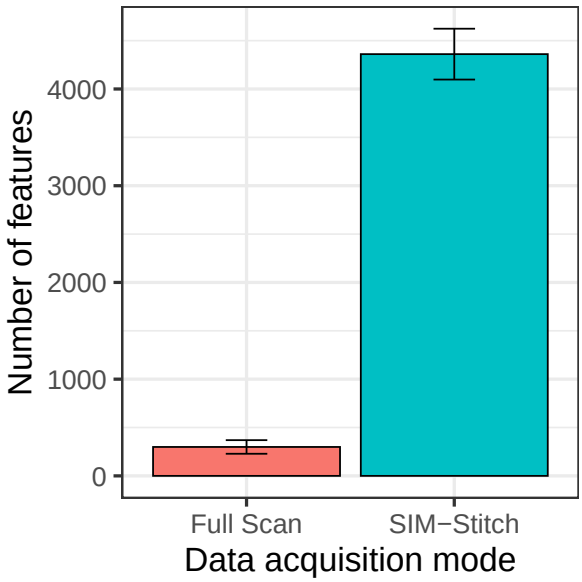


FIGURE 4.8: Comparison of the total number of features detected using a full scan and SIM-stitch acquisition mode.

To perform this a workflow was developed which involved utilising multiple SIM-stitch acquisitions within a single DI-MS infusion. To achieve the same effect of a replicate filter that is utilised in typical DI-MS data processing, peaks are filtered out instead during the Process Scans function during data Analysis.

A comparison of the number of features detected between full scan and SIM-stitch acquisition modes is shown in Figure 4.8. A clear increase in the number of features detected demonstrates that a modified SIM-stitch procedure as demonstrated in Figure 4.8 improves both the number of features detected but also increases the linear dynamic range overall. This is to be expected and has previously been shown by Southam et al. [121], however this is the first method which applies a spectral stitching method to the collection of LESA data.

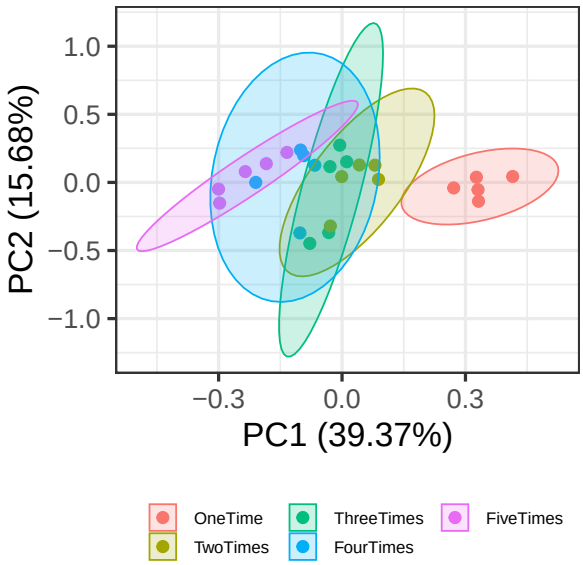


FIGURE 4.9: PCA plot displaying the effect of sampling the same DBS multiple times.

4.3.6 Repeated Sampling

Due to the inherent nature of surface sampling, it is critical to check whether or not sampling the same spot multiple times has an effect on the data collected. To check for this, the same spot was sampled multiple times using a 5 second dwell time. In total, five different spots were sampled 5 times each to investigate the difference made through multiple sampling.

Initially a PCA plot was produced from all high quality features and can be seen in Figure 4.9.

From Figure 4.9 we can see a distinct separation between groups that have been sampled for a set number of extractions. The greatest separation between groups is between samples that have been analysed for the first time and then subsequently analysed again. From this plot, it initially shows that the samples are most similar when they have been sampled between two and four times, before further variation

TABLE 4.4: Total number of features present after each subsequent extraction and the total number of features demonstrating an RSD of less than 30%.

Number of Times Sampled	Total Features present in at least 60% of all samples	Total Features with an RSD of less than 30%
One	4062	3855 (94.9%)
Two	3879	3662 (94.4%)
Three	3967	3464 (87.3%)
Four	3874	3725 (96.2%)
Five	3697	3623 (98.0%)

becoming apparent once sampled for a fifth consecutive time, potentially indicating that there may be a potential for re-sampling of the same spot.

Table 4.4 shows that there is a small loss in the total number of features that are present in at least 60% of all samples, when moving from a sample being analysed once against a sample that has been analysed two to four times. With a further subsequent drop in the number of features detected if a sample is analysed five times.

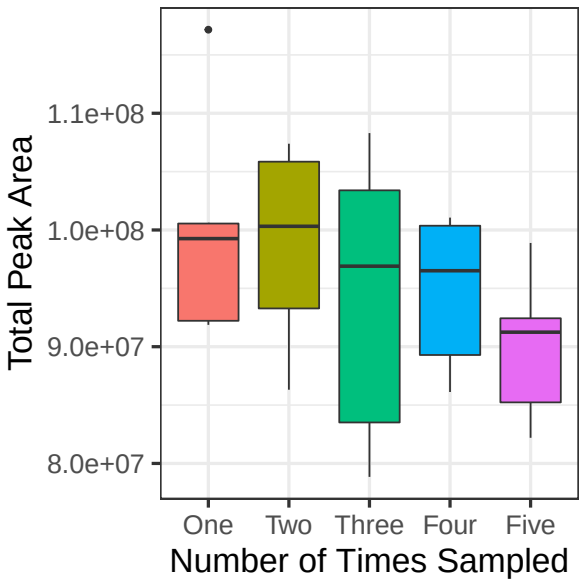


FIGURE 4.10: Total peak area of all features depending on the number of times a spot has been sampled.

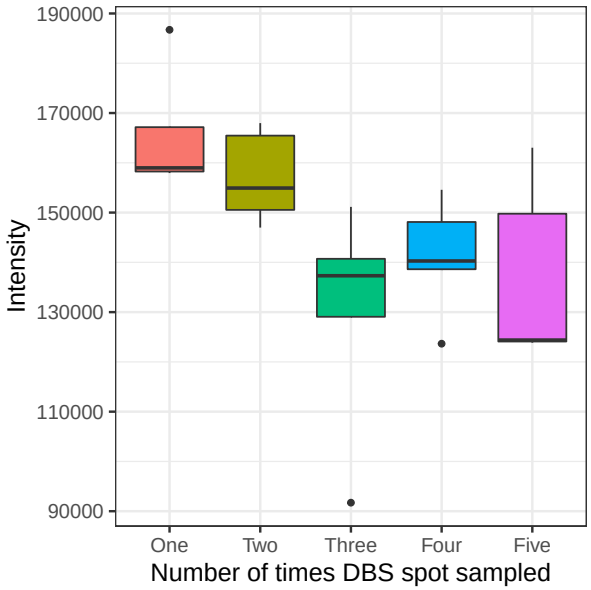


FIGURE 4.11: Intensity of tryptophan across sampling the same spot multiple times.

Furthermore, in Figure 4.10 a decrease in the total peak area depending on the number of times a spot has been sampled, shows agreement with Table 4.4 with regard to there being a gradual loss of data. Although the total peak area for samples which have been sampled once or twice is shown to be similar, when combining the gradual loss of features shown in Table 4.4 and the clear class separation in the PCA plot in Figure 4.9 between samples that have been analysed once against samples that have been analysed multiple times, it is clear that the surface of the spot is being modified too much by repeated sampling and DBS should only be sampled once. Otherwise, non-biological variation caused between samples which analysed multiple times, may mask genuine biological variation present between the samples.

Finally, in Figure 4.11 an example of the intensity of a specific feature is shown to demonstrate that after sampling multiple times, the intensity of features is dropping to the point that any sort of relative quantitation that is required for non-targeted

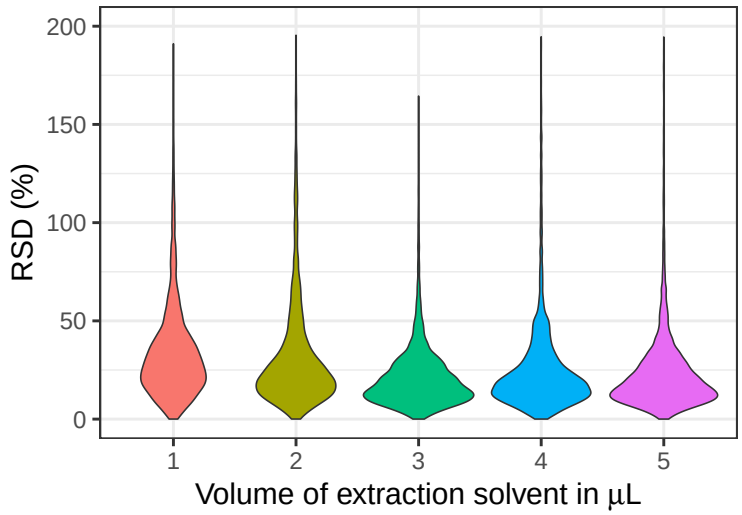


FIGURE 4.12: Violin plot of RSD's for all features depending on the volume dispensed onto the surface of the DBS.

metabolomics is not possible.

4.3.7 Solvent Dispense Volume

To decide on the total dispense volume to use, multiple spots were sampled with varying volumes of solvent deposited onto the surface of a DBS. Solvent was then allowed to dwell for 5 seconds before being aspirated back into the conductive pipette tip and infused directly into the MS. In total, 6 μL was aspirated from the solvent well and then five separate volumes were dispensed ranging from 1 μL to 5 μL.

Following the acquisition and filtering of data, the range of RSDs for all features was plotted and can be seen in Figure 4.12.

As the volume of solvent that has been deposited onto the spot increases, the average RSDs and distribution of RSDs lowers, indicating an increase in quality of the data. A potential cause for the increased RSDs at lower dispense volumes could be the decrease in surface area that is sampled from and therefore any slight fluctuations in

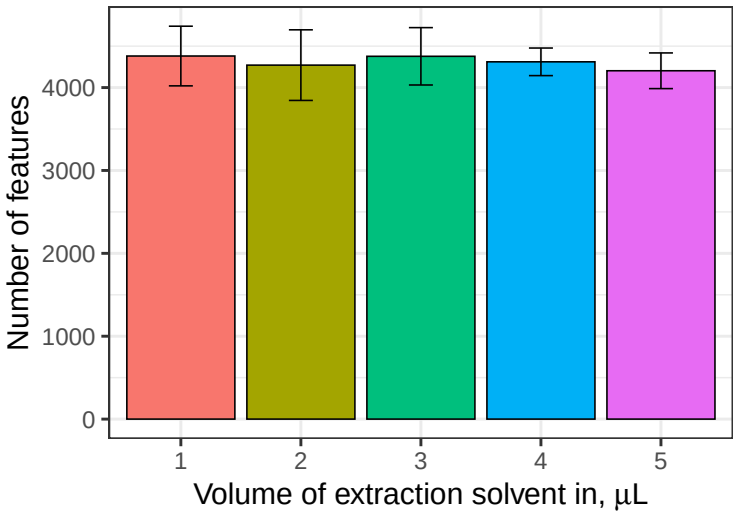


FIGURE 4.13: Boxplot describing the number of features detected depending on the volume dispensed onto the surface of the DBS.

the height of the card, even with the 3D printed template in use, could affect the total surface area of the DBS that is in contact with the extraction solvent.

In addition to RSD's the total number of features present in each group are described by Figure 4.13.

Figure 4.13 shows no clear statistically significant difference (p=0.969, one-way ANOVA) between the total number of features extracted from DBS and the volume dispensed onto the surface.

Whilst an increase in the number of features present was expected, this was not observed. An increase was expected due to the larger surface area of sample that is in contact with the LMJ during extractions from higher volumes of solvent. The reason for this expectation not be observed could be down to the minimal increase in the contact area between the DBS and LMJ. Whilst this contact area was not measured directly, visually, the increase in surface area that was extracted was not observed to be major when moving from a 1 μL extraction to 5 μL extractions.

Finally, as briefly mentioned in Section 4.3, stability of the nESI process is key in obtaining reliable and reproducible data. Whilst performing the above experiment, it was noticed that the current read back from the instrument was gradually decreasing as time progressed through the infusion of the sample into the MS. This is shown in Figure 4.14.

The drop in current could potentially be a problem for non-targeted studies due to it potentially representing the concentration of sample reducing as it is infused into the MS. Therefore due to a constantly decreasing ion signal, as each SIM-stitch 'loop' is undertaken, the concentration of any specific ion could be decreasing, thus an unreliable intensity value will be recorded.

The cause of this decrease in ion current could be due to the extracted metabolites being present in higher concentrations in the solvent closest to the beginning of the conductive pipette tip. The decrease in current would then be more substantial for the extractions where the total proportion of solvent which is dispensed onto the surface of the spot is less, for example 1 μL (16.7% of overall solvent volume) compared to 5 μL (83.3% of overall solvent volume).

To potentially overcome this issue a mixing procedure could be applied where the solvent is aspirated and re-aspirated but not onto a sample. However this was not available on the system used for this development therefore the technique was not used.

As a result, when higher proportions of the total volume are dispensed, the extracted metabolites are at a more consistent concentration throughout the total volume of solvent and therefore the intensity of ions infused into the MS will be more consistent.

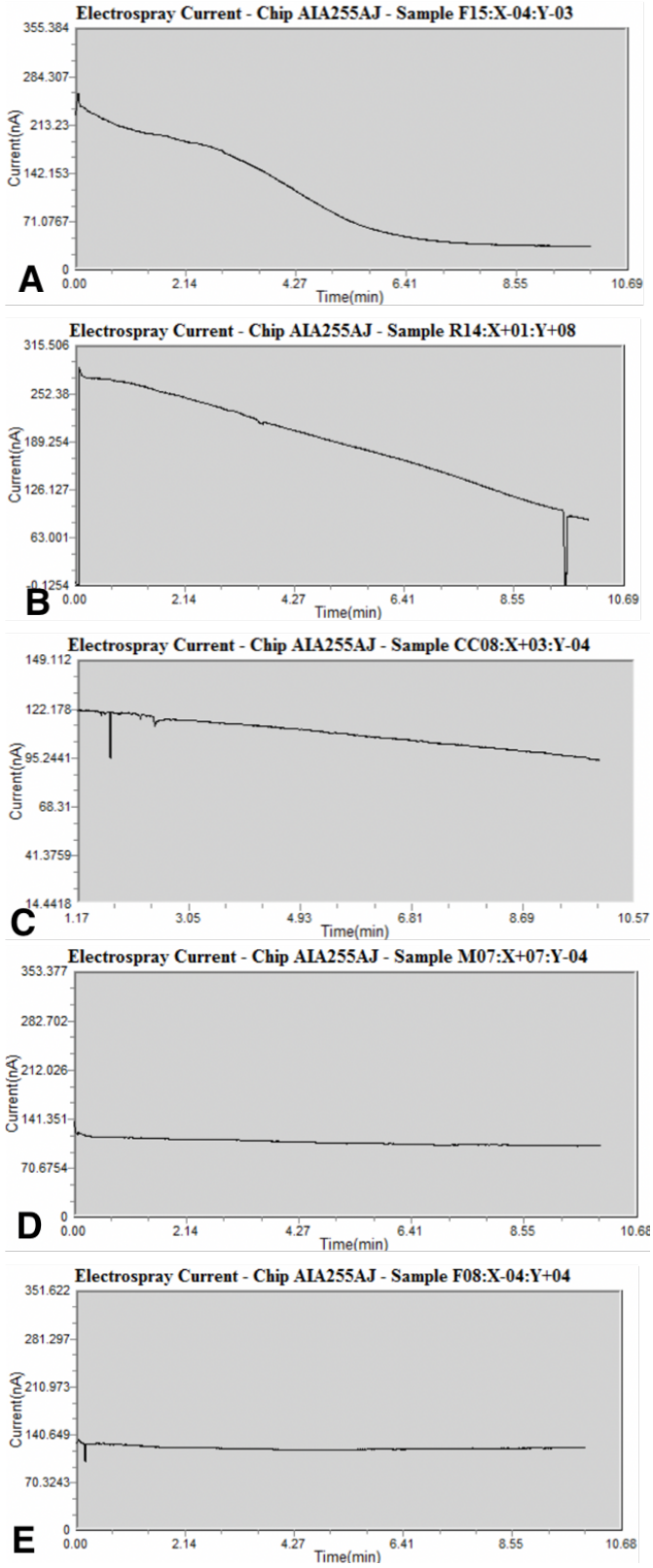


FIGURE 4.14: Current readings from instrument against time. A - 1 μL, B - 2 μL, C - 3 μL, D - 4 μL, E - 5 μL.

4.3.8 Solvent Dwell Time

Another key parameter to optimise for LESA is the length of dwell time on the surface of the spot. To check the effects on extraction, multiple dwell times ranging from 1 second up to 30 seconds in varying increments. Increasing dwell time would have the expected effect of increasing the intensity of metabolites that have been extracted from the DBS directly. However it is the opposite that was seen for the majority of metabolite related peaks seen within the sample. This could be described by the increase in dissolution of metabolites throughout the solvent plug within the conductive pipette tip as dwell time increases. Whereas during a 1 second extraction, the metabolites that have been extracted from the surface of the spot are 'frontloaded' within the solvent plug within the conductive pipette tip. Therefore, as this is the first part of the solvent plug that is infused into the MS, intensities are higher than those where metabolites have been allowed to diffuse into the remainder of the solvent due to the increased dwell time.

Figure 4.15 shows that there are a large number of features which show either a positive or negative correlation depending on whether or not a sample is analysed for a short or an extended duration.

Further investigation into this shows that if a feature shows a positive correlation with sampling duration, then this is likely to fall within a specific *m/z* region which can be seen in Figure 4.16 This is in comparison to features which show a negative correlation where they are more likely to be distributed over the whole *m/z* region which has been collected.

The features which demonstrate a positive correlation depending on the length of time sampled may be due to protein related small molecules such as heme of which the intensity has been plotted in Figure 4.18. To further investigate this, the distribution of

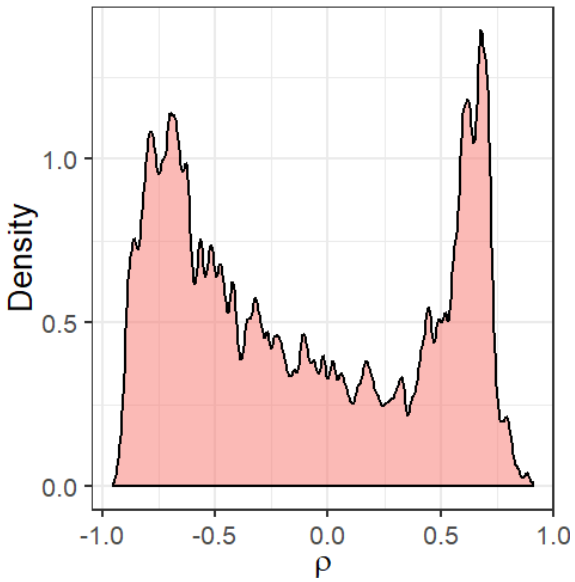


FIGURE 4.15: Distribution of Pearson's r correlation coefficients (ρ) across all reliable features.

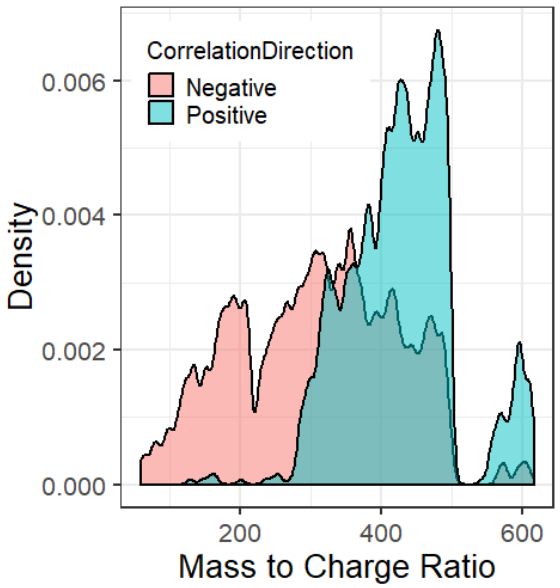


FIGURE 4.16: Distribution of positively and negatively correlated features across the *m/z* range (50 Da - 620 Da)

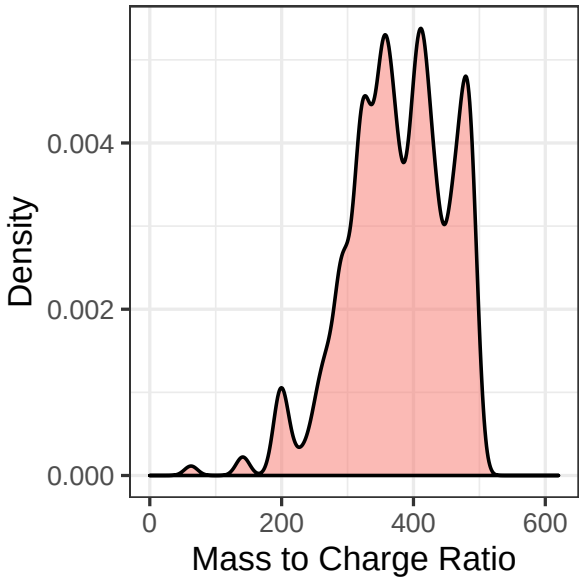


FIGURE 4.17: Distribution of features which are multiply charged across the *m/z* range measured.

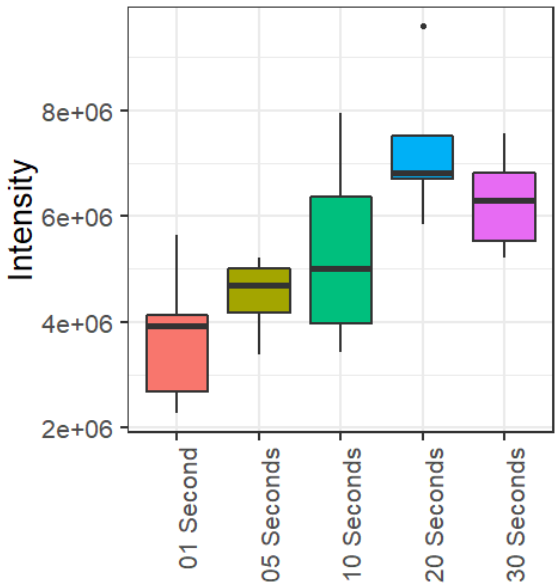


FIGURE 4.18: Intensity increase of heme as LESA dwell time on the surface of a DBS is increased.

multiply charged features present within the dataset has been plotted in Figure 4.17. What can be seen in Figure 4.17 is that the distribution of multiply charged features across the *m/z* range analysed, matches the distribution of positively correlated features as demonstrated in Figure 4.16. Therefore it can be assumed that the positively correlated features are protein related molecules.

Explaining the positive correlation could either be down to the high concentration of heme within the DBS as it has not been removed with prior sample preparation techniques or the release of heme from its protein complex as it is denatured by the organic extraction solvent present.

4.3.8.1 Method Recommendation

Due to the decreased RSD's as well as the minimal loss in intensity for metabolic features, five seconds has been chosen as a suitable dwell time moving forward throughout the optimisation.

4.3.9 Contact vs. Non-Contact LESA and Depth of Dispensation

LESA can be performed in one of two modes, Contact or Non-Contact. It has been shown before with bacterial colonies [236] that performing LESA applying a contact method where the conductive pipette tip is physically touching the sample, allows additional proteins to be extracted from the sample. However by performing a contact method of analysis, the surface area of the extraction is severely reduced and limited to the internal diameter of the conductive pipette tip itself.

Additionally, it is also key to check whether the height that the solvent is dispensed from the pipette tip to form the LMJ has any major effect on the extraction from the DBS. To determine the best mode to operate in, repeat infusions (n=5) of DBS were

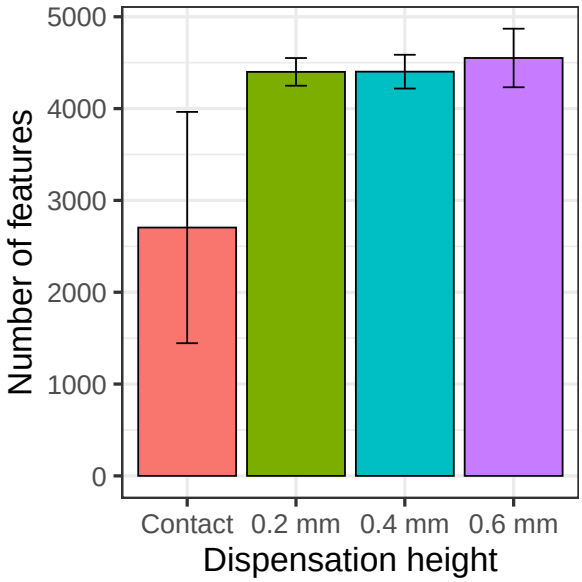


FIGURE 4.19: Total number of features detected depending on the height that the extraction solvent is dispensed onto the surface of the DBS.

made where the dispensation height of the extraction solvent above the DBS was adjusted from being in contact with the surface of the DBS to being 0.6 mm above, in 0.2 mm increments.

Figure 4.19 and 4.20 demonstrates that when the pipette tip is in contact with the surface of the DBS, the reliability of the extraction is compromised with widely variable number of features being extracted between replicate infusions and an RSD range which is unsuitable for non-targeted experiments.

To further confirm that the extraction efficiency when the conductive pipette tip is in contact with the surface of the sample is much less than when the tip is above the surface, the total peak area has been plotted and can be viewed in Figure 4.21.

With regard to the samples that were extracted using dispensation heights that were non in contact with the sample, the RSD's and number of features that were extracted were much more consistent. Applying a dispense height of 0.4 mm achieved the lowest

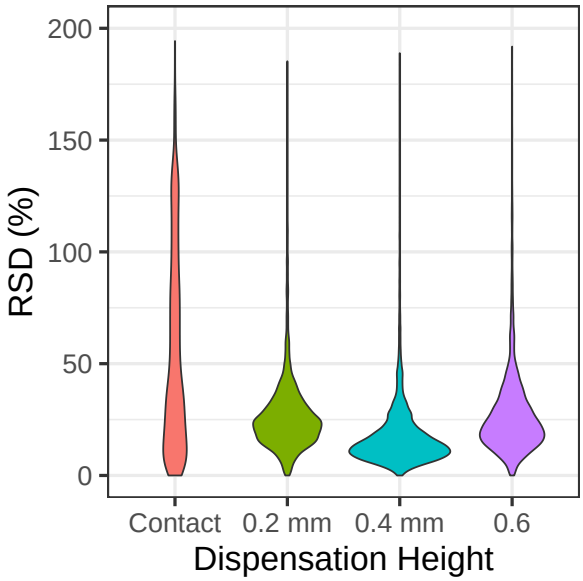


FIGURE 4.20: Violin plot describing the distribution of RSD's of all features detected depending on the height that the extraction solvent is dispensed onto the surface of the DBS.

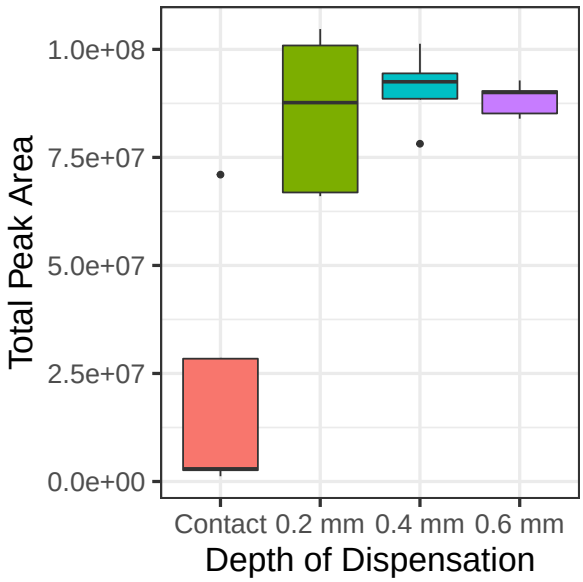


FIGURE 4.21: Total Peak Area observed depending on the height of the conductive pipette tip when extracting from the surface of the spot.

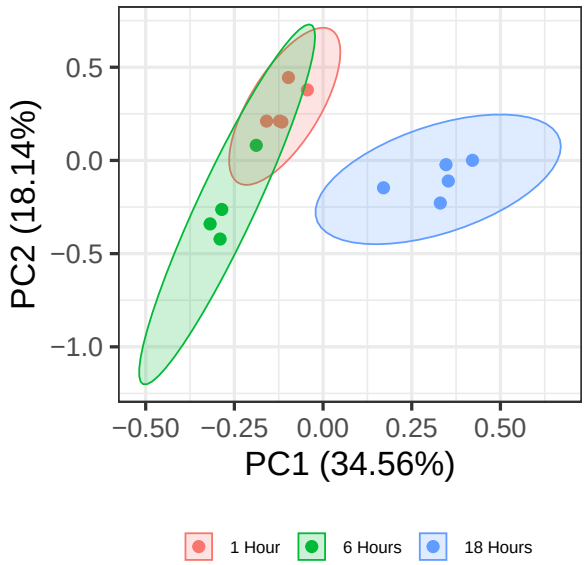


FIGURE 4.22: PCA scores plot demonstrating the large variance between samples analysed off cards left out in ambient conditions for different times prior to analysis.

range of RSD's accross all features and was therefore chosen as the dispense height used in the method.

4.3.10 Card effects

To demonstrate the need for DBS samples to be kept in similar conditions throughout storage before analysis, multiple sets of DBS were left out in open laboratory conditions at different times before being analysed alongside cards which were kept inside foil bags alongside dessicant.

Cards were left out on the bench top for 18 hours and 6 hours alongside cards which were only removed from packaging immediately before analysis.

After data acquisition and peak filtering, the following PCA scores plot was produced.

As shown in Figure 4.22 it can be seen that leaving cards out in different conditions has a major effect on the reproducibility of sampling spots. Therefore it is key to ensure that samples are collected and stored at the same time, temperature and conditions to ensure good reproducibility between samples.

The cause of this variation could potentially be down to a variety of reasons. Through being exposed to the ambient environment for significantly longer than samples which were kept in storage in controlled environments, samples have therefore been exposed to different conditions. DBS left in the environment with no procedures in place to control the humidity or temperature could lead to oxidation or changes of water content within the sample. This would not happen or would happen minimally in a controlled environment.

4.4 Results and Discussion - Comparison between two samples

To compare the developed LESA method against current MS approaches for metabolomics analysis of blood samples, two samples of rat whole blood were purchased from a commercial supplier (one sample collected from a male rat and another from a female rat). Comparisons were then made using UHPLC-MS and DI-MS and the LESA-MS method developed throughout this chapter.

TABLE 4.5: Number of features present in each LESA-MS and DI-MS datasets describing both the number of unique and common metabolic features

Analytical Method	Polarity	Total Features	Unique Features	Common Features
DI-MS	Positive	2694	1088	1606
LESA-MS		4211	2605	
DI-MS	Negative	4962	3715	1247
LESA-MS		4787	3540	

4.4.1 Results and Disussion

4.4.1.1 Presence of common and unique features between DI-MS and LESA-MS datasets

Following data pre-processing the total number of features present in each dataset after peak filtering is described in Table 4.5.

Between the two nESI based methods a direct comparison can be made between what can be assumed to be the same feature and the total number of common features can be calculated. Direct comparisons are only made between DI-MS and LESA-MS due to the similarity in which samples are analysed through the use of the same nESI procedure. However a point to note is that features detected using DI-MS may also be detected applying LESA-MS method except potentially as an adduct instead of the molecular ion.

From Table 4.5 it is shown that the total number of features that were common between the LESA-MS and DI-MS datasets was 1606 for positive ion mode and 1247 for negative ion mode. The difference in the total number of features detected can potentially be explained by several factors. Through using LESA-MS the sample itself has undergone no prior sample preparation before the direct extraction off the surface of the spot and subsequent analysis. Therefore additional metabolites can be expected

to be seen due to the presence of additional intercellular components of blood that are still present within the sample.

With regard to the unique features present within the DI-MS datasets which represent 40.4% of the total number of features present in positive ion mode and 74.9% of features in negative ion mode, these additional features can be explained through the additional sample preparation that is performed. This undoubtedly increases the extraction efficiency of metabolites from the sample whilst also removing protein related signal.

4.4.1.2 RSD performance

Finally a direct comparison can be made looking at the range of RSD's for every feature within the multiple datasets and for each class.

From Figure 4.23 a clear increase in the range of RSD's can be seen moving from UHPLC-MS to DI-MS and then finally to LESA-MS. This increase in average RSD's can be down to the different analysis procedures used for each analytical method. In UHPLC-MS, the flow rate through the ESI is much higher than that of the nESI techniques running at 0.5 μL/min compared to the high hundreds of nL/min which is used for nESI. This increase in flow rate leads to a much more stable electrospray.

4.4.1.3 Multivariate analysis

To ensure that the LESA method developed can detect the difference between the two different samples of DBS produced, a PCA scores plot of each method and polarity was produced. Only features which were present in both data sets were used to produce each PCA to ensure that any information which is unique to each method has no effect

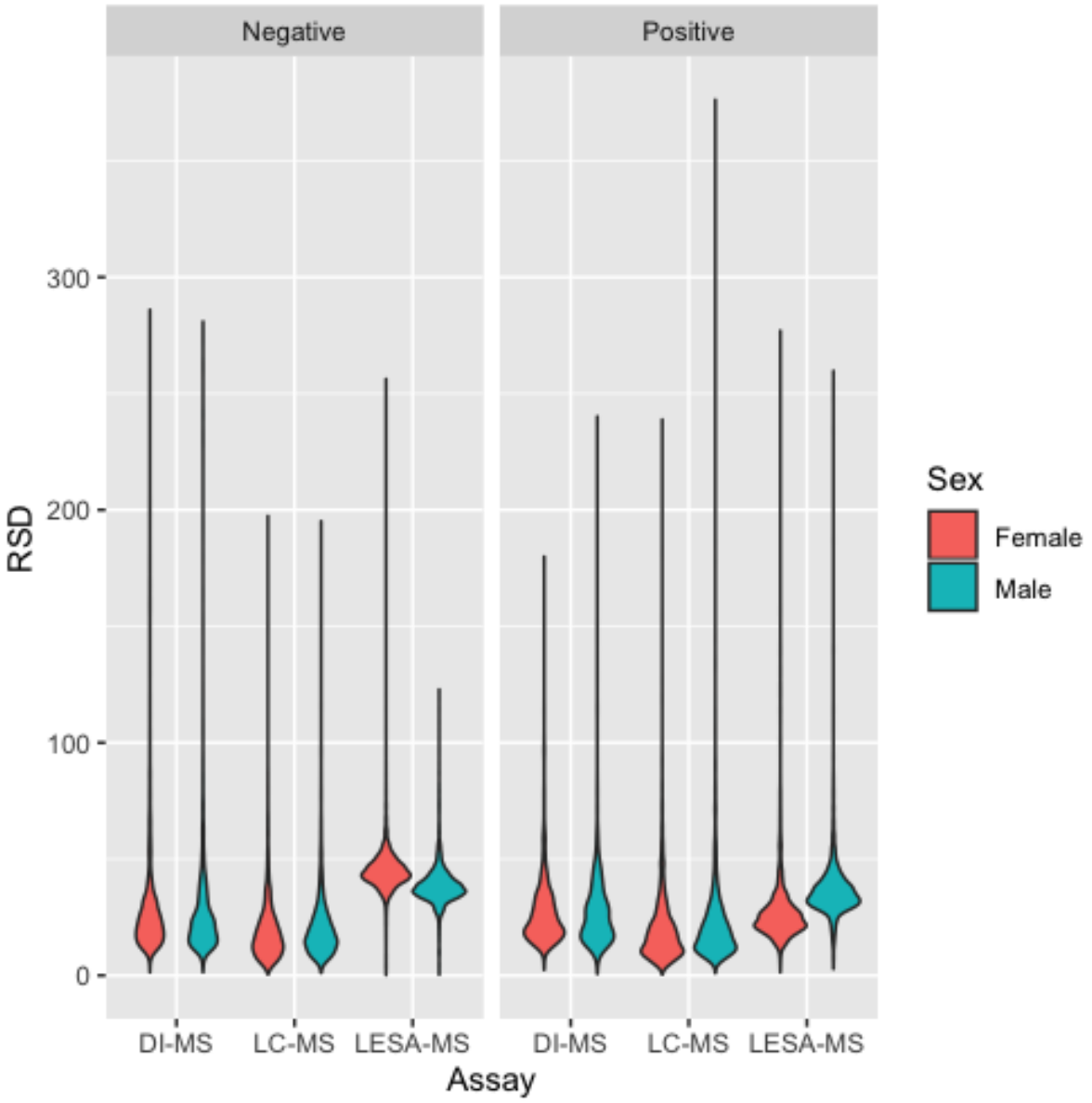


FIGURE 4.23: Distribution of RSD's in each dataset for each class applying all methods.

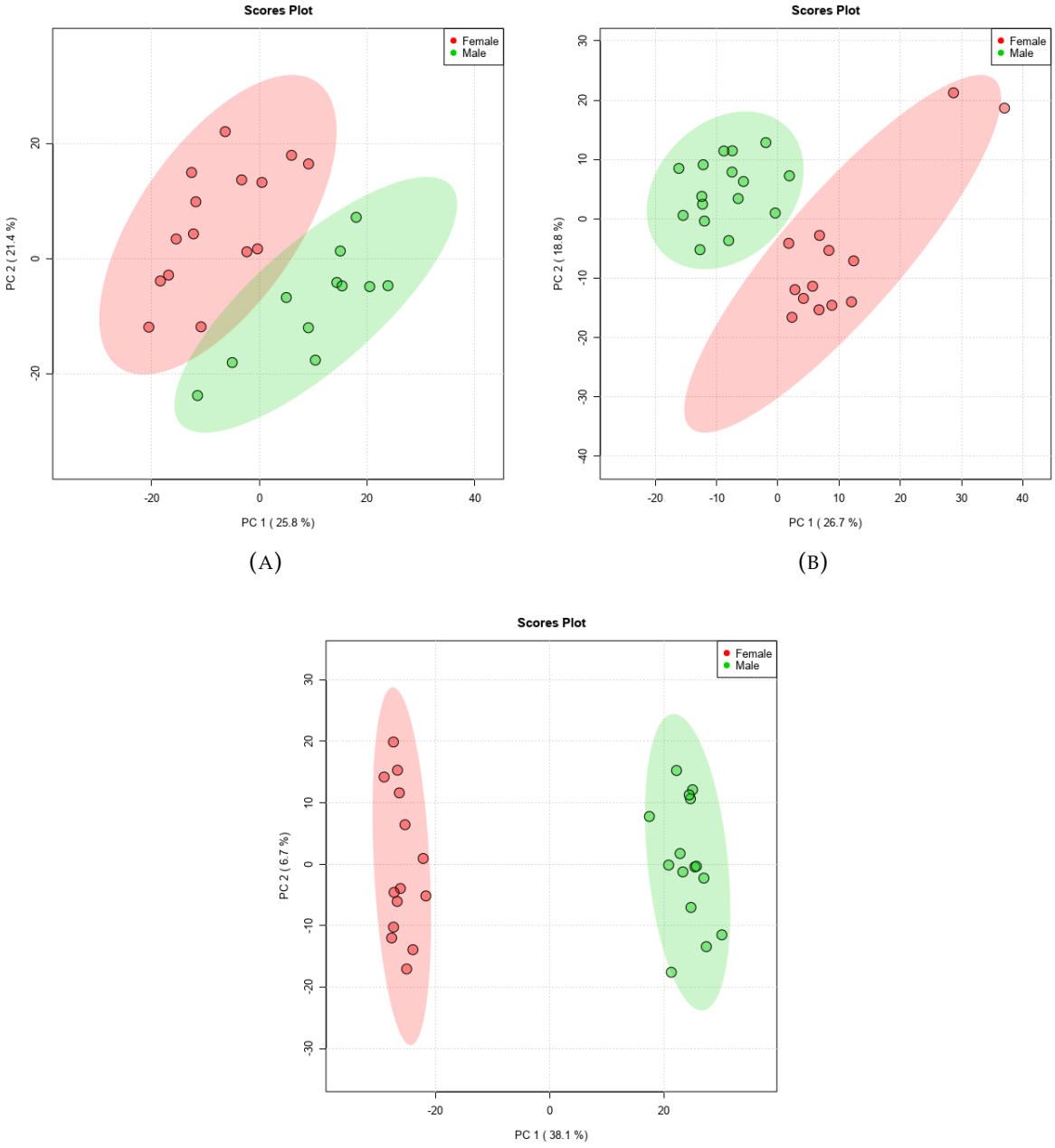
on the total variance explained by the PCA and also to ensure that a fair comparison between both the LESA-MS and DI-MS methods can be made.

As can be seen by Figure 4.24, clear separation between the two sample types (male and female) is seen meaning that the LESA-MS method developed is capable of detecting the same explanatory variable that produce the separation between the two groups as the DI-MS method.

However whilst the variance explained by PC1 is similar for both DI-MS and LESA-MS, the comparison against UHPLC-MS data is vastly different. This increased separation as shown in Figure 4.24 & 4.25 for UHPLC-MS data, demonstrates the benefit of applying chromatography, even though the technique may not be as sensitive as when applying the nESI method. Through the separation of compounds, the detection of features that are real features and not noise improves the quality of the data through increased confidence in the compounds detected. Additionally, features which may have been isomeric using a direct infusion method, are chromatographically separated and can be distinguished, giving additional information as to which features are driving the seperation between the two groups.

4.4.1.4 Univariate analysis

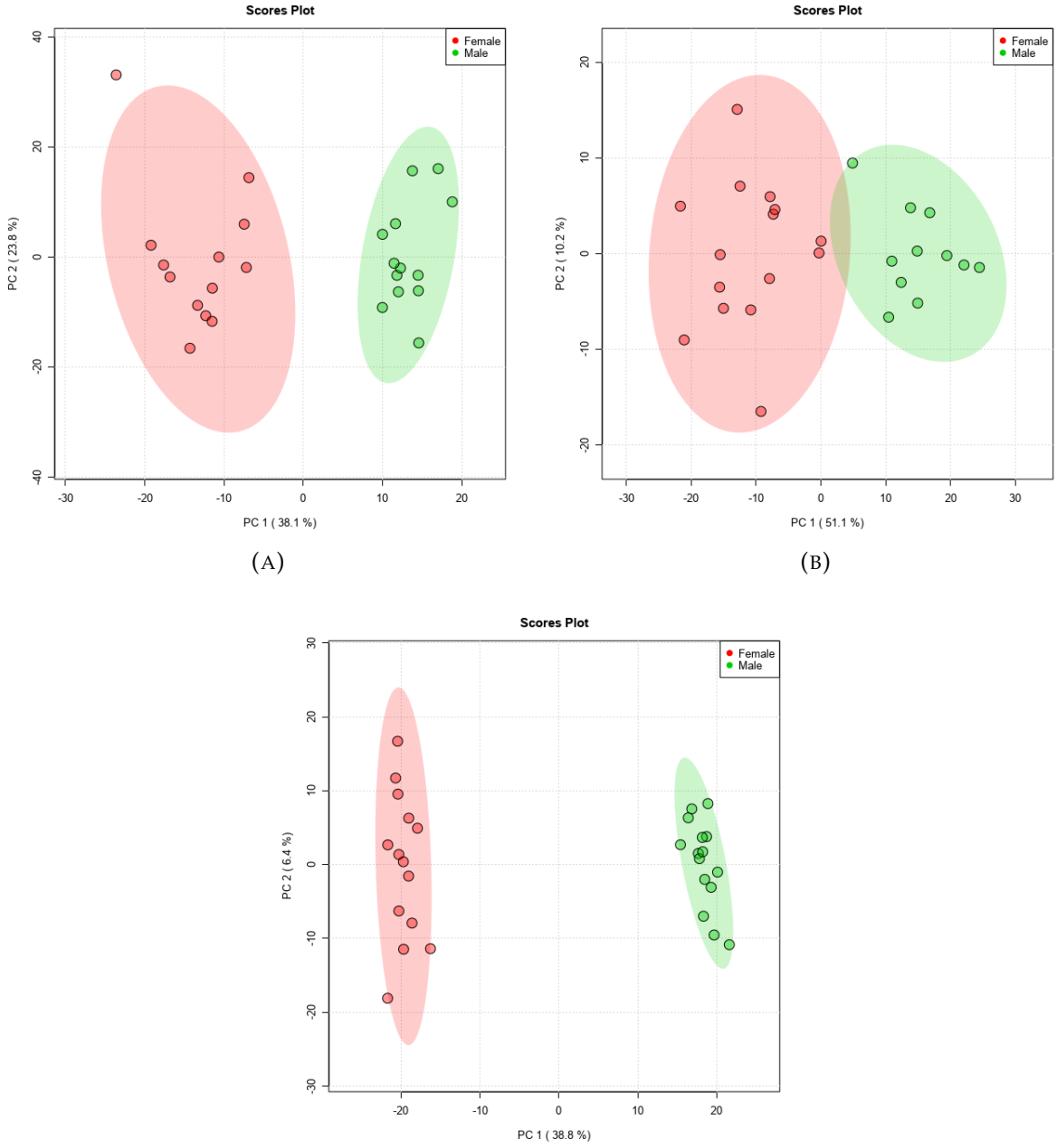
In addition to the multivariate analysis performed which determined that each of the variances between the male and female samples could be distinguished by the developed LESA-MS method and has comparable performance to DI-MS. A comparison between the number of significant features in each datset can be made. To determine the significant features, both a Wilcoxon signed-rank test was performed in addition to a volcano plot.



(A) (B)

(C)

FIGURE 4.24: PCA scores plots for all datasets. For LESA-MS and DI-MS only features which are present within both datasets are used to produce the scores plot. (A - DI-MS Positive, B - LESA-MS Positive, C - UHPLC-MS Positive). Unequal class sizes are down to samples being removed due to poor quality data at the filtering stage and failed infusions.



(A) (B)

(C)

FIGURE 4.25: PCA scores plots for all datasets. For LESA-MS and DI-MS only features which are present within both datasets are used to produce the scores plot. (A - DI-MS Negative, B - LESA-MS Negative, C - UHPLC-MS Negative). Unequal class sizes are down to samples being removed due to poor quality data at the filtering stage and failed infusions.

TABLE 4.6: Total number of features determined as significant by a Wilcoxon signed-rank test (p <0.05 (FDR corrected)) and the number of features with a p value <0.05 and a log2 fold change value of >2. Note: The total number of features for DI-MS and LESA-MS represent only those that are common between the two methods.

Analysis Method	Polarity	Number of Features	Significant Features	
			Wilcoxon signed-rank test	p <0.05 & log2 FC >2
LESA-MS	Positive	1606	650	158
	Negative	1247	720	227
DI-MS	Positive	1606	601	164
	Negative	1247	745	179
UHPLC-MS	Positive	5081	1876	628
	Negative	2956	1677	751

Table 4.6 demonstrates very similar performance between both the number of significant features as determined by a Wilcoxon signed-rank test and those features which have a p value of <0.05 and a Log2 fold change of greater than 2. However Table 4.6 also shows that UHPLC-MS detects a much greater number of significant features than both DI-MS and LESA-MS. This also explains the increased separation present in the PCA plots in Figures 4.24 & 4.25

4.4.1.5 LESA Analysis Pros and Cons

Through the development of the LESA method and the application of it to two separate samples it has been shown that LESA can be used to separate two separate sample classes. However associated with the analysis performed there are downsides which have been identified.

In Table 4.7 a SWOT analysis has been performed which discusses the strengths and weaknesses of the developed method whist also mentioning a few opportunities afforded by the method and threats which may prevent the method from gaining widespread use. Table 4.8 also compares the three methods used and looks at the

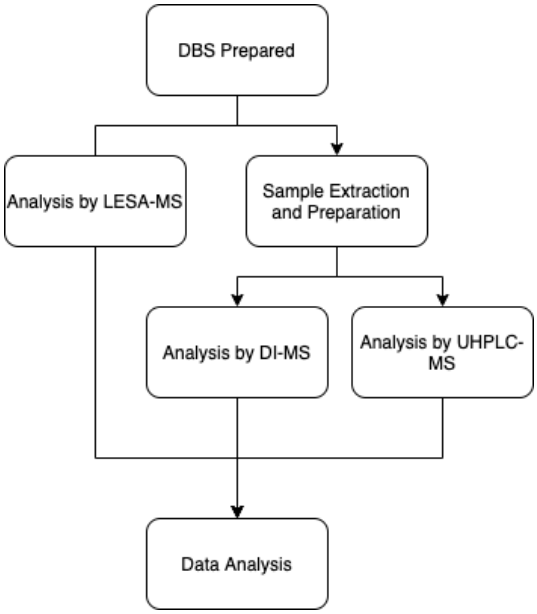


FIGURE 4.26: General workflow applying a LESA-MS method in comparison to DI-MS and UHPLC-MS demonstrating the advantage of no sample extraction and preparation before MS analysis.

individual pros and cons of each method separately.

4.4.2 Conclusion

In this chapter I have described an optimised LESA method capable of detecting thousands of metabolic features from a five second surface extraction.

The various different parameters that have been optimised have lead to various conclusions that should be used when applying LESA-MS to the extraction of DBS.

The use of an acetonitrile based extraction solution leads to a stable LMJ where other organic solvents have issues maintaining LMJ stability and thus compromising the ability to perform a LESA extraction. Applying a 'looped' SIM-Stitch MS method to collect reproducible data from a single spot and to increase the dynamic range and

TABLE 4.7: SWOT analysis of the developed LESA method with summaries of both the strengths and weaknesses of the method in addition to the opportunities afforded by the method and the threats that may prevent the method from gaining widespread use.

	Helpful (to achieving the objective)	Harmful (to achieving the objective)
Internal Origin	**Strengths** - Fast analysis times - Lack of requirement for prior sample extraction - Initial results show the ability to distinguish between multiple classes	**Weaknesses** - Samples cannot be frozen - Time required to allow a LMJ to form successfully - Replicate data is only collected from a single sample and not across multiple samples - Limitations in the types of solvent which can be used
External Origin	**Opportunities** - Could be used in clinical environments for fast screening of samples before further more in-depth analyses - Large scale studies in reduced times due to high throughput	**Threats** - No sample separation due to the direct infusion nature of analysis - Reproducibility not as high as other more in-depth analysis methods

TABLE 4.8: Pros and cons of the LESA-MS developed compared to DI-MS and UHPLC-MS

	Pros	Cons
LESA-MS	- No prior sample extraction necessary - Fast analysis times due to the lack of individual replicates, with all replicate data collected from a single sample - Highly sensitive due to the nESI source	- No external replication - Samples can only be analysed a single time - Limitation in the solvents that can be used - No separation leading to reduced confidence in metabolite annotation
DI-MS	- Fast analysis times - Highly sensitive due to the nESI source	- No separation - Due to replication, total analysis time per sample can reach run lengths that are matched by UHPLC-MS - No separation leading to reduced confidence in metabolite annotation
UHPLC-MS	- Slower individual sample analysis times when compared to DI-MS and LESA-MS - Separation of samples - Increased confidence in metabolite annotation	- Increased run times per sample - High costs involved - Peak broadening can occur over batches depending on age of columns used

sensitivity of the analysis over conventional full scan MS with this method being the first application of a SIM-stitch technique to the collection of LESA data.

Recommendations have also been made surrounding a number of key LESA parameters. Firstly it has been recommended that DBS should only be sampled once applying a LESA technique due to the surface changes and relative intensity changes that subsequently occur. Analysing a sample more than once therefore compromises the ability to have confidence in the relative quantification required to perform non-targeted metabolomics studies.

It has also been recommended that 5 μL of solvent is used for the sample extraction to ensure good reproducibility in any extraction as well as ensuring the sample is well mixed throughout the solvent plug in the conducive pipette tip before infusion into the mass spectrometer as smaller volumes of solvent demonstrate a decrease in spray current throughout analysis.

Whilst contact based LESA applications have been used in other 'omics' sciences, it has also been demonstrated within this chapter that a non-contact based LESA method is more applicable for non-targeted methods to ensure many metabolites are extracted from the surface of a DBS.

Through the application of LESA, high-throughput studies can be performed with each sample taking less than four minutes to analyse. Furthermore, without the requirement of sample extraction, preparation times are reduced and the potential areas for mistakes are limited.

The LESA-MS method applied here is able to detect a similar number of significant features between a single male and single female rat sample than what was detected by applying DI-MS, demonstrating that sufficient biological information is being extracted from the sample during the five second extraction.

Whilst the parameters presented here are recommendations, there is flexibility with some parameters. Depending on the type of sample analysed it may be useful to change the spray voltages or the height of dispensation, but with the nature of surface sampling and the fundamental change occurring when a sample is analysed, some recommendations are fixed such as only analysing a sample once.

However, there are limitations. With the loss of separation, the annotation of metabolites is made much more difficult and therefore the technique should only really be used as a rapid profiling technique rather than a confirmatory tool. Additionally, with the requirement that each sample can only be analysed once, multiple replicates are not possible thus there is no opportunity to check samples once again if the results require.

Further work is required however with a key step being the validation of the method in a biological application. Additionally it would also be useful to apply the method to other sample types, whether or not these are based on collection cards, to further understand the benefits and limitations of the method developed. It would also be interesting to investigate the effect of tertiary solvent systems involving additional solvents such as IPA. The addition of IPA may also be useful for the extraction of lipids at higher m/z ranges which was not investigates as part of this method development chapter. Also, through the development of the LESA technology which has occurred since the beginning of this project, it is now possible to apply separation to a LESA extraction, therefore it would also be interesting to compare the performance of LESA-LC-MS against UHPLC-MS in addition to the comparison of LESA-MS to DI-MS as was shown in this chapter. Finally, in the area of proteomics, the coupling of

high field asymmetric waveform ion mobility spectroscopy to LESA has been demonstrated [154], therefore it would be interesting to further develop a method for non-targeted metabolomics which uses this separation procedure whilst maintaining a high throughput method.

In summary, the LESA-MS method shows similar performance to traditional manual extraction followed by DI-MS. This could allow LESA-MS to be used in a metabolic profiling scenario where resources are limited and information is required quickly. However, it has been shown within this chapter, if the time and resources allow, then applying a UHPLC-MS experimental design provides more data than applying a direct infusion method (LESA-MS and DI-MS).

Chapter 5

Application and comparison of dried biofluid collection to whole biofluids in a non-targeted metabolomics investigation of Hypervitaminosis A

5.1 Introduction

Whilst the use of dried biofluid collection (DBC) is well documented for targeted metabolomics research [199, 237, 238], their application towards discovery based studies is largely limited. As mentioned throughout this thesis, the advantages gained from applying DBC allow for many applications where the cost of performing large scale population based studies is a barrier to many research questions[239, 240]. One area of research where population based studies are of huge importance are Vitamin A supplementation programs that are being executed in many developing countries around the world [241, 242].

Vitamin A deficiency in developing countries has been a major health concern since

the early 1990's and is on the list of the World Health Organisation (WHO) major nutrition topics to be investigated [243]. Due to the adverse health effects that being deficient in Vitamin A can bring, since the late 1990's the supplementation of Vitamin A into the diets of populations within developing world countries has been a major research focus. Supplementation programs are also delivered with large doses of Vitamin A given to children every 6/12 months at health and immunisation checks. Vitamin A is a key nutrient and has important functions in vision, immune competence and reproduction [244]. Evidence has suggested that these supplementation programs have been beneficial [245], however, whilst a person can be deficient in a specific group of nutrients, they can also be in excess leading to another set of adverse health effects on the other end of the scale.

Due to the method in which these supplementation programs are implemented, the population receiving the additional dose of Vitamin A, usually young children, receive infrequent, yet large doses of Vitamin A at a single point in time [244]. This supplementation of large doses of Vitamin A at a single point of time is potentially leading to Vitamin A levels increasing above the WHO recommended levels. Thus susceptibility leading to hypervitaminosis is of much greater concern where the metabolic consequences are not known.

Diet is a major contributing factor to the deficiency of Vitamin A [246]. With the body unable to synthesise Vitamin A directly, we rely on taking in Vitamin A either directly from our diet or through the intake of provitamins which are subsequently transformed into Vitamin A within the body. Many forms of Vitamin A that are present within food are present as retinol esters. These compounds need to be hydrolysed before absorption by the intestinal epithelial cells [247]. After absorption they are subsequently re-esterified where they are incorporated with other neutral lipid esters into

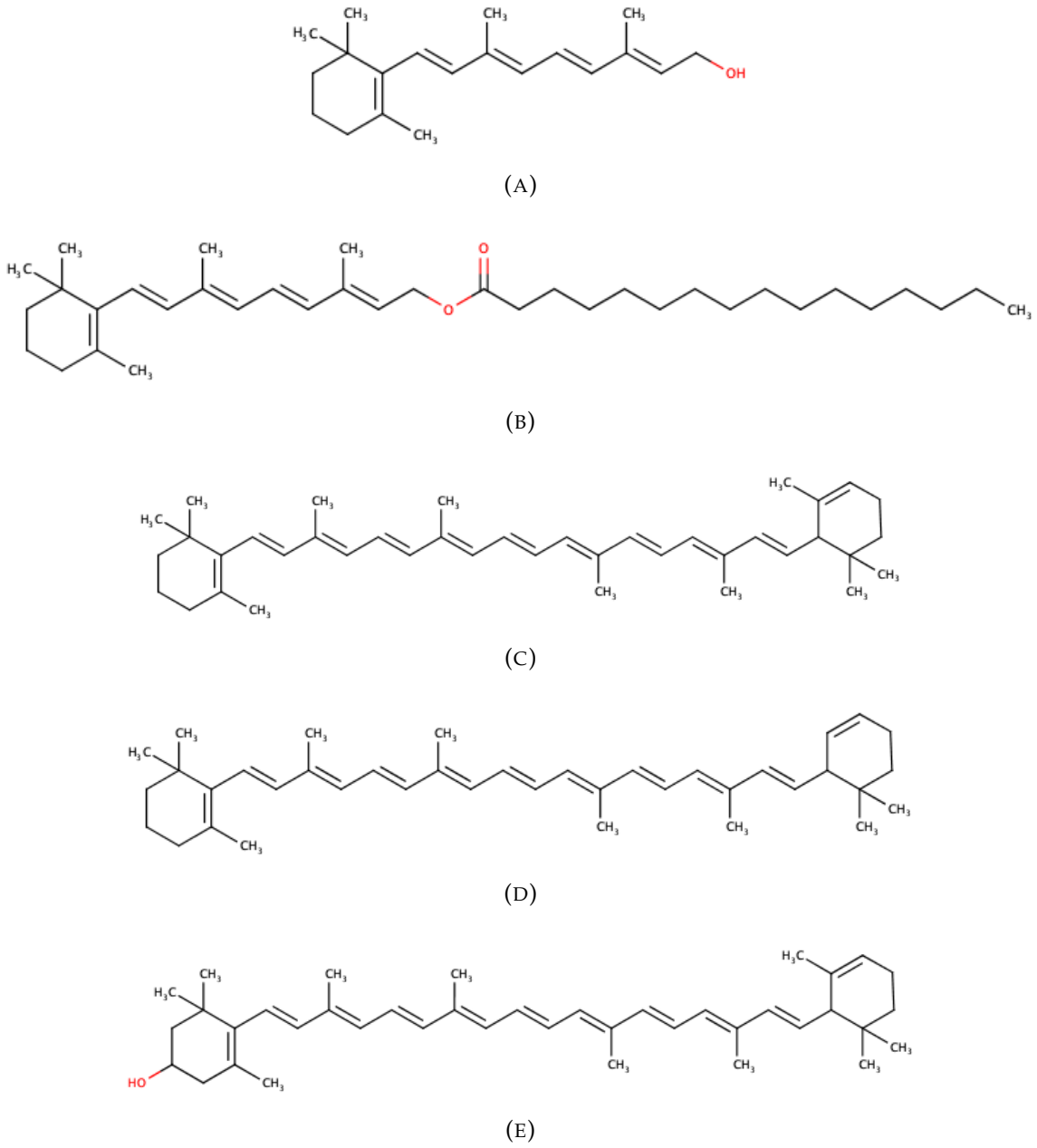


FIGURE 5.1: Structure of main forms of Vitamin A including the major sources of Vitamin A from dietary forms of provitamin A. (A) - Retinol, (B) - Retinyl Palmitate, (C) - α-Carotene, (D) - β-Carotene, (E) - β-Cryptoxanthin. Figure modified from [244].

chylomicra. The majority of the retinol esters are then transported and stored in the liver. [248, 249]

With the liver being the main tissue where Vitamin A products are deposited, due to its capacity to store and control the release of the lipid products, it is therefore of significant interest into looking at the effects of Hypervitaminosis A within populations. However, the collection of liver biopsies, for any analyses, not just metabolomics is problematic. The invasive nature of the collection means the cost of collecting samples is high and the number of people willing to go through with the procedure is very low, limiting the ability to perform large scale population based studies.

Alternatively, serum and plasma can be used as alternatives for non-targeted metabolomics analysis, however the cost of collecting, preparing and storing samples, can once again be prohibitive, if large scale studies are required. This is where dried biofluid spots can potentially aid in the research of large scale studies.

Whilst not identical to DBS, dried plasma and dried serum spots still represent some of the advantages which collection of DBS affords. Samples would still be collected in their native environment through venous blood draws, and then prepared into either serum or plasma as was necessary. Subsequently, accurate volume aliquots could then be deposited onto filter paper based collection cards and stored or transferred as needed. Providing facilities to prepare serum and plasma were available in sample collection locations, then many more replicates of a single sample of plasma or serum can be prepared and stored as dried biofluid spots in comparison to a single aliquot of serum/plasma. With the additional temporary stability of compounds imparted on the sample through storage as a dried form, cold storage is not needed andsending samples to collaborators in other locations, whether that be within or outside the country of collection is much cheaper, thus increasing the availability of analyses

that may be performed on a single sample.

Within this chapter two populations of young children (Philippines & Guatemala) which had undergone a Vitamin A supplementation program had serum or plasma collected and then subsequent dried biofluid spots prepared from the same samples. Non-targeted metabolomics applying a lipidomics UHPLC method was then undertaken with a comparison made between the serum and dried serum spots and conversely the plasma and dried plasma spots to identify any similarities in biological conclusion between collecting as a whole biofluid or as a dried spot. This study is a part of a larger scale study named GioVitas led by Georg Lietz at Newcastle University and funded by the Gates Foundation. Other studies falling under this larger ongoing project have been aimed towards defining the importance of lipids compared to polar metabolites and hence why lipids only were chosen for these dried samples. Finally, a comparison is made between the two populations to demonstrate that the biological outcomes of the two separate populations are demonstrating similar results indicating that the result is valid.

5.2 Materials and Methods

5.2.1 Materials

HPLC grade acetonitrile, isopropanol and water along with formic acid ($\geq$98%) was purchased from VWR Chemicals (Lutterworth, UK). Ammonium formate was purchased from Sigma Aldrich (Poole, UK).

5.2.2 Sample Collection and Preparation

5.2.2.1 Preparation of serum (Guatemala)

Whole blood was collected from each subject into BD Royal Blue Trace Element vacutainers before allowing the blood to clot at room temperature. After clotting, each vacutainer was centrifuged (1300g, 10 minutes, room temperature) and serum subsequently transferred into cryovials before being stored at -80°C until shipping back to the UK.

5.2.2.2 Preparation of plasma (Philippines)

Whole blood was collected from each subject into EDTA vacutainers before centrifuging (1300g, 10 minutes, room temperature). The plasma layer was then transferred into cryovials before storing at -80°C and subsequent shipping to the UK.

5.2.2.3 Preparation of dried plasma and dried serum spots

Following the preparation of the plasma and serum, 40 μL of each biofluid was accurately deposited onto Whatman 903 filter paper collection cards. Each of the dried samples were allowed to dry horizontally for 3 hours at ambient temperature before being placed into individual plastic zip-loc bags alongside desiccant packs.

All samples were shipped to the UK on dry ice before being stored at -80°C until extraction.

5.2.3 Sample Extraction for Philippines and Guatemala Dried Serum and Plasma Spots

Each sample was removed from storage at least 1 hour prior to extraction to defrost and then subsequently kept on ice for the majority of the extraction process. After defrosting, each spot was cut out of the collection card and into 8 smaller pieces to aid in the extraction of metabolites from the spot. 460 μL of IPA and 40 μL of water was then added to each sample followed by 20 minutes of gentle agitation using a vortex mixer. After agitation, each sample was centrifuged (20 minutes, 21000 g, 4°C) and 300 μL of supernatant subsequently transferred to a separate 2 mL Eppendorf tube. Samples were then dried down under vacuum at 35°C (Savant SPD111V SpeedVac Concentrator connected to a RVT5105 Refridgerated Vapout Trap, Thermo Scientific, UK) before being stored at -80°C until analysis.

5.2.4 Sample extraction for Philippines and Guatemala serum and plasma

As with DSS and DPS, samples were removed from storage 1 hour prior to extraction to defrost and kept on ice throughout the extraction process. After defrosting, 40 μL of each sample was transferred to an extraction Eppendorf where 460 μL of IPA was then added. Each sample was then thoroughly mixed for 30 seconds at high speed using a vortex mixer before being centrifuged (20 minutes, 21000 g, 4°C) and 300 μL of supernatant subsequently transferred to a separate 2 mL Eppendorf tube. Once more samples were then dried under dried down under vacuum at 35°C before being stored at -80°C until analysis. Alongside sample extraction, blank samples were extracted using identical volumes and solvents.

5.2.5 QC Sample Preparation

QC's were prepared from either a pooled sample of DPS and plasma extracts or DSS and serum extracts. From the QC pool, multiple QC samples were prepared by aliquoting out 300 μL of each sample into individual Eppendorf tubes before drying down along side extracted samples.

5.2.6 UHPLC-MS Analysis

When samples were ready to be analysed, all samples were removed from storage and allowed to defrost on ice for 1 hour prior to reconstitution. Each sample was then subsequently reconstituted in 96 μL of IPA/Water (75%/25%) before vortexing for 30 seconds at high speed using a vortex mixer. Samples were then transferred to low volume LC vials ready for analysis.

Samples was analysed in both positive and negative ion mode applying a lipidomic assay using a Vanquish UHPLC coupled to an electrospray Q Exactive Plus mass spectrometer (Thermo Scientific, San Jose, CA, US). Data was collected at mass resolution of 70,000 (FWHM at m/z 200). A 2μL injection volume was applied using a Thermo Hypersil GOLD column (2.1 x 100mm, 1.9μm), with a flow rate of 400μL/min flow rate and with two mobile phases, mobile phase A (10mM ammonium formate in 60% acetonitrile/water (v/v) + 0.1% formic acid) and mobile phase B (10mm ammonium formate in 90% IPA/acetonitrile (v/v) + 0.1% formic acid). A 14 minute gradient elution was applied as follows: 0 min, 20% B; 0.5 min, 20%B; 8.5 min, 100% B; 9.5 min, 100% B; 11.5 min, 20% B. QC's were injected every 5th sample, with 10 leading QC's prior to the collection of the first sample.

5.2.7 Data processing

Data was pre-processed from .raw format using to the open source .mzML format using ProteoWizard (version 3.0.1921). XCMS (version 3.8.0) was then used to perform peak picking and alignment (parameters are available in Appendix B). The resultant peak matrix was then filtered to remove low quality peaks and sampled. Each dataset was then inspected to identify if signal and batch correction was required. One dataset (dried serum spots - negative ion mode) required signal and batch correction which was performed using an in-house R package applying a QC-robust spline batch correction method as described by Kirwan et al. [233] Low quality samples were identified through visual identification and comparison of the total number of features detected and the sum of all intensities. For features to remain in the matrix, the following parameters had to be met; the mean of all QC samples had to be greater than the mean of the blank by a factor of 10, the RSD of all QC's had to be less than 30% and the feature had to be present in at least 60% of all QC's. Features were annotated using in house pre-release software applying only MS1 accurate mass identification along with adduct and isotope identification.

5.3 Results and Discussion

Due to the method in which samples were collected, a direct comparison can be made between samples which were collected either as the biofluid (e.g. serum or plasma) and those that were subsequently created as dried biofluids on filter paper collection cards (e.g. dried serum spots or dried plasma spots). Samples collected in Guatemala were either prepared as serum or dried serum spots, and those which were collected from the Philippines were prepared as plasma or dried plasma spots. Due to this, analytical

comparisons were only made between samples collected in the same country and same subjects, however biological conclusions can be compared to validate either study from the other.

5.3.1 Quality of UHPLC-MS data

The total number of features that were present in each dataset after the removal of low quality features can be found in Table 5.1.

From this table it is clear to see that there are similar numbers of features being extracted between both the dried plasma and serum spots, however there is a reduction in the number of features than what was initially expected for the plasma, positive ion mode dataset. This difference is difficult to explain and could be due to an issue with that occured throughout the analysis that was not identified. Whilst the total number of features present is slightly higher for the dried plasma spots when compared against dried serum spots, it is clear that a similar amount of biological information is being extracted, whether or not the spot was prepared with plasma or serum.

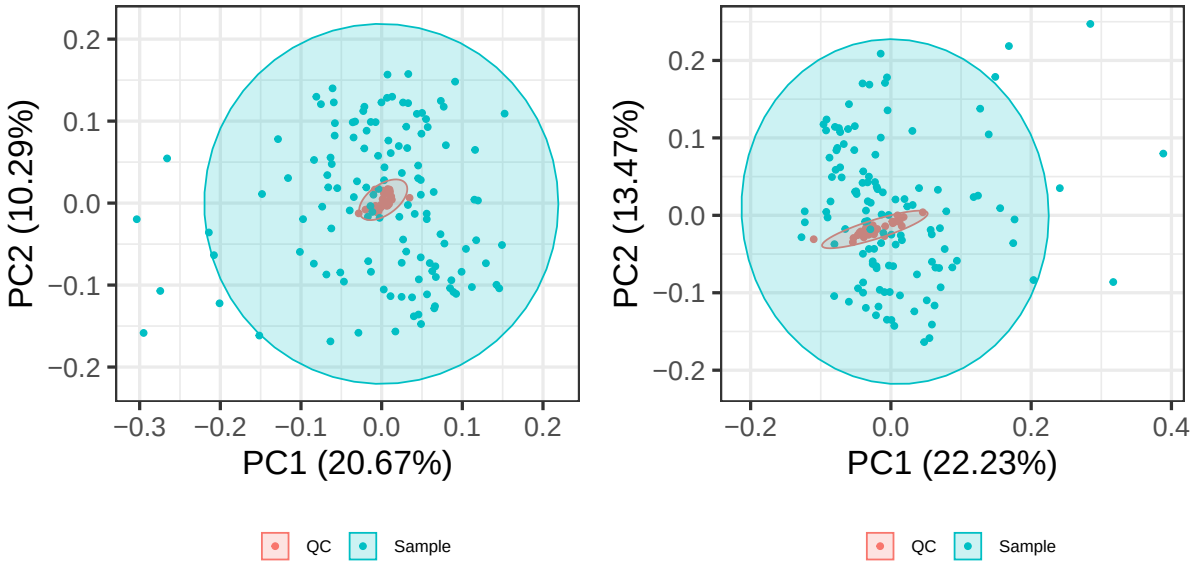
The extraction efficiency demonstrated through the total number of features present within each dataset demonstrate that there is a large amount of biological information that has been extracted, whether or not the sample was present in a dried form or not.

Initial inspection of the data after filtering demonstrates tight QC clustering in 7/8 datasets, with the Guatemala DSS Negative ion dataset showing lower than expected QC clustering which is unrelated to batch effects. This reduction in QC quality could

TABLE 5.1: Total number of features present in each dataset after filtering.

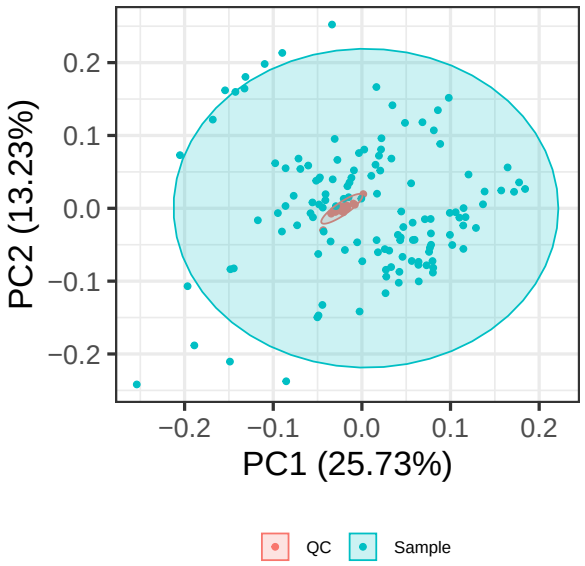
Polarity	DSS	Serum	DPS	Plasma
Positive	4448	5643	4790	2328
Negative	2374	2724	2723	2525

be related to the drop in performance of the LC column leading to broader peaks and lower quality data overall. PCA plots demonstrating QC performance can be seen in Figure 5.2 & 5.3.

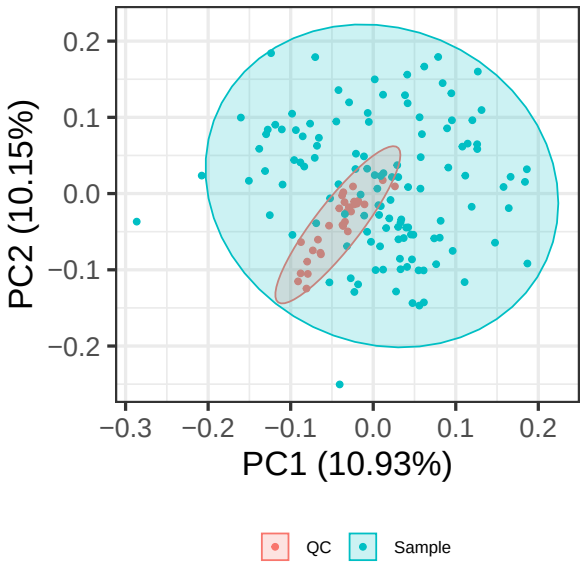


(A) Serum - Positive

(B) Serum - Negative

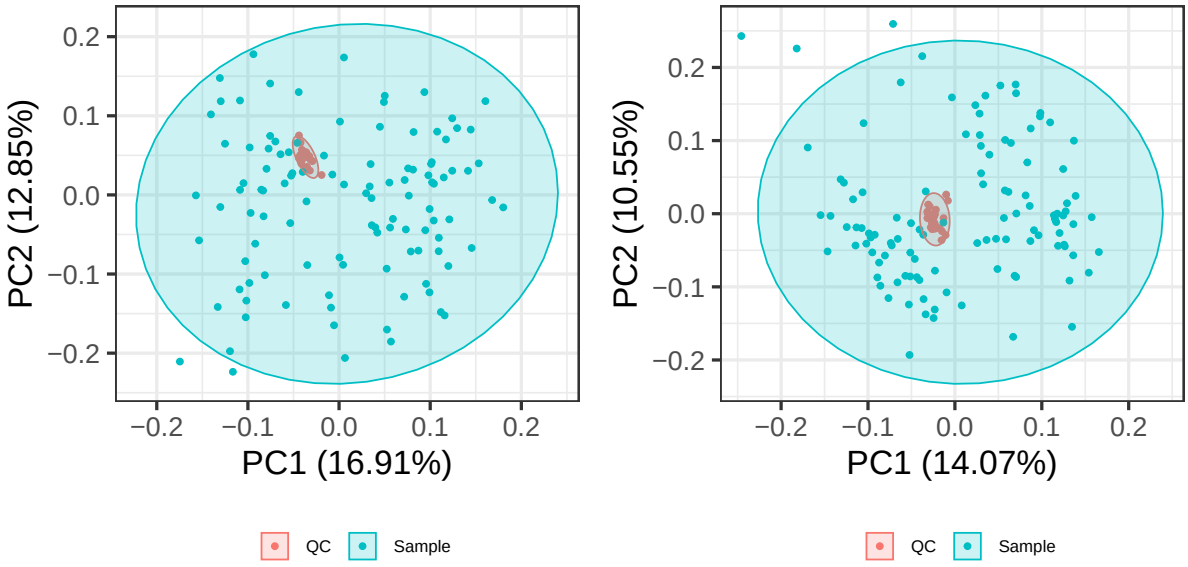


(C) DSS - Positive



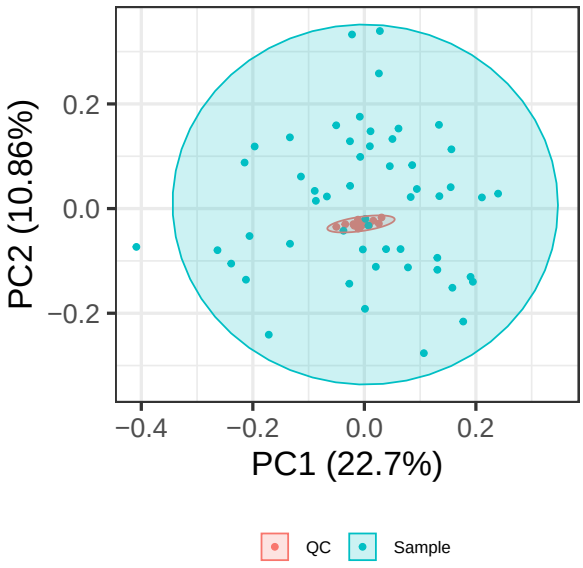
(D) DSS - Negative

FIGURE 5.2: PCA scores plot demonstrating quality of UHPLC-MS data for Guatemala study)

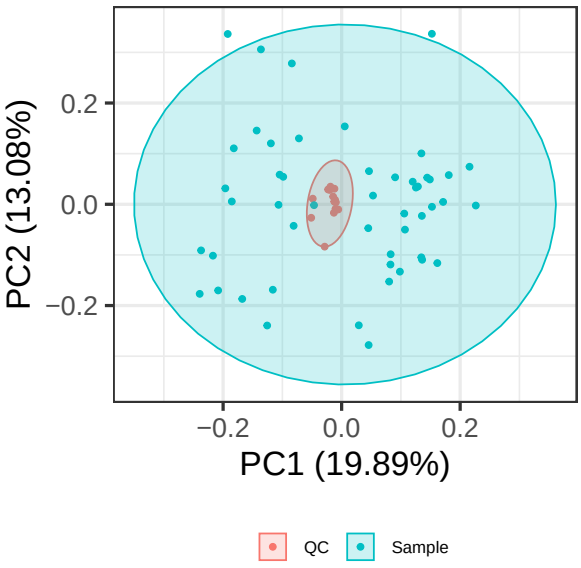


(A) Plasma - Positive

(B) Plasma - Negative



(C) DPS - Positive



(D) DPS - Negative

FIGURE 5.3: PCA scores plot demonstrating quality of UHPLC-MS data for the Philippines study)

5.3.2 Base Peak Chromatograms

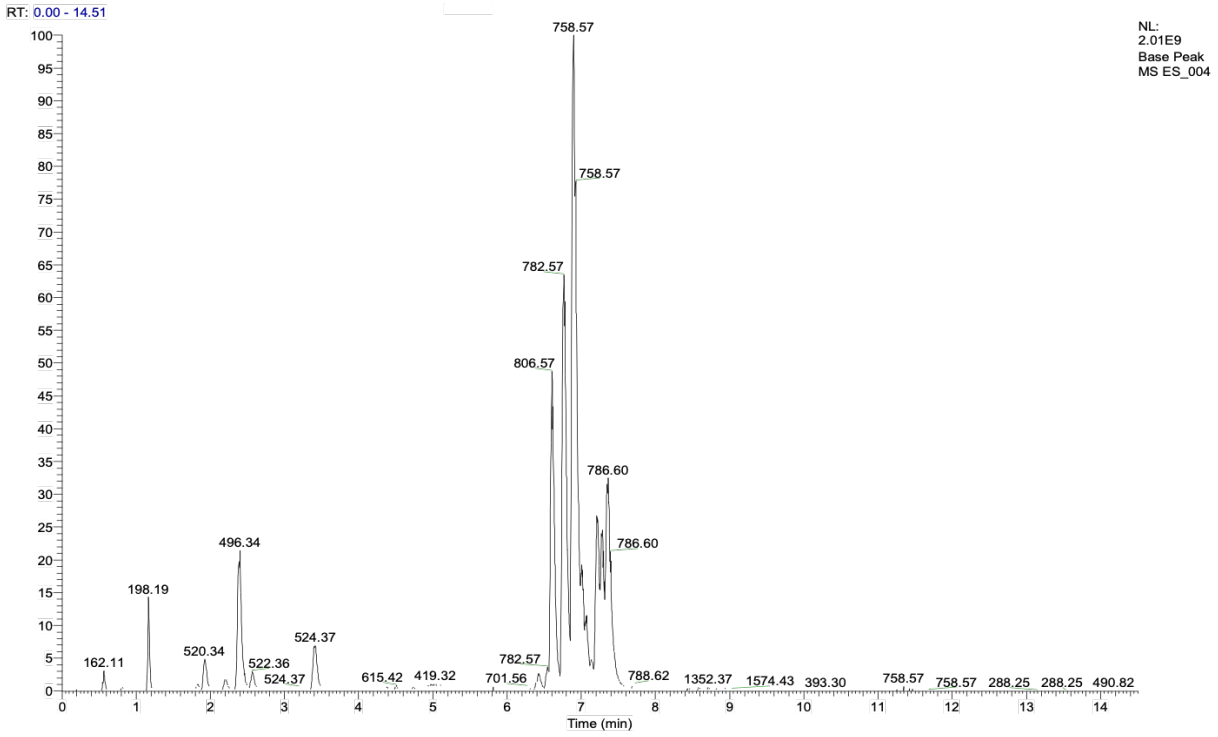
An initial visual comparison of the base peak chromatograms of both samples from serum and dried serum spot samples can be found in Figure 5.4 & 5.5.

Clear visual similarities between both serum and DSS samples can be identified with good correlation between major peaks present and retention times, as well as intensities of many peaks. However there are clear differences at certain regions of the chromatogram. When comparing DPS and plasma in Figure 5.4 there are additional and more intense peaks present towards the start of the plasma chromatogram prior to 4 minutes. Additionally, these peaks are also not present in the serum or DSS base peak chromatograms. The presence of these peaks in the chromatogram could be related to additional features present within plasma which are not present in serum such as clotting factors. Furthermore the peaks present could be related to compounds being extracted directly from the filter paper rather than from the sample directly.

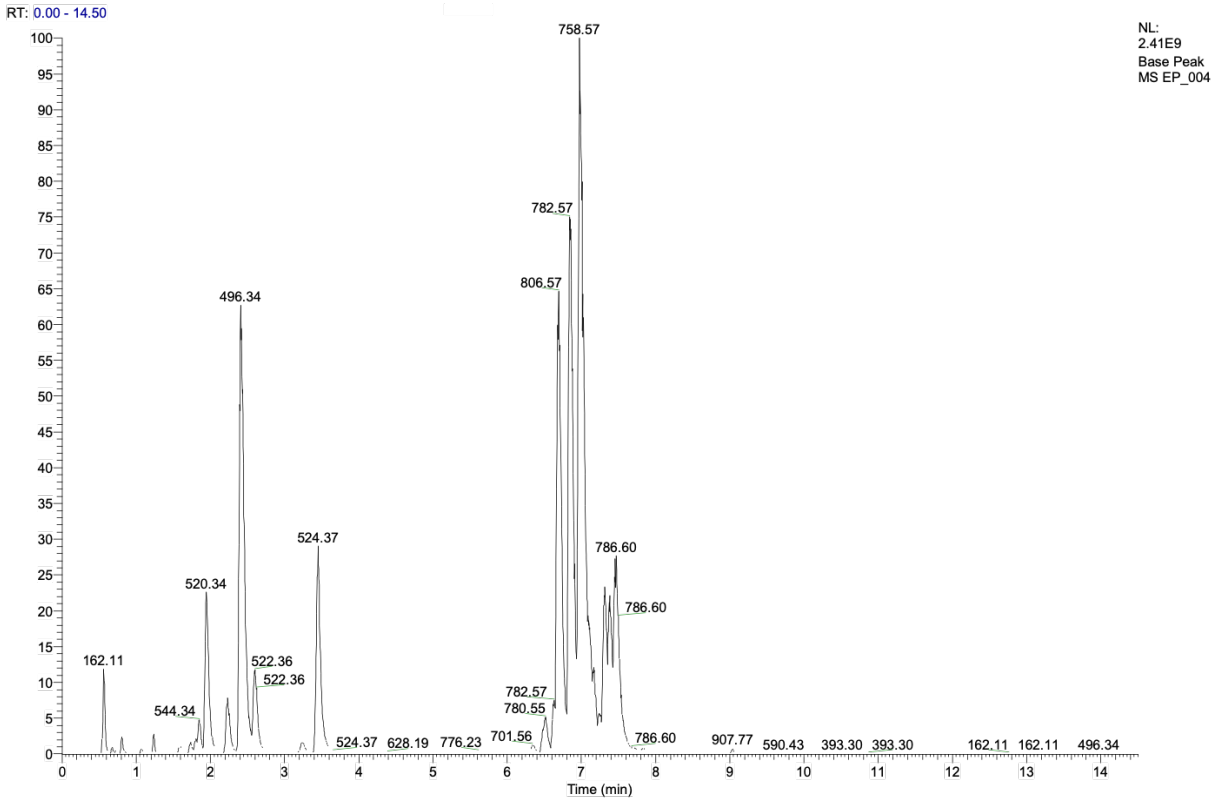
5.3.3 Comparison of extraction efficiency

To compare the extraction efficiency between extracting from whole biofluid and extraction from collection cards, the total peak area of each sample was compared for each dataset and ion mode. These values can be seen in the bar plot in Figure 5.6.

For the Philippines study, the total response shown in the sum of the peak areas, shows that there is no clear statistical significant (t-test, $p>0.05$) difference in the total analyte response between the two types of sample (whole or dried). However for the Guatemala study, there is a significant difference ($p < 0.005$) between the total sum of peak areas for dried and whole biofluid. Whilst this difference is significant, the extraction efficiency of the dried biofluid spots is still suitable for non-targeted metabolomics applications.

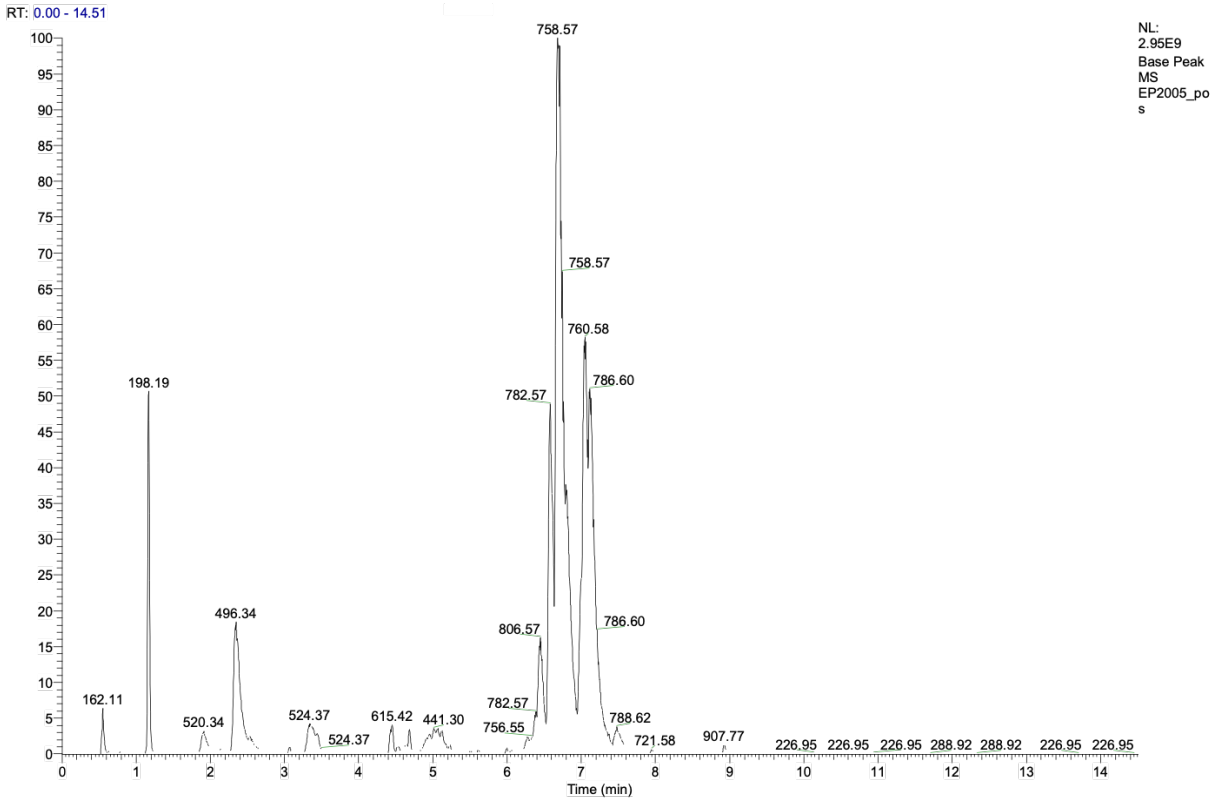


(A)

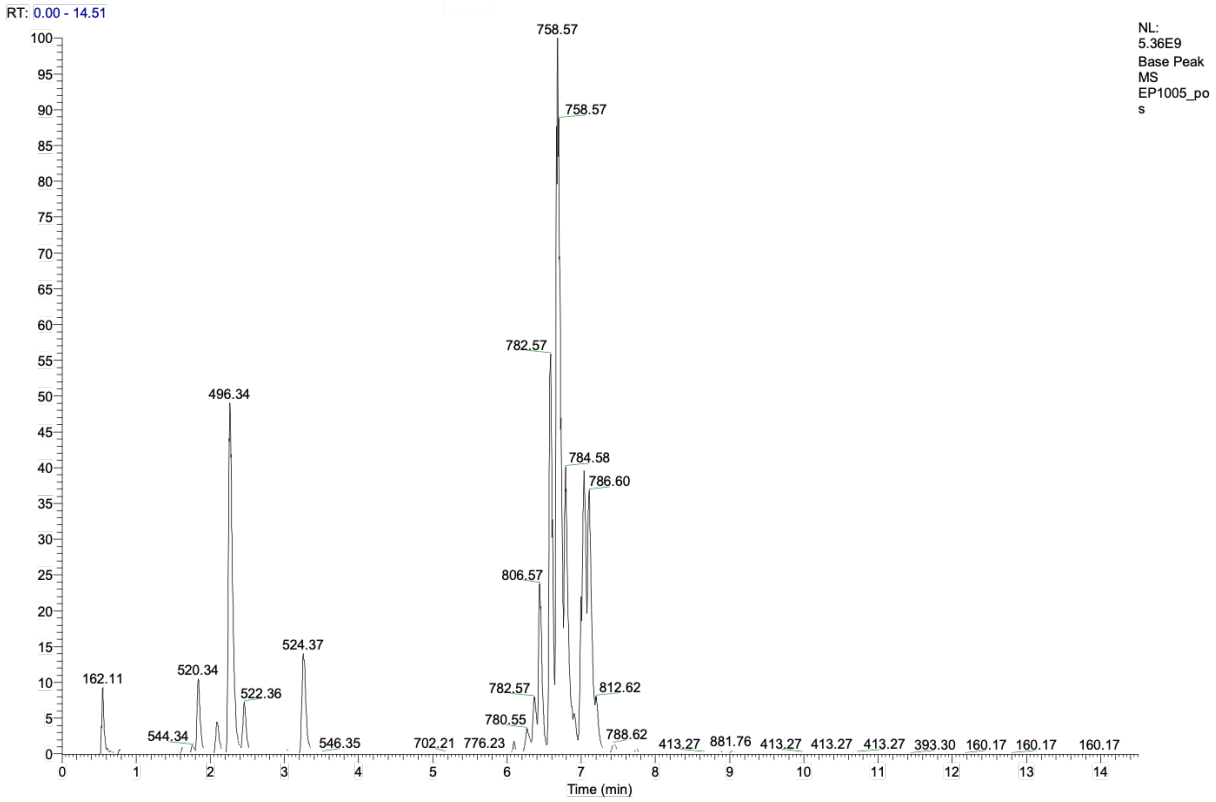


(B)

FIGURE 5.4: Base peak chromatograms of (A) dried plasma spots and (B) plasma collected from the Philippines.

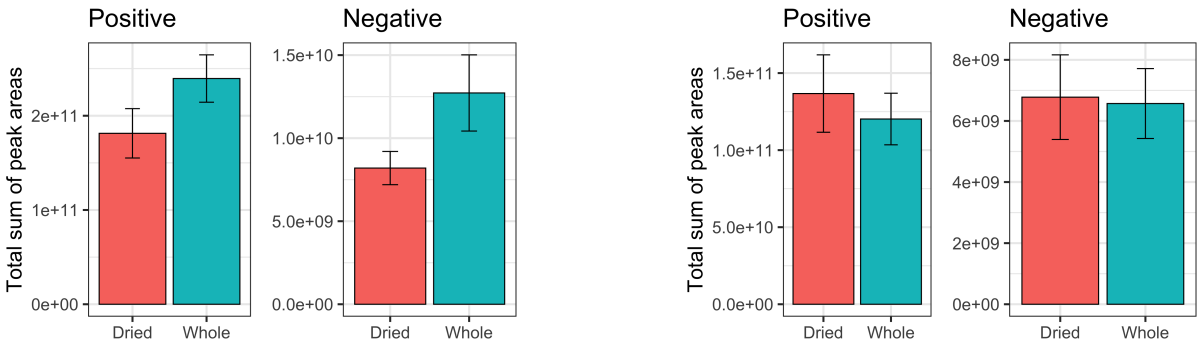


(A)



(B)

FIGURE 5.5: Base peak chromatograms of (A) dried serum spots and (B) serum.



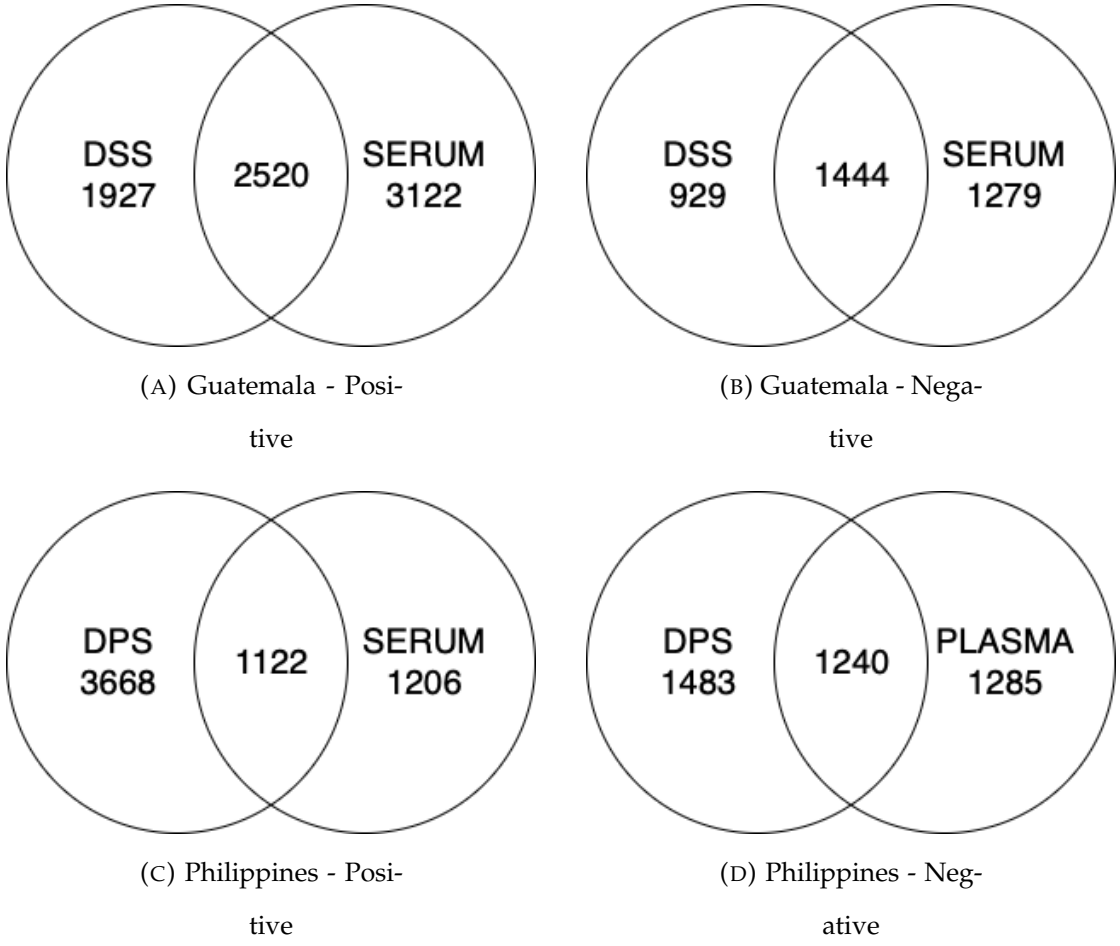
(A) Guatemala Study

(B) Philippines Study

FIGURE 5.6: Total sum of peak areas for all datasets.

5.3.4 Common Features

Comparison between the number of common features present between both DSS and serum and DPS and plasma was conducted. Features were determined as common if they were within a 5ppm mass difference, as well eluting from the column within 15 seconds of the other to account for any retention time shift that may have occurred between the two datasets. The common features are displayed in Figure 5.7



(A) Guatemala - Positive

(B) Guatemala - Negative

(C) Philippines - Positive

(D) Philippines - Negative

FIGURE 5.7: Number of common and unique features present within each dataset. Common features are match within a 5ppm mass difference and elute within 15 seconds of the other.

From Figure 5.7, whilst there are many ($\geq$1122) features that are common between both the biofluid and the dried sample, there are still a large number of features which are unique to each dataset.

The presence of features which are present in one form of the biofluid over the other may be attributed to several factors. Firstly, where features are present within the dried biofluid card over the whole biofluid, features may have leached from the card and into the extraction solution when the sample has undergone extraction and

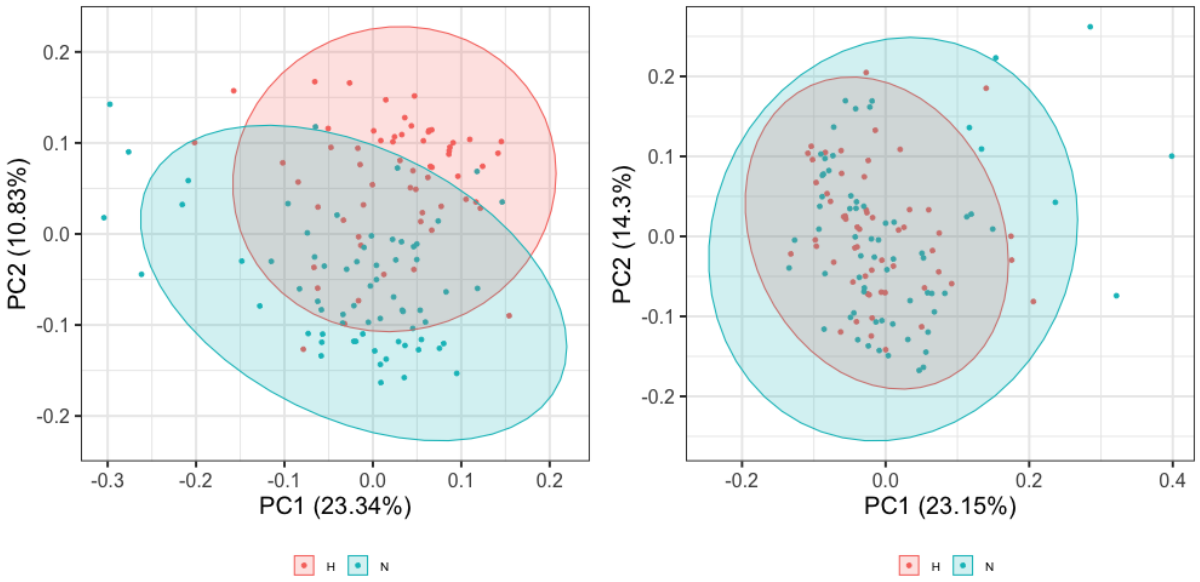
preparation. Whilst this is attempted to be controlled for through the extraction of blank cards, there potentially may be additional reactions and oxidation which may happen on the surface of the card which cannot be accounted for through a blank extraction of a collection card. Secondly, where features are present in the whole biofluid compared to its dried counterpart, an increase in the ability for features to partition into the extraction solution may be slightly easier than being extracted from the collection card. Another reason for unique features in each of the respective datasets could be attributed to an individual metabolite being present as multiple features present in separate datasets, therefore it would not have been detected as a common feature, but would still be related to the same metabolite.

However, with the extraction of more than 1000 biologically related features which are present within both datasets indicates that a large amount of biological information is present whether or not it is whole or dried.

5.3.5 Multivariate Analysis

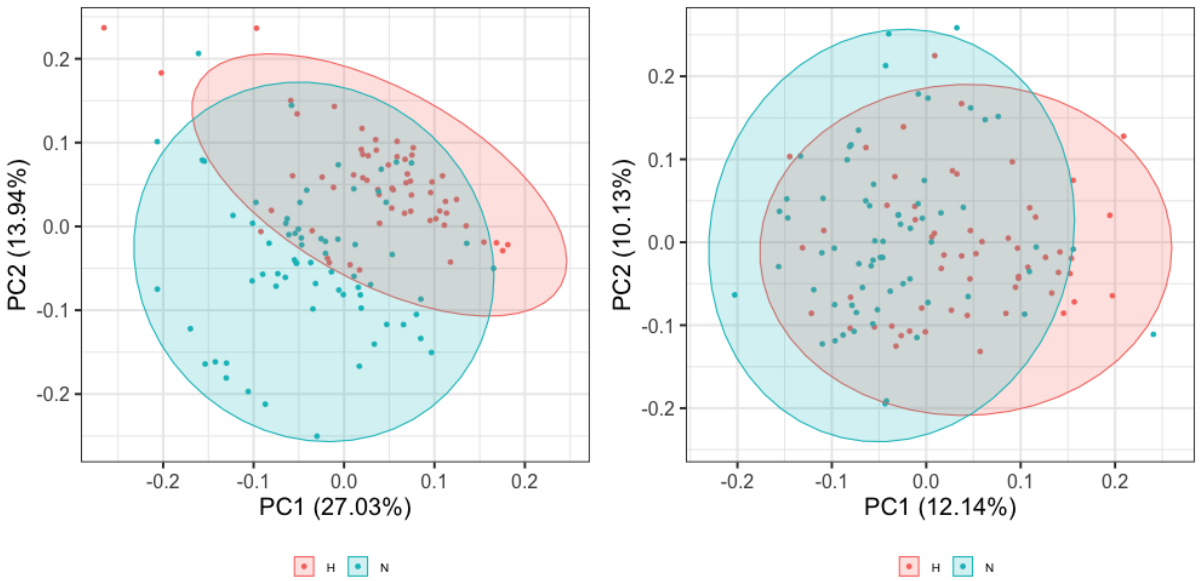
5.3.5.1 Principal Component Analysis

PCA plots were subsequently produced for each dataset and are shown in Figure 5.8 and 5.9. Samples from Guatemala were classified as either 'High' or 'Normal' intake of Vitamin A based on a food questionnaire completed by the subject prior to the sample being collected. For samples collected from the Philippines, a classification of 'High' or 'Normal' level of Vitamin A intake was determined from the calculated Vitamin A levels from the liver. Any sample which had a Vitamin A level of less than 1000 μg/kg of body weight were determined as 'Normal' and those which were higher classified as 'High'.



(A) Serum - Positive

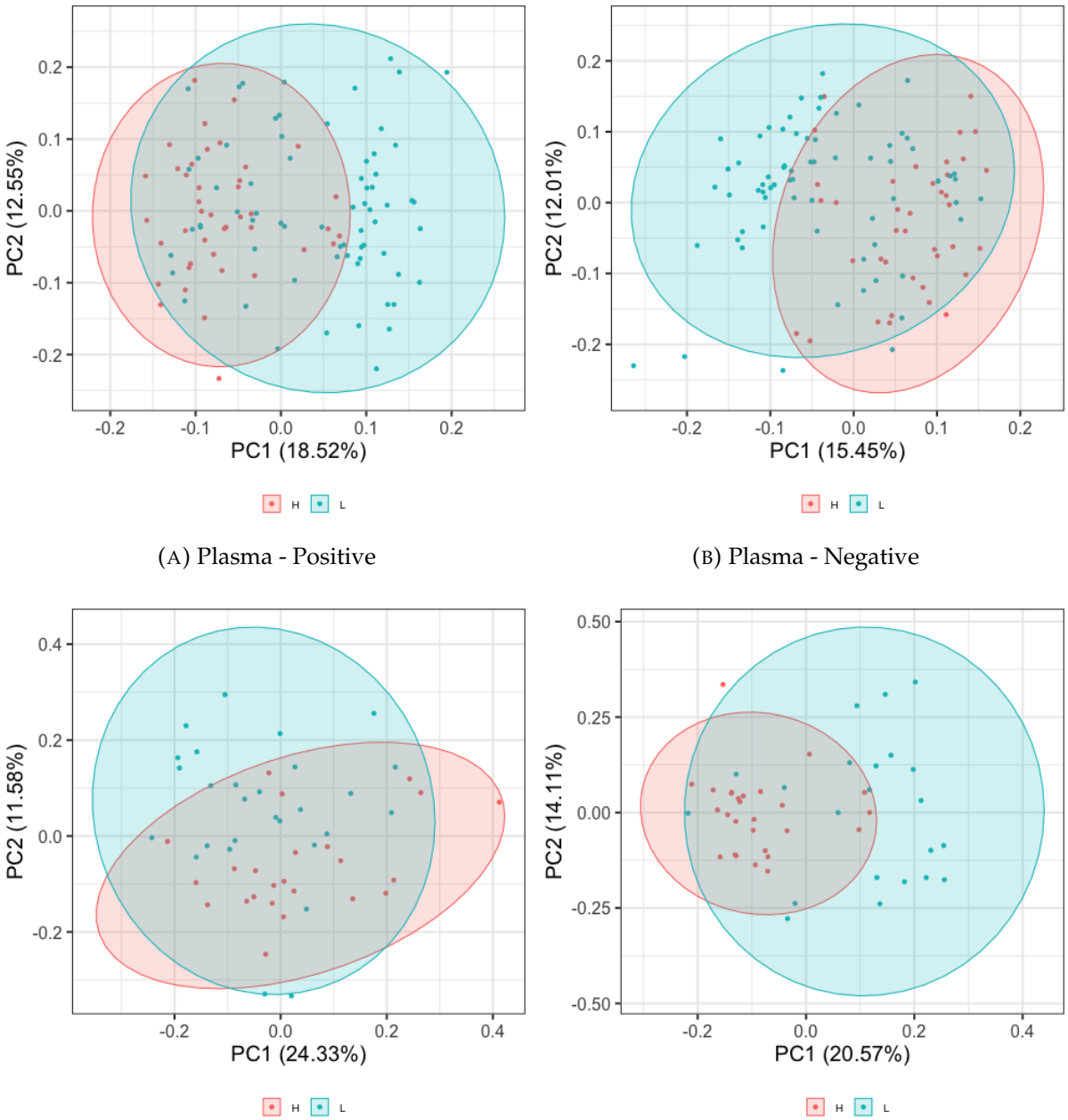
(B) Serum - Negative



(C) DSS - Positive

(D) DSS - Negative

FIGURE 5.8: PCA scores plot for samples in the Guatemala study with each samples classified as either 'High' intake of Vitamin A or 'Normal' intake of Vitamin A.)



(A) Plasma - Positive

(B) Plasma - Negative

(C) DPS - Positive

(D) DPS - Negative

FIGURE 5.9: PCA scores plot for samples in the Philippines study with each samples classified as either 'High' intake of Vitamin A or 'Normal' intake of Vitamin A.)

From the PCA plots it is clear to see that in most of the datasets, samples which had been classified as having a 'High' intake of Vitamin A are clustering together to a higher degree than those which are not. However, although there is some separation between the groups, there are samples present within the 'Normal' class which appear to be more biologically similar to the ''High' group than the 'Normal' group. This observation indicates that there may be an additional factor which is grouping samples together in addition to Vitamin A intake.

Additionally to observe the same pattern of grouping and separation between whole biofluid and the corresponding dried sample, indicates the same biological information is being extracted from the collection card as well as from the whole biofluid.

As a point to note, unlike in the Guatemala study where each whole biofluid sample has a corresponding dried serum sample, with the Philippines study, not every plasma sample has a corresponding dried plasma sample. Therefore there are fewer samples present within the DPS dataset compared to the plasma dataset. This is due to not enough plasma being present at the original time of preparation to prepare both plasma and corresponding dried plasma spots. This was unable to be corrected for as the samples had already been collected prior to the start of this research.

5.3.6 Univariate analysis

After multivariate analysis, a non-parametric Mann-Whitney U-Test was performed on each data set with the total number of significant features then summarised in Table 5.2. Comparisons were made between the first and the third tertile of total Vitamin A levels within the body.

TABLE 5.2: Total number of significant changes for each data set including the total number of each major lipid class defined as significantly changing by a Mann-Whitney U test ($p<0.05$, FDR corrected).

Lipid Class	Guatemala Study				Philippines Study			
	Serum		DSS		Plasma		DPS	
	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative
Total Number of Significant Features	157	53	327	19	81	87	283	215
Monoacylglycerols	-	1	2	2	-	-	-	-
Diacylglycerols	10	-	15	2	2	1	5	2
Tri(acyl | alkyl)glycerols	3	-	9	-	-	1	7	3
Fatty acyl carnitines	10	4	8	-	7	6	11	5
Phosphatidic acids	5	-	14	-	3	2	36	7
Phosphatidylcholines	18	5	39	1	25	7	32	10
Phosphatidylethanolamines	3	-	13	-	3	2	7	4
Phosphatidylglycerols	3	3	13	-	7	4	11	4
Phosphatidylinositols	-	13	16	-	4	5	5	10
Phosphatidylserines	52	10	67	5	10	12	19	18
Sphingomyelins	2	1	5	-	2	-	1	1

Major classes of lipids were then grouped together to give an indication into which classes of compounds were being up-regulated or down-regulated based on the Vitamin A intake level of the subject. In the Guatemala data set, several groups of lipids (phosphatidylserines (PS), phosphatidylcholines (PC) and diacylglycerols (DG)) are seen as to be significantly changing between the two intake groups of Vitamin A. These same groups of lipids are also shown as significantly changing with both datasets showing upregulation, not only in the dried serum spots prepared from the whole serum, but also in the independently collected samples from the Philippines, with both plasma and DPS also showing significant changes in the same groups.

The evidence that there are several lipid groups changing significantly in the serum and plasma, and the corresponding dried biofluid cards, indicate that the accumulation of vitamin A within the liver is having a biological effect on the patient which is then found in the serum and plasma metabolome.

Further to the above observation is the identification of the same lipid classes identified as significantly different in the study based in the Philippines. In this study, compared to other classes of lipids, it is the same groups in phosphatidylserines (PS) and phosphatidylcholines which are showing more changes.

The presence of similar classes of lipids which are significantly changing in separate populations demonstrates that a consistent metabolic change is occurring when between recommended Vitamin A levels and higher Vitamin A levels. This observation agrees with an unpublished pig study which was conducted along side this research. In this study, a controlled dose of retinyl proprionate was given to pigs to induce hypervitaminosis A. Subsequent results showed significant changes in the lipidome with many changes occuring in ceramides, glycerophospholipids, lysoglycerophospholipids, sterol lipids and triacylglycerides. Identifying common classes of

lipids such as glycerophospholipids that are changing between both the controlled pig model and the human studies performed in the Philippines and Guatemala opens up questions surrounding the potential for further targeted research.

With regard to the literature, there are many metabolic studies that look at the effects of Vitamin A deficiency, whereas the effects of being hypervitaminotic are limited. Suzuki et al. [250] reported that recieving high doses of Vitamin A coincided with increased concentrations of triglycerides in the serum. This agrees with the univariate analysis performed which demonstrates that triglycerides are also being significantly altered in the serum and plasma between subjects which have high and low levels of vitamin A intake.

Furthermore, Grolier et al [251] also reported that hypervitaminotic concentrations induced structural variations in the lipid membrane which were linked to enhancements of both phosphitdylcholines and sphingomyelins, whilst Sharma et al also reported that massive doeses of Vitamin A leads to phospholipid accumulation in rat livers.

Both of the above conclusions from Grolier et al and Sharma et al [252] agree with the statistical analysis in which there are large changes in phospholipids.

However as with any non-targeted study there are limitations, especially surrounding annotations. The annotations used to identify certain metabolites are only based on accurate mass-to-charge ratios and therefore there is a high probability of misattribution of a specific lipid being assigned to one class instead of another class, especially in cases of isobaric compounds.

Therefore for future studies it would be imperative that when focusing on the classes of lipids that could potentially be changing as described above, that further

MS/MS data is collected to ensure accurate annotations, and that a more in-depth targeted study towards particular classes is taken.

Due to the lack of literature that surrounds hypervitaminosis A and its effects on biochemical pathways, it is difficult to draw any robust biological conclusions.

5.4 Conclusion

In this chapter, the use of dried serum and plasma spots were applied to a biological non-targeted metabolomics application.

Although not identical, the use of dried plasma and serum spots compared to to dried blood spots still has advantages within non-targeted metabolomic applications. Particularly in cases where the cost of transporting samples under controlled temperature conditions across multiple countries may be prohibitive. However, with the lack of cellular components that are present within whole blood and therefore DBS, DPS and DSS will contain less biological information.

The use of DPS and DSS in a non-targeted metabolomics study has demonstrated similar analytical performance when compared to plasma and serum respectively. With sufficient extraction efficiency and over 1100 common features between the dried and whole form, there can be a high degree of confidence that there will be similar biological information present in both sample types.

Current metabolomic research surrounding Vitamin A and the effects it has on the liver when in excess is limited in the published literature. Therefore the work presented here represents the first investigation into the effect of high doses of Vitamin A in human subjects and the effects that are occurring in both the serum and plasma metabolome.

A point to note as well is the non-toxic concentrations of Vitamin A which have been supplied to the children taking part in both the Guatemala and Philippines study. To demonstrate that there is a metabolic effect occurring at increased levels of Vitamin A within the liver is an important conclusion with further analysis required to determine what physical effects this may have, both shorter and longer term.

Further targeted analysis focusing on lipid classes and pathways would potentially be a way in moving towards a more robust biological conclusion and determining any biochemical pathway effects which are occurring due to any increase in Vitamin A levels. Extra analysis could then be conducted using pathway analysis tools when more specific annotations have been made allowing the extraction of any biochemical pathways which are being altered either due to the up or down regulation of the various classes of lipids.

However, as just alluded to their are limitations to the study performed, primarily regarding the annotation of data. Due to only annotating based on *m/z*, putative class annotations are only possible and further work is required surrounding the more confident annotation of many of the lipid classes being detected as significantly different. Additionally, a more controlled study would increase the confidence in classifying whether or not a subject has high Vitamin A intake rather than relying on food based questionnaires which occurred for the Guatemala based study. Additional statistical analysis using supervised multivariate analysis techniques would also be of use when looking further into the data. Through the application of supervised methods such as PLS-DA, more confidence could be gained in the ability to identify metabolites which have significant effect on the models produced. Therefore allowing identification of metabolites which are perturbed through the consumption of excess Vitamin A.

Furthermore with the use of a pig model to look at the effects of hypervitamosis A,

additional samples could be taken from both adipose and liver tissues to build a more complete effect of the metabolic changes (due to being rich in lipids) that are occurring based on the high levels of Vitamin A.

In summary, DSS and DPS have been applied to a non-targeted metabolomics study where similar biological conclusions have been reached compared to the analysis of serum and plasma respectively. This shows that the use of collection cards is valid for the study of metabolic health and with the improvement in techniques and analysis, collection cards represent a cheaper alternative to population based studies both nationally and internationally.

Chapter 6

Conclusions

With the title of this thesis being 'The Investigation, Development and Application of Non-Targeted Metabolomic Methods Applying Dried Blood Spot Collection', I have aimed to investigate three main areas of DBS collection which can be viewed as limiting the wide scale use of collection cards in non-targeted research.

First, Chapters 2 and 3 are aimed at investigating two areas which are viewed as critical to the application of DBS in large scale studies. With stability and normalisation both key in ensuring the integrity of the data.

Chapter 2 - *Stability of DBS & DUS*

With Chapter 2, the stability of DBS in addition to DUS was investigated over a 12 month period at a variety of common temperatures applied in the storage of samples. Recommendations were made surrounding the longer term storage of DBS and DUS samples at reduced temperature, however it has been shown that samples were stable up to a period of one month before these recommendations should ideally be implemented. This provides plenty of time for samples to be collected remotely before being sent back to either the clinic or laboratory for longer term storage. This time period potentially allows patients to collect their own sample within their own homes or samples to be collected in remote locations where whole blood processing facilities may

not be available.

Further work surrounding the assessment of stability on other collection devices that are not filter paper based would be useful. With an increased number of commercial dried biofluid sampling kits supplied by companies, it would be beneficial to understand whether or not these can also be applied in non-targeted metabolomics scenarios.

Chapter 3 - *Normalisation*

Chapter 3 describes a novel implementation of a well established assay used in the determination of total protein content. Through the application of the Bradford assay to the measurement of total protein content in a DBS, a measurement value is obtained which can then be used to normalise the total blood volume which has been deposited onto the DBS. This method also has the benefit of being applied after the initial metabolomics extraction, thus only be performed on the remaining sample that was destined for waste. Applying normalisation both before the data was acquired as well as after data acquisition was shown to improve data quality, however better analytical performance was shown with normalisation prior to the acquisition of data through reconstitution volumes. It was also shown that applying a computational based parameter such as TIC normalisation to the normalisation of data was unsuitable when compared against the normalisation to total protein content.

Further work based on the results from this chapter could be expanding the range of samples types and collection cards which this method could be applied to as well as applying the assay to a biological question to further validate the implementation of the assay to the normalisation of DBS.

Chapter 4 - *LESA Method Development*

Chapter 4 was aimed at the development of a new method of analysis of DBS for use within non-targeted studies. This was achieved through the use of LESA-MS technology. Whilst LESA-MS has been applied to the analysis of DBS, this has primarily been in the area of proteomics. Additionally, LESA-MS has been used in non-targeted metabolomics, but not applied to DBS. Therefore there was an opportunity to fill this void in research and develop a new method of analysing DBS for non-targeted research. Through careful optimisation of parameters and solvents, a LESA-MS method was developed, with similar analytical performance to that of it's closest direct infusion technique. Multiple thousands of features were able to be extracted over five seconds from the surface of a DBS through a stable LMJ formed of a mixture of acetonitrile and water. The method was able to directly distinguish between two sample classes presented through PCA with a similar number of statistically significant metabolites being detected as by DI-MS. Whilst the performance of the method did not match techniques applying separation, the use of the method has future applicability in the high throughput screening of samples with out any sample preparation required. This reduces cost, increases throughput whilst still being able to extract a large amount of biological information from the surface of a DBS. More critically however is that LESA samples have to be left for a duration of time at ambient temperature to allow a clot to form which subsequently aids the formation of the LMJ. Therefore, for areas which require the immediate analysis of samples, the use of LESA could be limited for DBS due to the inability to form a LMJ. Furthermore, LESA-MS on DBS initially appears to be limited to a small variety of solvents, which could limit the groups of metabolites which can be extracted from DBS, and thus removing areas of research which may focus on specific areas of the metabolome such as lipidomics.

Potential further work for this area of research could be applying LESA to a biological study to allow for the high throughput screening of samples whilst requiring no sample preparation. Furthermore, with recent developments in the LESA platform it would also be beneficial to understand how well the method developed in this chapter compares to UHPLC through the ability to couple LESA to an LC column to perform separation on any extracted mixture.

Chapter 5 - *Application*

Finally, the application of collection cards in a non-targeted metabolomics study was demonstrated in Chapter 5. Through the use of DSS and DPS, and a direct comparison against serum and plasma respectively, two independent studies investigating the effects of Hypervitaminosis A were conducted. Results show similar biological changes accruing in both samples that were provided as whole or those which were present on collection cards. This similarity demonstrates the validity in applying collection cards in large population based studies. With large changes in glycerophospholipids being detected across both studies in Guatemala and the Philippines, there are biological effects from the high doses of Vitamin A that are given to children through each countries supplementation programs. Further investigation into the effects that high, but non-toxic concentrations of Vitamin A have on the subject is required. Especially seeing as the study performed only investigated serum and plasma, effects on adipose and liver tissues directly would further strengthen conclusions whilst also being able to build up a bigger picture of the metabolic health effects.

Through a combination of conclusions that have been reached across the various chapters presented in this thesis, the use of DBS in large scale, non-targeted metabolomics research in the future is possible. Providing recommendations surrounding the storage of samples, integrity of the data can be ensured up to a certain point. The normalisation of samples based on total protein content is a valid method to correct for non-biological sources of variation, and the high throughput of samples without prior extraction using LESA-MS, allows large amounts of biological data to be collected from a simple five second extraction.

6.1 Final Comments

Throughout this thesis I have demonstrated that the use of DBS and more broadly the application of using collection cards for the storage and transportation of samples has wide applicability in non-targeted metabolomics research areas.

With the advantages that collecting DBS has over samples collected via venipuncture, there is potentially a much broader area of research open to non-targeted metabolomics studies. Particularly in cases were the collection of samples is difficult based either on local geography or the facilities available to prepare, store and transport biofluids. With the use of collection cards, with relation to DBS and DUS, it has been shown that providing samples are collected and stored within one month of collection, then the integrity of the sample is not compromised and can still be applied to non-targeted studies.

Further development surrounding the use of DBS in non-targeted scenarios is no doubt required, especially surrounding whether or not there are alternative collection methods which remove the need to apply any potential normalisation procedures. Although applying a total protein content assay was shown in Chapter 3 to improve the quality of data obtained and improve the non-biological effects that different volumes have on the data. If commercial products which claim to collect accurate volumes of blood such as Neoteryx's Mitra kit are validated for non-targeted studies, then this removes the need for an additional step in the laboratory when dealing with real samples. However, applying a Bradford assay to samples which have been kept in long term storage in biobanks around the world, opens up a large volume of DBS samples to be analysed where blood volume will not have been recorded.

The LESA method developed in Chapter 4 demonstrates the potential to perform large scale screening of samples within non-targeted scenarios. With the lack of sample

preparation required, the use of LESA opens up the opportunity for 100's of samples to be analysed in a single day. Whilst the method was developed to maximise the number of features and ensure reproducibility of samples was within acceptable limits for non-targeted studies, it also has the potential to be applied in a targeted manner where only specific metabolites are of interest.

Finally, as mentioned earlier, the use of dried collection cards offer many opportunities for studies which may have previously been logistically impossible due to facility or financial constraints, and Chapter 5 demonstrates the use of collection cards in samples which have been collected in developing world countries such as Guatemala and the Philippines. With similar biological conclusions and similar biological classes of compounds being detected as significantly changing, it therefore stands to reason that the impact that using collection cards has on the overall study is minimal and should therefore be considered when designing experiments.

In conclusion, the use of collection cards, especially DBS, have a large potential future as collection devices for samples which are obtained for non-targeted metabolomics research. With improved logistical advantages, lower volumes and with a wide variety of methods available to analyse samples, DBS should be considered to be an appropriate sample collection strategy when performing non-targeted metabolomics experiments.

List of References

(1) Oliver, S. G.; Winson, M. K.; Kell, D. B.; Baganz, F. *Trends in Biotechnology* **1998**, *16*, 373–378.

(2) Williams, R. J. Biochemical Individuality: The Basis for the Genetotrophic Concept., New York, 1956.

(3) Mamas, M.; Dunn, W. B.; Neyses, L.; Goodacre, R. *Archives of Toxicology* **2011**, *85*, 5–17.

(4) Nicholson, J. K.; Lindon, J. C.; Holmes, E. *Xenobiotica* **1999**, *29*, 1181–1189.

(5) Johnson, C. H.; Gonzalez, F. J. *Journal of Cellular Physiology* **2012**, *227*, 2975–2981.

(6) Viant, M. R.; Kurland, I. J.; Jones, M. R.; Dunn, W. B. *Current Opinion in Chemical Biology* **2017**, *36*, 64–69.

(7) German, J. B.; Roberts, M.-A.; Watkins, S. M. *The Journal of Nutrition* **2003**, *133*, 2078S–2083S.

(8) German, J. B.; Watkins, S. M.; Fay, L.-B. *Journal of the American Dietetic Association* **2005**, *105*, 1425–1432.

(9) Van Ravenzwaay, B; Cunha, G. C.-P.; Leibold, E; Looser, R; Mellert, W; Prokoudine, A; Walk, T; Wiemer, J *Toxicology Letters* **2007**, *172*, 21–28.

(10) Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrometry Reviews* **2007**, *26*, 51–78.

(11) Holmes, E.; Loo, R. L.; Stamler, J.; Bictash, M.; Yap, I. K. S.; Chan, Q.; Ebbels, T.; De Iorio, M.; Brown, I. J.; Veselkov, K. A.; Daviglus, M. L.; Kesteloot, H.; Ueshima, H.; Zhao, L.; Nicholson, J. K.; Elliott, P. *Nature* **2008**, *453*, 396–400.

(12) Dumas, M.-E.; Maibaum, E. C.; Teague, C.; Ueshima, H.; Zhou, B.; Lindon, J. C.; Nicholson, J. K.; Stamler, J.; Elliott, P.; Chan, Q.; Holmes, E. *Analytical Chemistry* **2006**, *78*, 2199–2208.

(13) Nicholson, J. K.; Holmes, E.; Wilson, I. D. *Nature Reviews Microbiology* **2005**, *3*, 431–438.

(14) Golay, A.; Felber, J.; Jequier, E.; DeFronzo, R. A.; Ferrannini, E. *Diabetes/Metabolism Reviews* **1988**, *4*, 727–747.

(15) Fambrough, D. M. *Physiological Reviews* **1979**, *59*, 165–227.

(16) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N. *Nucleic Acids Research* **2017**, *46*, D608–D617.

(17) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill Jr, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W. *Nucleic Acids Research* **2006**, *35*, D527–D532.

(18) Kanehisa, M.; Goto, S. *Nucleic Acids Research* **2000**, *28*, 27–30.

(19) Nash, W. J.; Dunn, W. B. *Trends in Analytical Chemistry* **2018**.

(20) Breier, M.; Wahl, S.; Prehn, C.; Fugmann, M.; Ferrari, U.; Weise, M.; Banning, F.; Seissler, J.; Grallert, H.; Adamski, J.; Lechner, A. *PLOS ONE* **2014**, *9*, e89728.

(21) Beckonert, O.; Keun, H. C.; Ebbels, T. M. D.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. *Nature Protocols* **2007**, *2*, 2692.

(22) Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-Mcintyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Kell, D. B.; Goodacre, R. *Nature Protocols* **2011**, *6*, 1060–1083.

(23) Wang, X.; Zhang, A.; Han, Y.; Wang, P.; Sun, H.; Song, G.; Dong, T.; Yuan, Y.; Yuan, X.; Zhang, M.; Xie, N.; Zhang, H.; Dong, H.; Dong, W. *Molecular & Cellular Proteomics* **2012**, *11*, 370 LP –380.

(24) Ganti, S.; Weiss, R. H. *Urologic Oncology: Seminars and Original Investigations* **2011**, *29*, 551–557.

(25) Zhang, A.; Sun, H.; Wu, X.; Wang, X. *Clinica Chimica Acta* **2012**, *414*, 65–69.

(26) Mandal, R.; Guo, A. C.; Chaudhary, K. K.; Liu, P.; Yallou, F. S.; Dong, E.; Aziat, F.; Wishart, D. S. *Genome Medicine* **2012**, *4*, 38.

(27) Stoop, M. P.; Coulier, L.; Rosenling, T.; Shi, S.; Smolinska, A. M.; Buydens, L.; Ampt, K.; Stingl, C.; Dane, A.; Muilwijk, B.; Luitwieler, R. L.; Sillevis Smitt, P. A. E.; Hintzen, R. Q.; Bischoff, R.; Wijmenga, S. S.; Hankemeier, T.; van Gool, A. J.; Luider, T. M. *Molecular & Cellular Proteomics* **2010**, *9*, 2063 LP –2075.

(28) Sinclair, A. J.; Viant, M. R.; Ball, A. K.; Burdon, M. A.; Walker, E. A.; Stewart, P. M.; Rauz, S.; Young, S. P. *NMR in Biomedicine* **2010**, *23*, 123–132.

(29) Pieragostino, D.; D'Alessandro, M.; di Ioia, M.; Di Ilio, C.; Sacchetta, P.; Del Boccio, P. *Proteomics – Clinical Applications* **2015**, *9*, 169–186.

(30) Galbis-Estrada, C.; Martinez-Castillo, S.; Morales, J. M.; Vivar-Llopis, B.; Monleón, D.; Díaz-Llopis, M.; Pinazo-Durán, M. D. *BioMed Research International* **2014**, *2014*, 1–7.

(31) Römisch-Margl, W.; Prehn, C.; Bogumil, R.; Röhring, C.; Suhre, K.; Adamski, J. *Metabolomics* **2012**, *8*, 133–142.

(32) Lin, C. Y.; Wu, H.; Tjeerdema, R. S.; Viant, M. R. *Metabolomics* **2007**, *3*, 55–67.

(33) Wu, H.; Southam, A. D.; Hines, A.; Viant, M. R. *Analytical Biochemistry* **2008**, *372*, 204–212.

(34) Contrepois, K.; Jiang, L.; Snyder, M. *Molecular & Cellular Proteomics* **2015**, *14*, 1684–1695.

(35) Nicholson, J. K.; Lindon, J. C. *Nature* **2008**, *455*, 1054.

(36) Roberts, L. D.; Souza, A. L.; Gerszten, R. E.; Clish, C. B. In *Current Protocols in Molecular Biology*; SUPPL.98; John Wiley & Sons, Inc.: 2012; Vol. 1.

(37) Griffiths, W. J.; Koal, T.; Wang, Y.; Kohl, M.; Enot, D. P.; Deigner, H. *Angewandte Chemie International Edition* **2010**, *49*, 5426–5445.

(38) Viant, M. R.; Sommer, U. *Metabolomics* **2013**, *9*, 144–158.

(39) Goodacre, R.; Vaidyanathan, S.; Dunn, W. B.; Harrigan, G. G.; Kell, D. B. *Trends in Biotechnology* **2004**, *22*, 245–252.

(40) Sévin, D. C.; Kuehne, A.; Zamboni, N.; Sauer, U. *Current Opinion In Biotechnology* **2015**, *34*, 1–8.

(41) Naz, S.; Vallejo, M.; García, A.; Barbas, C. *Journal of Chromatography A* **2014**, *1353*, 99–105.

(42) Siskos, A. P.; Jain, P.; Romisch-Margl, W.; Bennett, M.; Achaintre, D.; Asad, Y.; Marney, L.; Richardson, L.; Koulman, A.; Griffin, J. L. *Analytical Chemistry* **2016**, *89*, 656–665.

(43) AbsoluteIDQ p400 HR Kit - https://www.biocrates.com/products/research-products/absoluteidq-p400-hr-kit.

(44) Gasperotti, M.; Masuero, D.; Guella, G.; Mattivi, F.; Vrhovsek, U. *Talanta* **2014**, *128*, 221–230.

(45) Álvarez-Sánchez, B; Priego-Capote, F; de Castro, M. D. L. *Trends in Analytical Chemistry* **2010**, *29*, 120–127.

(46) Lu, W.; Bennett, B. D.; Rabinowitz, J. D. *Journal of Chromatography B* **2008**, *871*, 236–242.

(47) Villas-Bôas, S. G.; Mas, S.; Åkesson, M.; Smedsgaard, J.; Nielsen, J. *Mass Spectrometry Reviews* **2005**, *24*, 613–646.

(48) Vuckovic, D. *Analytical and Bioanalytical Chemistry* **2012**, *403*, 1523–1548.

(49) Tulipani, S.; Llorach, R.; Urpi-Sarda, M.; Andres-Lacueva, C. *Analytical Chemistry* **2013**, *85*, 341–348.

(50) Lei, Z.; Huhman, D. V.; Sumner, L. W. Mass spectrometry strategies in metabolomics., 2011.

(51) Aharoni, A.; Ric de Vos, C. H.; Verhoeven, H. A.; Maliepaard, C. A.; Kruppa, G.; Bino, R.; Goodenowe, D. B. *Omics: A Journal of Integrative Biology* **2002**, *6*, 217–234.

(52) Horgan, R. P.; Kenny, L. C. *The Obstetrician & Gynaecologist* **2011**, *13*, 189–195.

(53) Gika, H.; Theodoridis, G. *Bioanalysis* **2011**, *3*, 1647–1661.

(54) Yin, P.; Lehmann, R.; Xu, G. *Analytical and Bioanalytical Chemistry* **2015**, *407*, 4879–4892.

(55) Lovelock, J. E. *Biochimica et Biophysica Acta* **1953**, *10*, 414–426.

(56) Sloviter, H. A. *Nature* **1962**, *193*, 884.

(57) Luque-Garcia, J. L.; Neubert, T. A. *Journal of Chromatography A* **2007**, *1153*, 259–276.

(58) Dettmer, K.; Almstetter, M. F.; Appel, I. J.; Nürnberger, N.; Schlamberger, G.; Gronwald, W.; Meyer, H. H. D.; Oefner, P. J. *Electrophoresis* **2010**, *31*, 2365–2373.

(59) Barri, T.; Dragsted, L. O. *Analytica Chimica Acta* **2013**, *768*, 118–128.

(60) Want, E. J.; Wilson, I. D.; Gika, H.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Holmes, E.; Nicholson, J. K. *Nature Protocols* **2010**, *5*, 1005–18.

(61) Chadha, V.; Garg, U.; Alon, U. S. *Pediatric Nephrology* **2001**, *16*, 374–382.

(62) Asiago, V. M.; Gowda, G. A. N.; Zhang, S.; Shanaiah, N.; Clark, J.; Raftery, D. *Metabolomics* **2008**, *4*, 328.

(63) Lauridsen, M.; Hansen, S. H.; Jaroszewski, J. W.; Cornett, C. *Analytical Chemistry* **2007**, *79*, 1181–1186.

(64) Chetwynd, A. J.; Abdul-Sada, A.; Holt, S. G.; Hill, E. M. *Journal of Chromatography A* **2016**, *1431*, 103–110.

(65) Warrack, B. M.; Hnatyshyn, S.; Ott, K.-H.; Reily, M. D.; Sanders, M.; Zhang, H.; Drexler, D. M. *Journal of Chromatography B* **2009**, *877*, 547–552.

(66) Calderón-Santiago, M.; Priego-Capote, F.; de Castro, M. D. L. *Electrophoresis* **2015**, *36*, 2179–2187.

(67) Bojko, B.; Reyes-Garcés, N.; Bessonneau, V.; Goryński, K.; Mousavi, F.; Souza Silva, E. A.; Pawliszyn, J. *Trends in Analytical Chemistry* **2014**, *61*, 168–180.

(68) Bligh, E. G.; Dyer, W. J. *Canadian Journal of Biochemistry and Physiology* **1959**, *37*, 911–917.

(69) Matyash, V.; Liebisch, G.; Kurzchalia, T. V.; Shevchenko, A.; Schwudke, D. *Journal of Lipid Research* **2008**, *49*, 1137–46.

(70) Sostare, J.; Di Guida, R.; Kirwan, J.; Chalal, K.; Palmer, E.; Dunn, W. B.; Viant, M. R. *Analytica Chimica Acta* **2018**, *1037*, 301–315.

(71) Fernández-Peralbo, M. A.; De Castro, M. D. L. *Trends in Analytical Chemistry* **2012**, *41*, 75–85.

(72) Issaq, H. J.; Van, Q. N.; Waybright, T. J.; Muschik, G. M.; Veenstra, T. D. *Journal of Separation Science* **2009**, *32*, 2183–2199.

(73) Guthrie, R.; Susi, A *Pediatrics* **1963**, *32*, 338–343.

(74) Blau, N.; van Spronsen, F. J.; Levy, H. L. *The Lancet* **2010**, *376*, 1417–1427.

(75) Jeanmonod, R.; Jeanmonod, D. In *StatPearls [Internet]*; StatPearls Publishing: 2018.

(76) Jager, N. G. L.; Rosing, H.; Linn, S. C.; Schellens, J. H. M.; Beijnen, J. H. *Therapeutic Drug Monitoring* **2015**, *37*, 833–836.

(77) Sakhi, A. K.; Bastani, N. E.; Ellingjord-Dale, M.; Gundersen, T. E.; Blomhoff, R.; Ursin, G. *BMC Cancer* **2015**, *15*, 265.

(78) Dried Blood Spot Sampling - https://cleancompetition.org/2016/05/11/whats-next-wednesday-dried-blood-spot-sampling/.

(79) Kadjo, A. F.; Stamos, B. N.; Shelor, C. P.; Berg, J. M.; Blount, B. C.; Dasgupta, P. K. *Analytical Chemistry* **2016**, *88*, 6531–6537.

(80) Freeman, J. D.; Rosman, L. M.; Ratcliff, J. D.; Strickland, P. T.; Graham, D. R.; Silbergeld, E. K. *Clinical Chemistry* **2018**, *64*, 656 LP –679.

(81) Solomon, S. S.; Solomon, S.; Rodriguez, I. I.; McGarvey, S. T.; Ganesh, A. K.; Thyagarajan, S. P.; Mahajan, A. P.; Mayer, K. H. *International Journal of STD & AIDS* **2002**, *13*, 25–28.

(82) Ngo-Giang-Huong, N.; Khamduang, W.; Leurent, B.; Collins, I.; Nantasen, I.; Leechanachai, P.; Sirirungsi, W.; Limtrakul, A.; Leusaree, T.; Comeau, A. M. *Journal of Acquired Immune Deficiency Syndromes* **2008**, *49*, 465.

(83) Lofgren, S. M.; Morrissey, A. B.; Chevallier, C. C.; Malabeja, A. I.; Edmonds, S.; Amos, B.; Sifuna, D. J.; Von Seidlein, L.; Schimana, W.; Stevens, W. S. *AIDS* **2009**, *23*, 2459.

(84) La Marca, G.; Malvagia, S.; Filippi, L.; Luceri, F.; Moneti, G.; Guerrini, R. *Epilepsia* **2009**, *50*, 2658–2662.

(85) Sadilkova, K.; Busby, B.; Dickerson, J. A.; Rutledge, J. C.; Jack, R. M. *Clinica Chimica Acta* **2013**, *421*, 152–156.

(86) Arai, N.; Narisawa, K.; Hayakawa, H.; Tada, K. *Pediatrics* **1982**, *70*, 426–430.

(87) O'Mara, M.; Hudson-Curtis, B.; Olson, K.; Yueh, Y.; Dunn, J.; Spooner, N. *Bioanalysis* **2011**, *3*, 2335–2347.

(88) Ren, X.; Paehler, T.; Zimmer, M.; Guo, Z.; Zane, P.; Emmons, G. T. *Bioanalysis* **2010**, *2*, 1469–1475.

(89) Holub, M.; Tuschl, K.; Ratschmann, R.; Strnadová, K. A.; Mühl, A.; Heinze, G.; Sperl, W.; Bodamer, O. A. *Clinica Chimica Acta* **2006**, *373*, 27–31.

(90) Denniff, P.; Spooner, N. *Bioanalysis* **2010**, *2*, 1385–1395.

(91) Mei, J. V.; Alexander, J. R.; Adam, B. W.; Hannon, W. H. *The Journal of Nutrition* **2001**, *131*, 1631S–1636S.

(92) Billett, H. H., *Hemoglobin and Hematocrit*, 1990.

(93) Thirup, P. *Sports Medicine* **2003**, *33*, 231–243.

(94) Capiau, S.; Stove, V. V.; Lambert, W. E.; Stove, C. P. *Analytical Chemistry* **2013**, *85*, 404–410.

(95) Meesters, R. J. W.; Hooff, G. P. *Bioanalysis* **2013**, *5*, 2187–2208.

(96) Holub, M.; Tuschl, K.; Ratschmann, R.; Strnadová, K. A.; Mühl, A.; Heinze, G.; Sperl, W.; Bodamer, O. A. *Clinica Chimica Acta* **2006**, *373*, 27–31.

(97) Oglesbee, D.; Sanders, K. A.; Lacey, J. M.; Magera, M. J.; Casetta, B.; Strauss, K. A.; Tortorelli, S.; Rinaldo, P.; Matern, D. *Clinical Chemistry* **2008**, *54*, 542–549.

(98) Otero-Fernández, M.; Cocho, J.; Tabernero, M. J.; Bermejo, A. M.; Bermejo-Barrera, P.; Moreda-Piñeiro, A. *Analytica Chimica Acta* **2013**, *784*, 25–32.

(99) Redondo, A. H.; Körber, C.; König, S.; Längin, A.; Al-Ahmad, A.; Weinmann, W. *Analytical and Bioanalytical Chemistry* **2012**, *402*, 2417–2424.

(100) Zheng, N.; Zeng, J.; Ji, Q. C.; Angeles, A.; Aubry, A.-F.; Basdeo, S.; Buzescu, A.; Landry, I. S.; Jariwala, N.; Turley, W. *Analytica Chimica Acta* **2016**, *934*, 170–179.

(101) Numako, M.; Takayama, T.; Noge, I.; Kitagawa, Y.; Todoroki, K.; Mizuno, H.; Min, J. Z.; Toyo'oka, T. *Analytical chemistry* **2015**, *88*, 635–639.

(102) Abdel-Rehim, A.; Abdel-Rehim, M. *Biomedical Chromatography* **2014**, *28*, 875–877.

(103) Rago, B.; Liu, J.; Tan, B.; Holliman, C. *Journal of Pharmaceutical and Biomedical Analysis* **2011**, *55*, 1201–1207.

(104) Sen, A.; Wang, Y.; Chiu, K.; Whiley, L.; Cowan, D.; Chang, R. C.-C.; Legido-Quigley, C. *Analytical Chemistry* **2013**, *85*, 7257–7263.

(105) Ulaszewska, M. M.; Weinert, C. H.; Trimigno, A.; Portmann, R.; Andres Lacueva, C.; Badertscher, R.; Brennan, L.; Brunius, C.; Bub, A.; Capozzi, F.; Cialiè Rosso, M.; Cordero, C. E.; Daniel, H.; Durand, S.; Egert, B.; Ferrario, P. G.; Feskens, E. J. M.; Franceschi, P.; Garcia-Aloy, M.; Giacomoni, F.; Giesbertz, P.; González-Domínguez, R.; Hanhineva, K.; Hemeryck, L. Y.; Kopka, J.; Kulling, S. E.; Llorach, R.; Manach, C.; Mattivi, F.; Migné, C.; Münger, L. H.; Ott, B.; Picone, G.; Pimentel, G.; Pujos-Guillot, E.; Riccadonna, S.; Rist, M. J.; Rombouts, C.; Rubert, J.; Skurk, T.; Sri Harsha, P. S. C.; Van Meulebroek, L.; Vanhaecke, L.; Vázquez-Fresno, R.; Wishart, D.; Vergères, G. *Molecular Nutrition & Food Research* **2019**, *63*, 1800384.

(106) Steuer, A. E.; Brockbals, L.; Kraemer, T. *Frontiers in Chemistry* **2019**, *7*, DOI: 10.3389/fchem.2019.00319.

(107) Mei, J. V.; Zobel, S. D.; Hall, E. M.; De Jesús, V. R.; Adam, B. W.; Hannon, W. H. *Bioanalysis* **2010**, *2*, 1397–1403.

(108) Luckwell, J.; Denniff, P.; Capper, S.; Michael, P.; Spooner, N.; Mallender, P.; Johnson, B.; Clegg, S.; Green, M.; Ahmad, S. *Bioanalysis* **2013**, *5*, 2613–2630.

(109) Smit, P. W.; Elliott, I.; Peeling, R. W.; Mabey, D.; Newton, P. N. *American Journal of Tropical Medicine and Hygiene* **2014**, *90*, 195–210.

(110) Kip, A. E.; Kiers, K. C.; Rosing, H; Schellens, J. H. M.; Beijnen, J. H.; Dorlo, T. P. C. *Journal of Pharmaceutical and Biomedical Analysis* **2017**, *135*, 160–166.

(111) HemaSpot-HF Blood Collection Device - Spot On Sciences - https://www.spotonsciences.com/products/hemaspot-hf/.

(112) Blood collection kit Mitra® Cartridge Neoteryx, LLC - https://www.medicalexpo.com/prod/neoteryx-llc/product-119251-847730.html.

(113) Brooks, K; DeLong, A; Balamane, M; Schreier, L; Orido, M; Chepkenja, M; Kemboi, E; D'Antuono, M; Chan, P. A.; Emonyi, W *Journal of Clinical Microbiology* **2016**, *54*, 223–225.

(114) Saracino, M. A.; Marcheselli, C.; Somaini, L.; Pieri, M. C.; Gerra, G.; Ferranti, A.; Raggi, M. A. *Analytical and Bioanalytical Chemistry* **2012**, *404*, 503–511.

(115) Kong, S. T.; Lin, H. S.; Ching, J.; Ho, P. C. *Analytical Chemistry* **2011**, *83*, 4314–4318.

(116) Barcenas, M.; Suhr, T. R.; Scott, C. R.; Turecek, F.; Gelb, M. H. *Clinica Chimica Acta* **2014**, *433*, 39–43.

(117) Markley, J. L.; Brüschweiler, R.; Edison, A. S.; Eghbalnia, H. R.; Powers, R.; Raftery, D.; Wishart, D. S. *Current Opinion in Biotechnology* **2017**, *43*, 34–40.

(118) Schrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod, S. D.; McLean, J. A. *Journal of The American Society for Mass Spectrometry* **2016**, *27*, 1897–1905.

(119) González-Domínguez, R.; Sayago, A.; Fernández-Recamales, *Bioanalysis* **2017**, *9*, 131–148.

(120) Koulman, A.; Tapper, B. A.; Fraser, K.; Cao, M.; Lane, G. A.; Rasmussen, S. *Rapid Communications in Mass Spectrometry* **2007**, *21*, 421–428.

(121) Southam, A. D.; Payne, T. G.; Cooper, H. J.; Arvanitis, T. N.; Viant, M. R. *Analytical Chemistry* **2007**, *79*, 4595–4602.

(122) Gika, H. G.; Wilson, I. D.; Theodoridis, G. A. *Journal of Chromatography B* **2014**, *966*, 1–6.

(123) Boelaert, J.; t'Kindt, R.; Schepers, E.; Jorge, L.; Glorieux, G.; Neirynck, N.; Lynen, F.; Sandra, P.; Vanholder, R.; Sandra, K. *Metabolomics* **2014**, *10*, 425–442.

(124) Broeckling, C. D.; Heuberger, A. L.; Prenni, J. E. *Journal of Visualized Experiments* **2013**, e50242.

(125) Tang, D.; Zou, L.; Yin, X.; Ong, C. N. *Mass Spectrometry Reviews* **2016**, *35*, 574–600.

(126) Cubbon, S.; Antonio, C.; Wilson, J.; Thomas-Oates, J. *Mass Spectrometry Reviews* **2010**, *29*, 671–684.

(127) Fei, F.; Bowdish, D. M. E.; McCarry, B. E. *Analytical and Bioanalytical Chemistry* **2014**, *406*, 3723–3733.

(128) Bakalyar, S. R.; Phipps, C.; Spruce, B.; Olsen, K. *Journal of Chromatography A* **1997**, *762*, 167–185.

(129) Swartz, M. E. *Journal of Liquid Chromatography & Related Technologies* **2005**, *28*, 1253–1263.

(130) Mazzeo, J. R.; D.Neue, U.; Kele, M.; Plumb, R. S. *Analytical Chemistry* **2005**, *77*, 460 A–467 A.

(131) Broeckhoven, K; Desmet, G *Trends in Analytical Chemistry* **2014**, *63*, 65–75.

(132) Gray, N.; Adesina-Georgiadis, K.; Chekmeneva, E.; Plumb, R. S.; Wilson, I. D.; Nicholson, J. K. *Analytical Chemistry* **2016**, *88*, 5742–5751.

(133) Zotou, A. *Central European Journal of Chemistry* **2012**, *10*, 554–569.

(134) Wilson, S. R.; Vehus, T.; Berg, H. S.; Lundanes, E. *Bioanalysis* **2015**, *7*, 1799–815.

(135) Shen, Y.; Tolić, N.; Masselon, C.; Paša-Tolić, L.; Camp, D. G.; Hixson, K. K.; Zhao, R.; Anderson, G. A.; Smith, R. D. *Analytical Chemistry* **2004**, *76*, 144–154.

(136) Huang, J. T.; McKenna, T.; Hughes, C.; Leweke, F. M.; Schwarz, E.; Bahn, S. *Journal of Separation Science* **2007**, *30*, 214–225.

(137) Chetwynd, A. J.; David, A. *Talanta* **2018**, *182*, 380–390.

(138) Jones, D. R.; Wu, Z.; Chauhan, D.; Anderson, K. C.; Peng, J. *Analytical Chemistry* **2014**, *86*, 3667–3675.

(139) Bleakney, W. *Physical Review* **1929**, *34*, 157–160.

(140) Hoffmann, E. d.; Stroobant, V., *Mass spectrometry: principles and applications*, 3rd ed; J. Wiley: Chichester, West Sussex, England ; Hoboken, NJ, 2007, p 489.

(141) Fiehn, O. *Trends in Analytical Chemistry* **2008**, *27*, 261–269.

(142) Munson, M. S. B.; Field, F. H. *Journal of the American Chemical Society* **1966**, *88*, 2621–2630.

(143) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64–71.

(144) Pramanik, B. N.; Ganguly, A. K.; Gross, M. L., *Applied electrospray mass spectrometry: practical spectroscopy series*; CRC Press: 2002; Vol. 32.

(145) Wood, T. D.; Moy, M. A.; Dolan, A. R.; Bigwarfe Jr, P. M.; White, T. P.; Smith, D. R.; Higbee, D. J. **2003**.

(146) Carroll, D. I.; Dzidic, I; Stillwell, R. N.; Haegele, K. D.; Horning, E. C. *Analytical Chemistry* **1975**, *47*, 2369–2373.

(147) Southam, A. D.; Payne, T. G.; Cooper, H. J.; Arvanitis, T. N.; Viant, M. R. *Analytical Chemistry* **2007**, *79*, 4595–4602.

(148) Castrillo, J. I.; Hayes, A.; Mohammed, S.; Gaskell, S. J.; Oliver, S. G. *Phytochemistry* **2003**, *62*, 929–937.

(149) Van Pelt, C. K.; Zhang, S.; Fung, E.; Chu, I.; Liu, T.; Li, C.; Korfmacher, W. A.; Henion, J. *Rapid Communications In Mass Spectrometry* **2003**, *17*, 1573–1578.

(150) Kertesz, V.; Van Berkel, G. J. *Journal of Mass Spectrometry* **2010**, *45*, 252–260.

(151) Kertesz, V.; Ford, M. J.; Van Berkel, G. J. *Analytical Chemistry* **2005**, *77*, 7183–7189.

(152) Eikel, D.; Vavrek, M.; Smith, S.; Bason, C.; Yeh, S.; Korfmacher, W. A.; Henion, J. D. *Rapid Communications in Mass Spectrometry* **2011**, *25*, 3587–3596.

(153) Bailey, M. J.; Randall, E. C.; Costa, C.; Salter, T. L.; Race, A. M.; de Puit, M.; Koeberg, M.; Baumert, M.; Bunch, J. *Analytical Methods* **2016**, *8*, 3373–3382.

(154) Griffiths, R. L.; Creese, A. J.; Race, A. M.; Bunch, J.; Cooper, H. J. *Analytical Chemistry* **2016**, *88*, 6758–6766.

(155) Edwards, R. L.; Griffiths, P.; Bunch, J.; Cooper, H. J. *Journal of the American Society for Mass Spectrometry* **2012**, *23*, 1921–1930.

(156) Martin, N. J.; Bunch, J.; Cooper, H. J. *Journal of The American Society for Mass Spectrometry* **2013**, *24*, 1242–1249.

(157) Montowska, M.; Alexander, M. R.; Tucker, G. A.; Barrett, D. A. *Analytical Chemistry* **2014**, *86*, 10257–10265.

(158) Wisztorski, M.; Desmons, A.; Quanico, J.; Fatou, B.; Gimeno, J.; Franck, J.; Salzet, M.; Fournier, I. *Proteomics* **2016**, *16*, 1622–1632.

(159) Paul, W.; Steinwedel, H. *Zeitschrift für Naturforschung A* **1953**, *8*, 448–450.

(160) Church, D. A. *Journal of Applied Physics* **1969**, *40*, 3127–3134.

(161) Ledford, E. B.; Rempel, D. L.; Gross, M. L. *Analytical Chemistry* **1984**, *56*, 2744–2748.

(162) Comisarow, M. B.; Marshall, A. G. *Chemical Physics Letters* **1974**, *25*, 282–283.

(163) Makarov, A. *Analytical Chemistry* **2000**, *72*, 1156–1162.

(164) Hardman, M.; Makarov, A. A. *Analytical Chemistry* **2003**, *75*, 1699–1705.

(165) Denisov, E.; Damoc, E.; Lange, O.; Makarov, A. *International Journal of Mass Spectrometry* **2012**, *325*, 80–85.

(166) *Planet Orbitrap - https://planetorbitrap.com*.

(167) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Analytical Chemistry* **2006**, *78*, 779–787.

(168) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *Bioinformatics* **2010**, *11*, 395.

(169) Sturm, M.; Bertsch, A.; Gröpl, C.; Hildebrandt, A.; Hussong, R.; Lange, E.; Pfeifer, N.; Schulz-Trieglaff, O.; Zerck, A.; Reinert, K. *BMC Bioinformatics* **2008**, *9*, 163.

(170) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nature Methods* **2015**, *12*, 523.

(171) Clasquin, M. F.; Melamud, E.; Rabinowitz, J. D. *Current Protocols in Bioinformatics* **2012**, *37*, 11–14.

(172) Di Guida, R.; Engel, J.; Allwood, J. W.; Weber, R. J.; Jones, M. R.; Sommer, U.; Viant, M. R.; Dunn, W. B. *Metabolomics* **2016**, *12*, DOI: 10.1007/s11306-016-1030-9.

(173) Steuer, R.; Morgenthal, K.; Weckwerth, W.; Selbig, J., *A Gentle Guide to the Analysis of Metabolomic Data BT - Metabolomics: Methods and Protocols*; Weckwerth, W., Ed.; Humana Press: Totowa, NJ, 2007, pp 105–126.

(174) Van den Berg, R. A.; Hoefsloot, H. C. J.; Westerhuis, J. A.; Smilde, A. K.; van der Werf, M. J. *BMC Genomics* **2006**, *7*, 142.

(175) Southam, A. D.; Weber, R. J.; Engel, J.; Jones, M. R.; Viant, M. R. *Nature Protocols* **2017**, *12*, 255–273.

(176) Madsen, R.; Lundstedt, T.; Trygg, J. *Analytica Chimica Acta* **2010**, *659*, 23–33.

(177) Yi, L.; Dong, N.; Yun, Y.; Deng, B.; Ren, D.; Liu, S.; Liang, Y. *Analytica Chimica Acta* **2016**, *914*, 17–34.

(178) Gromski, P. S.; Muhamadali, H.; Ellis, D. I.; Xu, Y.; Correa, E.; Turner, M. L.; Goodacre, R. *Analytica Chimica Acta* **2015**, *879*, 10–23.

(179) Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. *Metabolites* **2012**, *2*, 775–795.

(180) Dunn, W. B.; Erban, A.; Weber, R. J. M.; Creek, D. J.; Brown, M.; Breitling, R.; Hankemeier, T.; Goodacre, R.; Neumann, S.; Kopka, J. *Metabolomics* **2013**, *9*, 44–66.

(181) Wishart, D. S. *Bioanalysis* **2011**, *3*, 1769–1782.

(182) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. *Metabolomics* **2007**, *3*, 211–221.

(183) Wilson, I. *Bioanalysis* **2011**, *3*, 2255–2257.

(184) McDade, T. W.; Williams, S.; Snodgrass, J. J. *Demography* **2007**, *44*, 899–925.

(185) Oostendorp, M.; El Amrani, M.; Diemel, E. C.; Hekman, D.; van Maarseveen, E. M. *Clinical Chemistry* **2016**, *62*, 1534–1536.

(186) Capiau, S.; Wilk, L. S.; Aalders, M. C.; Stove, C. P. *Analytical Chemistry* **2016**, *88*, 6538–6546.

(187) Zakaria, R.; Allen, K. J.; Koplin, J. J.; Roche, P.; Greaves, R. F. *Electronic Journal of the International Federation of Clinical Chemistry and Laboratory Medicine* **2016**, *27*, 288–317.

(188) Neto, R.; Gooley, A.; Breadmore, M. C.; Hilder, E. F.; Lapierre, F. *Analytical and Bioanalytical Chemistry* **2018**, *410*, 3315–3323.

(189) Volani, C.; Caprioli, G.; Calderisi, G.; Sigurdsson, B. B.; Rainer, J.; Gentilini, I.; Hicks, A. A.; Pramstaller, P. P.; Weiss, G.; Smarason, S. V.; Paglia, G. *Analytical and Bioanalytical Chemistry* **2017**, *409*, 6263–6276.

(190) Youhnovski, N.; Bergeron, A.; Furtado, M.; Garofolo, F. In *Rapid Communications in Mass Spectrometry*, 2011; Vol. 25, pp 2951–2958.

(191) De Kesel, P. M.; Sadones, N.; Capiau, S.; Lambert, W. E.; Stove, C. P. *Bioanalysis* **2013**, *5*, 2023–2041.

(192) Hall, E. M.; Flores, S. R.; De Jesús, V. R. *International Journal of Neonatal Screening* **2015**, *1*, 69–78.

(193) Bowen, C. L.; Hemberger, M. D.; Kehler, J. R.; Evans, C. A. *Bioanalysis* **2010**, *2*, 1823–1828.

(194) Pupillo, D.; Simonato, M.; Cogo, P. E.; Lapillonne, A.; Carnielli, V. P. *Lipids* **2016**, *51*, 193–198.

(195) Alfazil, A. A.; Anderson, R. A. *Journal of Analytical Toxicology* **2008**, *32*, 511–515.

(196) Michopoulos, F.; Theodoridis, G.; Smith, C. J.; Wilson, I. D. *Journal of Proteome Research* **2010**, *9*, 3328–3334.

(197) Prentice, P.; Turner, C.; Wong, M. C.; Dalton, R. N. *Bioanalysis* **2013**, *5*, 1507–1514.

(198) Drolet, J.; Id, V. T.; Williams, B. A.; Id, B. P. G.; Hill, C.; Vishnudas, V. K.; Sarangarajan, R.; Narain, N. R.; Kiebish, M. A. *Metabolites* **2017**, *7*, 35.

(199) Haijes, H.; Willemsen, M.; van der Ham, M.; Gerrits, J.; Pras-Raves, M.; Prinsen, H.; van Hasselt, P.; de Sain-van der Velden, M.; Verhoeven-Duif, N.; Jans, J. *Metabolites* **2019**, *9*, 12.

(200) Rotter, M.; Brandmaier, S.; Prehn, C.; Adam, J.; Rabstein, S.; Gawrych, K.; Brüning, T.; Illig, T.; Lickert, H.; Adamski, J.; Wang-Sattler, R. *Metabolomics* **2017**, *13*.

(201) Di Guida, R.; Engel, J.; Allwood, J. W.; Weber, R. J. M.; Jones, M. R.; Sommer, U.; Viant, M. R.; Dunn, W. B. *Metabolomics* **2016**, *12*, 93.

(202) Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D. S.; Xia, J. *Nucleic Acids Research* **2018**, *46*, W486–W494.

(203) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Analytical Chemistry* **2006**, *78*, 779–787.

(204) Prentice, P.; Turner, C.; Wong, M. C.; Dalton, R. N. *Bioanalysis* **2013**, *5*, 1507–1514.

(205) Michopoulos, F.; Theodoridis, G.; Smith, C. J.; Wilson, I. D. *Bioanalysis* **2011**, *3*, 2757–2767.

(206) Dunn, W. B.; Lin, W.; Broadhurst, D.; Begley, P.; Brown, M.; Zelena, E.; Vaughan, A. A.; Halsall, A.; Harding, N.; Knowles, J. D.; Francis-McIntyre, S.; Tseng, A.; Ellis, D. I.; O'Hagan, S.; Aarons, G.; Benjamin, B.; Chew-Graham, S.; Moseley, C.; Potter, P.; Winder, C. L.; Potts, C.; Thornton, P.; McWhirter, C.; Zubair, M.; Pan, M.; Burns, A.; Cruickshank, J. K.; Jayson, G. C.; Purandare, N.; Wu, F. C. W.; Finn, J. D.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Goodacre, R.; Kell, D. B. *Metabolomics* **2015**, *11*, 9–26.

(207) Lewis, M. R.; Pearce, J. T. M.; Spagou, K.; Green, M.; Dona, A. C.; Yuen, A. H. Y.; David, M.; Berry, D. J.; Chappell, K.; Horneffer-van der Sluis, V.; Shaw, R.; Lovestone, S.; Elliott, P.; Shockcor, J.; Lindon, J. C.; Cloarec, O.; Takats, Z.; Holmes, E.; Nicholson, J. K. *Analytical Chemistry* **2016**, *88*, 9004–9013.

(208) Chen, Y.; Shen, G.; Zhang, R.; He, J.; Zhang, Y.; Xu, J.; Yang, W.; Chen, X.; Song, Y.; Abliz, Z. *Analytical Chemistry* **2013**, *85*, 7659–7665.

(209) Chetwynd, A. J.; Abdul-Sada, A.; Holt, S. G.; Hill, E. M. *Journal of Chromatography A* **2016**, *1431*, 103–110.

(210) Wu, Y.; Li, L. *Analytical Chemistry* **2012**, *84*, 10723–10731.

(211) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. *Analytical Chemistry* **2006**, *78*, 4281–4290.

(212) Sysi-Aho, M.; Katajamaa, M.; Yetukuri, L.; Orešič, M. *BMC Bioinformatics* **2007**, *8*, 93.

(213) Dunn, W. B.; Broadhurst, D.; Ellis, D. I.; Brown, M.; Halsall, A.; O'Hagan, S.; Spasic, I.; Tseng, A.; Kell, D. B. *International Journal of Epidemiology* **2008**, *37*, i23–i30.

(214) Kohl, S. M.; Klein, M. S.; Hochrein, J.; Oefner, P. J.; Spang, R.; Gronwald, W. *Metabolomics* **2012**, *8*, 146–160.

(215) Zukunft, S.; Prehn, C.; Röhring, C.; Möller, G.; Hrabě de Angelis, M.; Adamski, J.; Tokarz, J. *Metabolomics* **2018**, *14*, 18.

(216) Leuthold, L. A.; Heudi, O.; Déglon, J.; Raccuglia, M.; Augsburger, M.; Picard, F.; Kretz, O.; Thomas, A. *Analytical Chemistry* **2015**, *87*, 2068–2071.

(217) Lenk, G.; Sandkvist, S.; Pohanka, A.; Stemme, G.; Beck, O.; Roxhed, N. *Bioanalysis* **2015**, *7*, 2085–2094.

(218) Meesters, R. J.; Zhang, J.; van Huizen, N. A.; Hooff, G. P.; Gruters, R. A.; Luider, T. M. *Bioanalysis* **2012**, *4*, 2027–2035.

(219) Youhnovski, N.; Bergeron, A.; Furtado, M.; Garofolo, F. *Rapid Communications in Mass Spectrometry* **2011**, *25*, 2951–2958.

(220) Volani, C.; Caprioli, G.; Calderisi, G.; Sigurdsson, B. B.; Rainer, J.; Gentilini, I.; Hicks, A. A.; Pramstaller, P. P.; Weiss, G.; Smarason, S. V.; Paglia, G. *Analytical and Bioanalytical Chemistry* **2017**, *409*, 6263–6276.

(221) Capiau, S.; Wilk, L. S.; Aalders, M. C. G.; Stove, C. P. *Analytical Chemistry* **2016**, *88*, 6538–6546.

(222) Capiau, S.; Stove, V. V.; Lambert, W. E.; Stove, C. P. *Analytical Chemistry* **2013**, *85*, 404–410.

(223) Grintzalis, K.; Georgiou, C. D.; Schneider, Y.-J. *Analytical Biochemistry* **2015**, *480*, 28–30.

(224) Kertesz, V.; Van Berkel, G. J. *Journal of Mass Spectrometry* **2010**, *45*, 252–260.

(225) Edwards, R. L.; Creese, A. J.; Baumert, M.; Griffiths, P.; Bunch, J.; Cooper, H. J. *Analytical Chemistry* **2011**, *83*, 2265–2270.

(226) Martin, N. J.; Griffiths, R. L.; Edwards, R. L.; Cooper, H. J. *Journal of The American Society for Mass Spectrometry* **2015**, *26*, 1320–1327.

(227) Griffiths, R. L.; Dexter, A.; Creese, A. J.; Cooper, H. J. *Analyst* **2015**, *140*, 6879–6885.

(228) Rosting, C.; Yu, J.; Cooper, H. J. *Journal of Proteome Research* **2018**, *17*, 1997–2004.

(229) Van Berkel, G. J.; Kertesz, V.; Koeplinger, K. A.; Vavrek, M.; Kong, A.-N. T. *Journal of Mass Spectrometry* **2008**, *43*, 500–508.

(230) Bailey, M. J.; Randall, E. C.; Costa, C.; Salter, T. L.; Race, A. M.; de Puit, M.; Koeberg, M.; Baumert, M.; Bunch, J. *Analytical Methods* **2016**, *8*, 3373–3382.

(231) Davidson, R. L.; Weber, R. J. M.; Liu, H.; Sharma-Oates, A.; Viant, M. R. *GigaScience* **2016**, *5*, 10.

(232) Kirwan, J. A.; Weber, R. J. M.; Broadhurst, D. I.; Viant, M. R. *Scientific Data* **2014**, *1*, 140012.

(233) Kirwan, J. A.; Broadhurst, D. I.; Davidson, R. L.; Viant, M. R. *Analytical and Bioanalytical Chemistry* **2013**, *405*, 5147–5157.

(234) Wohlfarth, C., *Supplement to IV/16*; Lechner, M., Ed.; Landolt-Börnstein - Group IV Physical Chemistry, Vol. 24; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008.

(235) Southam, A. D.; Weber, R. J. M.; Engel, J.; Jones, M. R.; Viant, M. R. *Nature Protocols* **2017**, *12*, 310–328.

(236) Randall, E. C.; Bunch, J.; Cooper, H. J. *Analytical Chemistry* **2014**, *86*, 10504–10510.

(237) McDade, T. W.; Williams, S.; Snodgrass, J. J. *Demography* **2007**, *44*, 899–925.

(238) Zukunft, S.; Sorgenfrei, M.; Prehn, C.; Möller, G.; Adamski, J. *Chromatographia* **2013**, *76*, 1295–1305.

(239) Martial, L. C.; Aarnoutse, R. E.; Schreuder, M. F.; Henriet, S. S.; Brüggemann, R. J. M.; Joore, M. A. *PLOS ONE* **2016**, *11*, e0167433.

(240) Nichols, B. E.; Girdwood, S. J.; Shibemba, A.; Sikota, S.; Gill, C. J.; Mwananyanda, L.; Noble, L.; Stewart-Isherwood, L.; Scott, L.; Carmona, S.; Rosen, S.; Stevens, W. *Clinical Infectious Diseases* **2019**, 1–7.

(241) Imdad, A.; Yakoob, M. Y.; Sudfeld, C.; Haider, B. A.; Black, R. E.; Bhutta, Z. A. *BMC Public Health* **2011**, *11*, S20.

(242) Ray, J. *Archives of Disease in Childhood* **1995**, *73*, 275–275.

(243) Micronutrient deficiencies - https://www.who.int/nutrition/topics/vad/en/.

(244) Ross, D. A. *The Journal of Nutrition* **2002**, *132*, 2902S–2906S.

(245) Bruins, M.; Kraemer, K. *Community Eye Health* **2013**, *26*, 69–70.

(246) Tanumihardjo, S. A.; Russell, R. M.; Stephensen, C. B.; Gannon, B. M.; Craft, N. E.; Haskell, M. J.; Lietz, G.; Schulze, K.; Raiten, D. J. *The Journal of Nutrition* **2016**, *146*, 1816S–1848S.

(247) Harrison, E. H. *Annual Review of Nutrition* **2005**, *25*, 87–103.

(248) Goodman, D. S.; Blomstrand, R.; Werner, B.; Huang, H. S.; Shiratori, T. *Journal of Clinical Investigation* **1966**, *45*, 1615–1623.

(249) Riabroy, N.; Tanumihardjo, S. A. *The Journal of Nutrition* **2014**, *144*, 1188–1195.

(250) Suzuki, R.; Goda, T.; Takase, S. *The Journal of Nutrition* **1995**, *125*, 2074–2082.

(251) Grolier, P.; Cisti, A.; Daubeze, M.; Narbonne, J. F. *Nutrition Research* **1991**, DOI: 10.1016/S0271-5317(05)80347-1.

(252) Sharma, H. S.; Goyal, K. C.; Misra, U. K. *Nutrition Reports International* **1986**, *34*, 183–187.

Appendix A

Compound Discoverer Parameters

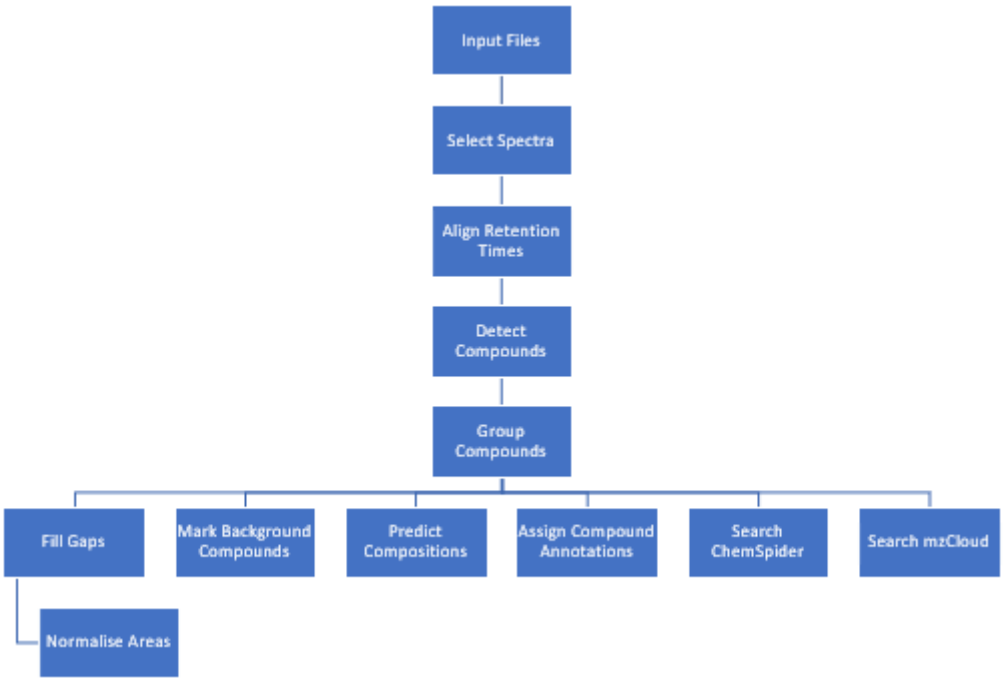


FIGURE A.1: Compound Discoverer 3.1 Workflow

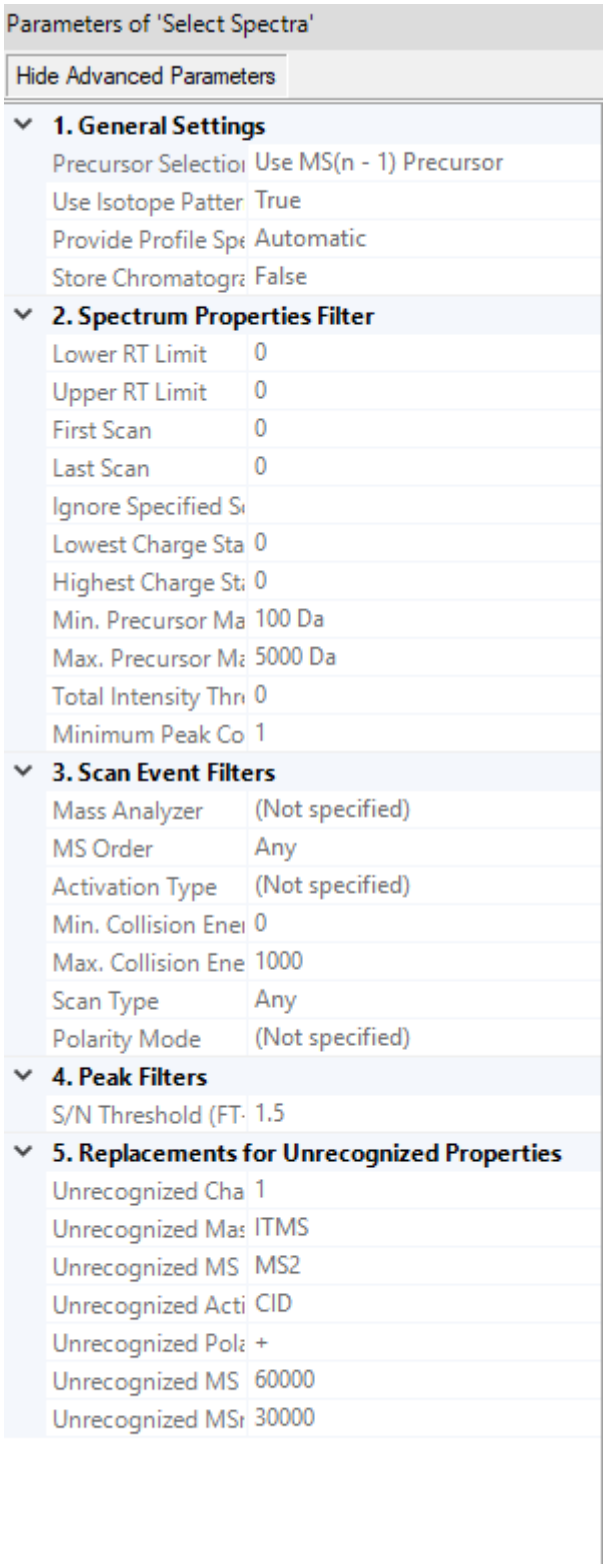


FIGURE A.2: Parameters applied for 'Select Spectra' module.

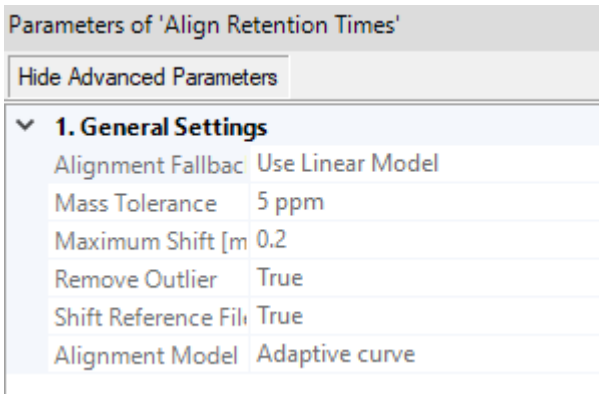


FIGURE A.3: Parameters applied for 'Align Rention Times' module.

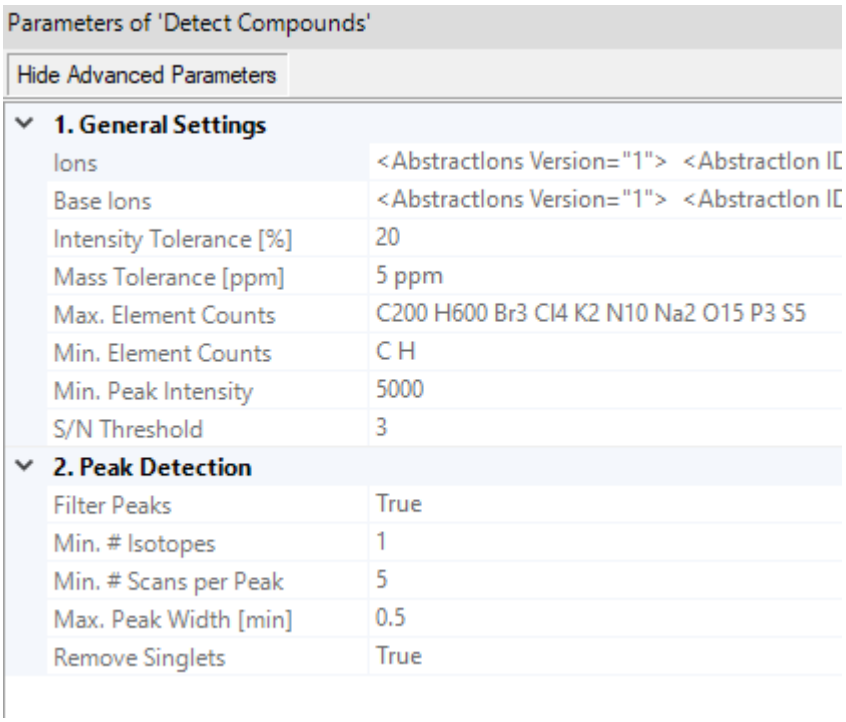


FIGURE A.4: Parameters applied for 'Detect Compounds' module.

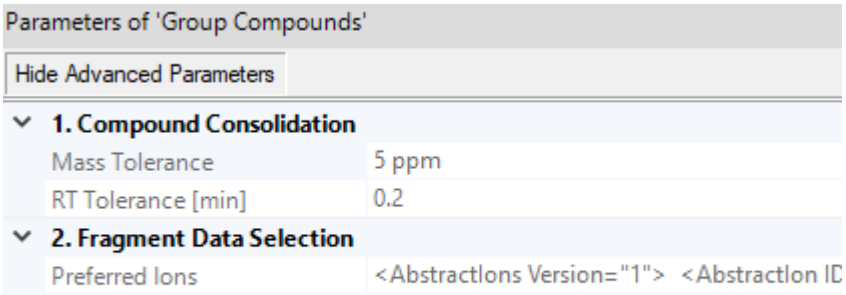


FIGURE A.5: Parameters applied for 'Group Compounds' module.

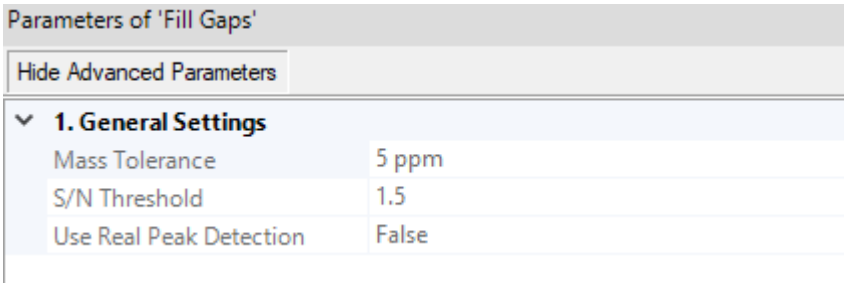


FIGURE A.6: Parameters applied for 'Fill Gaps' module.

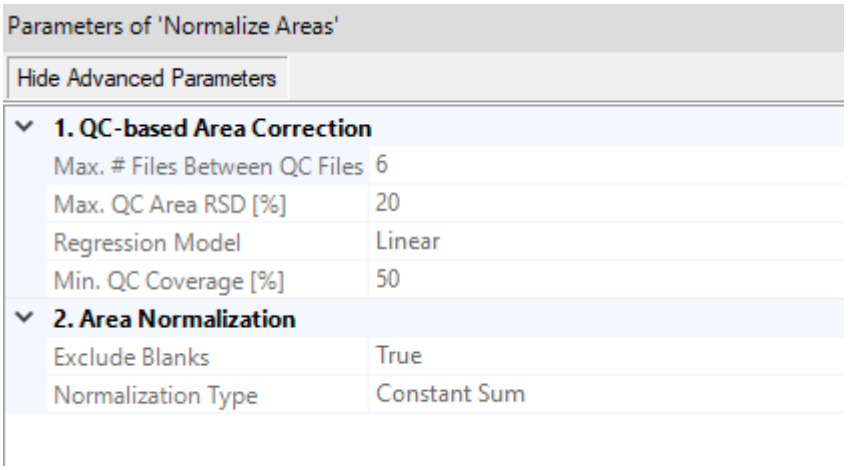


FIGURE A.7: Parameters applied for 'Normalise Areas' module.

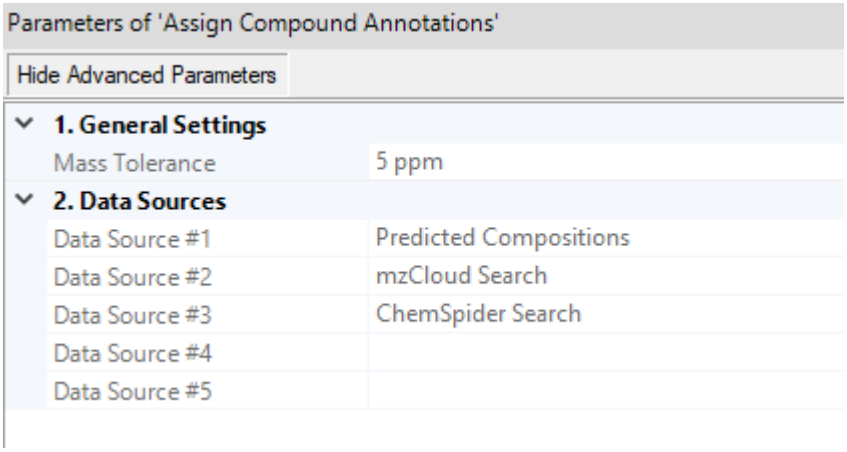


FIGURE A.8: Parameters applied for 'Assign Compound Annotations' module.

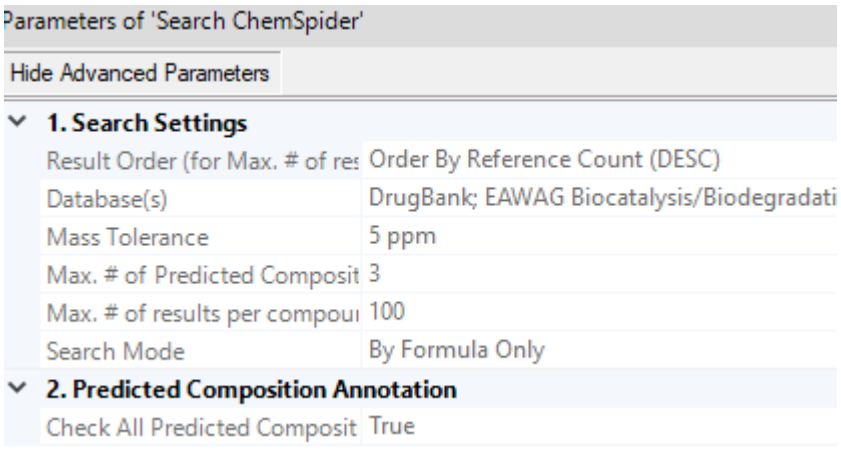


FIGURE A.9: Parameters applied for 'Search ChemSpider' module.

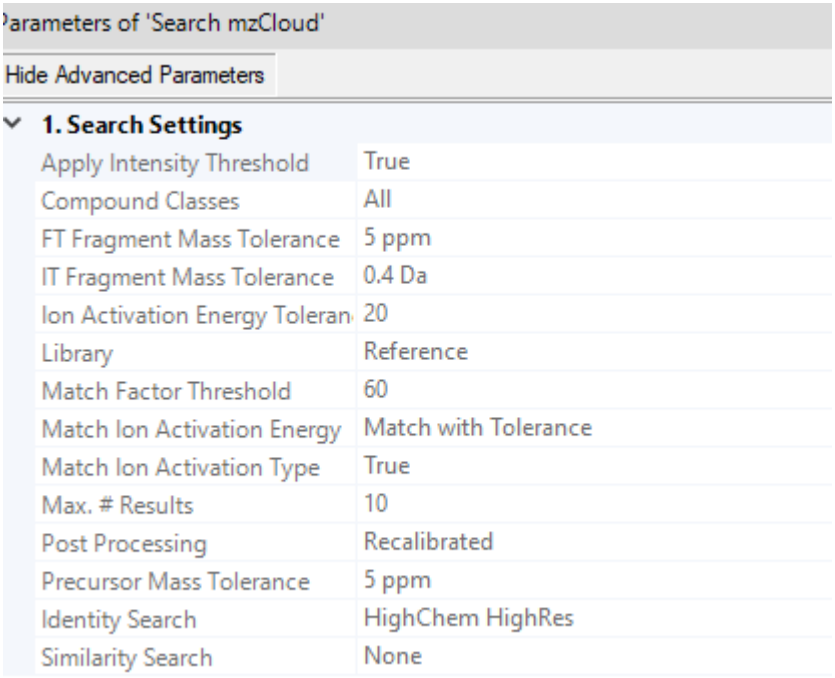


FIGURE A.10: Parameters applied for 'Search mzCloud' module.

Parameters of 'Predict Compositions'

Hide Advanced Parameters

1. Prediction Settings	
Mass Tolerance	5 ppm
Max. Element Counts	C90 H190 Br3 Cl4 N10 O18 P3 S5
Max. H/C	3.5
Max. # Candidates	10
Max. # Internal Candidates	200
Max. RDBE	40
Min. Element Counts	C H
Min. H/C	0.1
Min. RDBE	0
2. Pattern Matching	
Intensity Threshold [%]	0.1
Intensity Tolerance [%]	20
Min. Pattern Cov. [%]	80
Min. Spectral Fit [%]	30
S/N Threshold	3
Use Dynamic Recalibration	True
3. Fragments Matching	
Mass Tolerance	5 ppm
S/N Threshold	3
Use Fragments Matching	True

FIGURE A.11: Parameters applied for 'Predict Compositions' module.

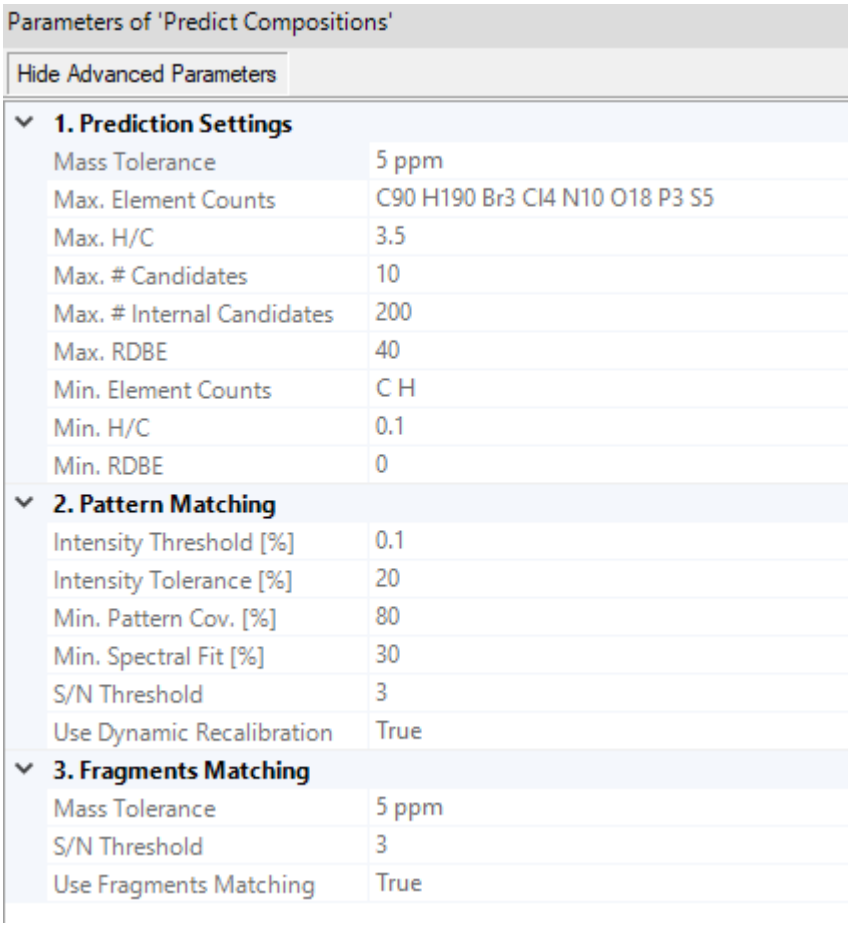


FIGURE A.12: Parameters applied for 'Mark Background Compounds' module.

Appendix B

XCMS Processing Parameters

```
#####HILIC XCMS PARAMETERS#####
xcmsParams <- list(
  'ppm' = 12,
  'peakwidth_low' = 4,
  'peakwidth_high' = 30,
  'snthresh' = 5,
  'prefilter_low' = 5,
  'prefilter_high' = 1000,
  'mzCenterFun' = "wMean",
  'integrate' = 1L,
  'mzdiff' = 0.001,
  'fitguass' = FALSE,
  'noise' = 1000,
  'gapinit' = 0.5,
  'gapExtend' = 2.4,
  'minfrac' = 0.5,
  'bw' = 0.25,
  'binsize' = 0.01,
  'verboseColumns' = TRUE
)
```

```
#####LIPIDS XCMS PARAMETERS#####
xcmsParams <- list(
  'ppm' = 14,
  'peakwidth_low' = 6,
  'peakwidth_high' = 30,
  'snthresh' = 5,
  'prefilter_low' = 5,
  'prefilter_high' = 1000,
  'mzCenterFun' = "wMean",
  'integrate' = 1L,
  'mzdiff' = 0.001,
  'fitguass' = FALSE,
  'noise' = 1000,
  'gapinit' = 0.4,
  'gapExtend' = 2.4,
  'minfrac' = 0.5,
  'bw' = 0.25,
  'binsize' = 0.01,
  'verboseColumns' = TRUE
)
#####REVERSED PHASE XCMS PARAMETERS#####
xcmsParams <- list(
  'ppm' = 11,
  'peakwidth_low' = 3,
  'peakwidth_high' = 30,
  'snthresh' = 5,
  'prefilter_low' = 5,
  'prefilter_high' = 1000,
  'mzCenterFun' = "wMean",
  'integrate' = 1L,
```

```
  'mzdiff' = 0.001,
  'fitguass' = FALSE,
  'noise' = 1000,
  'gapinit' = 0.4,
  'gapExtend' = 2.4,
  'minfrac' = 0.5,
  'bw' = 0.5,
  'binsize' = 0.01,
  'verboseColumns' = TRUE
)
```

Appendix C

Stability Study - PCA Plots

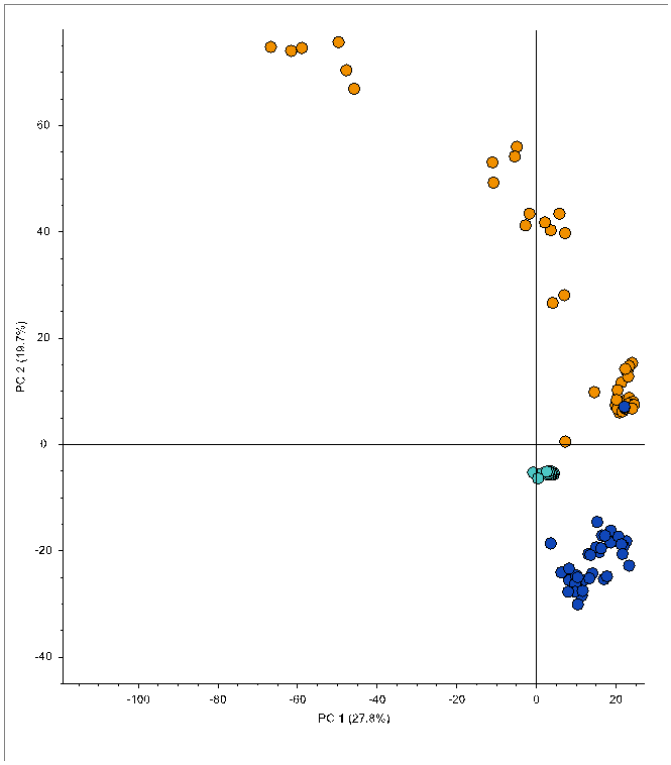


FIGURE C.1: PCA scores plot for DBS and plasma samples collected over 180 days at three different storage temperatures and analysed applying a lipidomic positive ion mode UHPLC-MS assay. DBS samples are dark blue in colour, plasma samples are orange in colour and QC samples are light blue in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

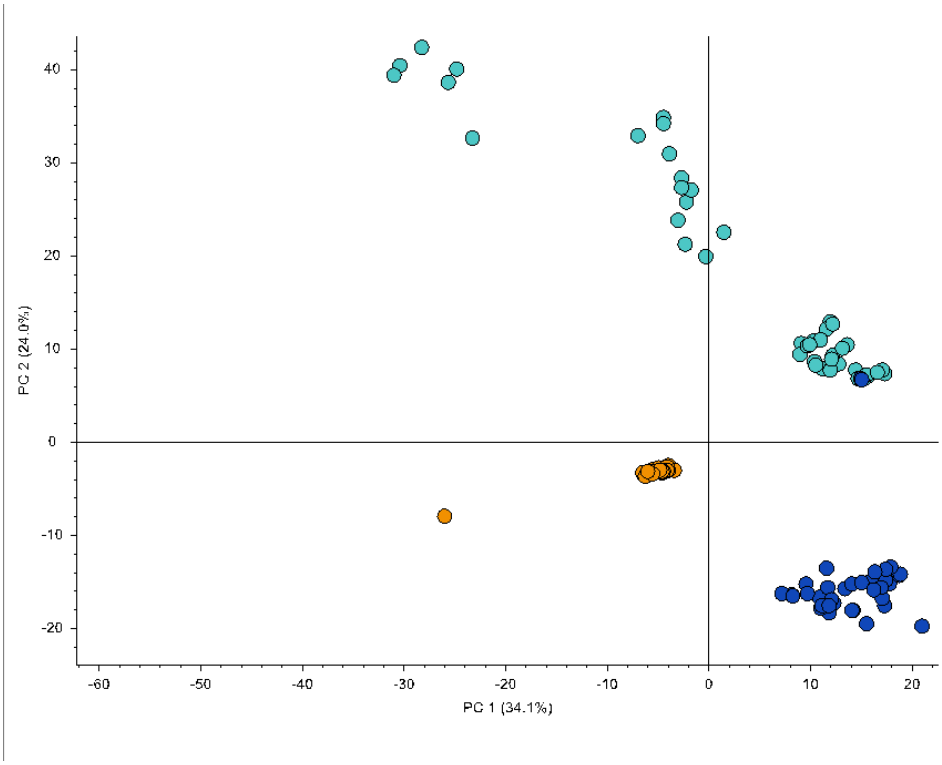


FIGURE C.2: PCA scores plot for DBS and plasma samples collected over 180 days at three different storage temperatures and analysed applying a lipidomic negative ion mode UHPLC-MS assay. DBS samples are dark blue in colour, plasma samples are light blue in colour and QC samples are orange in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

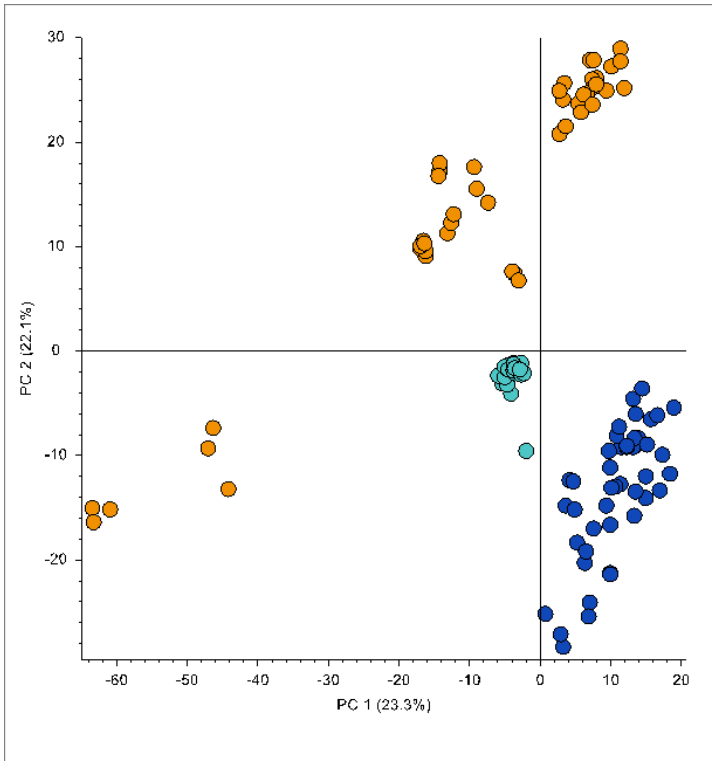


FIGURE C.3: PCA scores plot for DBS and plasma samples collected over 365 days at three different storage temperatures and analysed applying a HILIC positive ion mode UHPLC-MS assay. DBS samples are dark blue in colour, plasma samples are orange in colour and QC samples are light blue in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

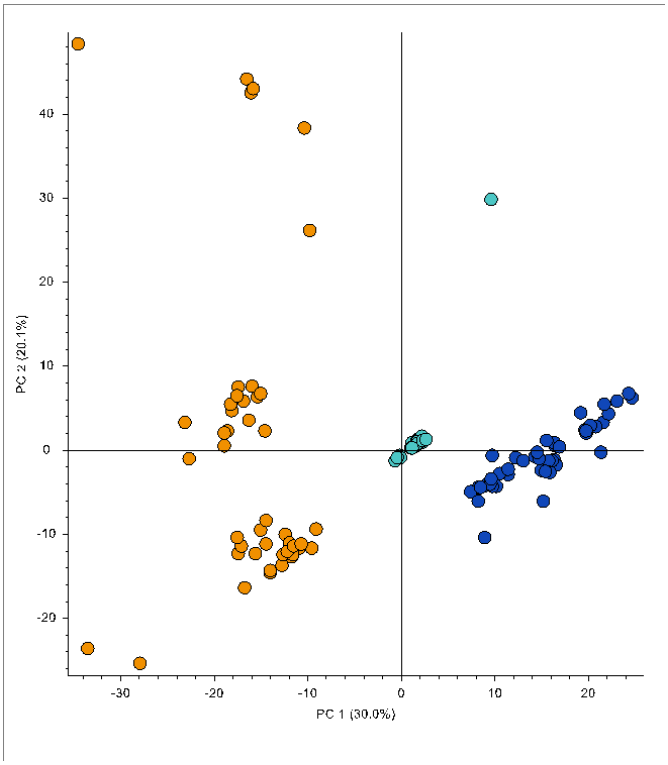


FIGURE C.4: PCA scores plot for DBS and plasma samples collected over 365 days at three different storage temperatures and analysed applying a HILIC negative ion mode UHPLC-MS assay. DBS samples are dark blue in colour, plasma samples are orange in colour and QC samples are light blue in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

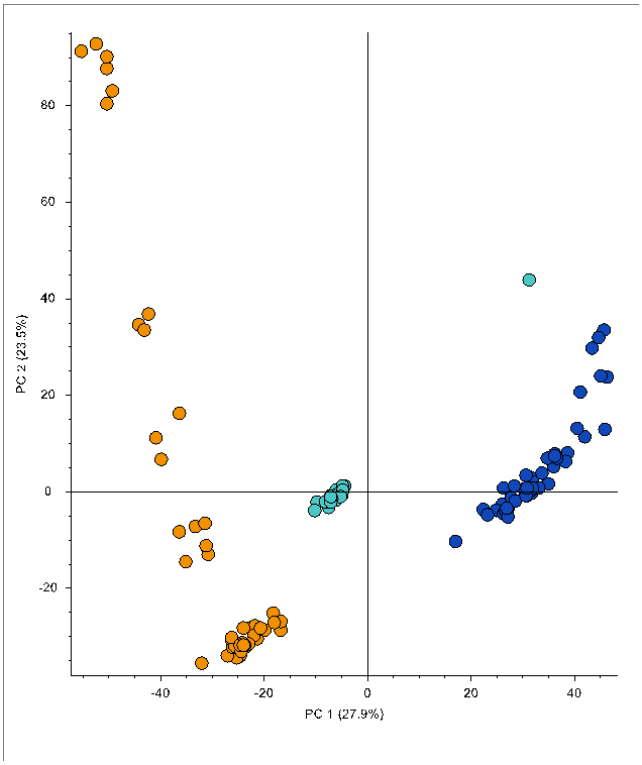


FIGURE C.5: PCA scores plot for DUS and urine samples collected over 365 days at three different storage temperatures and analysed applying a HILIC positive ion mode UHPLC-MS assay. DUS samples are dark blue in colour, urine samples are orange in colour and QC samples are light blue in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

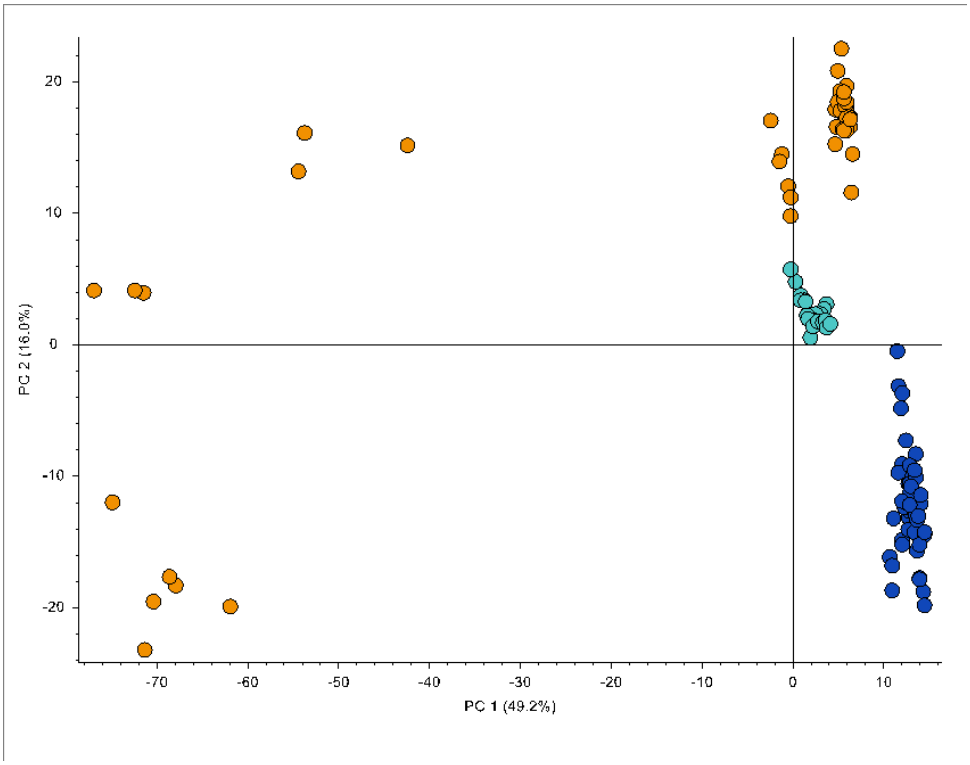


FIGURE C.6: PCA scores plot for DUS and urine samples collected over 365 days at three different storage temperatures and analysed applying a HILIC negative ion mode UHPLC-MS assay. DUS samples are dark blue in colour, urine samples are orange in colour and QC samples are light blue in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

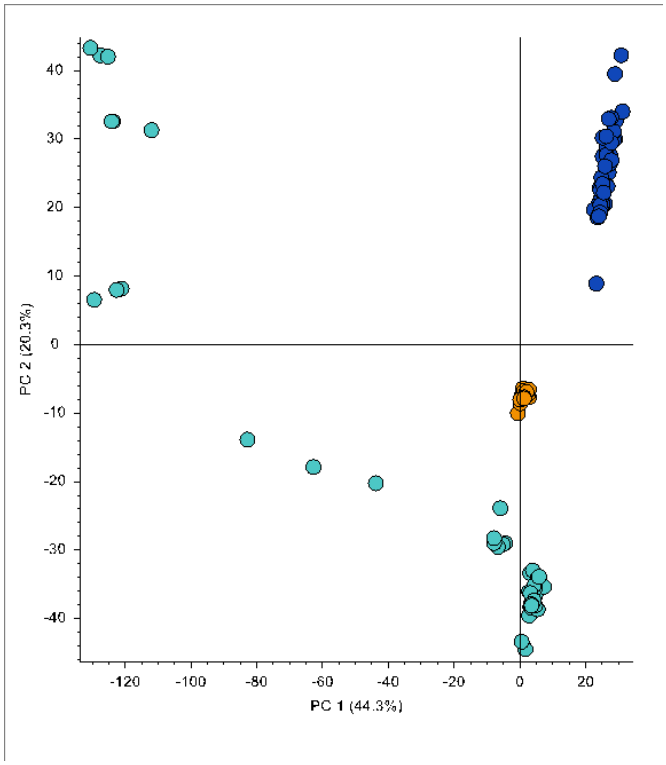


FIGURE C.7: PCA scores plot for DUS and urine samples collected over 365 days at three different storage temperatures and analysed applying a reversed phase positive ion mode UHPLC-MS assay. DUS samples are dark blue in colour, urine samples are light blue in colour and QC samples are orange in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

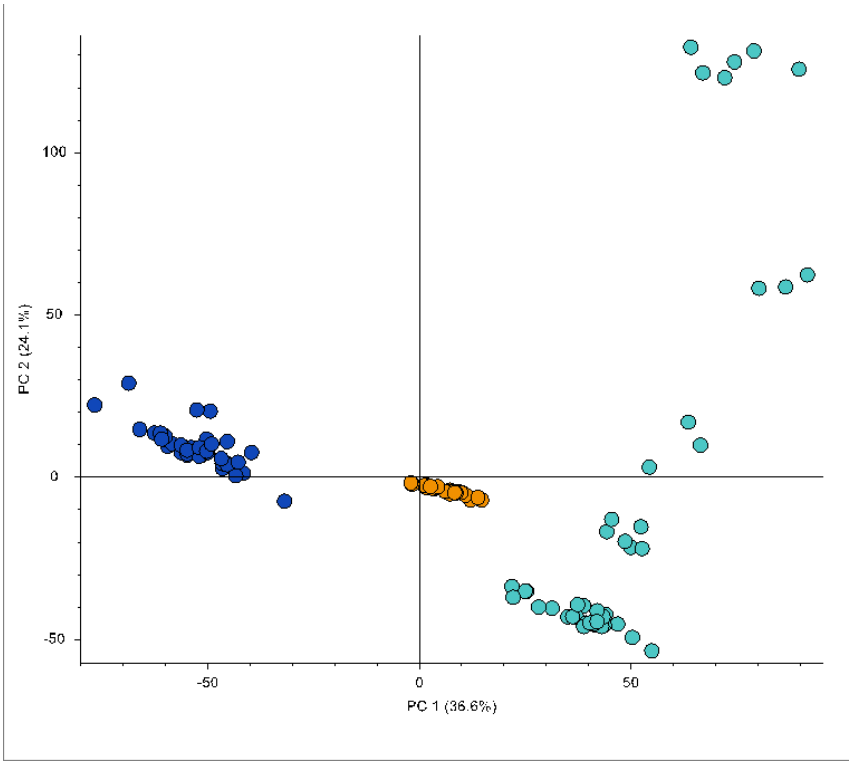


FIGURE C.8: PCA scores plot for DUS and urine samples collected over 365 days at three different storage temperatures and analysed applying a reversed phase negative ion mode UHPLC-MS assay. DUS samples are dark blue in colour, urine samples are light blue in colour and QC samples are orange in colour. Sample quantities may not be equal due to low quality samples which were removed in the filtering process due to high numbers of missing values.

Appendix D

Stability Study - Base Peak Chromatograms

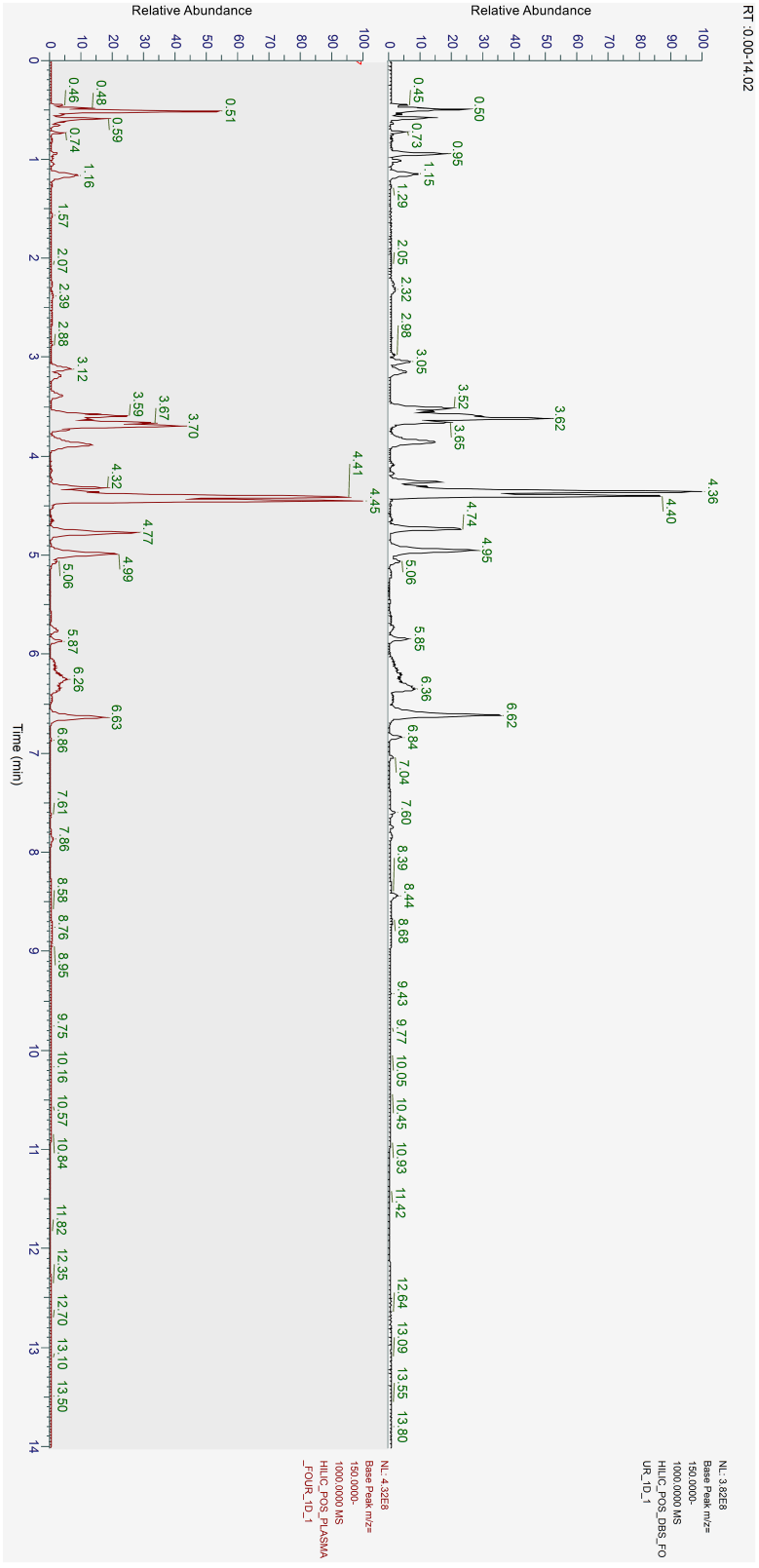


FIGURE D.1: Base peak chromatogram of both DBS (top) and plasma (bottom) applying a HILIC assay in positive ion mode.

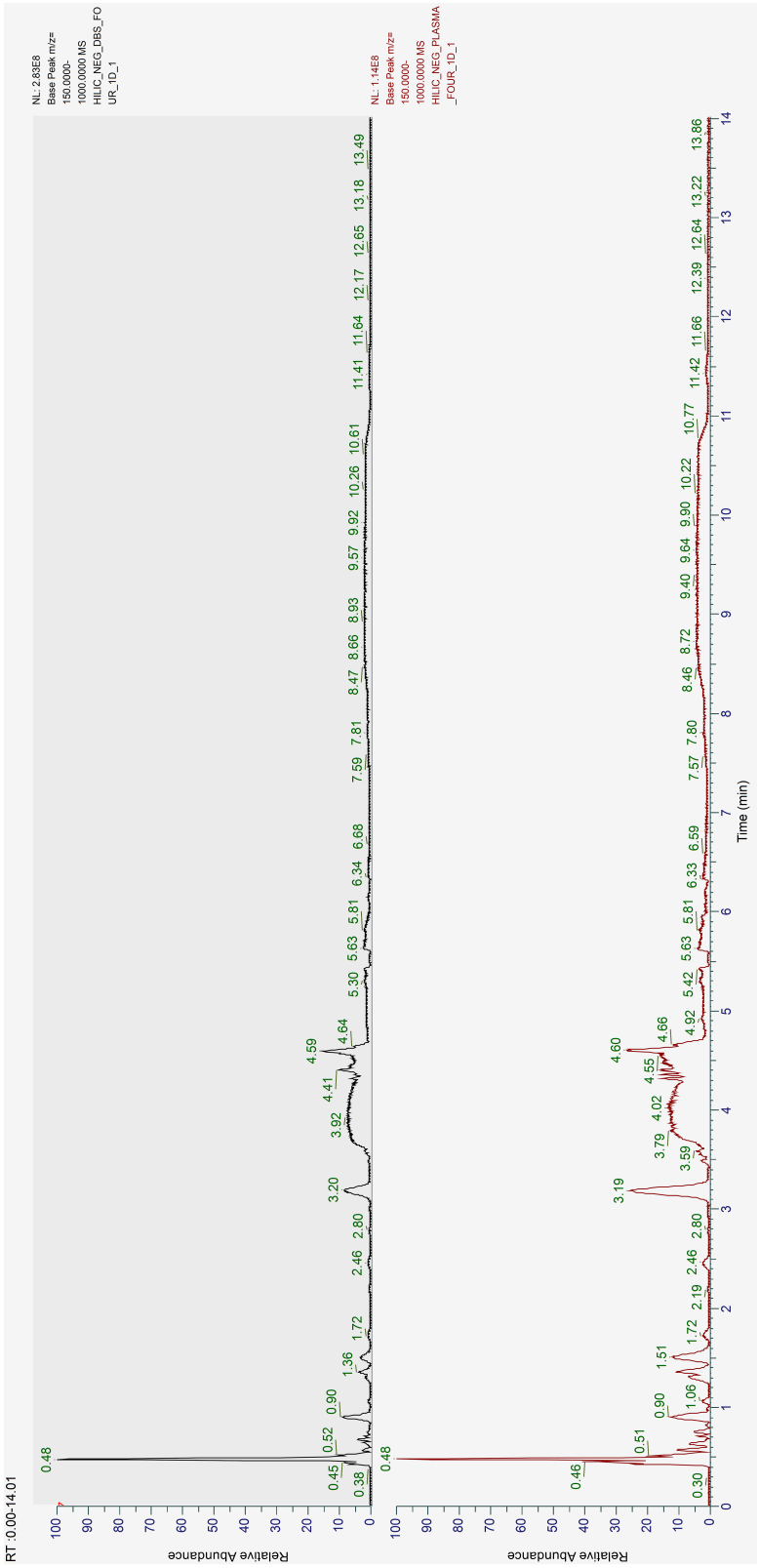


FIGURE D.2: Base peak chromatogram of both DBS (top) and plasma (bottom) applying a HILIC assay in negative ion mode.

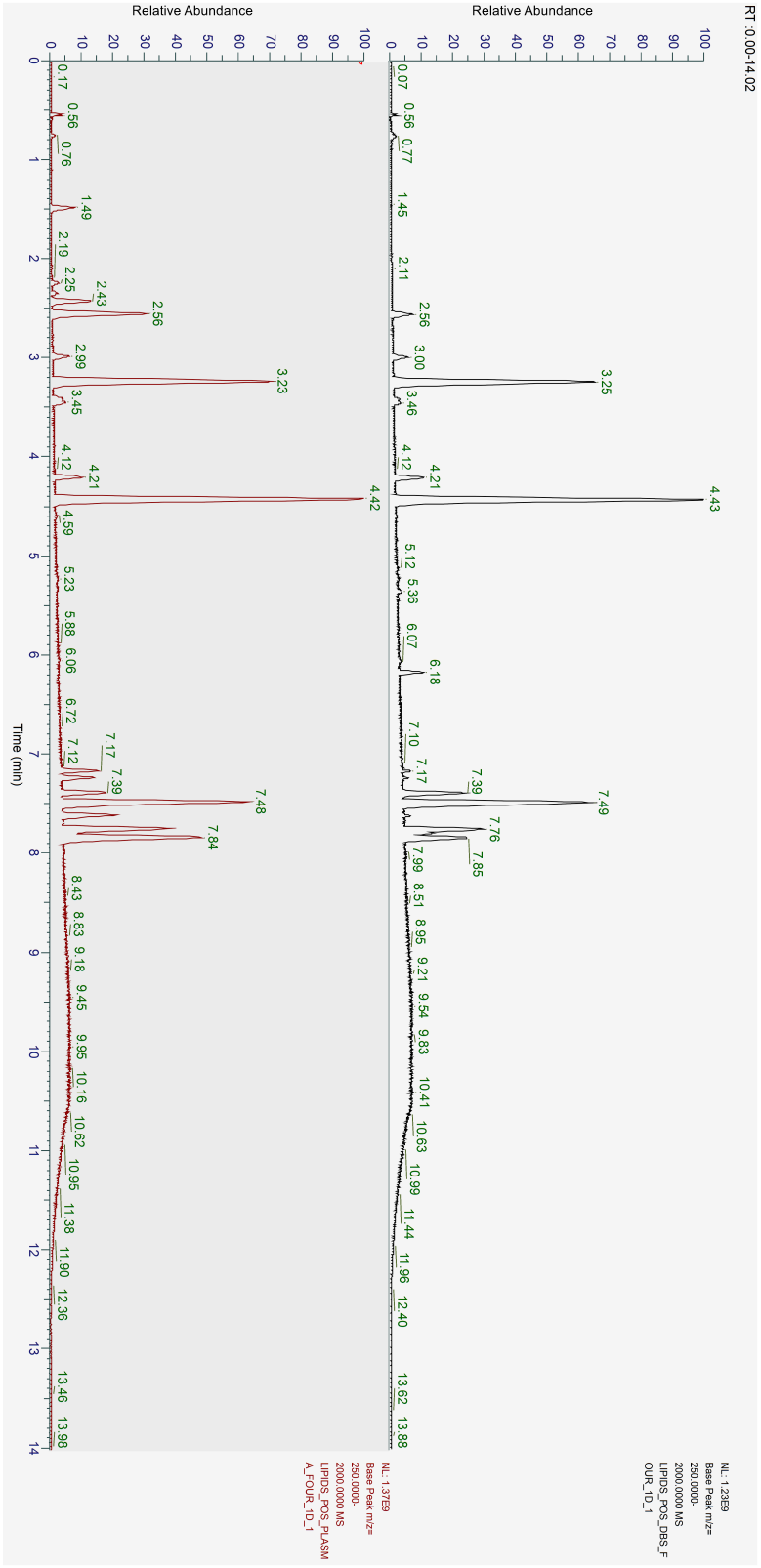


FIGURE D.3: Base peak chromatogram of both DBS (top) and plasma (bottom) applying a lipidomics assay in positive ion mode.

FIGURE D.4: Base peak chromatogram of both DBS (top) and plasma (bottom) applying a lipidomics assay in negative ion mode..

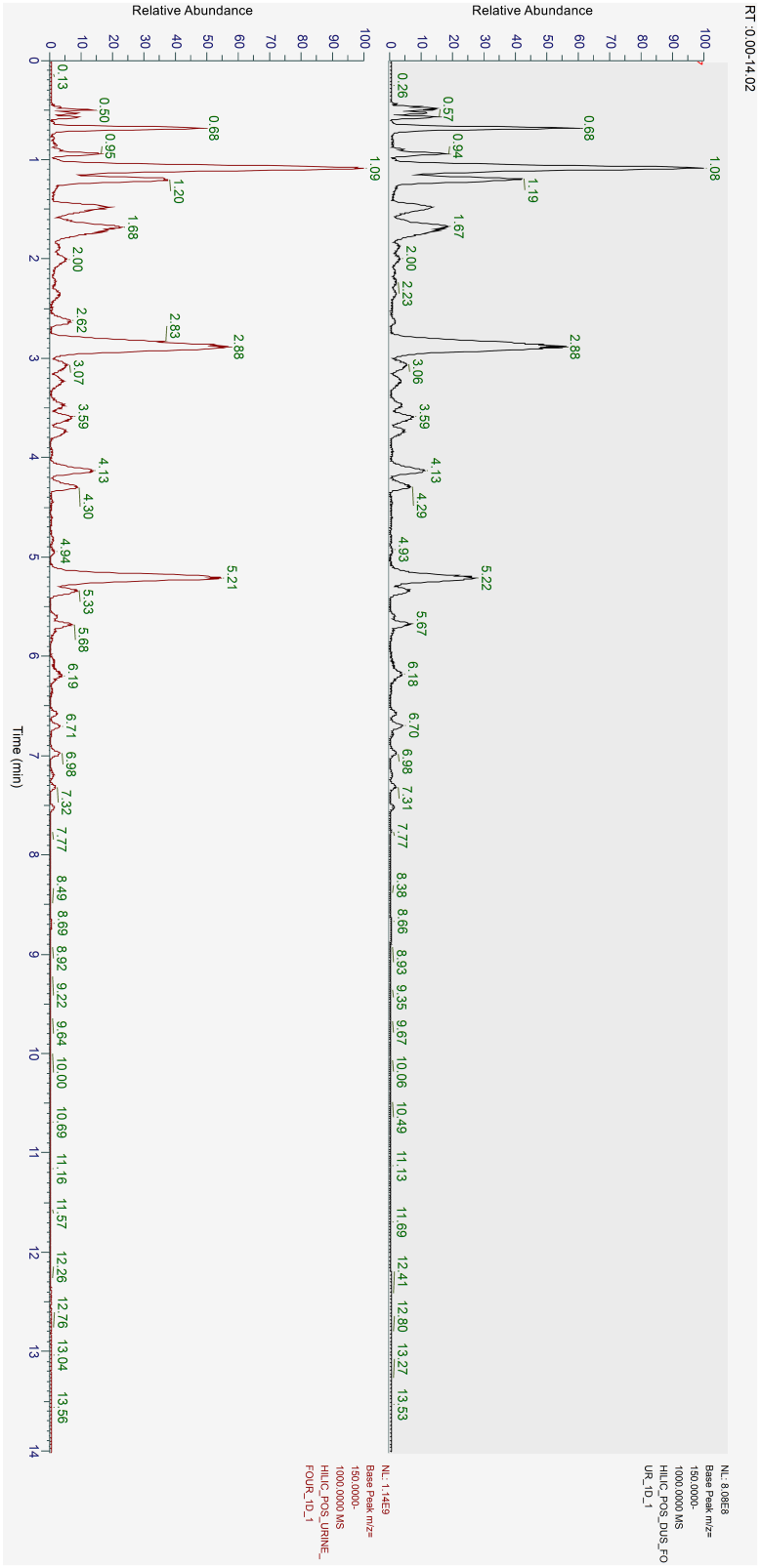


FIGURE D.5: Base peak chromatogram of both DUS (top) and urine (bottom) applying a HILIC assay in positive ion mode.

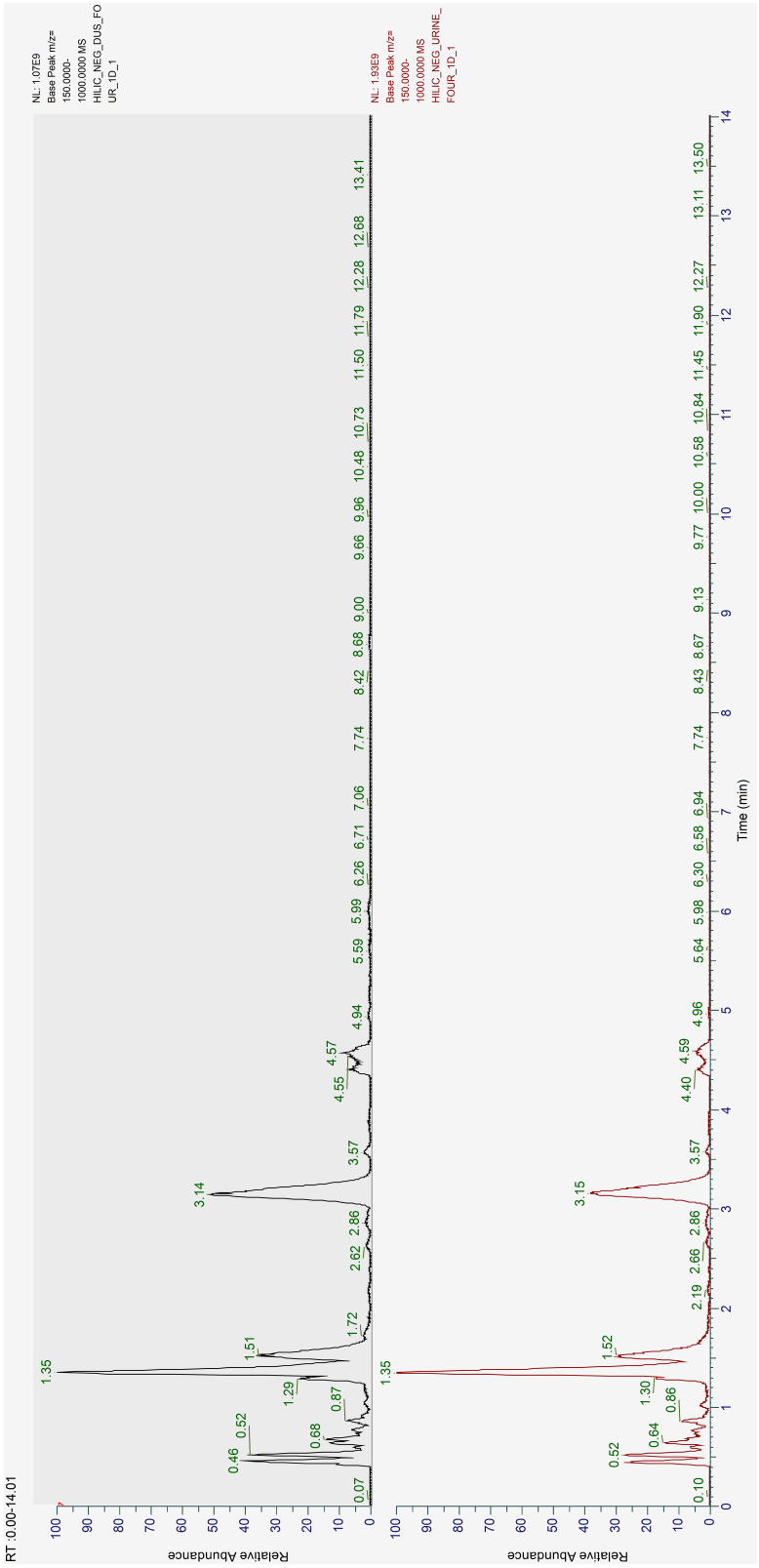


FIGURE D.6: Base peak chromatogram of both DUS (top) and urine (bottom) applying a HILIC assay in negative ion mode.

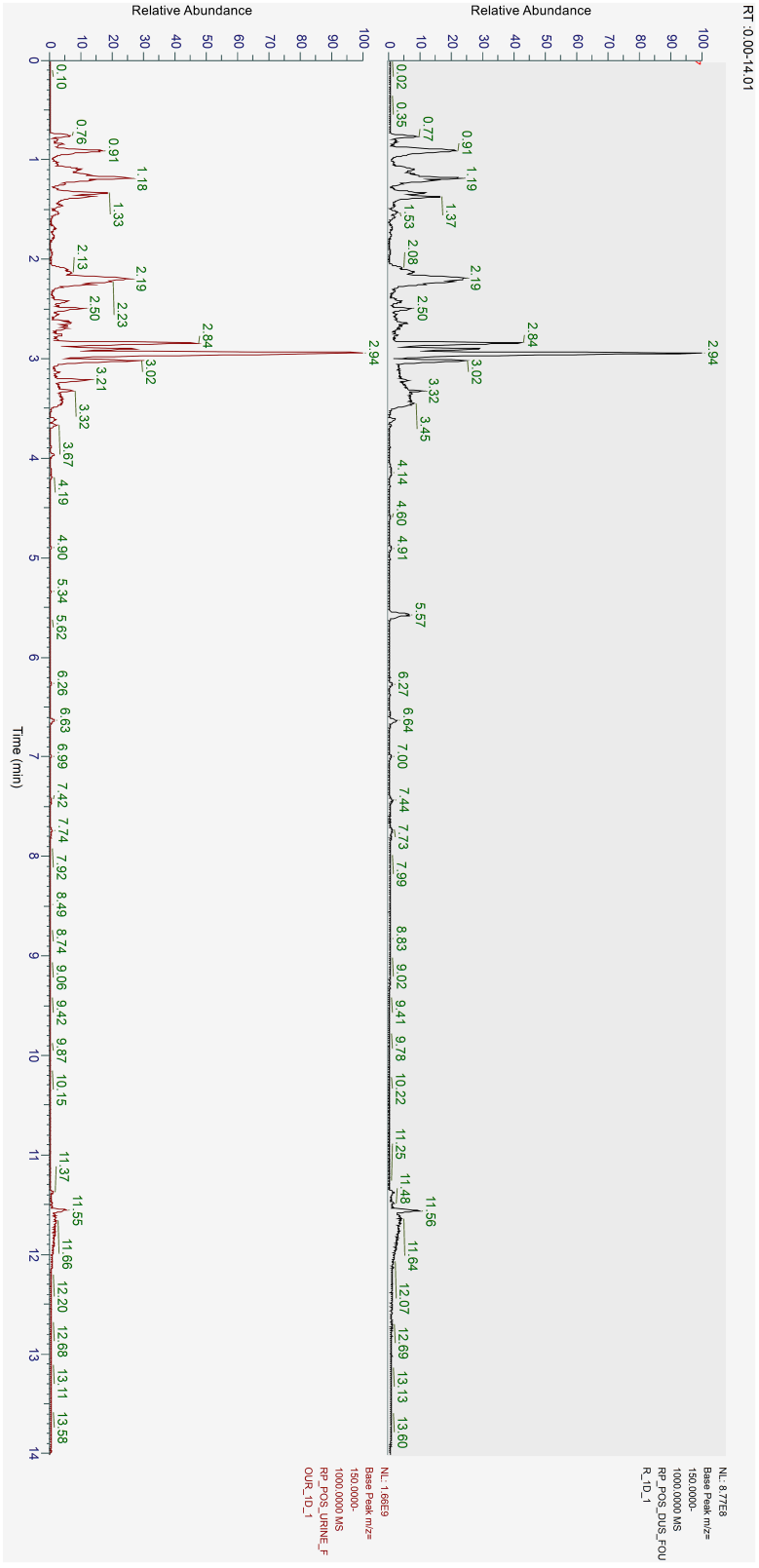


FIGURE D.7: Base peak chromatogram of both DUS (top) and urine (bottom) applying a reversed-phase assay in positive ion mode.

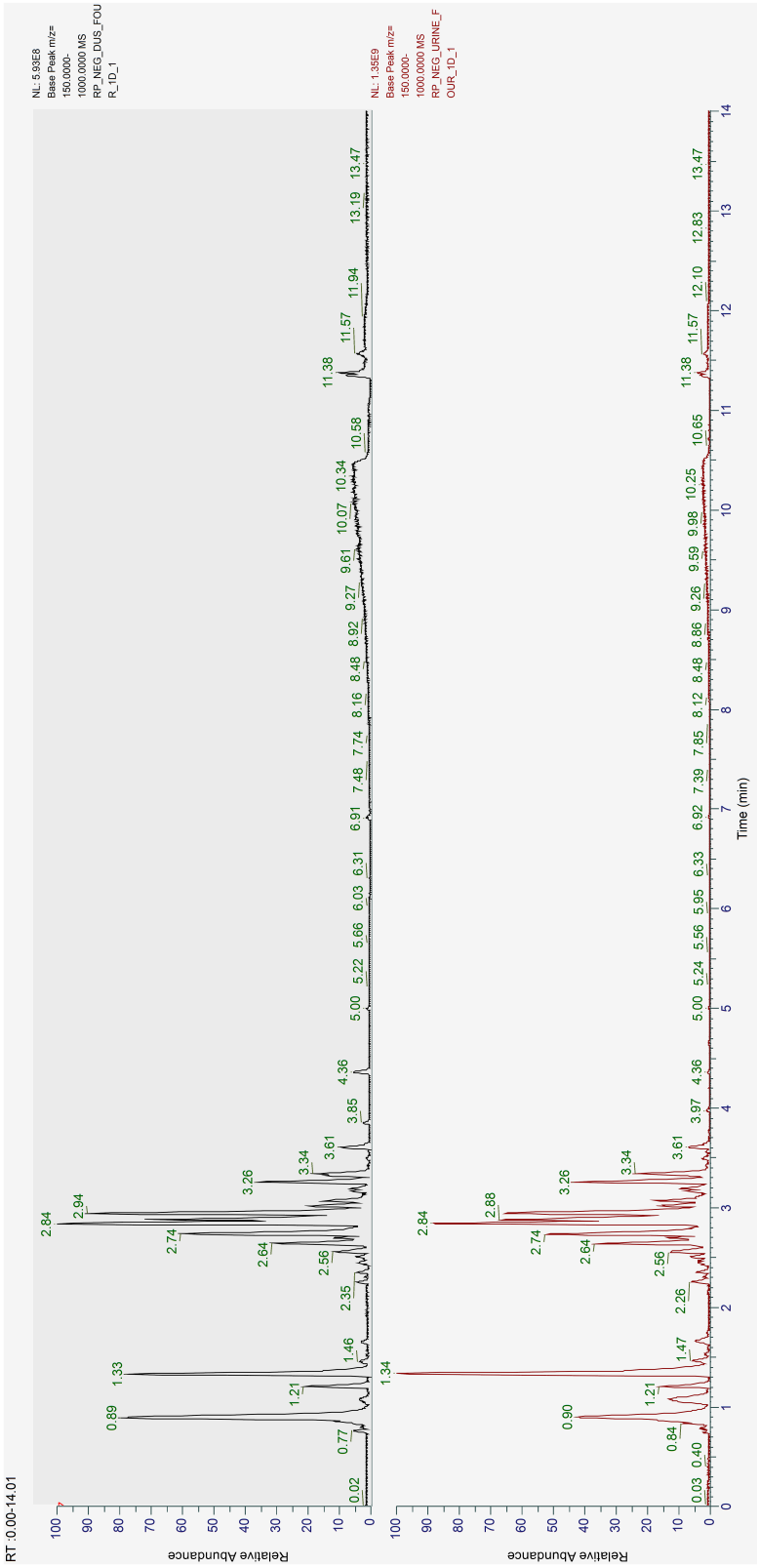


FIGURE D.8: Base peak chromatogram of both DUS (top) and urine (bottom) applying a reversed-phase assay in negative ion mode.

Appendix E

Stability Study - Significant feature plots

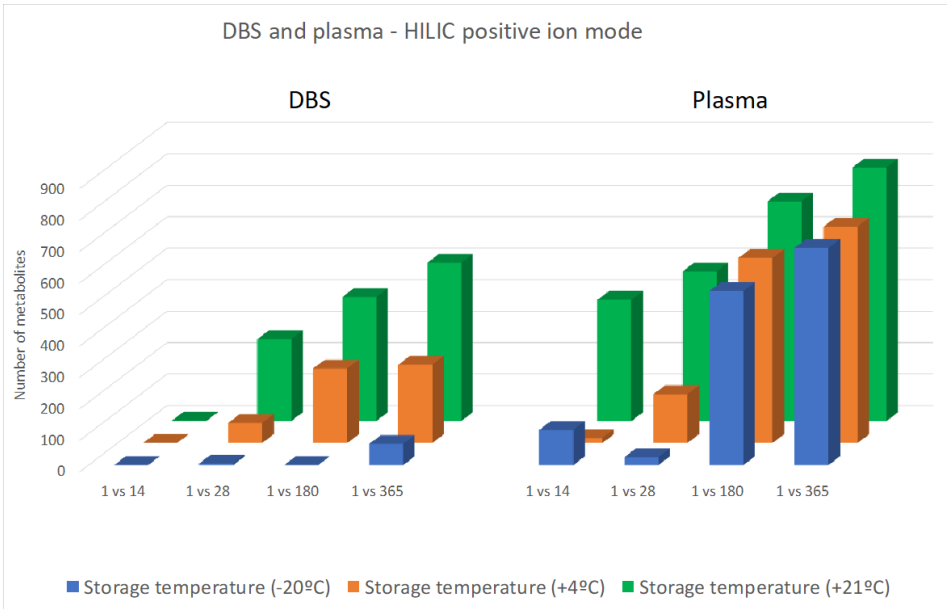


FIGURE E.1: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the HILIC positive results for the DBS and plasma comparison.

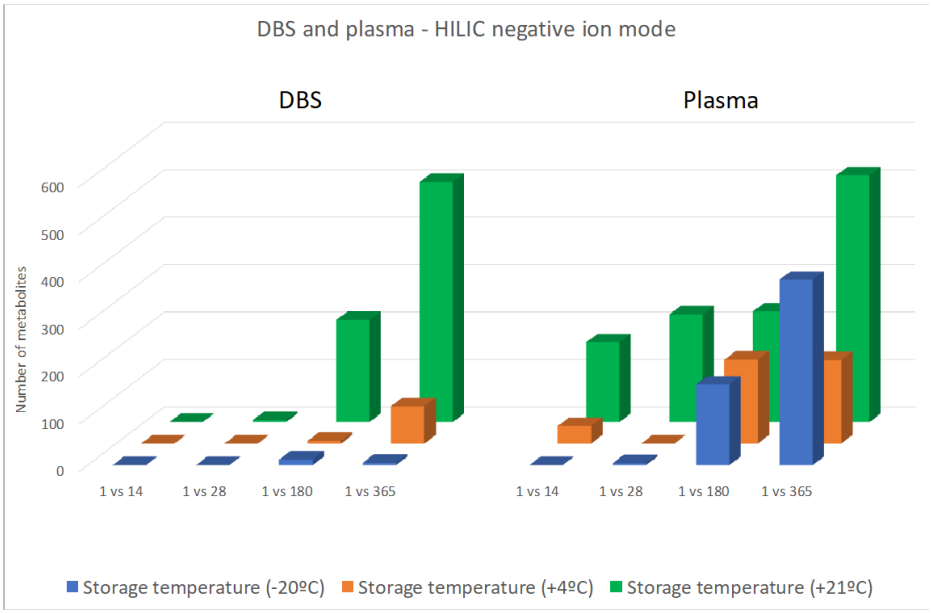


FIGURE E.2: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the HILIC negative results for the DBS and plasma comparison.

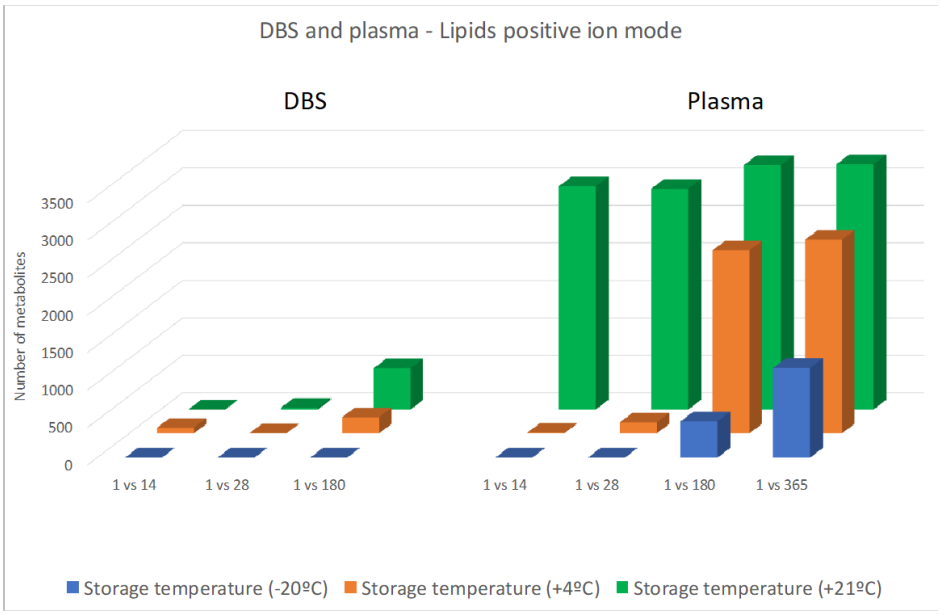


FIGURE E.3: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the lipidomics positive results for the DBS and plasma comparison.

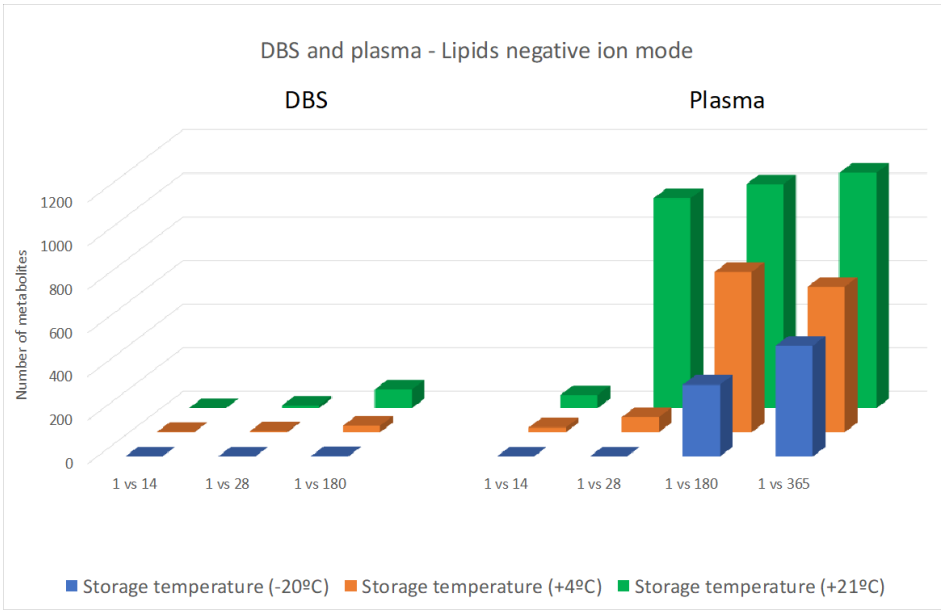


FIGURE E.4: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the HILIC positive results for the DBS and plasma comparison.

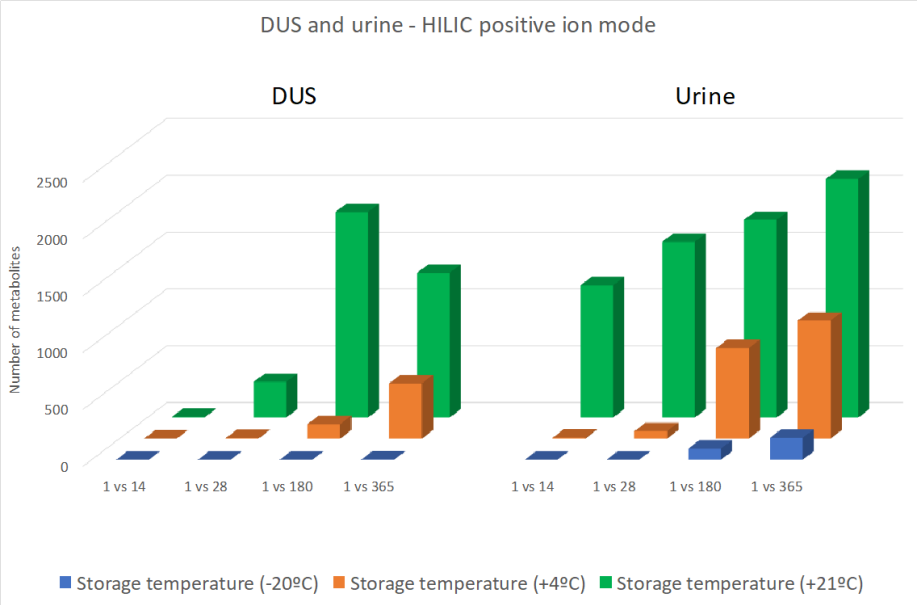


FIGURE E.5: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the HILIC positive results for the DUS and urine comparison.

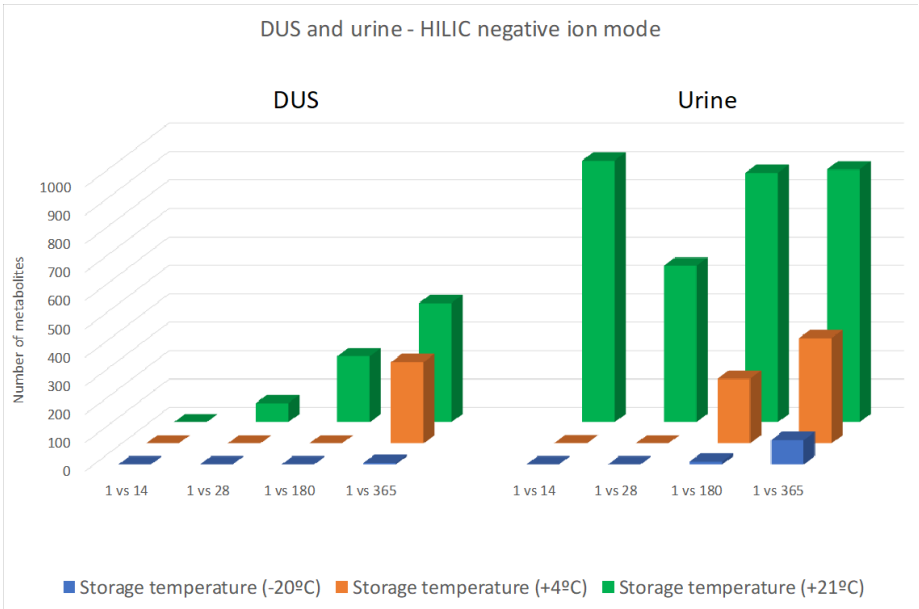


FIGURE E.6: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the HILIC negative results for the DUS and urine comparison.

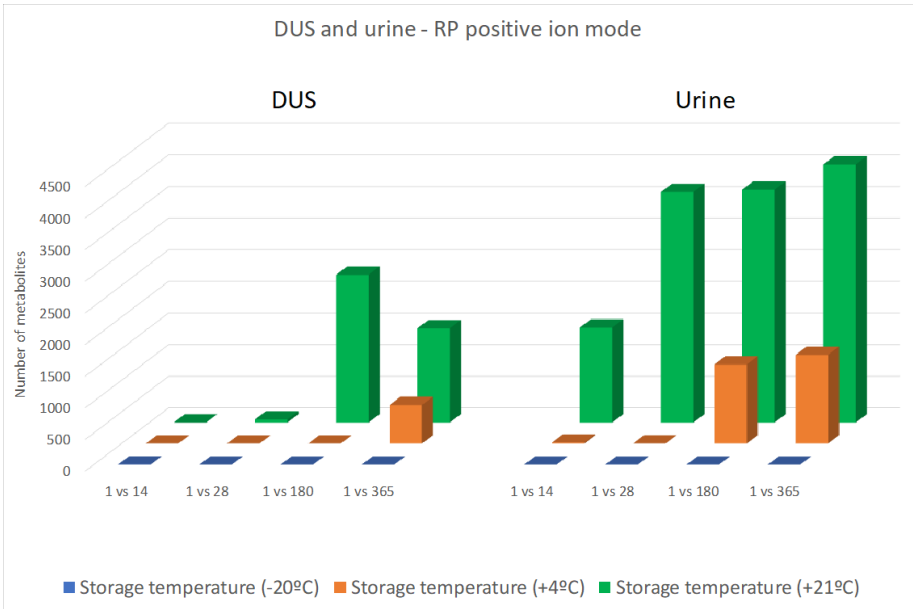


FIGURE E.7: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the reversed-phase positive results for the DUS and urine comparison.

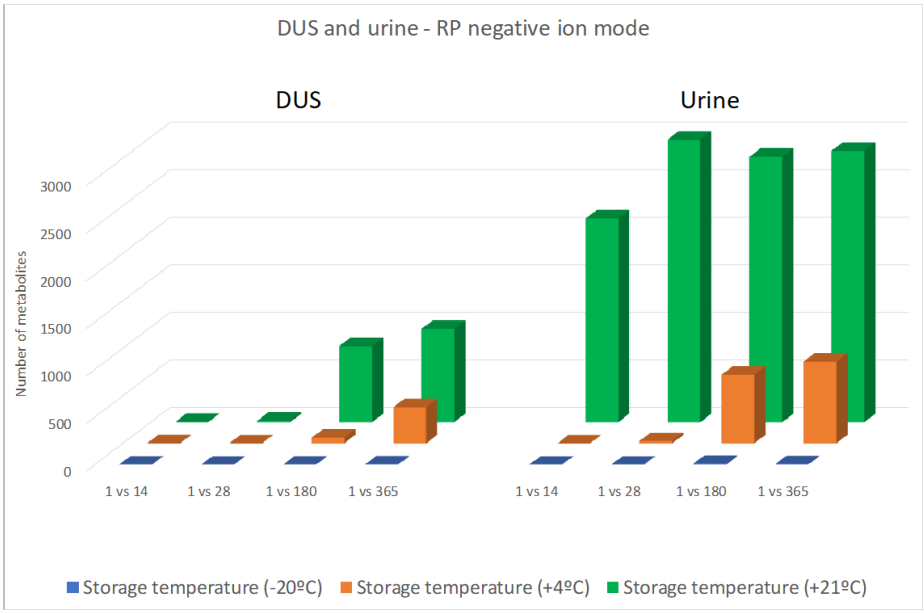


FIGURE E.8: Plot demonstrating the number of statistically significant metabolites at each time point and temperature. Plot represents the reversed-phase negative results for the DUS and urine comparison.

Appendix F

Stability Study - Significant Metabolites at Room Temperature

TABLE F.1: List of annotated metabolites which are statistically significant between day 28 and day 1 in DBS at room temperature.

P-value (following correction for multiple testing using Benjamin-Hochberg method)	Ratio between d28/d1	Metabolite name(s)	MW (g/mol)	RT (min)	Assay
0.04088908	2.792	Carboxymethylproline	173.0689	1.325	HILIC negative
0.013562818	0.195	Nonenoylglycine	213.13623	1.553	HILIC positive
0.028649307	1.51	Adenine	135.05443	2.025	HILIC positive
0.028649307	8.716	N1,N8-Diacetylspermidine	229.17871	7.567	HILIC positive
0.030135996	0.279	Tryptophan	204.08963	2.51	HILIC positive
0.030135996	0.382	5-Hydroxyindoleacetylglycine	248.0794	4.15	HILIC positive
0.030135996	1.762	Dehydrocholesterol	384.3389	0.468	HILIC positive
0.030135996	3.278	LysoCPA(18:0)	420.26327	1.297	HILIC positive
0.034093704	0.121	(d18:1) sphingosine	299.28195	2.59	HILIC positive
0.034093704	0.268	Acetyl-N-formyl-5-methoxykynurenamine	264.11054	0.613	HILIC positive
0.034093704	0.342	N5-Phenyl-L-glutamine	222.10012	2.445	HILIC positive
0.034093704	0.354	Isoleucyl-Valine	230.16268	0.738	HILIC positive

0.034093704	1.325	1,2-Diamino-4-nitrobenzene	153.05361	2.842	HILIC positive
0.034093704	1.53	Dihydroxyoxo-octadecenoic acid	328.22432	0.549	HILIC positive
0.034093704	1.756	Diamino-nitrotoluene	167.06926	1.01	HILIC positive
0.034093704	1.83	4-(beta-Acetylaminoethyl)imidazole	153.08974	0.592	HILIC positive
0.034093704	1.936	Streptamine phosphate	258.06111	5.363	HILIC positive
0.034093704	2.579	Delta-Guanidinovaleric acid	159.10057	2.846	HILIC positive
0.034093704	2.616	Methylpyrrole	81.0577	3.422	HILIC positive
0.034093704	2.631	xi-4,5-Dihydro-2,4(5)-dimethyl-1H-imidazole	98.08426	3.427	HILIC positive
0.034093704	2.78	Alanyl-Asparagine	203.09037	2.98	HILIC positive
0.034093704	3.557	4-O-(beta-L-Arabinofuranosyl)-(2S,4S)-4-hydroxyproline	263.10008	0.589	HILIC positive
0.034093704	3.715	N-Acetyl-L-2-amino-6-oxopimelate	231.0739	0.617	HILIC positive
0.034093704	4.027	4-Hydroxy-5-phosphooxypentane-2,3-dione	212.00823	6.637	HILIC positive
0.034093704	4.213	LysoPE(P-16:0)	437.28994	1.308	HILIC positive
0.034093704	4.59	N-(1-Deoxy-1-fructosyl)threonine	281.11203	1.488	HILIC positive
0.034093704	4.651	Vinylacetylglycine	143.05809	1.636	HILIC positive
0.034093704	4.892	6-Imino-5-oxocyclohexa-1,3-dienecarboxylate	151.02672	0.698	HILIC positive
0.034093704	5.313	Valyl-Valine	216.1471	7.377	HILIC positive
0.035400865	1.239	3,3′-Dimethoxybenzidine	244.12083	3.397	HILIC positive
0.038185897	0.741	Salsolinol	179.09435	1.6	HILIC positive
0.038185897	1.307	4-(Trimethylammonio)but-2-enoate	143.09434	6.992	HILIC positive
0.038185897	2.681	Betahistine	136.09992	2.796	HILIC positive
0.038185897	4.531	2,3-Diaminosalicylic acid	168.05333	0.694	HILIC positive
0.038821505	0.349	1,4-Dideoxy-1,4-imino-D-ribitol	133.07369	3.988	HILIC positive

0.042035837	4.454	Decatrienoylcarnitine	309.19346	3.715	HILIC positive
0.042082564	1.908	(+/-)11,12-Ep-15S-HETrE	336.22936	0.502	HILIC positive
0.042106997	0.766	Oxo-Tetradecadienoic acid	238.15654	0.631	HILIC positive
0.04228911	4.293	Glutamylthreonine	248.10046	0.621	HILIC positive
0.042738221	1.595	Pyrroline-5-carboxylate	113.04753	0.752	HILIC positive
0.043323605	0.405	Aminocyclopropanecarboxylic acid	101.04754	2.351	HILIC positive
0.043323605	1.683	Piperazine-2-carboxamide	129.09003	3.802	HILIC positive
0.044956722	1.563	17-Methyltestosterone	302.22398	0.495	HILIC positive
0.044956722	2.146	2,3-Butanediol glucoside	252.12186	3.031	HILIC positive
0.047585791	1.836	N-Nitrosoproline	144.05332	0.683	HILIC positive
0.049339137	0.093	Hydroxy-methylbutenoic acid	116.04706	0.939	HILIC positive
0.051478333	0.698	Epinephrine sulfate	263.04602	0.753	HILIC positive
0.051478333	1.668	N-Methylhexanamide	129.11524	2.559	HILIC positive
0.051478333	2.136	Mandelonitrile	133.05258	0.579	HILIC positive
0.051800892	1.693	Methyltryptophan	218.10522	0.623	HILIC positive
0.051800892	2.079	5-Aminopentanamide	116.0948	5.91	HILIC positive
0.051800892	2.098	2-Hydroxy-2-methylbutanenitrile	99.06826	5.914	HILIC positive
0.051800892	3.112	N-Methyl-N,4-dinitrosoaniline	165.0536	5.446	HILIC positive
0.051800892	5.105	Tyrosyl-Tyrosine	344.13664	0.94	HILIC positive
0.051975492	1.662	Hydroxytridecanoyl carnitine	373.2821	0.497	HILIC positive
0.053017773	0.728	Amino-methylpentanoic acid	131.09441	5.384	HILIC positive
0.054835237	1.457	Methylhistidine	169.08527	1.062	HILIC positive
0.054835237	2.072	3-(Pyrazol-1-yl)-L-alanine	155.06926	3.227	HILIC positive
0.056623851	1.765	PG(O-16:0/18:1)	734.54651	0.47	HILIC positive
0.057928423	1.882	4-Hydroxyphenylacetaldehyde oxime	151.06312	0.651	HILIC positive
0.057928423	2.511	Methylcytosine	125.05866	4.592	HILIC positive
0.058511189	1.455	beta-Tyrosine	181.07365	0.766	HILIC positive

0.063028496	1.64	Glycerophosphocholine	257.10254	7.622	HILIC positive
0.063076878	0.762	Cytidine	243.08516	5.084	HILIC positive
0.06384563	0.427	2-amino-6-oxohexanoate	145.07367	0.61	HILIC positive
0.064486446	1.556	N1-Acetylspermidine	187.16821	7.617	HILIC positive
0.065833234	1.375	3-[4-methoxy-3-(sulfooxy)phenyl]propanoic acid	276.03073	6.897	HILIC positive
0.065833234	1.807	2-Oxoarginine	173.07979	3.327	HILIC positive
0.065895836	1.434	Nicotinate D-ribonucleotide	336.04786	5.768	HILIC positive
0.067346392	0.514	Hydroxyhexadienoate	128.04711	0.606	HILIC positive
0.067346392	1.336	1H-Indole-3-acetamide	174.07907	0.716	HILIC positive
0.067346392	1.537	N-palmitoyl threonine	357.28753	0.505	HILIC positive
0.068750415	0.768	N1-Methyl-2-pyridone-5-carboxamide	152.05838	0.997	HILIC positive
0.068949175	9.707	3-Carboxy-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole-1-propanoic acid	288.11062	1.306	HILIC positive
0.06933875	0.777	4-Oxo-4-(3-pyridyl)-butanamide	178.07396	4.888	HILIC positive
0.06933875	1.429	Naphthylamine	143.07332	2.706	HILIC positive
0.06933875	1.792	4-Hydroxy-1H-indole-3-acetonitrile	172.06342	0.64	HILIC positive
0.06933875	1.801	3,4,5,6-Tetrahydrohippuric acid	183.08932	4.86	HILIC positive
0.06933875	2.065	Hexadecadienoylcarnitine	395.30288	3.194	HILIC positive
0.06933875	2.168	Methyl5-(but-3-en-1-yl)amino-1,3,4-oxadiazole-2-carboxylate	197.07974	5.439	HILIC positive
0.06933875	2.27	Pyridosine	254.12609	0.603	HILIC positive
0.06933875	2.775	Pyroglutamylglycine	186.06384	2.99	HILIC positive
0.071207367	0.407	Hydroxylinoleoylcarnitine	439.32939	3.484	HILIC positive
0.071207367	2.609	2-Hydroxyestrone sulfate	366.11201	7.078	HILIC positive
0.074178895	2.348	2,7-Anhydro-alpha-N-acetylneuraminic acid	291.09501	6.704	HILIC positive

0.074205423	1.356	L-Histidine trimethylbetaine	197.11618	7.651	HILIC positive
0.074270377	1.965	Dihydro-4-mercapto-3(2H)-furanone	118.00875	2.042	HILIC positive
0.076115942	0.523	13'-Hydroxy-alpha-tocopherol	446.37554	0.465	HILIC positive
0.076393965	0.614	Pyridoxamine	168.08961	1.39	HILIC positive
0.081529605	1.729	7,8-Diaminononanoate	188.1522	8.469	HILIC positive
0.081529605	4.01	4-(beta-Acetylaminoethyl)imidazole	153.09038	0.932	HILIC positive
0.08638167	0.637	15-methyl-15R-PGF2alpha	368.25567	0.536	HILIC positive
0.08638167	1.558	LysoPC(16:1)	493.31652	4.157	HILIC positive
0.08638167	2.086	Lysyl-Valine	245.17366	8.239	HILIC positive
0.086930116	1.465	Nonenoic acid	156.1147	0.509	HILIC positive
0.086930116	3.465	Alanyl-Proline	186.09999	5.877	HILIC positive
0.087014995	1.347	Alpha-linolenyl carnitine	421.3183	3.164	HILIC positive
0.087014995	1.371	(+/-)11,12-EpETrE	320.23451	0.489	HILIC positive
0.09055682	1.536	2-[(4-methoxyphenyl)formamido]acetic acid	209.06847	0.707	HILIC positive
0.09055682	2.236	Ethyl-methyloxazole	111.06822	0.72	HILIC positive
0.091455023	1.413	11,18-di-HEPE	334.21357	0.518	HILIC positive
0.097910096	1.31	Palmitoylcarnitine	399.33614	3.172	HILIC positive
0.099201128	1.296	11-Hydroxy-9,15,16-trioxooctadecanoic acid	342.20349	0.523	HILIC positive
0.099201128	1.436	Dimethylnonanoyl carnitine	329.256	0.513	HILIC positive
0.099738242	2.111	11,12-DiHETrE-EA	381.28709	0.502	HILIC positive
0.019799282	0.758	PS(O-18:0/22:6)	821.55527	8.07	Lipids negative
0.078082155	1.382	Hydroxy-Octadecanoic acid	300.26653	4.842	Lipids negative
0.078082155	1.949	Dihydroxy-dioxohexanoate	176.0319	0.603	Lipids negative
0.078005797	0.235	Docosahexaenoic acid	328.23928	8.035	Lipids positive
0.078005797	0.319	DG(34:2)	592.5067	8.023	Lipids positive
0.078005797	0.79	PE(18:2/P-18:1)	725.53592	7.73	Lipids positive
0.078005797	1.55	Succinylproline	215.07941	0.628	Lipids positive
0.078005797	1.962	24,25-didehydrovitamin D3	382.32325	6.756	Lipids positive
0.078005797	2.956	N-acetyl-aminoadipate	203.07933	0.6	Lipids positive

0.078005797	4.692	PA(27:0)	578.39479	6.897	Lipids positive
0.078005797	15.074	Arginyl-Isoleucine	287.19656	0.539	Lipids positive
0.085263958	0.723	PE(22:4/P-18:1)	777.56743	7.988	Lipids positive

TABLE F.2: List of annotated metabolites which are statistically significant between day 28 and day 1 in DBS at room temperature.

P-value (following correction for multiple testing using Benjamin-Hochberg method)	Ratio between d365/d1	Metabolite name(s)	MW (g/-mol)	RT (min)	Assay
0.0167	0.05	(9R,10S,12Z)-9,10-Dihydroxy-8-oxo-12-octadecenoic acid	328.2253	0.49	HILIC neg
0.0172	0.05	Ascorbate 2-sulfate	255.9891	0.64	HILIC neg
0.0377	0.06	hydroxy-heptadecatrienoic acid	280.2037	0.50	HILIC neg
0.0108	0.07	4-Carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate	216.0272	0.64	HILIC neg
0.0240	0.07	Trihydroxybenzoic acid	170.0218	0.64	HILIC neg
0.0351	0.09	Choline phosphate	184.0736	0.55	HILIC neg
0.0293	0.10	D-Arabinonate	166.0478	1.15	HILIC neg
0.0195	0.11	Vanillyl alcohol	154.0631	0.65	HILIC neg
0.0135	0.11	5-Hydroxyisourate	184.0232	1.49	HILIC neg
0.0503	0.11	Heptadiynediol	124.0526	0.53	HILIC neg
0.0518	0.13	Epoxy-hexadecenoic acid	268.2041	0.49	HILIC neg
0.0179	0.17	2-Hydroxyglutarate	148.0372	1.12	HILIC neg
0.0119	0.18	Dihydroxybutyric acid	120.0423	1.16	HILIC neg
0.0252	0.21	Undecanedioic acid	216.1363	0.50	HILIC neg
0.0187	0.23	2-Aminobut-2-enoate	101.0477	5.59	HILIC neg
0.0090	0.23	N4-Acetylcytidine	285.0963	2.48	HILIC neg
0.0197	0.23	Propylglutaric acid	174.0892	1.33	HILIC neg
0.0179	0.25	Methyl-pentenoic acid	114.0681	1.33	HILIC neg
0.0235	0.26	5,6-Dihydroxy-3-methyl-2-oxo-1,2-dihydroquinoline	191.0584	0.84	HILIC neg
0.0254	0.28	S-(N-Hydroxy-N-methylcarbamoyl)glutathione	380.0989	1.51	HILIC neg

0.0183	0.28	N-Acetyl-D-galactosamine	221.0901	5.47	HILIC neg
0.0400	0.29	Propionic acid	74.0367	1.34	HILIC neg
0.0186	0.30	Methyl-hexenoate	128.0837	1.34	HILIC neg
0.0834	0.31	Monomethyl sulfate	111.9831	0.86	HILIC neg
0.0115	0.31	Trihydroxybenzene	126.0318	0.78	HILIC neg
0.0267	0.32	Trihydroxybenzene	126.0318	0.58	HILIC neg
0.0584	0.32	26-Glucosyl-1,3,11,22-tetrahydroxyergosta-5,24-dien-26-oate	638.3648	0.52	HILIC neg
0.0231	0.32	3-Ureidoisobutyrate	146.0692	6.82	HILIC neg
0.0100	0.38	Deoxyuridine	228.0748	1.31	HILIC neg
0.0254	0.40	2-Hydroxy-2,4-pentadienoate	114.0319	1.33	HILIC neg
0.0179	0.40	Acrylic acid	72.0212	1.34	HILIC neg
0.0132	0.41	PC(31:1)	717.5319	3.01	HILIC neg
0.0293	0.41	Sulfolithocholylglycine	513.2763	2.85	HILIC neg
0.0275	0.47	(5Z,9E,12S,14Z)-8,11,12-Trihydroxyicosa-5,9,14-trienoate	354.2409	0.49	HILIC neg
0.0351	0.47	N-Acetylaspartylglutamic acid	304.0910	1.86	HILIC neg
0.0135	0.48	Hydroxydodecanoic acid	216.1726	0.54	HILIC neg
0.0318	0.48	4-Carboxypyrazole	112.0274	1.14	HILIC neg
0.0184	0.48	(22R)-22,25-dihydroxyvitamin D3 / (22R)-22,25-dihydroxycholecalciferol	408.2882	0.78	HILIC neg
0.0254	0.50	Butenoate	86.0367	1.02	HILIC neg
0.0092	0.51	Octadienol	126.1045	0.58	HILIC neg
0.0281	0.52	N-(1-Deoxy-1-fructosyl)histidine	317.1225	3.75	HILIC neg
0.0716	0.52	Taurallocholic acid	515.2920	3.72	HILIC neg
0.0252	0.53	hydroxybutyric acid	104.0474	1.08	HILIC neg
0.0395	0.54	hydroxybutyric acid	104.0474	0.96	HILIC neg
0.0186	0.54	N'-Phosphoguanidinoethyl methyl phosphate	277.0237	5.30	HILIC neg
0.0184	0.55	3-Nitrotyrosine	226.0592	1.14	HILIC neg
0.0365	0.56	3-Pyridinebutanoic acid	165.0790	4.93	HILIC neg
0.0219	0.56	Adenosine	267.0971	2.22	HILIC neg
0.0235	0.58	N-Ribosylhistidine	287.1120	3.50	HILIC neg
0.0651	0.59	2-Methylserine	119.0583	6.59	HILIC neg
0.0595	0.60	Octenol	128.1202	0.57	HILIC neg
0.0448	0.60	p-Cresol sulfate	188.0144	0.46	HILIC neg

0.0429	0.60	Pyrogallol-1-O-sulphate	205.9886	0.87	HILIC neg
0.0440	0.60	Octenediol	144.1151	0.65	HILIC neg
0.0398	0.60	(1-Ribosylimidazole)-4-acetate	258.0855	1.86	HILIC neg
0.0248	0.61	Parabanic Acid	114.0067	0.63	HILIC neg
0.0395	0.61	4-Carboxypyrazole	112.0274	1.03	HILIC neg
0.0599	0.61	Acetolactate	132.0423	2.54	HILIC neg
0.0379	0.61	3′-Amino-3′-deoxythimidine	241.1065	3.50	HILIC neg
0.0535	0.61	Vanillin 4-sulfate	232.0042	1.40	HILIC neg
0.0423	0.62	5,6-Dihydrouridine	246.0854	2.72	HILIC neg
0.0599	0.62	2-Deoxypentonic acid	150.0529	2.61	HILIC neg
0.0286	0.62	Taurochenodeoxycholic acid	499.2974	2.93	HILIC neg
0.0377	0.63	hydroxyicosanoic acid	328.2981	0.58	HILIC neg
0.0651	0.63	5-Carboxy-2′-deoxyuridine	272.0649	1.13	HILIC neg
0.0518	0.65	hydroxy-tetradecenoic acid	242.1884	0.53	HILIC neg
0.0503	0.65	Imidazoleacetic acid	126.0430	0.87	HILIC neg
0.0340	0.65	Mandelonitrile	133.0528	1.50	HILIC neg
0.0286	0.66	Taurochenodeoxycholic acid	499.2974	3.10	HILIC neg
0.0212	0.69	Oxoheptanoic acid	144.0787	0.63	HILIC neg
0.0557	0.69	Dihydroferulic acid 4-sulfate	276.0307	1.07	HILIC neg
0.0866	0.70	Phenylacetic acid	136.0526	2.09	HILIC neg
0.0941	0.70	(R)-3-(4-Hydroxyphenyl)lactate	182.0580	2.09	HILIC neg
0.0354	0.70	4-Amino-5-oxopentanoate	131.0583	2.04	HILIC neg
0.0254	0.75	3-Hydroxy-4-methylanthranilate	167.0583	0.85	HILIC neg
0.0357	0.78	PS(39:1)	831.6004	3.56	HILIC neg
0.0863	1.21	Heptenol	114.1045	0.60	HILIC neg
0.0725	1.26	Creatinine	113.0591	2.96	HILIC neg
0.0724	1.39	3,5-Dihydroxy-phenylglycine	183.0532	1.35	HILIC neg
0.0673	1.40	1H-Indole-3-carboxaldehyde	145.0528	0.49	HILIC neg
0.0630	1.51	3,4-Methylenedioxybenzoic acid	166.0266	2.75	HILIC neg
0.0426	1.55	Lactate	90.0315	3.99	HILIC neg
0.0149	1.56	N-Acetyl-L-2-amino-6-oxopimelate	231.0744	0.66	HILIC neg
0.0405	1.67	PG(O-16:0/18:0)	736.5651	0.47	HILIC neg

0.0354	1.94	2-[(4-methoxyphenyl)formamido]acetic acid	209.0689	0.64	HILIC neg
0.0328	2.01	3-hydroxybenzoic acid-3-O-sulphate	217.9883	0.82	HILIC neg
0.0254	2.08	1H-Indole-3-carboxaldehyde	145.0528	0.74	HILIC neg
0.0183	2.13	2-Indolecarboxylic acid	161.0477	0.55	HILIC neg
0.0826	2.16	4-Hydroxy-2-butenoic acid gamma-lactone	84.0213	0.73	HILIC neg
0.0395	2.16	Caffeic acid 3-sulfate	259.9992	0.71	HILIC neg
0.0977	2.27	(2-Methoxyethoxy)propanoic acid	148.0736	1.72	HILIC neg
0.0475	2.28	Tyrosol 4-sulfate	218.0249	0.42	HILIC neg
0.0377	2.44	Hydroxyhippuric acid	195.0532	1.35	HILIC neg
0.0184	2.58	Formyl-5-hydroxykynurenamine	208.0849	2.50	HILIC neg
0.0092	2.67	Pentenoic acid	100.0525	0.58	HILIC neg
0.0184	2.67	Methyl-hexenoate	128.0838	0.84	HILIC neg
0.0158	2.77	Trihydroxytoluene	140.0474	0.50	HILIC neg
0.0503	2.78	Benzeneacetonitrile	117.0580	2.73	HILIC neg
0.0159	2.81	Hydroxybenzaldehyde	122.0369	0.50	HILIC neg
0.0464	2.90	2-Oxosuccinamate	131.0220	0.64	HILIC neg
0.0236	3.02	3-hydroxy-3-(3-hydroxyphenyl)propanoic acid-O-sulphate	262.0148	0.68	HILIC neg
0.0286	3.04	Adenine	135.0546	1.89	HILIC neg
0.0187	3.30	2-Isopropyl-3-oxosuccinate	174.0528	1.36	HILIC neg
0.0167	3.37	2-Amino-2-deoxyisochorismate	225.0639	3.28	HILIC neg
0.0250	3.43	Phenol	94.0418	2.75	HILIC neg
0.0138	4.06	11-deoxy-16,16-dimethyl-PGE2	364.2618	0.48	HILIC neg
0.0145	4.18	Methylhexanoic acid	130.0995	0.81	HILIC neg
0.0135	4.66	3,4-Dihydroxybenzylamine	139.0634	0.68	HILIC neg
0.0139	5.48	3,5-Dihydroxy-phenylglycine	183.0532	0.68	HILIC neg
0.0539	5.80	trihydroxy palmitic acid	304.2253	0.70	HILIC neg
0.0128	6.00	2-Aminomuconate semialdehyde	141.0427	0.54	HILIC neg
0.0599	7.25	Isonicotinic acid	123.0322	0.58	HILIC neg
0.0135	8.20	2-Aminomuconate semialdehyde	141.0427	0.74	HILIC neg

0.0138	9.43	N,N-Dihydroxy-L-tyrosine	213.0638	1.36	HILIC neg
0.0090	10.29	3,4-Dihydroxymandelonitrile	165.0427	1.55	HILIC neg
0.0092	10.39	2,6-Dihydroxynicotinate	155.0220	0.99	HILIC neg
0.0090	10.60	2-Furoylglycine	169.0377	0.98	HILIC neg
0.0156	12.43	10,11-DiHDPE	362.2462	0.49	HILIC neg
0.0090	12.58	2-Aminobenzoic acid	137.0477	1.57	HILIC neg
0.0090	12.96	1-Pyrroline-5-carboxylate	113.0478	1.33	HILIC neg
0.0090	13.61	3-Hydroxykynurenine	224.0798	1.32	HILIC neg
0.0090	14.61	N-Acetyl-L-glutamate 5-semialdehyde	173.0689	1.33	HILIC neg
0.0090	16.30	2-Methylbutyrylglycine	159.0896	1.35	HILIC neg
0.0090	17.49	(S)-2-amino-6-oxohexanoate	145.0740	2.05	HILIC neg
0.0135	21.29	N-Formylmethionine	177.0460	2.29	HILIC neg
0.0090	22.94	2,5,6-Trihydroxy-5,6-dihydroquinoline	179.0583	0.75	HILIC neg
0.0090	43.92	Indoleacrylic acid	187.0634	0.54	HILIC neg
0.0090	44.37	2,6-Dihydroxynicotinate	155.0219	2.17	HILIC neg
0.0090	58.32	alhpa-tocopheronolactone	278.1510	0.68	HILIC neg
0.0090	88.96	2-Aminophenol	109.0529	0.52	HILIC neg
0.0130	0.00	1,4-Dideoxy-1,4-imino-D-ribitol	133.0737	3.99	HILIC pos
0.0187	0.02	2-Aminobut-2-enoate	101.0475	2.35	HILIC pos
0.0310	0.04	(S)-2-amino-6-oxohexanoate	145.0737	0.61	HILIC pos
0.0267	0.08	5-Hydroxyindoleacetylglycine	248.0794	4.15	HILIC pos
0.0173	0.09	amino-octadecanoic acid	299.2820	2.59	HILIC pos
0.0647	0.14	(-)-11-Hydroxy-9,15,16-trioxooctadecanoic acid	342.2035	0.52	HILIC pos
0.0947	0.14	1-pyrroline-3-hydroxy-5-carboxylic Acid	129.0424	5.93	HILIC pos
0.0562	0.15	Eicosanoyl-EA	355.3445	1.60	HILIC pos
0.0770	0.16	Vinylacetylglycine	143.0581	5.32	HILIC pos
0.0130	0.17	Oxohexenoic acid	128.0471	0.61	HILIC pos
0.0102	0.19	Formyl-N-acetyl-5-methoxykynurenamine	264.1105	0.61	HILIC pos
0.0667	0.21	Methylphenylalanine	179.0944	1.60	HILIC pos

0.0800	0.22	Trimethylamine N-oxide	75.0683	5.13	HILIC pos
0.0407	0.23	3'-CMP	323.0515	2.62	HILIC pos
0.0308	0.23	2-Aminoacrylic acid	87.0319	5.49	HILIC pos
0.0864	0.23	12-oxo-14,18-dihydroxy-9Z,13E,15Z-octadecatrienoic acid	324.1930	0.51	HILIC pos
0.0511	0.24	1H-Indole-2,3-dione	147.0318	0.43	HILIC pos
0.0287	0.28	Farnesyl diphosphate	382.1306	5.24	HILIC pos
0.0725	0.29	(3S,3'R,4xi)-beta,beta-Carotene-3,3',4-triol	584.4222	0.46	HILIC pos
0.0130	0.30	Streptamine phosphate	258.0611	5.36	HILIC pos
0.0469	0.31	Formiminoglutamic acid	174.0638	5.49	HILIC pos
0.0551	0.35	3-hydroxyundecanoyl carnitine	345.2509	0.58	HILIC pos
0.0287	0.44	4-Aminopyridine	94.0530	7.13	HILIC pos
0.0358	0.45	dodecenoic acid	198.1629	0.46	HILIC pos
0.0148	0.47	nonenoylglycine	213.1362	1.55	HILIC pos
0.0370	0.48	Oleoyl glycine	339.2768	0.50	HILIC pos
0.0287	0.49	6-Hydroxy-1H-indole-3-acetamide	190.0739	4.58	HILIC pos
0.0316	0.49	3,4-Dihydroxymandelonitrile	165.0424	0.55	HILIC pos
0.0615	0.50	PE(P-20:0/22:6)	803.5809	2.79	HILIC pos
0.0526	0.51	Hydroxyhexanoycarnitine	275.1729	5.43	HILIC pos
0.0797	0.51	(alpha-keto-dimethyl-delta-N,N-Dimethylguanidynol) valeric acid	201.1110	5.83	HILIC pos
0.0968	0.52	3-hydroxytridecanoyl carnitine	373.2821	0.50	HILIC pos
0.0756	0.55	LysoPE(20:1)	507.3322	4.13	HILIC pos
0.0589	0.57	Digallate	322.0311	5.28	HILIC pos
0.0358	0.61	2-Amino-4-methylpentanoic acid	131.0944	5.38	HILIC pos
0.0779	0.61	Norepinephrine sulfate	249.0304	1.40	HILIC pos
0.0860	0.62	3-Acetamidopropanal	115.0632	0.76	HILIC pos
0.0947	0.62	N-Decanoylglycine	229.1674	0.49	HILIC pos
0.0899	0.63	3-hydroxydecanoyl carnitine	331.2353	4.34	HILIC pos
0.0965	0.64	3-hydroxy-5-octenoylcarnitine	301.1885	5.53	HILIC pos
0.0858	0.66	Cytidine	243.0852	5.08	HILIC pos
0.0973	0.73	Isovalerylcarnitine	245.1623	4.54	HILIC pos
0.0968	0.75	(2E)-4-hydroxy-3-methylbut-2-enoic acid	116.0472	1.13	HILIC pos

0.0800	1.33	N-Nitrosodiethylamine	102.0792	6.03	HILIC pos
0.0894	1.34	Butenoate	86.0366	0.99	HILIC pos
0.0245	1.47	2,3-Diaminopropanoate	104.0585	1.82	HILIC pos
0.0993	1.50	N-(2-Methylpropyl)acetamide	115.0995	3.19	HILIC pos
0.0622	1.62	PE(20:0)	523.3268	4.11	HILIC pos
0.0666	1.79	midazole-4-acetaldehyde	110.0479	0.77	HILIC pos
0.0488	1.84	amino-octanoic acid	159.1257	0.59	HILIC pos
0.0929	1.85	Penicillenic acid	244.0514	1.74	HILIC pos
0.0271	1.89	3-Methylglutamyl-5-semialdehyde-N6-lysine	273.1684	2.12	HILIC pos
0.0404	1.90	11,12-DiHETrE-EA	381.2871	0.50	HILIC pos
0.0492	1.96	2-Amino-4-hydroxy-3-methylpentanoic acid	147.0893	6.50	HILIC pos
0.0889	1.97	Hydroxynicotine	178.1101	0.58	HILIC pos
0.0848	2.05	DG(42:8)	692.5369	0.46	HILIC pos
0.0534	2.11	3-Pyridinebutanoic acid	165.0788	0.52	HILIC pos
0.0662	2.16	N1-Acetylspermidine	187.1682	7.62	HILIC pos
0.0770	2.23	1,4-cholestadie24,25-didehydrovitamin D3 / 24,25-didehydrocholecalciferol	382.3231	0.47	HILIC pos
0.0666	2.26	LysoPC(18:1)	521.3475	4.11	HILIC pos
0.0204	2.37	6-Hydroxypseudooxynicotine	194.1053	0.72	HILIC pos
0.0308	2.59	2-Oxoarginine	173.0798	3.33	HILIC pos
0.0405	3.01	Methylphenylalanine	179.0943	3.50	HILIC pos
0.0316	3.07	2,3-Butanediol glucoside	252.1219	3.03	HILIC pos
0.0249	3.16	4-Guanidinobutanoate	145.0849	6.07	HILIC pos
0.0287	3.19	Pyrazinoic acid	124.0272	1.93	HILIC pos
0.0358	3.19	(E)-4-(Trimethylammonio)but-2-enoate	143.0943	6.99	HILIC pos
0.0332	4.01	N-Nitrosoproline	144.0533	0.68	HILIC pos
0.0331	4.18	Imidazoleacetic acid	126.0427	0.68	HILIC pos
0.0130	4.40	4-Guanidinobutanal	129.0900	3.80	HILIC pos
0.0287	4.69	4-(beta-Acetylaminoethyl)imidazole	153.0897	0.59	HILIC pos
0.0386	4.72	2-Aminoquinoline	144.0686	0.56	HILIC pos
0.0287	4.85	Isoleucyl-Valine	230.1628	7.40	HILIC pos
0.0666	5.18	2,3,4,5-tetranor-thromboxane B1	316.1882	3.38	HILIC pos
0.0132	5.61	1-Methylhistidine	169.0853	1.06	HILIC pos

0.0800	5.83	PA(O-18:0/19:1)	702.5574	0.47	HILIC pos
0.0285	5.95	Benzeneacetonitrile	117.0577	3.37	HILIC pos
0.0256	6.10	2-[(4-methoxyphenyl)formamido]acetic acid	209.0685	0.71	HILIC pos
0.0213	6.60	Histidine	155.0693	3.23	HILIC pos
0.0130	6.62	Alanyl-Proline	186.1000	5.88	HILIC pos
0.0308	6.92	Dopamine	151.0631	0.65	HILIC pos
0.0249	7.46	decatrienoylcarnitine	309.1935	3.72	HILIC pos
0.0506	7.66	N-Formyl-4-amino-5-aminomethyl-2-methylpyrimidine	166.0852	3.39	HILIC pos
0.0132	7.99	1-Methylhistidine	169.0852	1.42	HILIC pos
0.0102	8.20	1,1,3,3-tetramethylurea	116.0948	5.91	HILIC pos
0.0297	8.28	6-Methylquinoline	143.0733	2.71	HILIC pos
0.0287	8.92	Isonicotinylglycine	180.0532	1.13	HILIC pos
0.0177	9.00	PGH2-EA	393.2872	3.18	HILIC pos
0.0189	9.24	4,5-Dihydro-2,4(5)-dimethyl-1H-imidazole	98.0843	3.43	HILIC pos
0.0309	9.51	2,3-Diaminosalicylic acid	168.0533	0.69	HILIC pos
0.0356	9.96	6-Imino-5-oxocyclohexa-1,3-dienecarboxylate	151.0267	0.70	HILIC pos
0.0342	11.40	Glc-Cholesterol(22:0)	870.7290	0.45	HILIC pos
0.0189	11.52	beta-Tyrosine	181.0737	0.77	HILIC pos
0.0135	12.07	3-Methylene-indolenine	129.0577	2.93	HILIC pos
0.0190	12.23	Alanyl-Asparagine	203.0904	2.98	HILIC pos
0.0287	12.35	Mandelonitrile	133.0526	0.58	HILIC pos
0.0102	12.56	N-Methyl-N,4-dinitrosoaniline	165.0536	5.45	HILIC pos
0.0190	12.82	1H-Indole-3-acetamide	174.0791	0.58	HILIC pos
0.0130	14.25	2-O-Methylcytosine	125.0587	4.59	HILIC pos
0.0225	14.48	Pyroglutamylglycine	186.0638	2.99	HILIC pos
0.0532	18.00	Cyclo(L-tyrosyl-L-tyrosyl)	326.1260	0.50	HILIC pos
0.0187	18.84	Cyclo(L-Phe-L-Pro)	244.1209	4.22	HILIC pos
0.0509	19.16	CE(20:5)	670.5683	0.46	HILIC pos
0.0249	23.96	Tyrosyl-Tyrosine	344.1366	0.94	HILIC pos
0.0387	35.40	N-Stearoyl tyrosine	447.3342	3.00	HILIC pos
0.0130	38.67	Cervonyl carnitine	471.3343	2.98	HILIC pos
0.0130	39.37	Vinylacetylglycine	143.0581	1.64	HILIC pos
0.0130	39.39	Methyltryptophan	218.1052	0.62	HILIC pos
0.0130	45.88	Valyl-Valine	216.1471	7.38	HILIC pos
0.0130	65.13	3-hydroxyarachidonoylcarnitine	463.3291	3.42	HILIC pos
0.0130	78.78	N1,N8-Diacetylspermidine	229.1787	7.57	HILIC pos

0.0130	91.38	N-Acetylhistidine	197.0797	5.44	HILIC pos
0.0130	114.46	Phenylethylmalonamide	206.1053	3.39	HILIC pos
0.0925	153.00	3-Methyloxindole	147.0682	3.53	HILIC pos
0.0130	200.69	4-(beta-Acetylaminoethyl)imidazole	153.0900	1.07	HILIC pos
0.0022	0.00	PC(33:4)	739.5156	7.49	Lipids neg
0.0022	0.00	PC(35:3)	769.5627	8.02	Lipids neg
0.0020	0.00	PC(35:5)	765.5309	7.51	Lipids neg
0.0021	0.00	PE(18:3/P-18:1)	723.5207	7.69	Lipids neg
0.0022	0.00	PE(20:3/P-18:1)	751.5519	8.03	Lipids neg
0.0020	0.00	PS(39:3)	827.5679	7.42	Lipids neg
0.0022	0.00	GalNAcbeta1-4Galbeta1-4Glcbeta-Cer(d18:1/18:0)	1092.7322	4.44	Lipids neg
0.0022	0.00	LysoPC(15:0)	481.3168	3.24	Lipids neg
0.0020	0.00	LysoPC(17:0)	509.3480	4.44	Lipids neg
0.0059	0.00	LysoPE(20:2)	505.3168	2.56	Lipids neg
0.0041	0.00	LysoPS(21:0)	567.3539	3.41	Lipids neg
0.0484	0.00	PC(18:4/P-18:1)	763.5502	7.85	Lipids neg
0.0035	0.00	PC(33:1)	745.5620	7.82	Lipids neg
0.0020	0.00	PC(35:4)	767.5467	7.42	Lipids neg
0.0039	0.00	PC(37:5)	793.5619	7.42	Lipids neg
0.0022	0.00	PC(37:6)	791.5464	7.24	Lipids neg
0.0020	0.00	PC(39:6)	819.5781	7.65	Lipids neg
0.0022	0.00	PE(18:2/P-18:1)	725.5368	7.74	Lipids neg
0.0037	0.00	PE(20:4/P-18:1)	749.5362	7.70	Lipids neg
0.0032	0.00	PE(22:4/P-18:1)	777.5679	8.01	Lipids neg
0.0096	0.00	PE(40:9)	785.5002	6.70	Lipids neg
0.0022	0.00	PI(36:4)	858.5264	7.13	Lipids neg
0.0036	0.00	PS(37:0)	805.5833	7.82	Lipids neg
0.0022	0.00	PS(38:4)	811.5361	7.55	Lipids neg
0.0079	0.00	PS(39:2)	829.5831	7.53	Lipids neg
0.0021	0.00	PS(40:6)	835.5350	7.42	Lipids neg
0.0022	0.00	PS(P-18:0/22:6)	819.5398	8.03	Lipids neg
0.0020	0.00	LysoPC(P-15:0)	465.3222	4.96	Lipids neg
0.0034	0.00	LysoPE(20:4)	501.2861	2.35	Lipids neg
0.0022	0.00	PC(33:3)	741.5316	7.60	Lipids neg
0.0034	0.00	PC(37:4)	795.5781	8.01	Lipids neg
0.0020	0.00	PC(O-15:0/20:4)	753.5680	8.07	Lipids neg
0.0040	0.00	PE(20:5/P-18:1)	747.5209	7.54	Lipids neg
0.0022	0.00	PE(22:4/P-18:0)	779.5839	8.24	Lipids neg
0.0028	0.00	PE(22:5/P-18:1)	775.5518	7.89	Lipids neg
0.0020	0.00	PS(35:0)	777.5525	7.47	Lipids neg
0.0020	0.00	PS(38:1)	817.5837	7.70	Lipids neg
0.0039	0.00	PS(41:4)	853.5831	7.41	Lipids neg
0.0033	0.00	PS(43:4)	881.6144	7.77	Lipids neg
0.0033	0.00	LysoPE(18:1)	479.3016	3.58	Lipids neg
0.0022	0.00	PC(33:2)	715.5158	7.58	Lipids neg
0.0018	0.00	PC(P-16:0/15:1)	701.5369	8.07	Lipids neg

0.0022	0.00	PE(P-16:0)	437.2910	3.82	Lipids neg
0.0041	0.00	PI(36:2)	862.5572	7.58	Lipids neg
0.0030	0.00	PI(38:5)	884.5420	7.16	Lipids neg
0.0020	0.00	PS(36:2)	787.5369	7.63	Lipids neg
0.0102	0.00	PS(39:0)	833.6156	8.14	Lipids neg
0.0022	0.00	PS(40:3)	841.5836	7.61	Lipids neg
0.0020	0.00	PS(O-18:0/19:1)	789.5888	7.76	Lipids neg
0.0045	0.00	(8E,10S,12Z,15Z)-10-Hydroperoxyoctadeca-8,12,15-trienoate	310.2147	1.45	Lipids neg
0.0048	0.00	LysoPE(20:2)	505.3168	2.35	Lipids neg
0.0020	0.00	PE(16:1/P-18:1)	699.5205	7.79	Lipids neg
0.0064	0.00	PE(38:7)	761.5005	6.82	Lipids neg
0.0020	0.00	PS(38:6)	807.5028	7.48	Lipids neg
0.0075	0.00	PS(41:3)	855.5988	7.56	Lipids neg
0.0073	0.01	(8E,10S,12Z,15Z)-10-Hydroperoxyoctadeca-8,12,15-trienoate	310.2147	1.52	Lipids neg
0.0020	0.01	12-Oxo-c-LTB3	641.2997	4.44	Lipids neg
0.0051	0.01	9,10-epoxy-3E,12Z-Octadecadienoic acid	294.2198	2.12	Lipids neg
0.0019	0.01	LysoPC(16:0)	495.3328	3.87	Lipids neg
0.0020	0.01	N-palmitoyl-phosphoethanolamine	379.2491	2.29	Lipids neg
0.0022	0.01	PS(40:7)	833.5180	7.55	Lipids neg
0.0025	0.01	PS(O-16:0/20:5)	767.5110	7.35	Lipids neg
0.0085	0.01	9,10-epoxy-3E,12Z-Octadecadienoic acid	294.2198	2.25	Lipids neg
0.0020	0.01	PC(35:2)	771.5786	8.20	Lipids neg
0.0029	0.01	PS(O-18:0/20:5)	795.5419	7.72	Lipids neg
0.0076	0.01	9-oxo-12Z-Octadecenoic acid	296.2352	1.72	Lipids neg
0.0129	0.01	Eicosatetraenoic acid	304.2402	4.40	Lipids neg
0.0022	0.01	PC(P-16:0/17:2)	727.5528	8.10	Lipids neg
0.0020	0.01	PI(40:5)	912.5729	7.49	Lipids neg
0.0025	0.01	PI(40:6)	910.5583	7.37	Lipids neg
0.0027	0.01	PC(16:0/5:0(CHO))	593.3700	3.69	Lipids neg
0.0034	0.01	PC(35:6)	763.5158	7.35	Lipids neg
0.0023	0.01	PI(38:3)	888.5728	7.66	Lipids neg
0.0025	0.01	PS(35:1)	775.5371	7.11	Lipids neg
0.0020	0.01	PS(36:0)	791.5698	7.65	Lipids neg
0.0024	0.01	PS(36:3)	785.5212	7.29	Lipids neg
0.0207	0.01	PS(42:10)	855.5058	6.54	Lipids neg
0.0062	0.01	(9Z,11E)-12-((3S,5R)-5-((R)-1-hydroxypropyl)-1,2-dioxolan-3-yl)dodeca-9,11-dienoic acid	326.2097	1.03	Lipids neg

0.0057	0.01	Hydroperoxyoctadecadienoic acid	312.2304	2.12	Lipids neg
0.0020	0.01	LysoPC(15:0)	481.3172	4.56	Lipids neg
0.0035	0.01	PE(18:0/20:4(Ke))	781.5259	7.35	Lipids neg
0.0024	0.01	PI(37:4)	872.5421	7.32	Lipids neg
0.0024	0.01	PS(O-18:0/22:6)	821.5553	8.07	Lipids neg
0.0078	0.01	(+/-)-11-HEPE	318.2196	2.13	Lipids neg
0.0020	0.01	LysoPC(O-15:0)	467.3380	3.74	Lipids neg
0.0035	0.01	PC(32:4)	725.5003	7.28	Lipids neg
0.0024	0.01	(13E)-(15S)-15-Hydroxy-9-oxoprosta-10,13-dienoate	336.2304	1.50	Lipids neg
0.0123	0.01	PE(40:8)	787.5160	7.04	Lipids neg
0.0028	0.01	PI(O-16:0/22:4)	872.5795	7.71	Lipids neg
0.0020	0.01	PS(36:4)	783.5032	7.58	Lipids neg
0.0022	0.01	PS(41:0)	859.6311	8.18	Lipids neg
0.0070	0.01	LysoPE(16:0)	453.2861	3.12	Lipids neg
0.0066	0.01	PC(18:4/P-16:0)	737.5362	7.86	Lipids neg
0.0036	0.01	PI(40:4)	914.5908	7.74	Lipids neg
0.0015	0.01	PS(36:1)	789.5527	7.90	Lipids neg
0.0022	0.01	PS(P-18:0/17:2)	757.5264	6.24	Lipids neg
0.0055	0.01	10-Hydroperoxy-H4-neuroprostane	392.2207	0.81	Lipids neg
0.0035	0.01	LysoPE(18:1)	479.3015	3.36	Lipids neg
0.0389	0.01	LysoPE(24:6)	553.3171	2.25	Lipids neg
0.0023	0.01	PS(25:0)	637.3961	4.37	Lipids neg
0.0089	0.01	PS(39:4)	825.5531	6.85	Lipids neg
0.0020	0.01	Trypanothione	723.3034	3.25	Lipids neg
0.0031	0.02	LysoPS(18:1)	523.2914	2.93	Lipids neg
0.0054	0.02	PE(18:4/P-18:1)	721.5051	7.40	Lipids neg
0.0058	0.02	PI(O-18:0/20:5)	870.5631	7.37	Lipids neg
0.0024	0.02	PS(40:4)	839.5681	7.79	Lipids neg
0.0076	0.02	PS(43:6)	877.5835	7.27	Lipids neg
0.0020	0.02	PS(O-20:0/19:1)	817.6206	8.09	Lipids neg
0.0015	0.02	LysoPC(P-15:0)	465.3222	4.06	Lipids neg
0.0074	0.02	UDP-GlcNAc	607.0826	0.53	Lipids neg
0.0079	0.02	Ascorbyl palmitate	414.2623	1.62	Lipids neg
0.0020	0.02	N-palmitoyl threonine	357.2884	5.14	Lipids neg
0.0022	0.02	PC(31:1)	717.5309	7.87	Lipids neg
0.0067	0.02	(+/-)-11-HEPE	318.2196	2.83	Lipids neg
0.0022	0.02	LysoPE(16:0)	453.2860	3.36	Lipids neg
0.0020	0.02	Sphinganine 1-phosphate	381.2648	2.64	Lipids neg
0.0076	0.02	(13E)-(15S)-15-Hydroxy-9-oxoprosta-10,13-dienoate	336.2303	2.03	Lipids neg
0.0020	0.02	CDP-DG(38:6)	1025.5127	7.42	Lipids neg
0.0032	0.02	PC(16:0/5:0(COOH))	609.3630	3.27	Lipids neg

0.0044	0.02	PC(34:4)	753.5325	7.67	Lipids neg
0.0020	0.02	PI(36:1)	864.5770	7.85	Lipids neg
0.0041	0.02	(13E)-(15S)-15-Hydroxy-9-oxoprosta-10,13-dienoate	336.2304	1.72	Lipids neg
0.0034	0.02	12-oxo-14,18-dihydroxy-9Z,13E,15Z-octadecatrienoic acid	324.1940	0.84	Lipids neg
0.0034	0.02	LysoPE(18:2)	477.2858	2.45	Lipids neg
0.0025	0.02	Octadecadienoic acid	280.2401	4.61	Lipids neg
0.0111	0.02	PE(O-16:0)	439.3066	3.88	Lipids neg
0.0110	0.02	PG(36:1)	776.5553	7.69	Lipids neg
0.0048	0.02	9-oxo-12Z-Octadecenoic acid	296.2352	1.86	Lipids neg
0.0039	0.02	PI(38:6)	882.5268	6.98	Lipids neg
0.0026	0.02	PS(34:2)	759.5055	7.26	Lipids neg
0.0031	0.02	(+/-)11,12-EpETrE	320.2352	2.11	Lipids neg
0.0122	0.02	11,18-di-HEPE	334.2148	2.25	Lipids neg
0.0026	0.02	12-Oxo-c-LTB3	641.3001	4.21	Lipids neg
0.0039	0.02	LysoPS(18:2)	521.2757	2.07	Lipids neg
0.0020	0.02	PS(40:1)	845.6149	8.04	Lipids neg
0.0028	0.03	(+/-)11,12-DiHETrE	338.2458	2.11	Lipids neg
0.0022	0.03	1,25-Dihydroxyvitamin D3 3-glycoside	578.3825	5.75	Lipids neg
0.0088	0.03	15-methyl-15R-PGF2alpha	368.2569	1.62	Lipids neg
0.0048	0.03	Hydroxy-oxooctadecatrienoic acid	308.1991	1.18	Lipids neg
0.0103	0.03	(13E)-11a-Hydroxy-9,15-dioxoprost-13-enoic acid	352.2255	1.57	Lipids neg
0.0027	0.03	PC(16:0/9:0(COOH))	665.4277	3.90	Lipids neg
0.0047	0.03	Propane-1,2-diol 1-phosphate	156.0190	0.58	Lipids neg
0.0022	0.03	PS(44:7)	889.5811	7.77	Lipids neg
0.0129	0.03	(8E,10S,12Z,15Z)-10-Hydroperoxyoctadeca-8,12,15-trienoate	310.2147	1.92	Lipids neg
0.0392	0.03	D-erythro-L-galacto-Nonulose	270.0951	0.59	Lipids neg
0.0199	0.03	1-(8-[5]-ladderane-octanoyl)-2-(8-[3]-ladderane-octanyl)-sn-glycerophosphoethanolamine	771.5209	7.23	Lipids neg
0.0095	0.03	hydroxy-dodecenoic acid	214.1570	1.22	Lipids neg

0.0068	0.04	Cholesterol sulfate	466.3119	5.50	Lipids neg
0.0368	0.04	12,15-Epoxy-13,14-dimethyleicosa-10,12,14-trienoic acid	348.2665	3.03	Lipids neg
0.0060	0.04	PS(37:3)	799.5375	7.00	Lipids neg
0.0335	0.04	PS(42:8)	859.5371	7.04	Lipids neg
0.0020	0.04	PS(24:0)	623.3804	4.37	Lipids neg
0.0221	0.04	Octadecenoic acid	282.2559	5.36	Lipids neg
0.0020	0.04	LysoPI(18:0)	600.3281	3.93	Lipids neg
0.0123	0.04	Hydroxyprogesterone	332.1991	1.47	Lipids neg
0.0038	0.04	(Z)-15-Oxo-11-eicosenoic acid	324.2666	2.93	Lipids neg
0.0317	0.04	11-Dehydro-thromboxane B2	368.2203	0.82	Lipids neg
0.0059	0.05	Arginine	174.1115	0.55	Lipids neg
0.0165	0.05	(13E)-(15S)-15-Hydroxy-9-oxoprosta-10,13-dienoate	336.2304	2.18	Lipids neg
0.0025	0.05	(20S)-1alpha,20,25-trihydroxyvitamin D3 / (20S)-1alpha,20,25-trihydroxycholecalciferol	432.3243	1.84	Lipids neg
0.0062	0.05	10,13-epoxy-11,12-dimethyl-10,12-octadecadienoic acid	322.2509	2.63	Lipids neg
0.0156	0.05	S-Carboxymethyl-L-cysteine	179.0251	0.57	Lipids neg
0.0030	0.05	Glycerol 3-phosphate	172.0137	0.62	Lipids neg
0.0049	0.05	PI(42:11)	928.5140	8.24	Lipids neg
0.0020	0.05	Epoxy-hexadecenoic acid	268.2040	2.21	Lipids neg
0.0022	0.05	LysoPS(22:0)	581.3700	4.01	Lipids neg
0.0035	0.05	PS(29:0)	693.4589	4.83	Lipids neg
0.0059	0.05	2-C-Methyl-D-erythritol 4-phosphate	216.0401	0.58	Lipids neg
0.0034	0.05	PC(19:0)	551.3591	2.98	Lipids neg
0.0042	0.05	9-oxo-12Z-Octadecenoic acid	296.2352	3.17	Lipids neg
0.0210	0.05	11,18-di-HEPE	334.2148	1.44	Lipids neg
0.0021	0.05	LysoPS(18:0)	525.3070	3.96	Lipids neg
0.0022	0.06	9-oxo-12Z-Octadecenoic acid	296.2352	2.74	Lipids neg
0.0063	0.06	D-erythro-D-galacto-octitol	242.0991	0.55	Lipids neg
0.0020	0.06	hydroxy-oxo-hexadecanoic acid	286.2147	2.37	Lipids neg

0.0037	0.06	(22E)-(24R)-1alpha,24,25-trihydroxy-22,23-didehydrovitamin D3 / (22E)-(24R)-1alpha,24,25-trihydroxy-22,23-didehydrocholecalciferol	430.3086	1.97	Lipids neg
0.0495	0.06	PC(37:7)	789.5303	7.25	Lipids neg
0.0042	0.06	N-Adenylyl-L-phenylalanine	494.1318	4.61	Lipids neg
0.0028	0.07	1D-chiro-Inositol	180.0633	0.57	Lipids neg
0.0020	0.07	MG(16:0)	330.2772	5.14	Lipids neg
0.0300	0.07	Docosapentaenoic acid	330.2562	4.79	Lipids neg
0.0195	0.07	Phenylacetylglycine dimethylamide	220.1213	0.92	Lipids neg
0.0036	0.08	PC(15:0/P-18:1)	729.5689	8.35	Lipids neg
0.0155	0.08	Hydroperoxydocosahexaenoic acid	360.2305	1.40	Lipids neg
0.0036	0.08	Hydroxytetracosanoic acid	384.3606	7.17	Lipids neg
0.0140	0.08	PE(18:0/20:4(OH))	783.5421	6.73	Lipids neg
0.0339	0.08	PS(O-16:0/19:0)	763.5736	8.02	Lipids neg
0.0673	0.09	20-Trihydroxy-leukotriene-B4	384.2152	0.76	Lipids neg
0.0020	0.09	LysoPC(P-17:0)	493.3538	5.85	Lipids neg
0.0066	0.09	LysoPI(16:0)	572.2966	2.66	Lipids neg
0.0216	0.10	Octadecatrienoic acid	278.2246	3.89	Lipids neg
0.0203	0.10	1,5-Anhydro-D-glucitol	164.0684	0.57	Lipids neg
0.0029	0.10	hydroxy-pentacosanoic acid	398.3763	6.24	Lipids neg
0.0310	0.10	methoxy-hydroxy-octadecadienoic acid	326.2460	4.43	Lipids neg
0.0600	0.10	11-dehydro-TXB3	366.2046	0.82	Lipids neg
0.0217	0.11	(20R)-17alpha,20-dihydroxycholesterol	418.3450	7.23	Lipids neg
0.0285	0.11	7-Methyl-1,4,5-naphthalenetriol 4-[xylosyl-(1->6)-glucoside]	484.1597	0.78	Lipids neg
0.0400	0.11	LysoPS(22:6)	569.2759	1.83	Lipids neg
0.0030	0.11	9-oxo-12Z-Octadecenoic acid	296.2352	2.24	Lipids neg
0.0311	0.12	3-(2,4-dihydroxy-3,5-dimethoxyphenyl)propanoic acid	242.0799	0.78	Lipids neg

0.0022	0.13	2-O-(alpha-D-Glucopyranosyl)-D-glycerate	268.0793	0.57	Lipids neg
0.0254	0.13	(+/-)-11-HEPE	318.2196	1.82	Lipids neg
0.0028	0.13	LPA(16:0)	410.2436	3.20	Lipids neg
0.0020	0.13	(6R,7S)-6,7-Epoxyoctadecanoic acid	298.2509	4.00	Lipids neg
0.0605	0.14	LysoPI(20:4)	620.2967	1.96	Lipids neg
0.0022	0.14	LysoPC(O-17:0)	495.3694	5.92	Lipids neg
0.0022	0.15	hydroxy-hexadecenoic acid	270.2196	2.84	Lipids neg
0.0041	0.15	Glycerophosphoglycerol	246.0505	0.57	Lipids neg
0.0076	0.15	hydroxy-decanoic acid	188.1413	1.03	Lipids neg
0.0039	0.15	LPA(18:2)	434.2435	2.53	Lipids neg
0.0031	0.15	Manno-2-heptulose	210.0738	0.57	Lipids neg
0.0285	0.18	Hexacosanedioic acid	426.3715	6.29	Lipids neg
0.0033	0.19	1alpha,25-dihydroxy-23-azavitamin D3 / 1alpha,25-dihydroxy-23-azacholecalciferol	417.3244	4.42	Lipids neg
0.0088	0.19	Tetranorbiotin	188.0264	0.58	Lipids neg
0.0386	0.20	11-dehydro-2,3-dinor-TXB2	340.1890	0.74	Lipids neg
0.0051	0.21	MG(18:4)	350.2462	0.89	Lipids neg
0.0074	0.21	Tocopheronic acid	294.1472	0.78	Lipids neg
0.0867	0.21	methyl-heptadecanoic acid	284.2715	6.08	Lipids neg
0.0032	0.21	Dityrosine	360.1324	0.64	Lipids neg
0.0511	0.21	dimethyl-tetradecanoic acid	256.2403	5.25	Lipids neg
0.0512	0.22	LysoPA(20:4)	458.2438	2.27	Lipids neg
0.0284	0.23	(25S)-5alpha-cholestan-3beta,6beta,15alpha,16beta,26-pentol	452.3482	5.91	Lipids neg
0.0040	0.24	10-keto myristic acid	242.1883	2.06	Lipids neg
0.0099	0.25	3-(Pyrazol-1-yl)-L-alanine	155.0694	0.56	Lipids neg
0.0271	0.27	SQMG(16:1)	554.2734	5.91	Lipids neg
0.0066	0.29	Glycyl-Phenylalanine	222.1004	0.61	Lipids neg
0.0222	0.29	MG(22:1)	412.3557	5.97	Lipids neg
0.0335	0.32	11,18-di-HEPE	334.2148	1.58	Lipids neg
0.0419	0.33	10,13-diepi-10-F4c-NeuroP	378.2411	1.22	Lipids neg
0.0140	0.34	Thymidine	242.0902	0.59	Lipids neg
0.0264	0.34	2-hydroxy-tricosanoic acid	370.3450	6.90	Lipids neg

0.0421	0.35	5a-Dihydrotestosterone sulfate	370.1815	5.68	Lipids neg
0.0390	0.37	hydroxy-hexadecanoic acid	272.2354	2.00	Lipids neg
0.0097	0.40	11beta,12beta-Dihydroxy-5beta-cholan-24-oic Acid	392.2931	1.63	Lipids neg
0.0020	0.40	(5Z,9E,12S,14Z)-8,11,12-Trihydroxyicosa-5,9,14-trienoate	354.2409	1.33	Lipids neg
0.0031	0.44	eicosapentaenoic acid	302.2247	2.52	Lipids neg
0.0020	0.45	LysoPA(22:6)	482.2438	2.22	Lipids neg
0.0066	0.46	LysoPA(18:1)	436.2592	3.41	Lipids neg
0.0573	0.46	N-palmitoyl tyrosine	419.3040	3.84	Lipids neg
0.0021	0.51	N-linoleoyl valine	379.3090	4.17	Lipids neg
0.0331	0.51	hydroxyoctadecanoic acid	300.2665	3.17	Lipids neg
0.0285	0.51	N-palmitoyl valine	355.3090	4.77	Lipids neg
0.0015	0.51	N-oleoyl phenylalanine	429.3246	5.00	Lipids neg
0.0022	0.52	Coenzyme Q10	862.6804	5.69	Lipids neg
0.0321	0.55	Deoxycytidine	227.0906	0.57	Lipids neg
0.0528	0.57	methyl-tridecanedioic acid	258.1834	1.11	Lipids neg
0.0084	0.61	Tryptophan	204.0898	0.61	Lipids neg
0.0699	0.65	Ethyl glucuronide	222.0739	0.57	Lipids neg
0.0339	1.51	hydroxy-heneicosanoic acid	342.3136	6.29	Lipids neg
0.0089	1.52	N-lactoyl-Phenylalanine	237.1001	0.64	Lipids neg
0.0329	1.99	Dihydroxydocosanoic acid	372.3243	3.78	Lipids neg
0.0696	2.03	D-Altritol	182.0787	0.57	Lipids neg
0.0531	2.76	Vanillyl alcohol	154.0629	0.64	Lipids neg
0.0164	2.79	D-Apiitol	152.0683	0.57	Lipids neg
0.0170	2.93	(10R,13R,16R)-delta14-9-PhytoF[9R,12R]	344.2201	0.67	Lipids neg
0.0197	3.23	N-Acetyl-leucyl-leucine	286.1897	0.72	Lipids neg
0.0226	3.31	1D-5-O-methyl-chiro-inositol	194.0789	0.57	Lipids neg
0.0388	3.37	3-carboxy-4-methyl-5-pentyl-2-furanpropanoic acid	268.1311	0.66	Lipids neg
0.0183	3.50	hydroxy-decenoic acid	186.1257	0.83	Lipids neg
0.0088	3.59	11beta,12beta-Dihydroxy-5beta-cholan-24-oic Acid	392.2937	1.05	Lipids neg
0.0080	3.59	N-D-Glucosylarylamine	255.1106	0.64	Lipids neg

0.0135	3.76	Quinic acid	192.0631	0.59	Lipids neg
0.0060	4.17	5,6-Dihydroxy-3-methyl-2-oxo-1,2-dihydroquinoline	191.0580	0.65	Lipids neg
0.0247	4.51	(10R,13R,16R)-delta11-14-NeuroF[14R,17R]	394.2365	0.91	Lipids neg
0.0099	4.59	Hexanoylcarnitine	259.1784	0.68	Lipids neg
0.0125	4.73	Homoveratric acid	196.0733	0.66	Lipids neg
0.0082	5.11	Pseudouridine	244.0694	0.58	Lipids neg
0.0573	5.74	S-Acetyldihydrolipoamide	249.0848	0.57	Lipids neg
0.0045	5.81	Dihydroxyheptenedioate	190.0476	0.59	Lipids neg
0.0041	5.95	oxo-decynoic acid	182.0944	0.65	Lipids neg
0.0167	6.07	7-Methylinosine	283.1055	0.65	Lipids neg
0.0064	6.50	(2S)-2-amino-3-(4-hydroxy-3-methoxyphenyl)-2-methylpropanoic acid	225.1000	0.65	Lipids neg
0.0296	6.68	2, 3, 4-Trimethoxycinnamate	238.0842	0.64	Lipids neg
0.0032	6.96	3-Methyldioxyindole	163.0632	0.65	Lipids neg
0.0049	7.25	3-Phenylpropionylglycine	207.0894	0.65	Lipids neg
0.0184	7.28	1-Methylseleno-N-acetyl-D-galactosamine	299.0287	0.65	Lipids neg
0.0026	7.74	Mandelate	152.0473	0.66	Lipids neg
0.0022	7.94	11-Deoxycortisol	346.2150	1.04	Lipids neg
0.0099	7.99	Glycocholic Acid	465.3096	0.79	Lipids neg
0.0044	8.03	Dodecenedioic acid	228.1362	0.79	Lipids neg
0.0133	8.13	Dihydrolipoate	208.0582	0.58	Lipids neg
0.0055	8.29	Isopropylmaleate	158.0577	0.64	Lipids neg
0.0040	8.62	Dimethyl-decadiendioic acid	226.1207	0.74	Lipids neg
0.0071	9.22	(22R)-22,25-dihydroxyvitamin D3 / (22R)-22,25-dihydroxycholecalciferol	408.2881	1.07	Lipids neg
0.0028	9.63	4-O-(beta-L-Arabinofuranosyl)-(2S,4S)-4-hydroxyproline	263.1005	0.59	Lipids neg
0.0022	9.85	Dodecenedioic acid	228.1361	0.66	Lipids neg
0.0039	9.88	Aldehydo-N-acetyl-D-glucosamine	221.0899	0.57	Lipids neg
0.0441	9.92	decenoic acid	170.1306	0.76	Lipids neg

0.0050	10.15	11beta,12beta-Dihydroxy-5beta-cholan-24-oic Acid	392.2932	1.78	Lipids neg
0.0028	10.52	3-Hydroxydodecanedioic acid	246.1466	0.65	Lipids neg
0.0061	10.70	2-Phenylethanol glucuronide	298.1051	0.64	Lipids neg
0.0024	10.83	TG(9:0)	260.1258	0.63	Lipids neg
0.0036	10.95	Phaseolic acid	262.1415	0.64	Lipids neg
0.0062	11.39	2-Carboxy-D-arabinitol	196.0581	0.57	Lipids neg
0.0120	11.49	N-(1-Deoxy-1-fructosyl)serine	267.0956	0.57	Lipids neg
0.0022	11.84	3-Indolehydracrylic acid	205.0736	0.66	Lipids neg
0.0022	12.16	4'-phosphonatopantothenate	299.0772	0.59	Lipids neg
0.0075	12.19	Vanilloloside	316.1159	0.64	Lipids neg
0.0698	12.21	15-methyl-15R-PGD2	366.2410	0.76	Lipids neg
0.0021	12.29	2-Methylquinolin-4-ol	159.0685	0.69	Lipids neg
0.0021	12.65	3-ethylphenyl Sulfate	202.0299	0.66	Lipids neg
0.0220	13.81	3-Imidazole-2-oxopropanoate	154.0385	0.58	Lipids neg
0.0036	14.19	Dimethylglutaric acid	160.0734	0.65	Lipids neg
0.0034	14.73	hydroxy-decenoic acid	186.1256	0.64	Lipids neg
0.0040	16.69	oxo-decenoic acid	184.1099	0.66	Lipids neg
0.0126	17.44	Glutamylproline	244.1059	0.58	Lipids neg
0.0025	17.78	2,3-Dihydroxybenzoate	154.0265	0.63	Lipids neg
0.0015	18.16	Indolylacryloylglycine	244.0846	0.65	Lipids neg
0.0015	20.56	cis-Cyclo(aspartylphenylalanyl)	262.0952	0.62	Lipids neg
0.0020	21.61	N-Acetyl-D-galactosamine 6-sulfate	301.0467	0.56	Lipids neg
0.0028	21.62	hydroxy-Octadienoic acid	156.0785	0.64	Lipids neg
0.0026	22.60	Glycylproline	172.0847	0.58	Lipids neg
0.0208	22.93	LysoPC(10:0)	412.2467	0.77	Lipids neg
0.0037	23.06	Dimethyl-octadienedioic acid	198.0891	0.66	Lipids neg
0.0037	24.43	hydroxy-dodecadienoic acid	212.1414	0.77	Lipids neg
0.0022	24.50	3-Phospho-D-erythronate	216.0027	0.58	Lipids neg
0.0061	26.07	2,2'-Iminobispropanoic acid	161.0687	0.59	Lipids neg
0.0027	28.69	hydroxyoctenoylglycine	215.1157	0.61	Lipids neg
0.0034	29.21	1-Octen-3-yl glucoside	290.1728	0.65	Lipids neg

0.0025	29.28	1-Keto-D-chiro-inositol	178.0475	0.58	Lipids neg
0.0030	29.34	2,6-Dimethyl-1,8-octanedioic acid	202.1204	0.66	Lipids neg
0.0028	29.97	Aspartyl-Threonine	234.0851	0.57	Lipids neg
0.0020	33.63	Oxooctanoic acid	158.0941	0.65	Lipids neg
0.0043	33.67	N-butanoyl-lhomoserine lactone	171.0895	0.66	Lipids neg
0.0022	33.77	2,5,6-Trihydroxy-5,6-dihydroquinoline	179.0581	0.64	Lipids neg
0.0045	34.31	2-Dehydro-D-galactonate	194.0425	0.60	Lipids neg
0.0028	36.97	Ferulic acid 4-sulfate	274.0148	0.59	Lipids neg
0.0022	37.02	N-(1-Deoxy-1-fructosyl)threonine	281.1123	0.56	Lipids neg
0.0022	37.78	3,4-Dihydroxy-L-phenylalanine	197.0688	0.64	Lipids neg
0.0026	37.88	(22R)-22,25-dihydroxyvitamin D3 / (22R)-22,25-dihydroxycholecalciferol	408.2883	0.82	Lipids neg
0.0044	38.46	dihydroxydecanoic acid	204.1358	0.65	Lipids neg
0.0026	40.95	Dihydroxy-dioxohexanoate	176.0319	0.60	Lipids neg
0.0047	41.95	3-Methoxy-4-Hydroxyphenylglycol sulfate	264.0302	0.60	Lipids neg
0.0027	43.58	(22E)-3alpha,6beta,7beta-Trihydroxy-5beta-chol-22-en-24-oic Acid	406.2726	0.84	Lipids neg
0.0022	43.99	Indole-3-carboxilic acid-O-sulphate	241.0045	0.59	Lipids neg
0.0022	44.73	3-Isopropylmalate	176.0683	0.62	Lipids neg
0.0025	51.59	4-(Glutamylamino) butanoate	232.1059	0.57	Lipids neg
0.0056	52.05	1-Deoxy-D-xylulose 5-phosphate	214.0244	0.58	Lipids neg
0.0039	52.13	4-Sulfobenzoate	201.9935	0.60	Lipids neg
0.0022	53.51	Hydroxydecanedioic acid	218.1152	0.64	Lipids neg
0.0023	55.59	Serylserine	192.0746	0.57	Lipids neg
0.0066	56.62	Isocitric acid	192.0269	0.60	Lipids neg
0.0065	58.89	3-hydroxy-3-(3-hydroxyphenyl)propanoic acid-O-sulphate	262.0146	0.59	Lipids neg
0.0022	59.36	Hydroxypropionylcarnitine	233.1263	0.62	Lipids neg

0.0020	59.55	Methylcrotonylglycine	157.0737	0.64	Lipids neg
0.0025	59.79	Nonenoic acid	156.1149	0.65	Lipids neg
0.0024	61.85	5-Hydroxydopamine	169.0739	0.62	Lipids neg
0.0020	65.20	Dehydroepiandrosterone sulfate	368.1657	0.68	Lipids neg
0.0040	84.87	nonynoic acid	154.0993	0.67	Lipids neg
0.0020	88.85	2-Methylbutyrylglycine	159.0894	0.65	Lipids neg
0.0020	92.76	(22E)-3alpha,6beta,7beta-Trihydroxy-5beta-chol-22-en-24-oic Acid	406.2726	0.98	Lipids neg
0.0024	94.27	2-O-Benzoyl-D-glucose	284.0895	0.62	Lipids neg
0.0020	105.43	3,4,5,6-Tetrahydrohippuric acid	183.0895	0.64	Lipids neg
0.0022	111.56	Methylhippuric acid	193.0737	0.65	Lipids neg
0.0029	113.99	hexadecatetraenoic acid	248.1777	1.74	Lipids neg
0.0078	132.15	Indoxyl sulfate	213.0095	0.61	Lipids neg
0.0022	137.47	4-Sulfobenzyl alcohol	188.0142	0.64	Lipids neg
0.0022	151.37	3-Hydroxy-5-methyl-L-tyrosine	211.0843	0.65	Lipids neg
0.0022	248.28	O-methoxycatechol-O-sulphate	204.0091	0.64	Lipids neg
0.0022	581.73	Glutamyl-taurine	254.0579	0.68	Lipids neg
0.0035	0.00	NeuAcalpha2-3Galbeta-Cer(d18:1/20:0)	1046.7262	4.45	Lipids pos
0.0053	0.00	PC(34:3)	755.5464	7.19	Lipids pos
0.0086	0.00	PC(36:2)	785.5957	7.40	Lipids pos
0.0072	0.00	PC(40:9)	827.5435	7.26	Lipids pos
0.0060	0.00	LysoPA(19:0)	452.2903	2.33	Lipids pos
0.0049	0.00	PC(32:0)	733.5644	7.19	Lipids pos
0.0057	0.00	PC(36:4)	809.5926	7.77	Lipids pos
0.0035	0.00	PC(36:4)	781.5613	7.41	Lipids pos
0.0073	0.00	PC(38:5)	807.5758	7.40	Lipids pos
0.0035	0.00	PC(38:6)	805.5615	7.26	Lipids pos
0.0049	0.00	(3-sulfo)Galbeta-Cer(d18:1/18:0(2OH))	823.5484	7.88	Lipids pos
0.0144	0.00	LysoPS(20:0)	553.3384	1.03	Lipids pos
0.0057	0.00	PA(39:3)	740.5349	7.19	Lipids pos
0.0103	0.00	PC(20:2/P-18:1)	795.6139	8.00	Lipids pos
0.0063	0.00	PC(34:2)	757.5610	7.50	Lipids pos
0.0046	0.00	PC(36:3)	783.5787	7.52	Lipids pos
0.0043	0.00	PC(40:6)	833.5926	7.63	Lipids pos
0.0096	0.00	1,3,7-Trioxo-5beta-cholan-24-oic Acid	402.2409	1.07	Lipids pos
0.0049	0.00	DG(44:8)	720.5664	7.52	Lipids pos
0.0063	0.00	Hexadecadienylcarnitine	423.3343	2.57	Lipids pos

0.0058	0.00	LysoPC(16:0)	495.3316	3.26	Lipids pos
0.0035	0.00	LysoPC(18:0)	523.3630	4.45	Lipids pos
0.0044	0.00	LysoPC(22:4)	571.3610	4.59	Lipids pos
0.0046	0.00	PC(40:5)	835.6073	7.76	Lipids pos
0.0049	0.00	PC(40:8)	829.5581	7.41	Lipids pos
0.0057	0.00	PS(O-16:0/18:0)	749.5594	6.09	Lipids pos
0.0046	0.00	PS(O-16:0/20:3)	771.5411	6.09	Lipids pos
0.0075	0.01	PA(O-16:0/21:0)	704.5734	7.66	Lipids pos
0.0103	0.01	PC(20:3/P-18:1)	793.5977	7.67	Lipids pos
0.0039	0.01	PC(20:4/P-18:1)	791.5801	7.51	Lipids pos
0.0046	0.01	PC(36:6)	777.5282	7.19	Lipids pos
0.0053	0.01	PS(O-16:0/20:2)	773.5563	6.10	Lipids pos
0.0050	0.01	PS(O-16:0/20:3)	771.5411	6.24	Lipids pos
0.0049	0.01	1-(6-[3]-ladderane-hexanoyl)-2-(8-[3]-ladderane-octanyl)-sn-glycerophosphoethanolamine	745.5024	7.68	Lipids pos
0.0044	0.01	15-HETE-DA	455.3015	2.71	Lipids pos
0.0250	0.01	3-Deoxyvitamin D3	368.3438	7.19	Lipids pos
0.0261	0.01	Cholesterol	386.3546	7.19	Lipids pos
0.0067	0.01	LysoPC(16:0)	495.3316	3.01	Lipids pos
0.0047	0.01	LysoPC(17:0)	509.3480	3.88	Lipids pos
0.0085	0.01	LysoPS(20:1)	551.3223	1.02	Lipids pos
0.0106	0.01	PC(16:1/2:0)	535.3275	1.03	Lipids pos
0.0043	0.01	PC(37:5)	793.5593	7.69	Lipids pos
0.0049	0.01	PI(34:2)	834.5276	7.40	Lipids pos
0.0062	0.01	PS(O-18:0/17:2)	759.5425	7.59	Lipids pos
0.0065	0.01	(3-sulfo)Galbeta-Cer(d18:1/16:0(2OH))	795.5158	8.01	Lipids pos
0.0132	0.01	(3beta,17alpha,23S)-17,23-Epoxy-3,29-dihydroxy-27-norlanosta-7,9(11)-diene-15,24-dione	470.3036	1.09	Lipids pos
0.0052	0.01	LysoPE(22:0)	537.3795	4.96	Lipids pos
0.0076	0.01	MG(18:0)	358.3078	3.19	Lipids pos
0.0218	0.01	N-Stearoyl tyrosine	447.3345	2.46	Lipids pos
0.0062	0.01	PC(32:2)	729.5308	7.10	Lipids pos
0.0035	0.01	PC(34:1)	787.6087	7.76	Lipids pos
0.0133	0.01	PI(36:5)	856.5067	7.20	Lipids pos
0.0046	0.01	PS(O-16:0/22:2)	801.5883	6.60	Lipids pos
0.0049	0.01	LysoPA(18:0)	424.2951	4.97	Lipids pos
0.0055	0.01	LysoPC(16:1)	493.3167	2.26	Lipids pos
0.0267	0.01	LysoPC(20:4)	543.3319	2.45	Lipids pos
0.0100	0.01	LysoPC(22:5)	569.3474	2.60	Lipids pos
0.0124	0.01	PI(38:4)	886.5557	7.49	Lipids pos
0.0201	0.01	CL(78:6)	1537.1032	7.51	Lipids pos
0.0077	0.01	LysoPC(16:1)	493.3167	2.34	Lipids pos

0.0046	0.01	LysoPE(20:3)	503.2987	2.62	Lipids pos
0.0057	0.01	MG(18:2)	354.2765	3.68	Lipids pos
0.0064	0.01	Palmitoylcarnitine	399.3345	3.19	Lipids pos
0.0059	0.01	PE(22:6/P-18:1)	773.5337	8.01	Lipids pos
0.0052	0.01	LysoPC(14:0)	467.3013	2.05	Lipids pos
0.0043	0.01	PS(38:7)	805.4868	7.15	Lipids pos
0.0049	0.01	PS(O-18:0/20:5)	795.5389	7.18	Lipids pos
0.0211	0.01	LysoPS(20:2)	549.3065	0.95	Lipids pos
0.0065	0.01	PC(40:7)	831.5739	7.56	Lipids pos
0.0054	0.01	PS(40:8)	831.5023	7.18	Lipids pos
0.0226	0.01	PS(42:11)	853.4869	6.69	Lipids pos
0.0043	0.01	PS(O-20:0/20:5)	823.5719	7.11	Lipids pos
0.0168	0.01	nitrooctadecenoic Acid	327.2408	1.91	Lipids pos
0.0074	0.01	PC(18:2/P-18:1)	767.5829	7.44	Lipids pos
0.0049	0.01	PI(36:3)	860.5411	7.21	Lipids pos
0.0044	0.01	PI(38:7)	880.5078	7.11	Lipids pos
0.0070	0.01	PS(O-16:0/22:2)	801.5880	6.69	Lipids pos
0.0057	0.01	CPA(16:0)	392.2328	3.83	Lipids pos
0.0388	0.01	PI(40:7)	908.5374	7.48	Lipids pos
0.0060	0.01	PS(31:2)	717.4576	4.61	Lipids pos
0.0120	0.01	PS(O-18:0/20:3)	799.5727	6.72	Lipids pos
0.0043	0.01	TG(57:9)	914.7344	6.54	Lipids pos
0.0043	0.02	LysoPC(16:0)	495.3321	2.56	Lipids pos
0.0049	0.02	LysoPC(20:2)	547.3638	3.71	Lipids pos
0.0049	0.02	LysoPC(P-18:0)	507.3688	3.94	Lipids pos
0.0076	0.02	CerP(d18:1/8:0)	505.3525	3.00	Lipids pos
0.0158	0.02	LysoPC(20:3)	545.3482	2.95	Lipids pos
0.0130	0.02	15-methyl-15R-PGF2alpha	368.2563	1.78	Lipids pos
0.0240	0.02	1alpha,25-dihydroxy-24-oxo-23-azavitamin D2 / 1alpha,25-dihydroxy-24-oxo-23-azaergocalciferol	445.3193	1.85	Lipids pos
0.0065	0.02	LysoPC(20:0)	551.3949	5.41	Lipids pos
0.0059	0.02	LysoPE(22:5)	527.3012	2.53	Lipids pos
0.0159	0.02	PC(36:5)	779.5465	6.84	Lipids pos
0.0049	0.02	PS(38:2)	815.5672	6.18	Lipids pos
0.0128	0.02	LysoPA(16:1)	408.2280	1.58	Lipids pos
0.0174	0.02	LysoPA(24:0)	522.3682	7.19	Lipids pos
0.0043	0.02	LysoPC(14:0)	467.3013	1.86	Lipids pos
0.0104	0.02	LysoPC(P-16:0)	479.3376	3.68	Lipids pos
0.0051	0.02	PC(33:5)	737.4997	7.19	Lipids pos
0.0142	0.02	PS(40:9)	829.4868	6.81	Lipids pos
0.0137	0.02	eicosadienoylcarnitine	451.3661	3.59	Lipids pos
0.0063	0.02	LysoPS(22:1)	579.3534	3.37	Lipids pos
0.0291	0.02	PE-Cer(d14:2/23:0)	700.5518	7.11	Lipids pos
0.0128	0.02	11′-Carboxy-gamma-chromanol	404.2923	1.85	Lipids pos

0.0049	0.02	PI(34:1)	836.5391	7.39	Lipids pos
0.0305	0.02	(8E,10S,12Z,15Z)-10-Hydroperoxyoctadeca-8,12,15-trienoate	310.2142	1.92	Lipids pos
0.0055	0.02	7,8-Diaminononanoate	188.1524	0.55	Lipids pos
0.0115	0.02	Hexadecenoylcarnitine	397.3191	2.28	Lipids pos
0.0112	0.02	15-methyl-15R-PGD2	366.2406	1.16	Lipids pos
0.0180	0.02	MG(20:4)	378.2769	3.77	Lipids pos
0.0127	0.02	PA(16:0)	424.2226	1.06	Lipids pos
0.0093	0.02	PS(29:2)	689.4246	4.48	Lipids pos
0.0064	0.02	PS(42:9)	857.5172	7.39	Lipids pos
0.0043	0.02	LysoPE(22:1)	535.3638	4.02	Lipids pos
0.0061	0.02	N-stearoyl serine	371.3035	3.96	Lipids pos
0.0137	0.03	(22,23-dinor)-24-vinyl-cholest-5-en-3beta,24-diol	400.3338	5.29	Lipids pos
0.0251	0.03	3-hydroxylinoleoylcarnitine	439.3298	0.99	Lipids pos
0.0080	0.03	LysoPC(18:2)	519.3318	2.37	Lipids pos
0.0526	0.03	LysoPC(20:1)	549.3795	4.34	Lipids pos
0.0068	0.03	LysoPE(22:6)	525.2858	2.20	Lipids pos
0.0161	0.03	PG(29:2)	676.4312	4.61	Lipids pos
0.0069	0.03	LysoPC(18:3)	517.3171	1.89	Lipids pos
0.0191	0.03	9,10-epoxy-3E,12Z-Octadecadienoic acid	294.2193	2.25	Lipids pos
0.0105	0.03	LysoPC(O-18:0)	509.3846	4.91	Lipids pos
0.0039	0.03	PC(42:8)	857.5892	7.90	Lipids pos
0.0055	0.03	PS(34:0)	763.5392	7.26	Lipids pos
0.0069	0.03	PS(41:7)	847.5397	7.77	Lipids pos
0.0043	0.03	SM(d18:2/18:1)	726.5671	7.20	Lipids pos
0.0055	0.03	1-(14-methyl-pentadecanoyl)-2-(8-[3]-ladderane-octanyl)-sn-glycerol	602.5272	7.57	Lipids pos
0.0507	0.03	PC(42:11)	851.5443	6.94	Lipids pos
0.0076	0.03	PS(42:1)	873.6427	4.22	Lipids pos
0.0104	0.03	LysoPC(22:0)	579.4263	6.18	Lipids pos
0.0043	0.03	PS(O-18:0/22:6)	821.5564	6.07	Lipids pos
0.0086	0.03	LysoPC(18:4)	515.2988	2.35	Lipids pos
0.0080	0.03	LysoPC(O-20:0)	537.4160	5.81	Lipids pos
0.0081	0.03	LysoPC(P-18:0)	507.3690	4.82	Lipids pos
0.0049	0.04	LysoPE(20:5)	499.2676	2.67	Lipids pos
0.0043	0.04	PS(39:7)	819.5026	7.35	Lipids pos
0.0129	0.04	(+/-)-10-HDoHE	344.2353	5.25	Lipids pos
0.0143	0.04	octadecadiynoic acid	276.2086	2.48	Lipids pos
0.0062	0.04	PI(40:8)	906.5235	7.14	Lipids pos
0.0320	0.04	CDP-DG(28:0)	897.4506	6.69	Lipids pos
0.0059	0.04	docosadienoylcarnitine	479.3973	6.31	Lipids pos

0.0062	0.04	LysoPC(O-16:0)	481.3531	3.76	Lipids pos
0.0052	0.04	Isolimonic acid	506.2170	3.96	Lipids pos
0.0195	0.04	O-(13-carboxytridecanoyl)carnitine	401.2775	1.23	Lipids pos
0.0137	0.04	PC(42:7)	859.6112	7.67	Lipids pos
0.0044	0.04	(10R)-1alpha,19,25-trihydroxy-10,19-dihydrovitamin D3 / (10R)-1alpha,19,25-trihydroxy-10,19-dihydrocholecalciferol	434.3393	5.82	Lipids pos
0.0152	0.04	DG(40:8)	664.5047	7.94	Lipids pos
0.0218	0.04	LysoPC(24:0)	607.4578	6.81	Lipids pos
0.0057	0.04	PC(31:0)	719.5489	7.21	Lipids pos
0.0221	0.04	PE(42:10)	811.5148	7.24	Lipids pos
0.0058	0.04	(23Z)-25-hydroxy-16,17,23,24-tetradehydrovitamin D3 / (23Z)-25-hydroxy-16,17,23,24-tetradehydrocholecalciferol	396.3031	1.85	Lipids pos
0.0187	0.04	hydroxyoctadecatrienoylcarnitine	437.3141	1.13	Lipids pos
0.0060	0.04	PC(16:0/9:0(CHO))	649.4319	4.79	Lipids pos
0.0359	0.04	6alpha-Glucuronosylhyodeoxycholate	567.3176	0.65	Lipids pos
0.0074	0.04	LysoPC(22:1)	577.4107	5.47	Lipids pos
0.0128	0.04	(22,23-dinor)-24-vinyl-cholest-5-en-3beta,24-diol	400.3337	5.47	Lipids pos
0.0057	0.04	Heptadecanoyl carnitine	413.3504	3.74	Lipids pos
0.0395	0.04	LysoPC(22:6)	567.3321	2.11	Lipids pos
0.0055	0.04	LysoPE(O-16:0)	439.3061	3.89	Lipids pos
0.0064	0.05	1-Palmitoyl-2-(5-hydroxy-8-oxo-6-octenedioyl)-sn-glycero-3-phosphocholine	665.3902	3.18	Lipids pos
0.0049	0.05	3-hydroxyeicosanoylcarnitine	471.3921	4.38	Lipids pos
0.0079	0.05	LysoPA(P-16:0)	394.2485	4.74	Lipids pos
0.0085	0.05	PS(33:5)	739.4397	4.62	Lipids pos
0.0375	0.05	carboxyheptadecenoylcarnitine	455.3244	0.79	Lipids pos
0.0411	0.05	hydroxy-octadecadienynoic acid	292.2037	1.53	Lipids pos

0.0196	0.05	13'-Hydroxy-alpha-tocopherol	446.3759	6.54	Lipids pos
0.0059	0.05	MG(18:1)	356.2925	4.74	Lipids pos
0.0057	0.05	PA(P-16:0/13:0)	590.4312	7.30	Lipids pos
0.0221	0.05	PC(41:7)	845.5896	7.93	Lipids pos
0.0199	0.05	(25S)-24S,26-cyclo-5alpha-cholestan-3beta-ol	384.3392	5.47	Lipids pos
0.0096	0.05	PC(22:4/P-18:0)	821.6286	7.91	Lipids pos
0.0503	0.05	hydroxy-octadecadienynoic acid	292.2036	1.91	Lipids pos
0.0052	0.05	PS(28:1)	677.4266	4.75	Lipids pos
0.0352	0.06	(8E,10S,12Z,15Z)-10-Hydroperoxyoctadeca-8,12,15-trienoate	310.2142	1.52	Lipids pos
0.0097	0.06	PG(O-16:0/12:0)	652.4679	6.44	Lipids pos
0.0165	0.06	PS(O-18:0/20:5)	795.5389	6.10	Lipids pos
0.0057	0.06	(17S,20R)-1alpha,25-dihydroxy-17,20-methano-21-norvitamin D3 / (17S,20R)-1alpha,25-dihydroxy-17,20-methano-21-norcholecalciferol	414.3136	1.85	Lipids pos
0.0097	0.06	(22R)-1alpha,22,25-trihydroxy-26,27-dimethyl-23,23,24,24-tetradehydro-24a,24b-dihomovitamin D3 / (22R)-1alpha,22,25-trihydroxy-26,27-dimethyl-23,23,24,24-tetradehydro-24a,24b-dihomocholecalciferol	484.3529	6.21	Lipids pos
0.0261	0.06	LysoPA(17:0)	424.2587	3.50	Lipids pos
0.0050	0.06	PC(20:1)	563.3598	1.32	Lipids pos
0.0068	0.06	11-Dehydrocorticosterone	344.1983	4.62	Lipids pos
0.0053	0.06	LysoPS(20:4)	545.2743	2.94	Lipids pos
0.0103	0.06	PC(38:8)	801.5296	7.01	Lipids pos
0.0090	0.06	PS(30:1)	705.4578	4.60	Lipids pos
0.0053	0.07	Hydroxyhexadecanoylcarnitine	415.3300	3.38	Lipids pos
0.0097	0.07	PS(27:1)	663.4112	4.77	Lipids pos
0.0087	0.07	PS(42:7)	861.5483	6.07	Lipids pos

0.0166	0.07	Tetradecenoylcarnitine	369.2878	1.54	Lipids pos
0.0218	0.07	4,8 Dimethylnonanoyl carnitine	329.2565	1.32	Lipids pos
0.0117	0.07	PA(40:8)	744.4710	7.34	Lipids pos
0.0049	0.07	PI-Cer(d20:1/16:0)	807.5625	4.62	Lipids pos
0.0053	0.07	28-Glucosyl-3b-hydroxy-12-oleanene-30-methoxy-28-oic acid 3-[arabinosyl-(1->3)-glucuronide]	970.4777	7.11	Lipids pos
0.0046	0.07	LysoPS(22:4)	573.3047	1.24	Lipids pos
0.0049	0.07	PA(27:0)	578.3948	6.90	Lipids pos
0.0057	0.07	PA(O-16:0/13:0)	592.4473	7.89	Lipids pos
0.0517	0.07	DG(38:6)	640.5063	7.62	Lipids pos
0.0162	0.07	hydroxytetradecanoylcarnitine	387.2987	1.55	Lipids pos
0.0901	0.07	PC(42:10)	853.5622	6.78	Lipids pos
0.0509	0.07	DG(40:7)	666.5201	8.23	Lipids pos
0.0035	0.07	LysoPC(22:2)	575.3952	4.74	Lipids pos
0.0083	0.07	PC(22:1)	591.3878	5.24	Lipids pos
0.0091	0.08	PA(36:2)	700.5048	7.21	Lipids pos
0.0183	0.08	1alpha,25-dihydroxy-2beta-(6-hydroxyhexyl)vitamin D3 / 1alpha,25-dihydroxy-2beta-(6-hydroxyhexyl)cholecalciferol	516.4179	6.18	Lipids pos
0.0059	0.08	PI(42:9)	932.5392	7.36	Lipids pos
0.0497	0.08	MG(22:6)	402.2769	3.54	Lipids pos
0.0683	0.08	PA(O-16:0/21:0)	704.5740	8.38	Lipids pos
0.0127	0.08	6a-hydroxy DHEA 3-sulfate	384.1589	3.90	Lipids pos
0.0248	0.08	N-oleoyl alanine	353.2927	2.36	Lipids pos
0.0094	0.08	(E)-2-Butenyl-4-methyl-threonine	187.1215	3.21	Lipids pos
0.0057	0.08	Kinetensin 4-7	571.2869	2.63	Lipids pos
0.0397	0.08	PE(42:11)	809.4964	7.02	Lipids pos
0.0046	0.08	LysoPE(16:1)	451.2695	2.35	Lipids pos
0.0236	0.08	PC(32:1)	731.5452	7.76	Lipids pos
0.0128	0.08	Pentadecanoylcarnitine	385.3188	2.61	Lipids pos
0.0875	0.08	DG(44:10)	716.5394	7.18	Lipids pos
0.0207	0.09	5′-Methylthioadenosine	297.0895	0.62	Lipids pos
0.0207	0.09	hydroxyoctadecenoylcarnitine	411.3445	2.45	Lipids pos
0.0077	0.09	PA(39:2)	742.5514	6.83	Lipids pos
0.0311	0.09	PE(6:0/6:0)	411.2019	0.61	Lipids pos
0.0053	0.09	LysoPS(O-18:0)	511.3274	1.21	Lipids pos
0.0088	0.09	MG(18:3)	352.2612	3.21	Lipids pos
0.0480	0.09	N-palmitoyl tyrosine	419.3032	3.54	Lipids pos

0.0583	0.09	N-stearoyl serine	371.3035	2.08	Lipids pos
0.0084	0.09	PE(P-16:0/13:0)	633.4734	6.80	Lipids pos
0.0063	0.09	DG(36:8)	608.4418	7.30	Lipids pos
0.0095	0.09	DG(40:6)	668.5355	8.32	Lipids pos
0.0334	0.09	4,8 Dimethylnonanoyl carnitine	329.2565	1.38	Lipids pos
0.0151	0.09	O-(17-carboxyheptadecanoyl)carnitine	457.3404	0.85	Lipids pos
0.0188	0.10	(20R)-17alpha,20-dihydroxycholesterol	418.3443	6.45	Lipids pos
0.0049	0.10	LysoPE(22:2)	533.3482	3.14	Lipids pos
0.0115	0.10	Cer(d14:2/20:1(2OH))	549.4760	6.31	Lipids pos
0.0294	0.10	2alpha-(Hydroxymethyl)-17-methyl-5alpha-androstane-3beta,17beta-diol	336.2662	2.25	Lipids pos
0.0719	0.10	LysoPC(20:5)	541.3171	1.12	Lipids pos
0.0157	0.10	PS(35:4)	769.4885	5.25	Lipids pos
0.0059	0.10	1-Naphthylacetylspermine	370.2719	2.42	Lipids pos
0.0163	0.10	1,4-cholestadie24,25-didehydrovitamin D3 / 24,25-didehydrocholecalciferol	382.3237	7.22	Lipids pos
0.0656	0.10	(23R)-1alpha,23,25-trihydroxy-24-oxovitamin D3 / (23R)-1alpha,23,25-trihydroxy-24-oxocholecalciferol	446.3034	1.15	Lipids pos
0.0268	0.10	(8E,10S,12Z,15Z)-10-Hydroperoxyoctadeca-8,12,15-trienoate	310.2142	1.65	Lipids pos
0.0077	0.10	DG(42:7)	694.5533	7.18	Lipids pos
0.0151	0.11	Hepteneoylglycine	185.1045	4.62	Lipids pos
0.0079	0.11	PA(30:1)	618.4262	7.18	Lipids pos
0.0060	0.11	Arachidyl carnitine	455.3973	5.19	Lipids pos
0.0283	0.11	(25S)-24S,26-cyclo-5alpha-cholestan-3beta-ol	384.3392	6.11	Lipids pos
0.0055	0.11	PI(39:7)	894.5232	7.31	Lipids pos
0.0219	0.11	DG(38:4)	644.5381	8.22	Lipids pos

0.0200	0.11	1-Palmitoyl-2-(5-keto-6-octenedioyl)-sn-glycero-3-phosphocholine	663.3750	2.85	Lipids pos
0.0053	0.11	1alpha,25-dihydroxy-24-oxo-23-azavitamin D2 / 1alpha,25-dihydroxy-24-oxo-23-azaergocalciferol	445.3188	4.04	Lipids pos
0.0079	0.11	LysoPC(24:1)	605.4420	6.06	Lipids pos
0.0150	0.12	Amino acid(Arg-)	174.1116	0.54	Lipids pos
0.0288	0.12	(20R)-17alpha,20-dihydroxycholesterol	418.3444	4.88	Lipids pos
0.0076	0.12	PG(34:2)	746.5096	6.54	Lipids pos
0.0089	0.12	PS(O-16:0/20:3)	771.5411	6.53	Lipids pos
0.0642	0.12	11-deoxy-PGF1a	340.2610	3.10	Lipids pos
0.0050	0.12	Alpha-linolenyl carnitine	421.3173	3.85	Lipids pos
0.0069	0.12	Docosahexaenoic acid	328.2393	8.04	Lipids pos
0.0256	0.13	DG(38:5)	642.5224	7.93	Lipids pos
0.0228	0.13	2-Methylbutyroylcarnitine	245.1634	0.94	Lipids pos
0.0212	0.13	Oleoylethanolamide	325.2978	4.27	Lipids pos
0.0124	0.13	PS(30:2)	703.4421	4.61	Lipids pos
0.0076	0.13	LysoPI(20:3)	622.3099	3.94	Lipids pos
0.0081	0.13	PC(16:0)	509.3121	1.38	Lipids pos
0.0057	0.13	PI(44:12)	954.5214	7.35	Lipids pos
0.0120	0.14	LysoPS(18:3)	519.2574	2.70	Lipids pos
0.0059	0.14	PC(28:0)	677.4992	7.00	Lipids pos
0.0125	0.14	LysoPS(20:3)	547.2888	3.95	Lipids pos
0.0077	0.14	PI(42:6)	934.5535	7.46	Lipids pos
0.0051	0.14	Stearoylcarnitine	427.3659	4.14	Lipids pos
0.0284	0.14	N-propyl-16,16-dimethyl-5Z,8Z,11Z,14Z-docosatetraenoyl amine	401.3654	6.11	Lipids pos
0.0044	0.14	PS(26:1)	649.3953	3.62	Lipids pos
0.0072	0.14	Butyrylcarnitine	231.1469	0.60	Lipids pos
0.0079	0.14	CerP(d18:1/14:0)	589.4472	5.96	Lipids pos
0.0537	0.15	(22,23-dinor)-24-vinyl-cholest-5-en-3beta,24-diol	400.3337	5.80	Lipids pos
0.0086	0.15	4-(3-Methylbut-2-enyl)-L-tryptophan	272.1525	0.86	Lipids pos
0.0535	0.15	Anandamide (18:4, n-3)	319.2508	2.37	Lipids pos
0.0043	0.15	27-Nor-5b-cholestane-3a,7a,12a,24,25-pentol	438.3341	5.49	Lipids pos

0.0128	0.15	LysoPS(16:0)	497.2756	2.71	Lipids pos
0.0331	0.15	PA(37:1)	716.5358	7.18	Lipids pos
0.0077	0.16	17alpha-(N-Acetyl-D-glucosaminyl)-estradiol 3-D-glucuronide	651.2886	3.88	Lipids pos
0.0789	0.16	LysoPS(22:2)	577.3384	1.25	Lipids pos
0.0080	0.16	PC(18:0)	537.3434	2.94	Lipids pos
0.0106	0.16	PS(30:3)	729.4577	4.85	Lipids pos
0.0119	0.16	Histidine trimethylbetaine	197.1163	0.55	Lipids pos
0.0321	0.16	2alpha-(Hydroxymethyl)-17-methyl-5alpha-androstane-3beta,17beta-diol	336.2661	3.95	Lipids pos
0.0072	0.16	LysoPC(P-20:0)	535.4003	5.73	Lipids pos
0.0076	0.16	(20S)-20-cyclopropyl-1alpha,25-dihydroxy-16,17-didehydro-21-norvitamin D3 / (20S)-20-cyclopropyl-1alpha,25-dihydroxy-16,17-didehydro-21-norcholecalciferol	440.3290	5.25	Lipids pos
0.0508	0.16	Dodecenoylcarnitine	341.2565	1.91	Lipids pos
0.0141	0.16	PC(41:6)	847.6049	6.08	Lipids pos
0.0394	0.16	Alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptanedioate	390.1756	3.90	Lipids pos
0.0228	0.16	PS(31:4)	713.4266	3.66	Lipids pos
0.0662	0.16	(22E)-1alpha-hydroxy-22,23-didehydrovitamin D3 / (22E)-1alpha-hydroxy-22,23-didehydrocholecalciferol	398.3183	3.94	Lipids pos
0.0351	0.16	hydroxyoctadecanoylcarnitine	443.3606	3.44	Lipids pos
0.0159	0.16	Icosadienoic acid	308.2714	4.30	Lipids pos
0.0060	0.17	DG(34:2)	592.5067	8.02	Lipids pos
0.0147	0.17	Decaprenol phosphate	778.6001	7.42	Lipids pos
0.0082	0.17	Propanoylcarnitine	217.1312	0.59	Lipids pos
0.0059	0.18	DG(36:3)	618.5223	8.04	Lipids pos

0.0564	0.18	(9Z,11E)-12-((3S,5R)-5-((R)-1-hydroxypropyl)-1,2-dioxolan-3-yl)dodeca-9,11-dienoic acid	326.2094	0.96	Lipids pos
0.0602	0.19	(4E,8E,10E-d18:3)sphingosine	295.2508	2.06	Lipids pos
0.0241	0.19	1alpha,25-dihydroxy-19-nor-22-oxavitamin D3 / 1alpha,25-dihydroxy-19-nor-22-oxacholecalciferol	406.3081	4.55	Lipids pos
0.0250	0.19	9,10-epoxy-3E,12Z-Octadecadienoic acid	294.2194	2.35	Lipids pos
0.0577	0.20	11'-Carboxy-gamma-chromanol	404.2925	3.90	Lipids pos
0.0320	0.20	octadecapentaenoic acid	274.1935	1.13	Lipids pos
0.0076	0.20	PE(42:9)	813.5298	7.08	Lipids pos
0.0059	0.20	Dodecanoylcarnitine	343.2722	1.29	Lipids pos
0.0057	0.20	CerP(d18:1/12:0)	561.4158	5.16	Lipids pos
0.0874	0.21	16:0 Sitosteryl ester	652.6186	3.01	Lipids pos
0.0389	0.21	DG(28:0)	372.2877	1.86	Lipids pos
0.0108	0.21	hydroxy cholesterol	402.3497	5.71	Lipids pos
0.0691	0.21	LysoPS(20:5)	543.2601	0.97	Lipids pos
0.0204	0.21	PC(22:0)	593.4054	5.35	Lipids pos
0.0062	0.21	LysoPS(O-20:0)	539.3591	1.63	Lipids pos
0.0184	0.21	decadienoic acid	168.1150	1.91	Lipids pos
0.0162	0.21	LysoPC(O-17:0)	495.3689	4.36	Lipids pos
0.0470	0.21	(3R)-11-cis-3-Hydroxyretinal	300.2088	1.52	Lipids pos
0.0148	0.21	2-(8-[3]-ladderane-octanyl)-sn-glycero-3-phosphocholine	529.3516	4.83	Lipids pos
0.0620	0.21	DG(38:7)	638.4912	7.34	Lipids pos
0.0079	0.22	1-Aminopropan-2-ol O-phosphate	155.0348	0.56	Lipids pos
0.0265	0.22	14-methyl-20,14-retro-retinoic acid	314.2244	1.52	Lipids pos
0.0328	0.23	2-Hexaprenyl-3-methyl-6-methoxy-1,4-benzoquinol	562.4362	7.93	Lipids pos
0.0090	0.23	N-Nonanoylglycine	215.1530	5.16	Lipids pos
0.0322	0.23	Tetradecadiencarnitine	367.2724	1.16	Lipids pos
0.0043	0.23	Docosanoylcarnitine	483.4284	5.95	Lipids pos

0.0104	0.23	(22E)-1alpha-hydroxy-22,23-didehydrovitamin D3 / (22E)-1alpha-hydroxy-22,23-didehydrocholecalciferol	398.3186	5.09	Lipids pos
0.0432	0.24	Carnitine	161.1051	0.57	Lipids pos
0.0201	0.24	LysoPC(O-16:0)	481.3531	3.61	Lipids pos
0.0125	0.24	LysoPE(24:0)	565.4110	5.82	Lipids pos
0.0101	0.24	PS(31:3)	715.4402	4.83	Lipids pos
0.0107	0.24	PS(32:4)	727.4402	4.61	Lipids pos
0.0199	0.24	5-Aminoimidazole ribonucleotide	295.0584	0.55	Lipids pos
0.0321	0.25	PG(O-16:0/18:1)	734.5464	7.19	Lipids pos
0.0322	0.25	hydroxy cholesterol	402.3497	6.11	Lipids pos
0.0288	0.25	PS(26:0)	651.4112	3.70	Lipids pos
0.0282	0.26	(22,23-dinor)-24-vinyl-cholest-5-en-3beta,24-diol	400.3342	2.18	Lipids pos
0.0197	0.26	LysoPI(13:0)	530.2499	2.36	Lipids pos
0.0203	0.26	PC(39:7)	817.5595	8.05	Lipids pos
0.0563	0.26	2-methylcitric acid	206.0433	0.55	Lipids pos
0.0225	0.26	PC(20:0)	565.3745	3.85	Lipids pos
0.0136	0.27	Glycerophosphocholine	257.1026	0.55	Lipids pos
0.0490	0.27	PS(O-16:0/20:3)	771.5410	7.04	Lipids pos
0.0360	0.27	PS(35:2)	773.5199	5.23	Lipids pos
0.0464	0.28	LysoPE(20:1)	507.3327	4.67	Lipids pos
0.0306	0.29	Cer(d14:2/18:1(2OH))	521.4425	5.82	Lipids pos
0.0255	0.29	3-Hexaprenyl-4,5-Dihydroxybenzoic acid	562.4001	7.19	Lipids pos
0.0318	0.29	13'-Hydroxy-alpha-tocopherol	446.3760	6.90	Lipids pos
0.0064	0.29	1,4-cholestadie24,25-didehydrovitamin D3 / 24,25-didehydrocholecalciferol	382.3235	6.99	Lipids pos
0.0183	0.30	PG(O-16:0/20:5)	754.5128	6.69	Lipids pos
0.0201	0.30	1-(O-alpha-D-glucopyranosyl)-25-keto-(1,3R,27R)-octacosanetriol	618.4735	5.88	Lipids pos
0.0351	0.31	1,4-cholestadie24,25-didehydrovitamin D3 / 24,25-didehydrocholecalciferol	382.3233	7.18	Lipids pos

0.0428	0.32	PI(O-20:0/21:0)	922.6870	7.47	Lipids pos
0.0214	0.32	PS(28:0)	679.4424	4.68	Lipids pos
0.0152	0.33	hydroxy-heneicosanoic acid	342.3130	5.32	Lipids pos
0.0476	0.33	Phorbol 12-tiglate 13-decanoate	600.3638	5.75	Lipids pos
0.0148	0.36	3-Hexaprenyl-4,5-Dihydroxybenzoic acid	562.3997	7.23	Lipids pos
0.0357	0.37	DG(34:4)	616.5066	7.73	Lipids pos
0.0096	0.37	PC(36:8)	773.4975	7.34	Lipids pos
0.0697	0.37	PC(O-16:0/21:0)	789.6636	8.00	Lipids pos
0.0072	0.37	13'-Hydroxy-alpha-tocopherol	446.3757	7.45	Lipids pos
0.0077	0.38	(20R)-17alpha,20-dihydroxycholesterol	418.3443	5.66	Lipids pos
0.0796	0.38	PS(37:2)	801.5500	6.34	Lipids pos
0.0778	0.39	DG(36:5)	614.4907	7.46	Lipids pos
0.0051	0.39	1-(8-[3]-ladderane-octanoyl-2-(8-[3]-ladderane-octanyl)-sn-glycerol	650.5277	7.36	Lipids pos
0.0294	0.40	LysoPA(17:1)	422.2429	6.08	Lipids pos
0.0586	0.40	LysoPA(14:0)	382.2118	1.40	Lipids pos
0.0913	0.41	LysoPE(22:1)	535.3638	3.79	Lipids pos
0.0741	0.41	Type III cyanolipid 22:0 ester	419.3762	6.11	Lipids pos
0.0960	0.44	Plastoquinone-1	204.1144	0.93	Lipids pos
0.0493	0.45	(25S)-24S,26-cyclo-5alpha-cholestan-3beta-ol	384.3391	7.10	Lipids pos
0.0567	0.46	CPA(18:2)	416.2319	2.54	Lipids pos
0.0047	0.46	Eicosatrienoic acid	306.2553	2.94	Lipids pos
0.0574	0.47	Protoporphyrin	562.2579	3.92	Lipids pos
0.0295	0.48	Hydroxyquinine	340.1779	4.62	Lipids pos
0.0560	0.48	PS(O-16:0/17:2)	731.5103	7.19	Lipids pos
0.0316	0.53	DG(34:3)	590.4914	6.79	Lipids pos
0.0358	0.54	trimethoxybenzoic acid	212.0674	0.56	Lipids pos
0.0260	0.54	PS(33:3)	743.4730	4.35	Lipids pos
0.0428	0.55	(22,23-dinor)-24-vinyl-cholest-5-en-3beta,24-diol	400.3343	4.86	Lipids pos
0.0234	0.57	N-Octanoyl-L-homoserine lactone	227.1528	4.62	Lipids pos
0.0531	0.58	LysoPC(O-14:1)	451.3082	4.21	Lipids pos
0.0662	0.65	PA(O-16:0/14:1)	604.4443	4.62	Lipids pos
0.0078	0.68	12-HETE-Gly	377.2565	1.41	Lipids pos

0.0874	1.27	Sphinganine	301.2978	2.96	Lipids pos
0.0680	1.39	12-Oxo-20-carboxy-leukotriene B4	364.1884	0.83	Lipids pos
0.0983	1.45	N-Docosahexaenoyl phenylalanine	475.3107	4.01	Lipids pos
0.0832	1.47	Ketosphingosine	297.2667	4.62	Lipids pos
0.0874	1.48	(22E)-(24R)-1alpha,24-dihydroxy-26,27-cyclo-22,23-didehydro-20-epivitamin D3 / (22E)-(24R)-1alpha,24-dihydroxy-26,27-cyclo-22,23-didehydro-20-epicholecalciferol	412.2973	4.06	Lipids pos
0.0817	1.49	N-(3E-hexadecenoyl)-deoxysphing-4-enine-1-sulfonate	599.4605	4.63	Lipids pos
0.0810	1.61	anhydroretinol	268.2191	4.12	Lipids pos
0.0508	1.67	LysoPC(O-14:1)	451.3058	5.01	Lipids pos
0.0670	1.68	3-hydroxyoctanoyl carnitine	303.2045	0.60	Lipids pos
0.0396	1.69	Cer(d18:0/22:0)	623.6217	8.32	Lipids pos
0.0125	1.75	Anandamide (18:2, n-6)	323.2821	3.45	Lipids pos
0.0577	1.83	Prenylated FMNH2	526.1807	6.07	Lipids pos
0.0314	1.86	LysoPA(20:5)	456.2277	1.11	Lipids pos
0.0422	1.90	11-cis retro-gamma-retinal	284.2132	3.43	Lipids pos

Appendix G

Normalisation - Total Peak Area plots

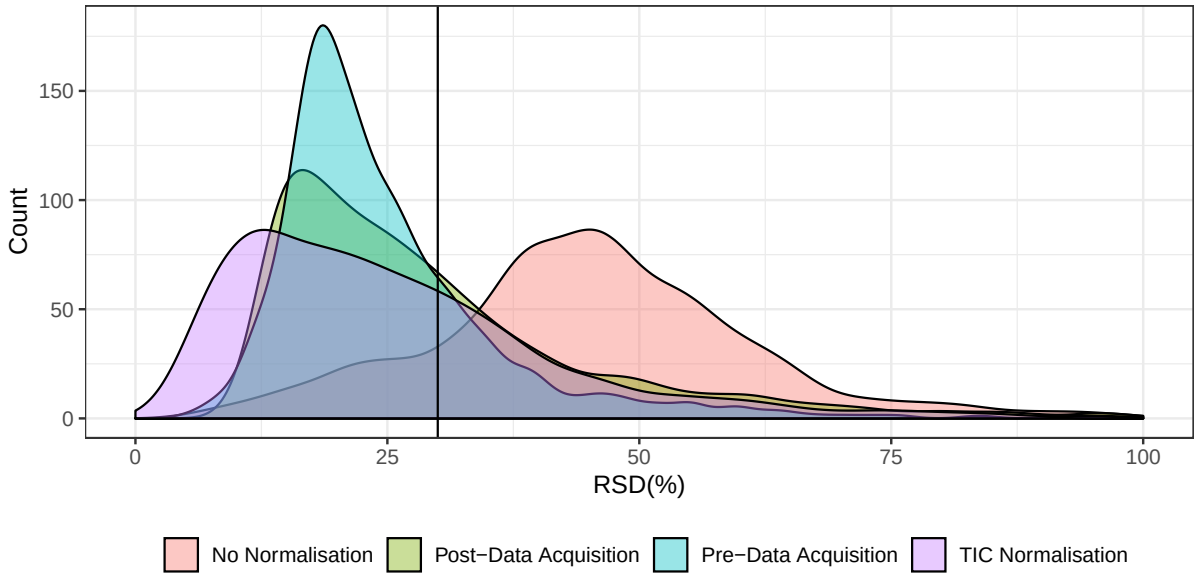


FIGURE G.1: RSD of all high quality features in the lipidomics, negative ion mode assay before and after normalisation either pre or post data acquisition or normalising applying total ion current alone. Vertical line indicates a 30% RSD cut off.

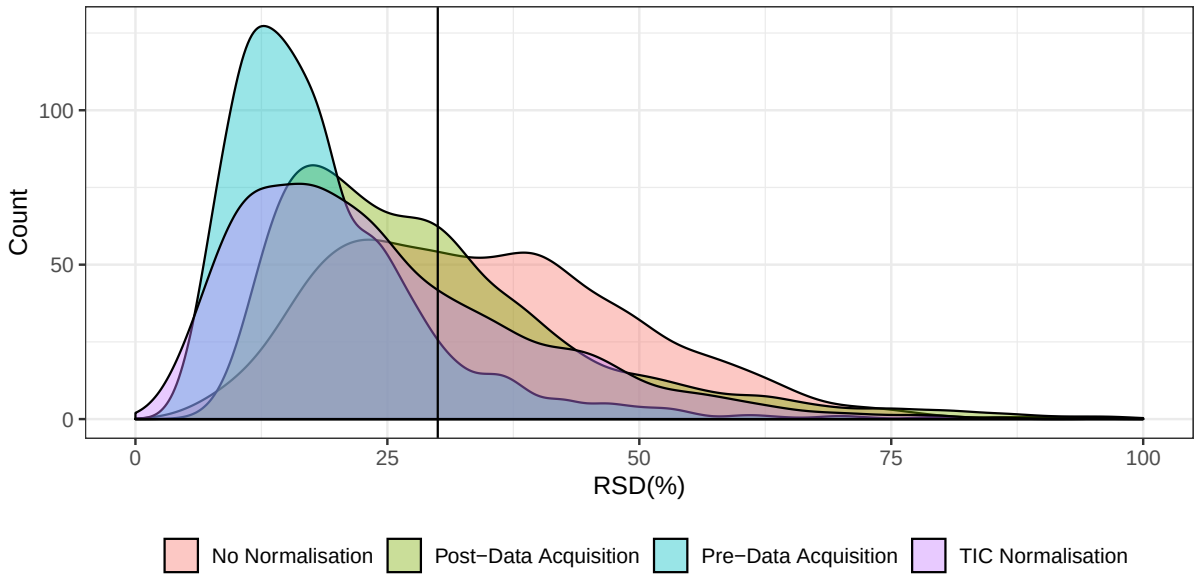


FIGURE G.2: RSD of all high quality features in the HILIC, positive ion mode assay before and after normalisation either pre or post data acquisition or normalising applying total ion current alone. Vertical line indicates a 30% RSD cut off.

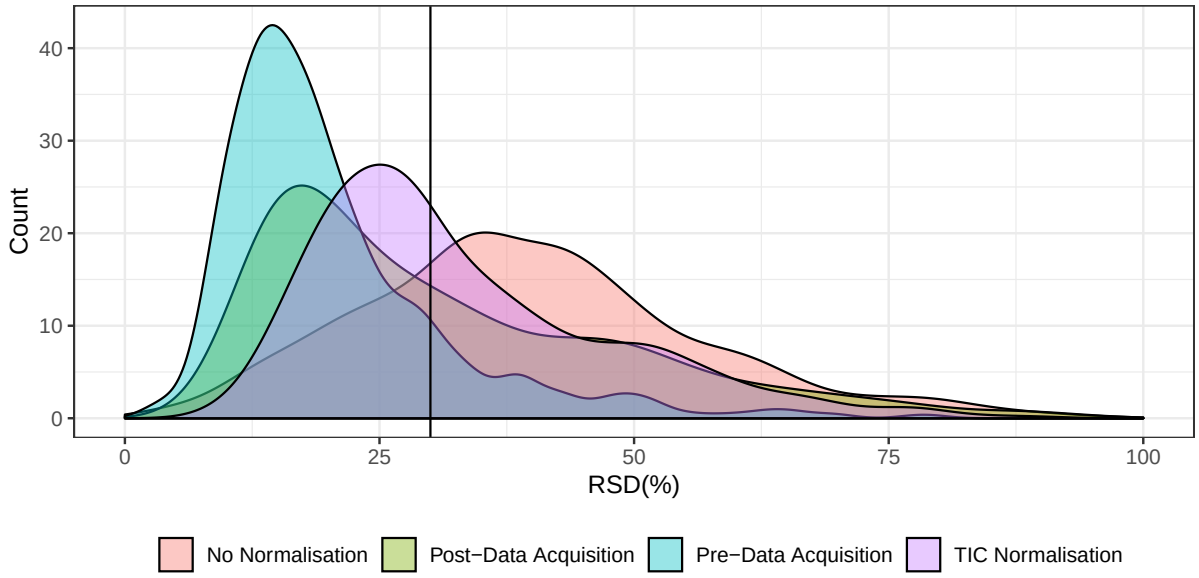
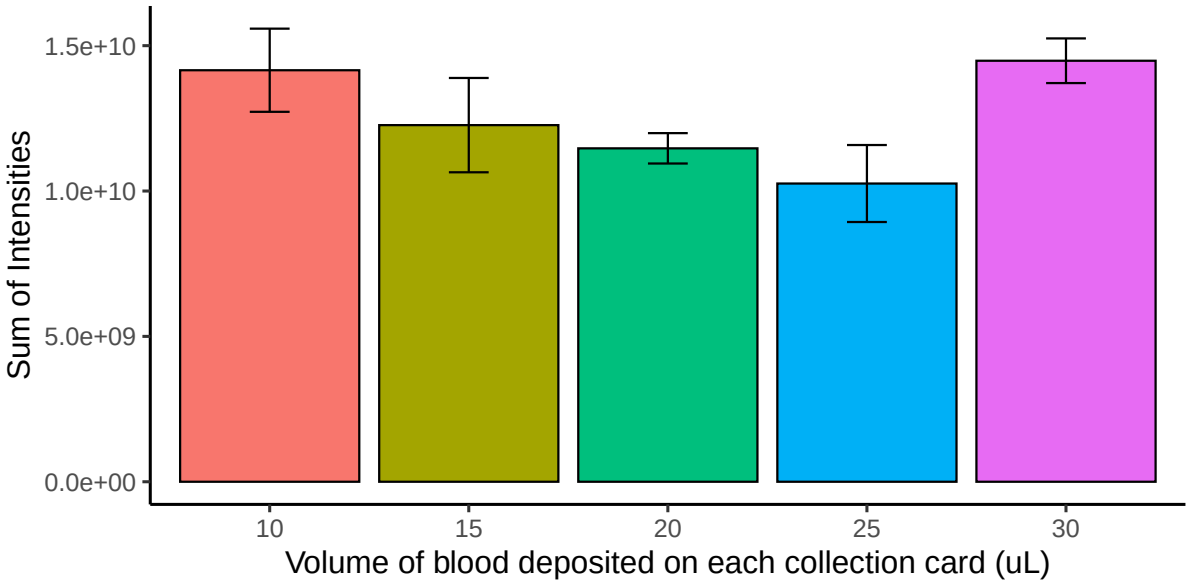
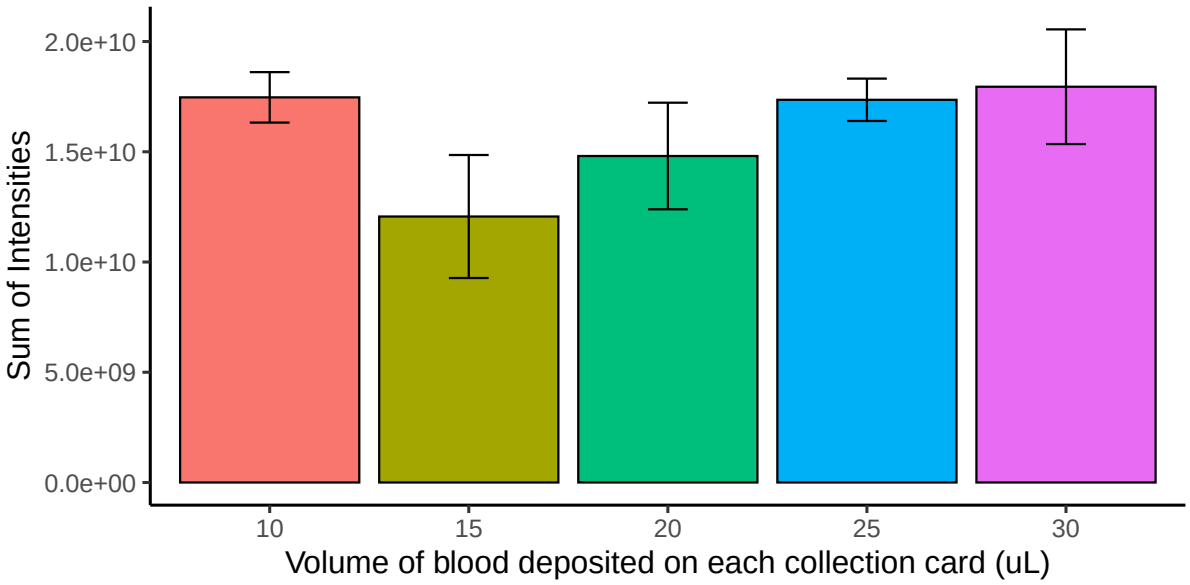


FIGURE G.3: RSD of all high quality features in the HILIC, negative ion mode assay before and after normalisation either pre or post data acquisition or normalising applying total ion current alone. Vertical line indicates a 30% RSD cut off.

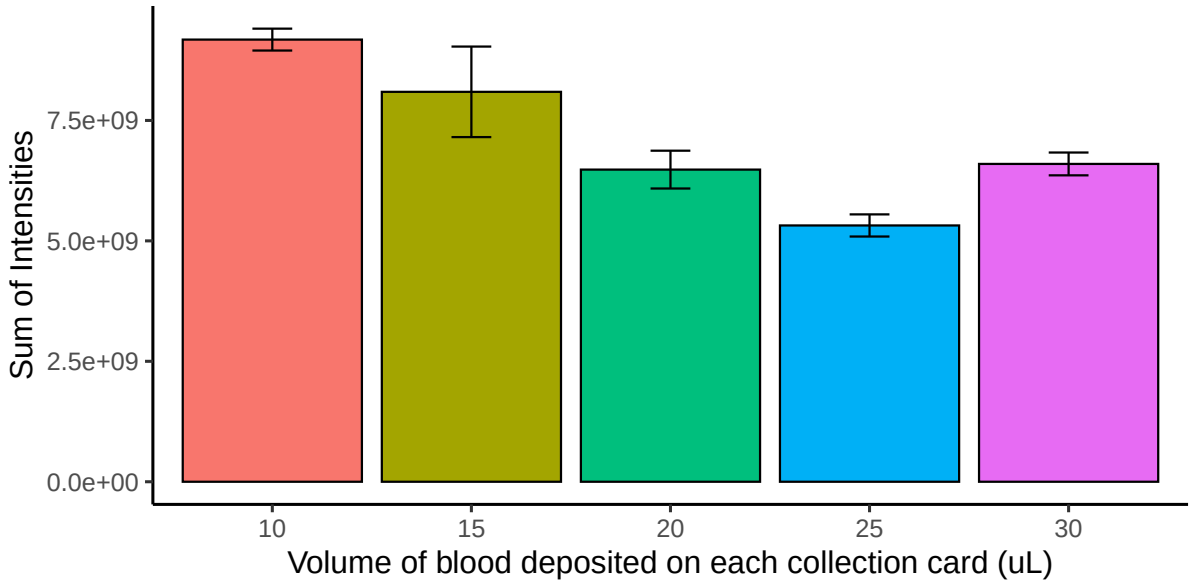


(A) Post-data acquisition

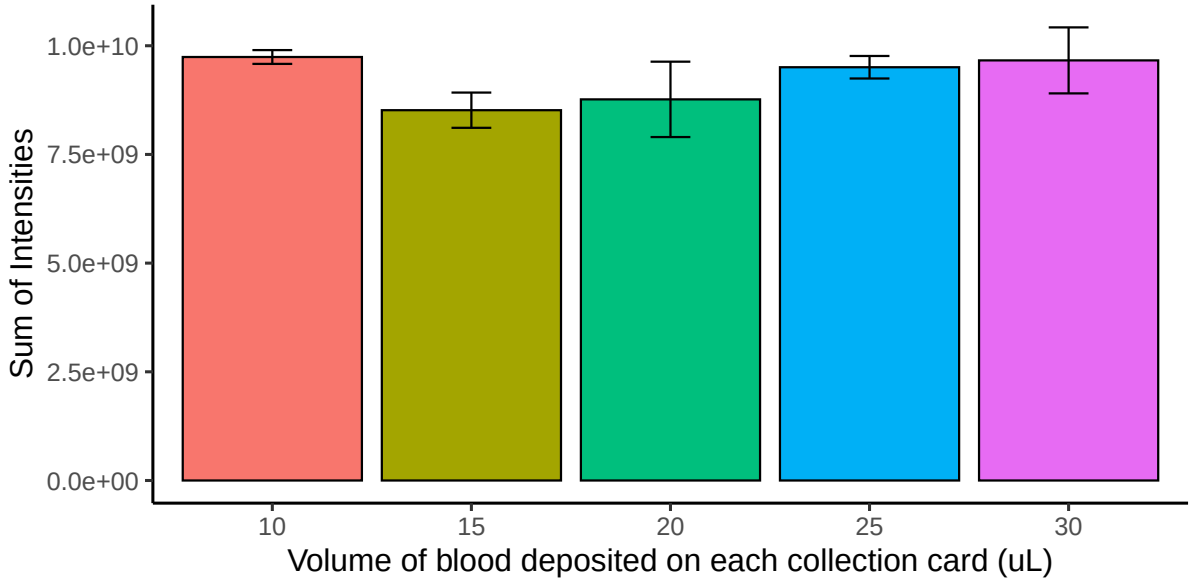


(B) Pre-data acquisition

FIGURE G.4: Sum of all intensities from features that are present in at least 75% of all replicates within every group for the Lipidomics negative ion mode assay. (A) Post-data acquisition normalisation. (B) Pre-data acquisition normalisation.

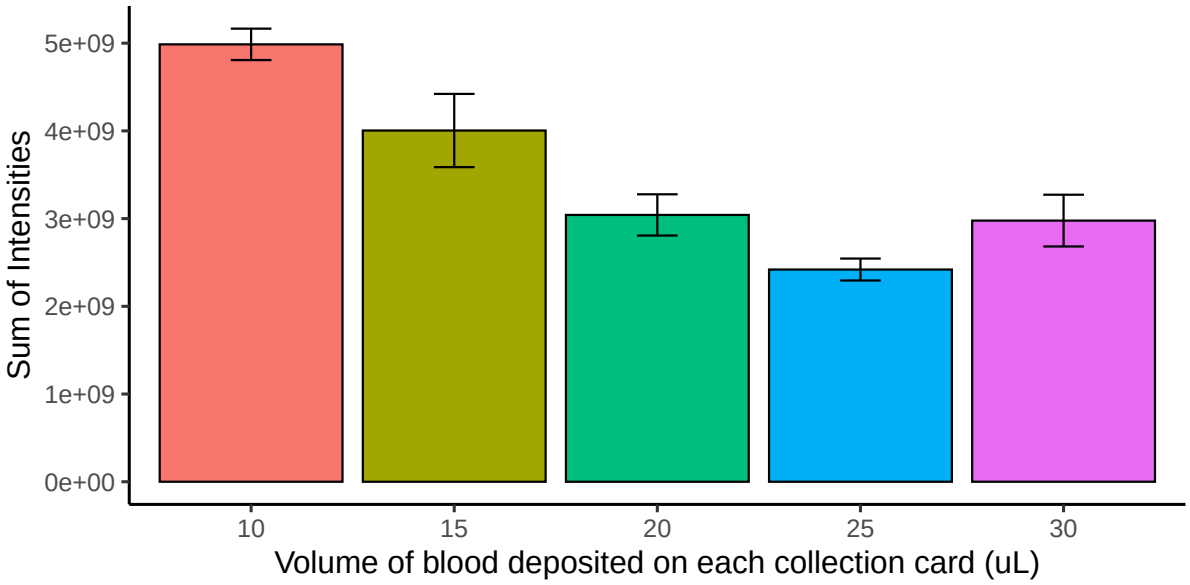


(A) Post-data acquisition

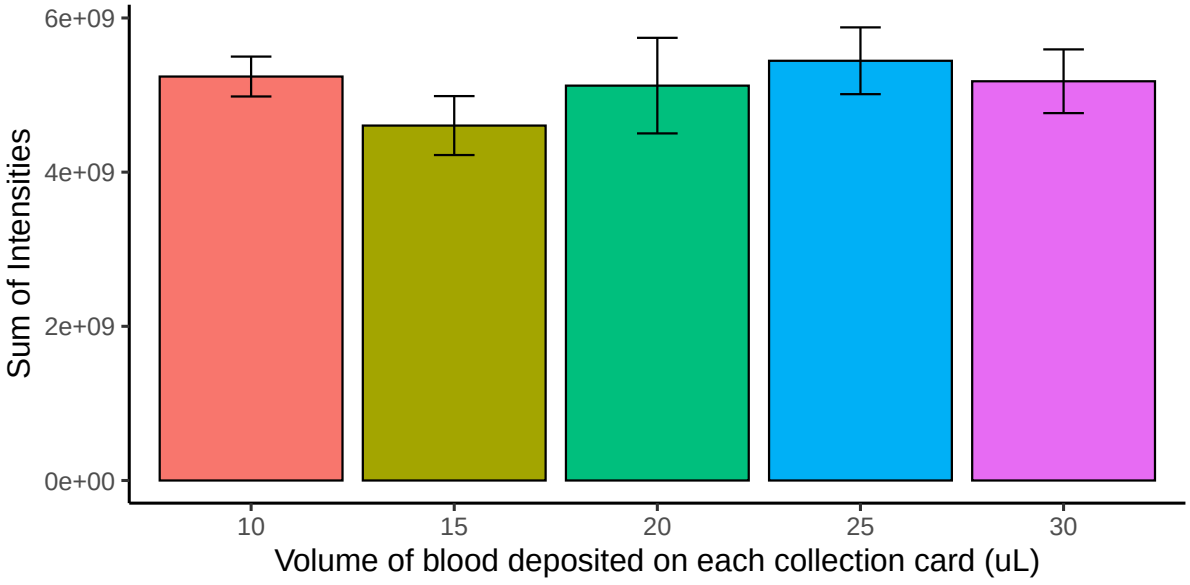


(B) Pre-data acquisition

FIGURE G.5: Sum of all intensities from features that are present in at least 75% of all replicates within every group for the HILIC positive ion mode assay. (A) Post-data acquisition normalisation. (B) Pre-data acquisition normalisation.



(A) Post-data acquisition



(B) Pre-data acquisition

FIGURE G.6: Sum of all intensities from features that are present in at least 75% of all replicates within every group for the HILIC negative ion mode assay. (A) Post-data acquisition normalisation. (B) Pre-data acquisition normalisation.