STRATHCLYDE INSTITUTE OF PHARMACY & BIOMEDICAL SCIENCES

University of Strathclyde Science

The Application of Metabolomics in Human Health and Disease

By

Ahmed Alshehri

Supervisors:

Dr. David G. Watson & Prof. Alex Mullen

A Thesis Submitted in Partial Fulfilment of the Requirement for the Award of a Degree of Doctor of Philosophy in the Strathclyde Institute of Pharmacy and Biomedical Sciences at Strathclyde University.

May, 2019

Declaration

‘This thesis is the result of the author’s original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.’

Signed:

Date:

Acknowledgements

Firstly, I would like to express my gratitude to my supervisors Dr. David Watson and Prof Alex Mullen for guiding me throughout this work. Your dedication, advice and support have been so important to me.

I would like to thank all my friends and colleagues in Dr Dave Watson's laboratory with whom I enjoyed some good times throughout this period. I also thank my family for the immense emotional support and prayers. May Allah bless you abundantly!

Finally, I thank the Saudi Cultural Bureau for the ongoing financial support towards my PhD studies.

Dedication

This work is dedicated to my family.

Table of Contents

DECLARATION II

ACKNOWLEDGEMENTS III

DEDICATION IV

LIST OF TABLES IX

LIST OF FIGURES X

ABBREVIATIONS AND ACRONYMS XIII

PUBLISHED WORK XV

ABSTRACT XVI

1 GENERAL INTRODUCTION 2

1.1 WHAT IS METABOLOMICS? 2

1.2 APPROACHES TO METABOLOMICS STUDIES 3

1.2.1 Targeted approaches 3

1.2.2 Untargeted approaches 4

1.3 ANALYTICAL TECHNIQUES 4

1.3.1 LC-MS 5

1.3.2 GC-MS 12

1.4 DATA PROCESSING 15

1.5 DATA ANALYSIS 18

1.6 DATA VISUALISATION 18

1.6.1 Unsupervised Techniques 18

1.6.2 Principal Component Analysis (PCA) 19

1.6.3 Hierarchical Clustering Analysis (HCA) 20

1.6.4 Supervised Techniques 20

1.7 MODEL VALIDATION 21

1.7.1 Cross validated ANOVA (CV-ANOVA) *23*

1.7.2 Biomarkers identification using an S-plot *23*

1.7.3 Variable importance in the projection (VIP) *23*

1.8 AIMS AND OBJECTIVES 24

1.8.1 Aim 1: *24*

1.8.2 Aim 2: *24*

1.8.3 Aim 3: *24*

1.8.4 Aim 4: *25*

2 MATERIALS AND METHODS 27

2.1 PARTICIPANTS AND STUDY SAMPLES 27

2.2 SOLVENTS AND CHEMICALS 27

2.3 INSTRUMENTAL TECHNIQUES AND COLUMNS 27

2.4 MOBILE PHASES 28

2.5 THE MS RUN CONDITIONS 28

2.6 DATA EXTRACTION AND ANALYSIS 29

2.7 OTHER EQUIPMENT USED 29

3 METABOLOMICS EFFECTS OF ULTRAMARATHON EXERCISE 32

3.1 INTRODUCTION 32

3.2 MATERIALS AND METHODS 34

3.2.1 Chemicals and Solvents *34*

3.2.2 Participants *35*

3.2.3 Plasma samples *35*

3.2.4 Sample preparation *36*

3.2.5 LC-MS conditions *36*

3.2.6 Data extraction and analysis *37*

3.3 RESULTS 38

3.3.1 Physiological response to the marathon *38*

3.3.2 Variation of metabolic profile with exercise *38*

3.3.3 Univariate comparisons *41*

3.4 DISCUSSION 45

3.5 CONCLUSIONS 54

4 EVALUATION OF THE METABOLOMICS EFFECTS OF *E. COLI* INCUBATION IN DIFFERENT CARBON SOURCES 56

4.1 INTRODUCTION 56

4.2 MATERIALS AND METHODS 59

4.2.1 Chemicals and Solvents *59*

4.2.2 Study samples *60*

4.2.3 Sample preparation *60*

4.2.4 LC-MS conditions *61*

4.2.5 Data extraction and analysis *62*

4.3 RESULTS 62

4.3.1 Unsupervised analysis *62*

4.3.2 Supervised analysis *63*

4.3.3 Metabolite alterations *69*

4.4 DISCUSSION 87

5 METABOLOMIC ANALYSIS OF THE EFFECTS OF DIFFERENT DIETARY FIBRES IN CROHN'S DISEASE 90

5.1 INTRODUCTION 90

5.2 MATERIALS AND METHODS 94

5.2.1 Chemicals and Solvents *94*

5.2.2 Participants *95*

5.2.3 Sample preparation *95*

5.2.4 LC-MS conditions *96*

5.2.5 Data extraction and analysis *96*

5.3 RESULTS 97

5.3.1 Unsupervised analysis 97

5.3.2 Supervised analysis 103

5.3.3 Metabolite alterations 108

5.4 DISCUSSION 124

5.4.1 Metabolites in Crohn's disease vs. healthy controls at all time points 125

5.4.2 Metabolites at 0 vs. 24/48 hrs in CD faecal incubations 126

5.4.3 Metabolites at 0 vs. 24/48 hrs in HC faecal incubations 128

5.5 CONCLUSION 129

6 PREDICTION OF CANCER ASSOCIATED MUSCLE WASTING FROM HUMAN PLASMA METABOLITES 132

6.1 INTRODUCTION 132

6.2 MATERIALS AND METHODS 134

6.2.1 Chemicals and Solvents 134

6.2.2 Participants 134

6.2.3 Sample collection and storage 134

6.2.4 Sample preparation 135

6.2.5 LC-MS conditions 135

6.2.6 Data extraction and analysis 136

6.3 RESULTS 136

6.4 DISCUSSION 144

6.5 CONCLUSION 151

7 GENERAL DISCUSSION AND CONCLUSION 153

8 REFERENCES 159

List of Tables

Table 3.1: *Participant biographic information and metadata 35*

Table 3.2: *Table showing all the significant metabolites affected by the marathon exercise.* Matches retention time of standard. ‡Data from runs on ACE C4 column. 41*

Table 4.1: *Description summary of the test sample groupings and control 60*

Table 4.2: Showing the significant metabolites between Control-G vs meat proteinB. 73

Table 4.3: Significant metabolite based on the comparisons of Control vs Fibre-M 79

Table 4.4: The significant metabolite for the comparisons of Control vs Fibre-O 84

Table 5.1: *The below table showing the comparisons of healthy patients with Crohn's disease (CD vs HC)..... 110*

Table 5.2: *Significant metabolite differences based on time points (0 hrs vs. 24+48 hrs) among the Crohn's disease (CD)* ***Table 5.2:*** *Significant metabolite differences based on time points (0 hrs vs. 24+48 hrs) among the Crohn's disease (CD)..... 114*

Table 6.1: *Patient details 137*

List of Figures

Figure 1.1: *A schematic diagram to illustrate the components of an HPLC system. 6*

Figure 1.2: A schematic diagram to show the main components of a mass spectrometer. 9

Figure 1.3: A schematic diagram representing the Orbitrap Mass Spectrometer. 11

Figure 1.4: A schematic diagram showing a GC-MS system. 14

Figure 3.1: *PCA separation of pre- 80K samples (C n=9) and post 80K (E n=8) samples based on polar metabolites analysed on a ZICpHILIC column. P= pooled samples. One post sample in the set is missing due to a technical failure. The data was Pareto scaled and log transformed. 39*

Figure 3.2: *PCA separation of pre- 80K samples (C n=9) and post 80K (E n=9) samples based on lipophilic metabolites analysed on an ACE C4 column. P= pooled samples. The data was Pareto scaled. 40*

Figure 3.3: *Hydroxy linoleic acids in a pre-80 K sample and a post 80K sample run on an ACE C4 column. 50*

Figure 3.4: *Heat map showing the relative abundance of the 30 most abundant fatty acids in plasma for the pre- and post-80K samples and two post-exercise samples. Red = highest value (3.93 x 10^7), Yellow = 1x 10^5 and blue = 5 x 10^3. ..... 50*

Figure 3.5: *Changes in the 40 most abundant acylcarnitines in plasma following an ultramarathon analysed by RP method. 53*

Figure 4.1: *PCA-X analysis of the metabolomics footprint of the 36 samples from E. coli cultures in different carbon sources. Circles coloured according to the time of sample collection post diet. 63*

Figure 4.2: *OPLS-DA analysis to compare B samples (Fibre = 1% cooked meat medium) with controls (G samples, representing 1% glucose). There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 9.9604e-017. ..... 64*

Figure 4.3: *Cross validation of the OPLS-DA model comparing B samples (Fibre = 1% cooked meat medium) with controls (G samples, representing 1% glucose). 65*

Figure 4.4: *OPLS-DA analysis to compare M samples (Fibre = 1% maize carbohydrate meal) with controls (G samples, representing 1% glucose). There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 6.26598e-011.* *66*

Figure 4.5: *Cross validation of the OPLS-DA model comparing M samples (Fibre = 1% maize carbohydrate meal) with controls (G samples, representing 1% glucose).* *67*

Figure 4.6: *OPLS-DA analysis to compare O samples (Fibre = 1% Olive kernel oil as C source) with controls (G samples, representing 1% glucose). There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 4.7657e-012.* *68*

Figure 4.7: *Cross validation of the OPLS-DA model comparing O samples (Fibre = 1% Olive kernel oil as C source) with controls (G samples, representing 1% glucose).* *69*

Figure 5.1: *PCA-X analysis of the metabolomics footprint of the 63 samples from healthy controls showing that time of sample collection plays a key role in sample classification according to the model. Circles coloured according to the time of sample collection post diet.* *98*

Figure 5.2: *PCA-X analysis of the metabolomics footprint of the 63 samples from HC participants showing that diet does not significantly affect sample clustering as each cluster contains all the 7 dietary fibres. Circles coloured according to the dietary fibres.* *99*

Figure 5.3: *PCA-X analysis of the metabolomics footprint of the 63 samples from CD patients in which both time of sample collection and individual patient peculiarities combine to define the model.* ... *100*

Figure 5.4: *PCA-X analysis of the metabolomics footprint of the 63 samples from CD patients. Each of the five natural groupings has representations from all 7 diets, implying that diet did not have much contribution to the scatter. As observed in Figure 3, time and individual patient peculiarities had more significant contribution to the observed clustering pattern.* *101*

Figure 5.5: *PCA-X analysis of all 126 samples showing near complete separation of CD and HC samples. There are two outliers in the model (sample 49 and sample 75) which are misclassified. Samples coloured according to whether they are CD or HC.* *102*

Figure 5.6: *OPLS-DA analysis to compare samples at 0h with those at 24 and 48 h in healthy controls. CA-ANOVA =1.9506e-029* *103*

Figure 5.7: *OPLS-DA analysis to compare samples at 0h with those at 24 and 48 h in CD patients. CV-ANOVA=3.08413e-035 ..... 104*

Figure 5.8: *Cross-validation of the model for healthy controls. 105*

Figure 5.9: *Cross-validation of the model for CD patients. 106*

Figure 5.10: *OPLS-DA analysis of the 126 samples showed clear separation between CD and HC groupings but with subgroupings based on time of collection (especially in CD samples) and possibly due to patient differences. 106*

Figure 6.1: *PCA-X analysis of the metabolomics footprint of the 18 plasma samples from cancer patients. Circles colored according to the weight loss categorization as weight stable or weight losing. The red circles represent pooled samples. Green circles (group 1) = weight stable. Blue circles (group 2) = weight losing. 138*

Figure 6.2: *PCA-X analysis of the metabolomics footprint of the 18 plasma samples from cancer patients with the pooled samples removed. Green circles (group 1) = weight stable. Blue circles (group 2) = weight losing 139*

Figure 6.3: *OPLS-DA analysis to compare weight losing samples with weight stable samples. There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 0.0477247 ..... 140*

Figure 6.4: *Cross validation of the OPLS-DA model comparing the weight losing samples with weight stable samples..... 141*

Figure 6.5: *Heat map showing relative levels of lysolipids purple= > 30%, yellow >1% blue >0.1% .. 145*

Abbreviations and Acronyms

1HNMR	Proton Nuclear Magnetic Resonance
ACN	Acetonitrile
ANOVA	Analysis of Variance
APCI	Atmospheric Pressude Chemical Ionisation
CV-ANOVA	Cross Validated Analysis of Variance
Da	Daltons
DAD	Diode Array Detector
DIMS	Direct Infusion Mass Spectrometry
ELSD	Evaporative Light Scattering Detection
ESI	Electrospray Ionisation
ESI	Electrospray Ionisation
FDR	False Discovery Rate
FT	Fourier Transformation
FT-IR	Fourier Transformation Infrared
GC	Gas Cromatogrepahy
GC-MS	Gas Chromatography-Mass Spectrometry
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
IR	Infrared

LC	Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
LTQ	Linear Trap Quadrupole
m/z	Mass to Charge Ratio
MALDI	Matrix Assisted Laser Desorption/Ionisation
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
NMR	Nuclear Magnetic Resonance
NP	Normal Phase
OPLS-DA	Orthogonal Projections to Latent Structures Discriminant Analysis
PCA-X	Principal Component Analysis
PDA	Photodiode Array detector
QToF	Quadrupole-Time of Flight
ROC	Receiver Operator Characteristic
RP	Reversed phase
SIMCA	Soft Independent Modelling by Class Analogies
ToF	Time of Flight
UV/Vis	Ultraviolet/Visible
VIP	Variable Importance in the Projection

Published Work

Christopher C. F. Howe, **Ahmed Alshehri**, David Muggeridge, Alexander B. Mullen, Marie Boyd, Owen Spendiff, Hannah J. Moir and David G. Watson. Untargeted Metabolomics Profiling of an 80.5 km Simulated Treadmill Ultramarathon, *Metabolites* 2018, **8**, 14.

Abstract

Metabolomics remains one of the rapidly growing tools for the identification of new disease diagnostic biomarkers. Mass spectrometry (MS) coupled to a liquid chromatographic (LC) system and nuclear magnetic resonance (NMR) are the two major analytical plafroms currently employed in metabolomic profiling studies of complex biofluid samples. However, due to its inherent higher sensitivity and fast data acquisition, MS remains one of the most dominant analytical techniques used in metabolomics.

In this project, metabolomics was employed in various studies to assess metabolite biomarkers associated with health and disease. All the studies employed liquid chromatography-mass spectrometry (LC-MS) on an Orbitrap Exactive mass analyser, and using ZIC-pHILIC or/and C18 analytical columns. Data was acquired using XCalibur software and metabolite identification was ascertained based on accurate mass detection, retention time comparisons with authentic external standards, and database searching. The acquired data was analysed using both unsupervised (PCA-X) and supervised (OPLS-DA) models in SIMCA in order to determine discriminating metabolite biomarkers responsible for the observed clustering patterns.

Investigation of metabolomic effects of an 80 km ultramarathon exercise among healthy volunteers on a treadmill gave clear separation between the pre- and post-80km samples. The study revealed that many of the amino acids were lowered in plasma post-exercise but the clearest impact of endurance exercise observed was on fatty acid metabolism with respect to formation of medium chain unsaturated and partially oxidised fatty acids and conjugates of fatty acids with carnitines, which suggested that exercise may have led to increased peroxisomal metabolism.

It is becoming increasingly clear that human health is strongly impacted by the gut microbiome. Evaluation of metabolomics effects of *E. coli* incubation *in vivo* with different carbon sources of 1% cooked meat, 1% maize meal and 1% olive kernel oil revealed that there were significant effects on amino acid, lipid, carbohydrate, and nucleotide metabolism. In addition, there were effects on intermediates of peptide and polyketide biosynthesis, as well as on xenobiotic breakdown products and vitamin cofactors. These findings suggested that the *E. coli* metabolome is closely associated with the type of fibre that the microorganism is exposed to and this was consistent with a number of other previous studies.

The study of metabolomic effects of dietary fibres on urinary metabolites from patients with Crohn's disease revealed that each of the 7 dietary fibres did not induce any significant differences in Crohn's disease patients relative to the controls. On the other hand, it was found that metabolites were affected by the time of sample collection post treatment, and the overall effect was that the levels of specific metabolites tended to increase post treatment. The most common pathways affected were

those of amino acid metabolism, lipid metabolism, nucleotide metabolism, polyketides, vitamins and cofactors, and xenobiotics, but the effect on carbohydrate metabolism was minimal.

Finally, the study to ascertain whether it was possible to predict cancer associated muscle wasting from plasma metabolites in patients with upper gastrointestinal cancer (oesophageal, gastric, pancreatic) revealed that the levels of significantly altered metabolites were generally higher in patients who had lost so much weight (>7.6 kg weight loss). The discriminating metabolites belonged mainly to the lipid metabolic pathways where long chain fatty acids and lysolipids were affected. The observed effects on lipid metabolisms in cancer cachexia suggests that there is an increased tendency towards peroxisomal proliferation in patients who had lost significant muscle mass.

Based on these findings, it can be concluded that LC-MS based metabolomics is a valuable tool in discriminating various disease states from the normal physiological state. Although the studies presented in this thesis considered vastly differing physical states ranging from healthy participants performing a simulated ultramarathon exercise on a treadmill to diseased participants suffering from either Crohn's disease or gastrointestinal cancer, the technique was capable of determining the metabolic alterations associated with each disease state. This further reinforces the capacity for metabolomics in discovering new biomarkers for various diseases that could be crucial in the diagnosis, monitoring disease progression, therapeutic efficacy evaluation of novel treatment, and detecting relapses following treatment.

Chapter One:

Introduction to Metabolomics

1 GENERAL INTRODUCTION

1.1 What is metabolomics?

Metabolomics is the comprehensive analysis of small molecule metabolites (Mol. Wt. <1,000 Daltons) in a given organism or a specific biological sample. It is considered one of the latest of the so-called "omic" technologies and is concerned with intermediates and products of metabolism which include amino acids, carbohydrates, nucleotides, fatty acids, organic acids, vitamins, antioxidants, pigments, among others. The complete set of metabolites synthesised within a particular organism form its "metabolome" in analogy with related terms such as "genome", "proteome" and "transcriptome". The growing popularity of metabolomics has coincided with the increasing need for better understanding of disease aetiology, particularly in the assessment of the influence of genetic and environmental factors on the disease state, and in the elucidation of individualised therapeutic interventions (Nicholson, 2006, Holmes *et al*., 2008).

Thus metabolomics is important in the understanding of effects of new therapeutic interventions in certain diseases, elucidating the impact of exercise regimens or food intake on the body's metabolism, determining the mechanisms of action of new drug molecules, identifying biomarkers for detection and diagnosis of disease states, etc (Wishart, 2008). However, although the field of metabolomics is considered to be relatively new in systems biology, the first reported application of metabolic studies were reported in ancient China, where ants were employed in the detection of diabetes based on glucose

levels in urine (van der Greef and Smilde, 2005). In addition, "urine charts" for correlating the smell, taste, and color of urine were employed in the Middle Ages to diagnose metabolic diseases (Nicholson and Lindon, 2008). However, it was not until recently, in the 1940's, that the idea that individuals might have distinctive "metabolic profiles" that could be "fingerprinted" was proposed by Williams *et al*. in the 1940s (Gates and Sweeley, 1978).

1.2 Approaches to metabolomics studies

The metabolic alterations associated with a given disease state or treatment can be studied using targeted, semi-targeted or untargeted metabolomic approaches (Dunn *et al*., 2013). The choice between any of these approaches depends on the need for quantitatification, expected levels of precision and accuracy, sample complexity, and the number of metabolites involved.

1.2.1 Targeted approaches in metabolomics studies

Targeted metabolomic approaches deal with a few known metabolites which allows the analytical techniques to be optimised for high precision, accuracy and selectivity. For this reason, targeted metabolomics is mainly quantitative. Conversely, semi-targeted or untargeted approaches are less quantitative but their main aim is to identify all the metabolites detected in a given sample for hyphothesis generation. For this reason, targeted approaches utilise hypotheses generated from untargeted or semi-targeted studies to obtain conclusive evidence on the observed biological phenomena.

1.2.2 Untargeted approaches in metabolomics studies

Untargeted metabolomics deals with detection of thousands of metabolites in a given set of samples with limited or no prior knowledge of the expected metabolite profiles. The observations made from these studies enable the investigator to derive hypotheses for testing at a later stage during targeted analyses. As there are normally a lot of metabolites involved, untargeted methods normally are less quantitative and in most cases some of the metabolites cannot be easily identified.

1.3 Analytical techniques

Metabolites constitute a diverse group of chemical compounds of varying molecular weights and functional groups. Although their analysis may be performed using the same techniques employed in routine chemical analyses, there is need for higher sensitivity and selectivity to enable identification of individual metabolites in complex mixtures (Dunn and Ellis, 2005). The most commonly used techniques include those based on chromatographic separation such as liquid chromatography (LC), gas chromatography (GC), and capillary electrophoresis (CE). These techniques can be coupled to suitable detectors and the commonest of these are mass spectrometers (MS) (Katajamaa and Orešič, 2007). Thus, LC-MS and GC-MS are some of the commonly encountered analytical platforms in metabolomic profiling studies. The MS coupled chromatographic systems enable the detection of hundreds of metabolites in a single run (Budczies *et al*., 2012, Dettmer *et al*., 2007, Takahashi *et al*., 2011, Johnson *et al*., 2003). Other analytical plaforms previously employed in metabolomics include Fourier transform infrared (FT-IR)

spectroscopy (Kim *et al*., 2010), nuclear magnetic resonance (NMR) spectroscopy (Serkova *et al*., 2005), and direct infusion mass spectrometry (DIMS) (Kaderbhai *et al*., 2003). Certainly, the choice of technique depends on sample type/complexity and the required levels of selectivity, sensitivity, accuracy and speed of the analysis.

1.3.1 LC-MS

Chromatography is a technique of physical separation of compounds in a mixture based on differential affinities of the analytes for two phases, namely, the mobile phase and the stationary phase (Watson, 2012). Chromatographic techniques can be classified based on the stationary and mobile phases used. In liquid chromatography (LC), the mo-lie phase is a liquid, while in gas chromatography (GC), the mobile phase is a gas. LC can be further categorized as reversed phase (RP), normal phase (NP), and hydrophilic interaction liquid chromatography (HILIC) depending on the nature of the stationary phase. Generally, RP is the commonest of the three LC techniques used because of its wider applicability. It utilizes a hydrophobic stationary phase (for example, a C18) and a hydrophilic mobile phase in which water is mixed with a miscible organic modifier such as methanol, ethanol, acetonitrile, or tetrahydrofuran (Harris, 2010). Despite its populalrity, RP is not well suited for highly polar metabolites that are normally encountered in metabolomics because these are not well retained in non polar stationary phases. For this reason, there has been a growing trend in the use of HILIC-based techniques in metabolomic studies in recent years. HILIC's retention properties are orthogonal to

those of RP, so that it shows higher retention of hydrophilic metabolites and low retention for hydrophobic ones.

The modern LC-MS system consists a high performance liquid chromatograph (HPLC) connected to a mass spectrometer (MS). The main parts of a HPLC itself include a solvent reservoir, an online degasser, a pump, an autosampler, a column compartment, and a suitable detection system (Figure 1.1). The most commonly used detectors in HPLC are ultraviolet (UV) detectors, diode array detectors (DAD), evaporative light scattering detectors (ELSD), and mass spectrometers (Harris, 2010). Each of these detectors are associated with certain strengths and limitations. For instance, UV detectors are unable to detect compounds lacking chromophores while ELSD has limited quantitative capacity due to nonlinearity of the detector response signal with increasing concentration. The mass spectrometer in combination with a HPLC offers a very powerful and reliable analytical platform for metabolomics studies (Watson, 2012).

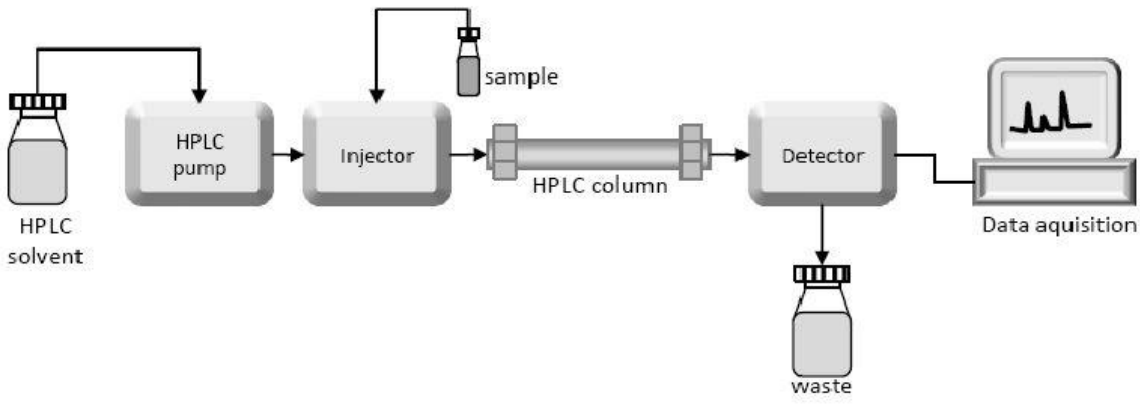

Figure 1.1: A schematic diagram to illustrate the components of an HPLC system. The figure was been accessed from https://laboratoryinfo.com/hplc/ on 05 February 2019.

Although reversed phase chromatography is the most popular in pharmaceutical analysis, both in industry and quality control laboratories, it is not very popular in metabolomics profiling studies due to the fact that it does not retain most metabolites. Instead, metabolomics analysis employs stationary phases capable of separation by HILIC since these are designed to retain polar compounds that constitute most metabolomics metabolites. Various HILIC columns are currently available with varying chemistries including bare silica gel, silicon hydride, and derivatised phases such as phenylhydride, but zwitterionic phases such as ZICHILIC columns have also been popular in metabolomics profiling studies due to their high reproducibility (Zhang *et al*., 2014). In general HILIC separations employ reversed phase type mobile phases with normal phase columns; thus water, rather than the organic component such as acetonitrile, is the stronger solvent in HILIC unlike in reversed phase (Watson, 2012).

The mechanisms of separation of analytes in HILIC are not fully understood but it is thought that partitioning into a stationary layer of water established on the surface of the stationary phase plays an essential role in achieving resolution between analytes (Santali *et al*., 2014, Bawazeer *et al*., 2012). However, other mechanisms have also been described which include ionic interactions, dipole-dipole interactions, van der Waals

forces, etc. The number of possible interactions in a HILIC column depends on the chemistry. For instance, ZIC-HILIC columns contain oppositely charged groups (zwitterions) which can provide additional sites of interactions through repulsion (similar charges) or attraction (opposite charges) of any charged analytes thus affecting their retention on the column. The strength of these ionic interactions can be modified in the presence of competing ions from mobile phase additives such as buffers which might modulate the retention behaviour. On the other hand, the mobile phase additives can modify the thickness of the layer on the stationary phase surface thus affecting the strength of partitioning and subsequently its retentivity. In general, increase in the ionic strength of the salts increases the thickness of the water layer depending on the hydration energies of the counter ions in the mobile phase. Thus, analyte separation is based on differences in analyte polarity, molecular weight, shape, and charge.

The role of the mass spectrometer as a detector in LC-MS is to measure the mass-to-charge ratio (m/z) of the analytes present in a sample. Mass spectrometry is the technique of choice in metabolomics studies and as such it is the most commonly used tool (Katajamaa and Orešič, 2007). The MS is superior to other common methods of detection such as UV, ELSD, fluorimetry, and NMR due to a combination of its high sensitivity, selectivity, resolution, and ability to give accurate mass data depending on the instrument used. The mass spectrometer (MS) has three main components: an ionisation chamber, a mass analyser, and a detector (Figure 2). The ion source is used to produce gas phase ions from sample. The ionization processes used in MS vary in their techniques and they

include those which operate under vacuum such as electron impact (EI) and chemical ionization (CI), and those which operate at atmospheric pressure such as electrospray ionisation (ESI) and atmospheric pressure chemical ionization (APCI) (Kraj *et al.*, 2008, Watson and Sparkman, 2007). The ions produced by the ion source are accelerated through a region of electric and magnetic fields so that only those ions with m/z in a given range can reach the analyser and be detected.

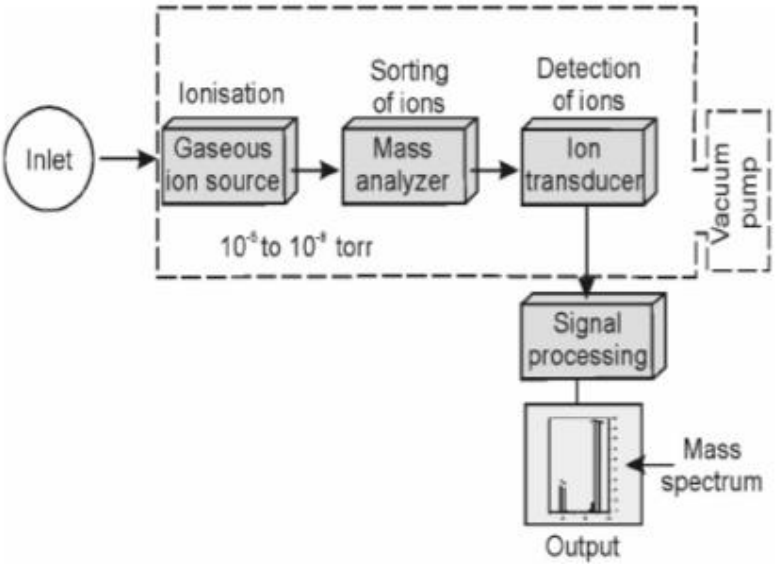

Figure 1.2: A schematic diagram to show the main components of a mass spectrometer. The figure has been accessed from Bijaya U. Kumar (2014) available at https://www.slideshare.net/bijayauprety/mass-spectrometry-41214136 on 05 February 2019.

The second main component of MS instrument is the analyzer, where the ions are separated based on their mass-to-charge (m/z) ratios. There are different mass analyzers

which are currently used in MS systems to separate the ions in time or space and they include quadrupoles (Q), ion traps (IT), time of flight (TOF), Fourier transform ion cyclotron resonance (FT-ICR), and the Orbitrap analyser (Hu *et al*., 2005). The main differences between these mass analyzers arise from their resolving power, mass accuracy, sensitivity, dynamic range and fragmentation capabilities for MS^n studies. Recent developments have led to hybrid mass systems that combine strengths of various mass analysers so that a single mass spectrometer can have various capabilities based on the ion separation techniques being encompassed in the hybrid system. Examples of such hybridized mass analysers include triple-Q, Q-IT, TOF-TOF, Q-TOF, IT-Orbitraps, LTQ-Orbitraps and Q-Exactives (Michalski *et al*., 2011). The current MS systems were made suitable for application to metabolomics by the addition of soft ionization techniques such as APCI or ESI which form mainly molecular ions without fragmentation, allowing the compounds to be identified based on their databases constructed specifically using accurate mass data of the common metabolites (Watson and Sparkman, 2007, Kraj *et al*., 2008).

The third part of MS is the detector. In this region, the mass-to-charge ratios (m/z) of the detected ions and their abundances are measured. In the Orbitrap, for example, detection is based on image current of the ions in the mass analyser (Makarov and Scigelova, 2010). The conversion of the image current into the mass spectrum utilizes mathematical algorithms such as Fourier transformation (FT) which is also employed in other FT instruments such as FT-ICR (Michalski *et al*., 2012).

An ideal MS system should be able to detect very low concentrations of a given analyte in a sample (sensitivity), achieve high mass resolution between very closely related masses (selectivity), provide high mass accuracy, and should have a high dynamic scan range. These attributes can all be achieved in some modern mass spectrometers such as the Orbitraps and Orbitrap Exactives (Hu *et al*., 2005) (Figure 3). Additionally, mass spectrometers allow the analyst to tailor the conditions of the analysis to the specific analytes in the sample which improves the robustness of the detection method. Generally, LCMS systems have better sensitivity in the analysis of metabolites than GCMS without need for prior derivatisation. The capabilities of LCMS can be expanded through MS/MS studies and high-energy C-trap dissociation (HCD) on the Orbitrap to enhance parent ion characterisation and to elucidate the structures of fragment from analyte breakdown (Kamleh *et al*., 2009a, Holčapek *et al*., 2012).

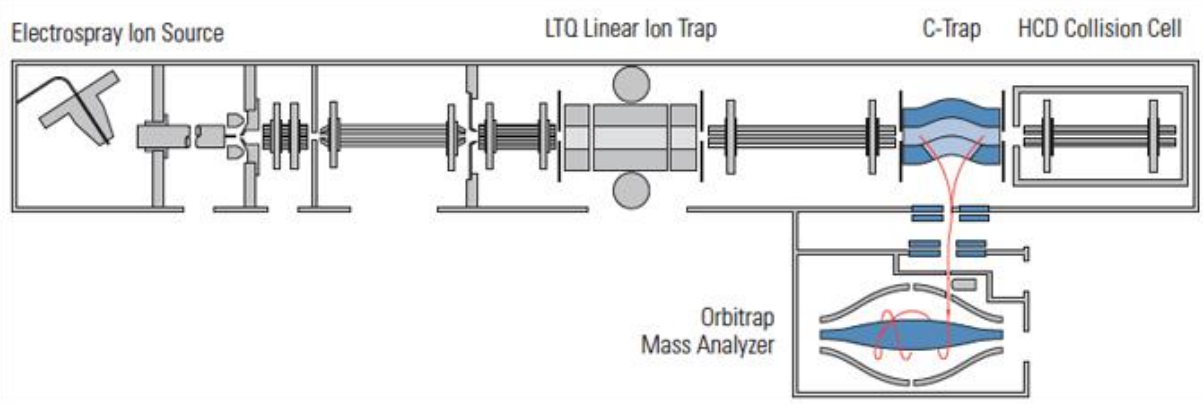

Figure 1.3: A schematic diagram representing the Orbitrap Mass Spectrometer. The figure has been obtained from Hu *et al*., 2005. The Orbitrap: a new mass spectrometer. *J Mass Spectr*, 40, 430-443.

Recent advancements in mass analyser technologies have transformed mass spectrometers into essential analytical tools for biological research. The high sensitivities and mass accuracy of Orbitrap analysers, for example, has increased data richness and the value of metabolomics data that can be achieved from the analysis of a single sample compared with earlier technologies (Hu *et al*., 2005). Currently, there are different generations of Orbitrap mass analysers of which the LTQ Orbitrap is the first generation first commercialized around 2005 (Makarov *et al*., 2006). This MS combines a linear ion trap (LTQ) and a Fourier transform mass analyser (Orbitrap) which gives it capabilities for MS^n studies in addition to the enhanced sensitivity, mass accuracy and resolution of the Orbitrap technology. The high mass accuracy and resolving power of Orbitrap mass analysers permit accurate mass measurements which are essential for the determination of elemental composition of metabolites, which facilitates their identification (Makarov, 2000, Makarov *et al*., 2006, Hu *et al.*, 2005).

1.3.2 GC-MS

Gas chromatography (GC) uses an inert gas as mobile phase, hence its name. The most commonly used gases in GC include nitrogen, helium, and hydrogen. The use of gas as the mobile phase is the key difference between the GC and LC as the latter uses liquid. The columns used in modern GC systems are long capillary columns which contain stationary phases coated onto the internal wall. The stationary phase coatings vary in their polarity, polar ones include carbowax phases while methyl silicone phases are less polar. There is a range of commercial GC columns available whose stationary phases vary in

terms of their polarity and suitability for the analysis of different analytes. Samples can be injected in split or splitless mode depending on their concentration; the splitless mode is preferable for samples containing very low levels of the metabolites being analysed (Watson, 2012). The main compartments of the GC such as column and injection port are maintained at high temperatures in an oven maintained to keep analytes in the gas phase. The temperature gradient of the column compartment is employed to modify retention times of the analytes. Apart from temperature, retention times depend also on the molecular weight and polarity of the compound which in turn affect its volatility.

A GC-MS system consists of a GC with a MS as the detector, as illustrated in Figure 4. Different types of mass spectrometers can be interfaced to the GC but since analytes entering the MS are in gas phase, only MS ion sources that are capable of dealing with gas-phase analytes are employed. These ion sources include chemical ionisation (CI) and electron impact (EI) ionisation. The latter uses high collisional energy (70eV) with fast moving electrons to ionize analytes, which results in extensive fragmentation of the compound that can facilitate identification procedures. GC-MS can be carried out with quadrupole and time of flight (TOF) mass analysers. GC-quadrupole systems have high dynamic range and sensitivity but mass accuracy and scan speed are quite low. On the other hand, GC-TOF/MS has higher mass resolution and mass accuracy (Bedair and Sumner, 2008).

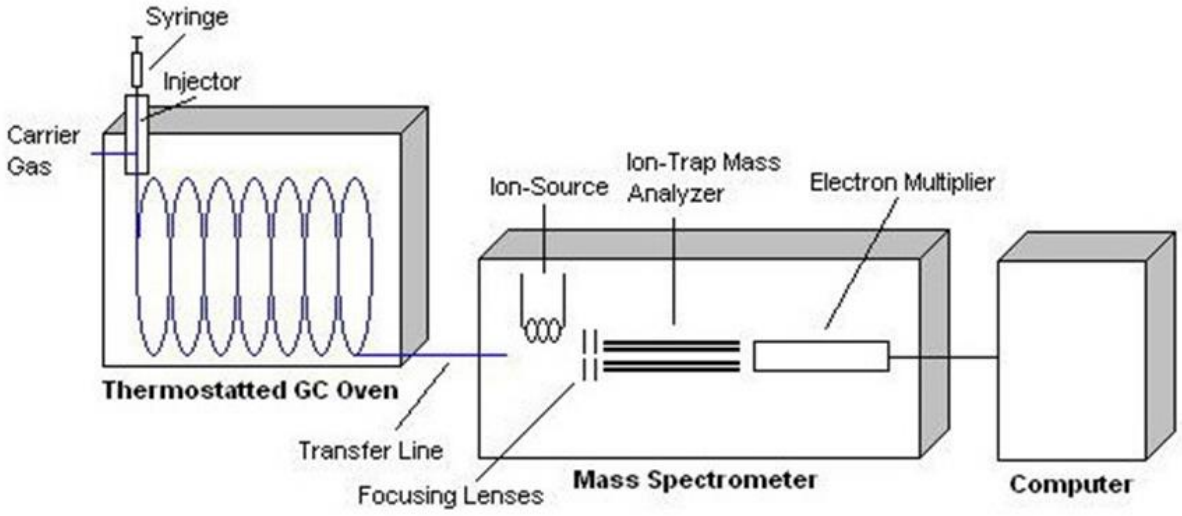

Figure 1.4: A schematic diagram showing a GC-MS system. The figure was accessed on 05 February 2019 at https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_(Analytical_Chemistry)/Instrumental_Analysis/Chromatography/Gas_Chromatography.

GC-MS is one of the preferred techniques applied in metabolomics research because it combines the high separation efficiency (chromatographic resolution) of a capillary GC column and the high sensitivity and robustness of the mass spectrometer (Kopka, 2006). The availability of GC-MS spectral libraries also makes the job of metabolite identification easier. Compared to LC-MS, GC-MS has a limited application because of the need for the samples to be volatile and thermally stable. For some non-volatile analytes such as fatty acids, volatility can be achieved by derivatisation, for instance fatty acids can be derivatised through methylation to form esters which are volatile (Schauer *et al*., 2005, Kopka, 2006). Another common method for derivatisation is by oxime/silylation derivatisation.

Silylation introduces a trimethylsilyl (TMS) group onto the non-volatile compound resulting into volatility and the method can be applied to alcohols, amides, amino acids and thiols (Roessner *et al*., 2000, Dunn *et al*., 2005). The TMS derivatives are less polar than the parent metabolites and have less dipole-dipole interactions which increase their volatility that is suitable for analysis by GC-MS (Dunn *et al*., 2005). It should be noted that the derivatisation process can be time-consuming and the additional sample preparation might introduce extra technical errors into the experiment thus increasing the total variability in the samples. In the case of thermal stability, except small molecular weight hydrocarbons, short chain alcohol and esters, most metabolites are affected by the high temperatures employed in GC-MS which can be as high as 350°C (Bedair and Sumner, 2008).

1.4 Data processing

During metabolomic studies, huge amounts of complex data are generated even for a few samples, and this requires statistical software for analysis before interpretation. The processing of metabolomics data enables the extraction of all relevant information about the analytes present in a sample and at the same time it minimizes background noise in order to facilitate subsequent data analysis and interpretation. Depending on the techniques employed, such as GC-MS, LC-MS, direct infusion MS, and NMR (Cui *et al*., 2008; Kuhn *et al* al., 2008), data processing commences with its acquisition from the analytical instrument. For example, the data resulting from LC-MS analysis as collected by XCalibur software requires pre-processing prior to multivariate statistical analysis (MVA). This

pre-processing is intended to remove background ions and to correct for retention time drifts.

In general, there are four considerations which should be made during data processing of LC-MS metabolomics data. To begin with, the LC/MS technique captures data of the whole metabolome thus it contains a large number of diverse metabolites varying in their molecular structures and concentrations in the sample. Secondly, data processing prior to MVA can be a significant source of errors that may lead to false signals being produced by peak picking, alignment, and noise. Thus, these parameters should to be optimised for a given set of samples during pre-processing in order to obtain the most accurate results from MVA (Dettmer *et al*., 2007).

Thirdly, LC-MS can generate adduct and fragment ions that may be generated by a single metabolite during the process of ionization in the ESI. This signal redundancy interferes with the process of metabolite identification by increasing the number of variables that can produce a false positive in MVA. Another matter of concern is that the identification step of acquired signals in metabolomic experiments consumes time and effort. It is usually done by matching databases with the filtered mass spectra. Further experiments such as MS/MS fragmentation are needed to rigorously confirm the results (Werner *et al*., 2008).

Broadly, two methods for extracting information from metabolomics data which are currently employed are chemometric (or non-targeted) and quantitative (or targeted) methods. Chemometrics is defined as “the application of mathematical, statistical, graphical or symbolic methods to maximize the information which can be extracted from chemical or spectral data”. It can be used to distinguish between sample categories under different conditions by reducing the number of dimensions to simplify the dataset display significant differences between the samples being analysed. It is useful for identifying phenotypes and drawing conclusions without the need to identify and quantify the specific metabolites. Indeed, the chemometric analysis approach presents two most widely utilized approaches for pattern recognition, namely, principal components analysis (PCA) and partial least squares discrimination analysis (PLS-DA). PCA allows the clustering tendency to be easily detected by visualization in an unsupervised manner. The chemometric tools for metabolomics data processing must be selected based on the purpose of the study. If the purpose is sample discrimination, and prior information on sample identity is unknown, then unsupervised methods such as PCA are used. Supervised methods such as PLS may be used when the class of some samples is known. Contrary to the chemometric methods, the quantitative methods require the metabolites in a given biological sample to be identified and quantified prior to analysis. These methods can be applied to the data as acquired by NMR or MS via available reference databases to identify their signals. Chemometric approaches may be less applicable to LC-MS data because the

number of peaks present in a sample run can be > 20,000, and because the alignment of chromatograms drifts from run to run (Trygg *et al*., 2007, Werner *et al*., 2008).

1.5 Data analysis

Metabolomics data are highly rich in information which requires processing by software or web-based tools in order to facilitate the interpretation. Commonly used software for this purpose include: MZMatch, MZMine and SIEVE. Both MZMatch and MZMine are non-commercial open-source software that allows for peak detection, filtering, normalization and identification based on local and online databases. One key drawback of MZMatch is the need for the user to possess some knowledge of computer programming before proper use of this software (Pluskal *et al*., 2010). MzMine was first introduced in 2005 for the processing of mass spectrometry based profile data (Katajamaa *et al*., 2006). It has been applied to various metabolomic analyses (Guan *et al*., 2009, Macintyre *et al*., 2014, Muhsen Ali *et al*., 2016).

1.6 Data visualisation

1.6.1 Unsupervised Techniques

Due to the nature of the information contained in biological data sets (such as metabolomics data), LC-MS can generate very large amounts of data. As in this research, it is required to establish possible relationships (or correlations) among the various subjects or variables; the greater the amount of information there is to analyse, the higher the difficulty and complexity of obtaining the required results will be. It would be almost

impossible to properly examine and analyse the data without appropriate statistical software. Hence, it is necessary to apply suitable statistical methods to increase the chance of identifying any potential similarities or differences among the various samples in the data, by reducing the dimensionality of the input space of the data to a small number of dimensions.

To classify the samples into groups of similar characteristics, which can give an insight in the situation under investigation, statistical methods such as Principal Components Analysis (PCA) and Cluster Analysis such as Hierarchical Cluster Analysis (HCA) can be used. Samples classified in a group will have similar characteristics, but be different from those in other groups. No information about the groups is known beforehand and no assumptions are necessary concerning the group into which a sample may be classified. These unsupervised pattern recognition techniques aim to reduce the amount of data complexity and afterwards present in a graphical form the patterns or clusters identified in the data (Prelorendjos, 2014).

1.6.2 Principal Component Analysis (PCA)

This is an unsupervised model employed to explore how variables cluster regardless to which class an observation is belongs to (Kirwan *et al*., 2012). It considered the main tool used by analysts for data reduction to extract meaningful information (Yamamoto *et al*., 2009). This achieved by combining variables that correlate with each other into few latent variables (components). The higher the correlation among variables is the smaller

the number of components that will be needed with components < observations, without losing an important amount of the total variation of the data (Prelorendjos, 2014). PCA is normally employed as the first step in the analysis of metabolomics data (Kirwan *et al*., 2012, Trygg *et al*., 2007) in order to visualize data and detect outliers.

1.6.3 Hierarchical Clustering Analysis (HCA)

The concept of HCA or dendrogram –both are used interchangeably- as a clustering analysis tool is to try to find a natural grouping of a data set, so that there is high similarity (low variability) of observations within clustered groups and less similarity (high variability) of observations between clustered groups. In HCA clustering, the two closest clusters or observations are merged, thereafter the two closest clusters or points are again merged, etc, until one super cluster remains (Lozano *et al*., 2014). HCA is extensively used when a study is done with no previous knowledge about grouping, and is considered a preface for supervised multivariate techniques.

1.6.4 Supervised Techniques

PCA provides an overview of the dataset but it does not relate the phenotype-disease state for instance- of an individual to the measured parameters. Partial least squares-discriminant analysis (PLS-DA) performs a PCA analysis on the Y-matrix (observations/samples) to yield a small number of latent variables, and then constructs a series of latent variables from the X-matrix (descriptors/variables/metabolites) which explain the maximum variance in these Y latent variables.

Orthogonal partial least squares - discriminanat analysis (OPLS-DA) is an extension of PLS-DA modelling, and has an advantage over the PLS-DA that it can separate variation in *X* that correlates to *Y* (horizontally) called predictive variation, and also separate variation in *X* that is uncorrelated to *Y* (orthogonal). OPLS-DA is most powerful technique that is employed to examine the difference between groups (Kirwan *et al*., 2012), it can identify reliable biomarkers that have a strong association with separation between groups (Trygg *et al*., 2007) and relate disease to perturbations in metabolic pathways (Goodacre, 2007) and thus help expanding our understanding of the pathophysiology and of future therapeutic targets.

The quality of a supervised model is assessed by R^2 (the goodness of fit) and Q^2 (the goodness of prediction), and *P* CV-ANOVA (the p-value of the model) from cross-validation procedures which determine the degree of significance of the model (Triba *et al*., 2015) and are called quality parameters (Wheelock and Wheelock, 2013).

1.7 Model validation

During analysis, the quality parameters R^2 and Q^2 are the most powerful tools for validating any applied model. R^2 is a quantitative measure of the goodness of fit, it relates *y* (observations) to *x* (variables), by quantifying the fraction of *y* (observations) explained by the variation in *x* (variables). The issue with such a parameter is that it can be made arbitrarily close to one, the maximal value, as long as we increase the number of components. This might lead to over-fitting the data due to the large number of variables compared to small number of observations and thus give too optimistic results. However,

this can be controlled by the goodness of prediction parameter Q^2, it is obtained via cross validation (CV) (Kirwan *et al*., 2012) by which predefined number of observations should be left out and followed by refitting the model, this process applied to all the data until all have been kept out only once (Eriksson *et al*., 2013a). Then the average value of the refitted models Q^2 are compared to the R^2 of that model which provides an indication that it predicts much better than chance.

For the purpose of cross validation SIMCA P software - by default - leaves $1/7^{th}$ of the data out. An "observed" vs "predicted" plot is employed to examine the efficiency of CV, by which the R^2 of the regression line should be improved. Moreover, in order to evaluate whether the specific grouping of the observations in the two designed classes is significantly better than any other random grouping in two arbitrary classes, permutations test applied (Westerhuis *et al*., 2008).

In this test, the R^2 and Q^2 parameters obtained from the original model are compared to newly permuted R^2 and Q^2, this process can be repeated to generate new quality parameters. The new parameters generated from this permutation should all be lower in value than the original values. In addition to that, the regression line of the predictive model should cross the horizontal zero line (Eriksson *et al*., 2013b). In order to test the significance of the variation predicted by the supervised model, ANOVA of the cross validated residuals is employed (CV-ANOVA). Once the predictive ability of the model is validated, then the accuracy of the model in discriminating observations based on their metabolic

profile should be assessed and reported using area under the receiver operating characteristic (ROC) curve.

1.7.1 Cross validated ANOVA (CV-ANOVA)

The validity of a supervised model can be well assessed using cross validated ANOVA (CV-ANOVA) which tests the variation predicted by the model against the null hypothesis (H_0) of equal cross validated predictive residuals around the mean (Eriksson *et al*., 2008).

1.7.2 Biomarkers identification using an S-plot

The S-plot is a tool used to identify biomarkers based on a supervised model. The metabolites in the upper right and lower left corners of the plot are highly associated with the differences between the two groups being considered. However, there are no clear cut-offs that can be relied upon when selecting metabolites using an S-plot, which may lead to some of the significant metabolites being inadvertently neglected. In order to overcome this shortcoming, it would be appropriate to employ univariate analysis so that all metabolites are afforded equal chance of selection without losing potential biomarkers.

1.7.3 Variable importance in the projection (VIP)

The contribution a given metabolite in a model is examined by considering its variable importance in the projection (VIP). This parameter estimates and ranks the importance of each variable (metabolite) in the projection and it is often used for variable selection during metabolomics (Chong and Jun, 2005). Metabolites are generally considered to

have a high contribution in the model if VIP > 1 (Eriksson *et al*., 2013d, Zhang *et al*., 2016). VIP provides two VIP values for each metabolite one denoting its contribution to the difference between groups (VIPpred) and the second denoting its contribution to the within group variability (VIPortho). Any metabolite with high VIPpred and low VIPortho values is sensitive and specific.

1.8 Aims and Objectives

The project was aimed at addressing four major objectives as follows:

1.8.1 Aim 1:

To investigate the metabolomic effects of an 80 km ultramarathon exercise simulated on a treadmill in healthy adults.

1.8.2 Aim 2:

To evaluate the metabolomic effects of *E. coli* incubation in different carbon sources using three types of fibres: 1% cooked meat, 1% maize meal and 1% olive kernel oil in comparison with a negative control of 1% D glucose.

1.8.3 Aim 3:

To investigate the effects of incubating fecal samples from Crohn's disease patients and controls with different dietary fibres on the metabolomic profiles of these samples. .

1.8.4 Aim 4:

To investigate the capacity for prediction of cancer associated muscle wasting from plasma metabolites of adult patients with gastrointestinal cancer.

Chapter Two:
Materials and Methods

2 MATERIALS AND METHODS

2.1 Participants and study samples

All study samples (plasma or urine) presented in this work were obtained from collaborating hospitals and research groups in the U.K. as described in detail in the relevant sections. The samples were stored frozen at -20°C until required for analysis.

2.2 Solvents and chemicals

HPLC grade Acetonitrile (ACN) was purchased from Fisher Scientific (Loughborough, UK) and HPLC grade water was produced by a Direct-Q3 UltrapureWater System (Millipore, Watford, UK). AnalaR-grade formic acid (98%) was obtained from BDH-Merck (Poole, UK). Authentic stock standard metabolites (Sigma-Aldrich, Poole, U.K.) were prepared as previously described [19] and diluted four times with ACN before LC-MS analysis. Ammonium acetate was purchased from Sigma-Aldrich (Poole, UK).

2.3 Instrumental techniques and columns

Liquid chromatographic separation was carried out on an Accela HPLC system interfaced to an Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany)

using either a ZIC-pHILIC column (150 × 4.6 mm, 5 µm, HiChrom, Reading UK) or/and a reversed phase column (ACE C4, 150 × 3.0 mm, 3 µm, HiChrom Reading UK).

2.4 Mobile phases

The ZICpHILIC column was eluted with a mobile phase consisting of 20 mM ammonium carbonate in HPLC-grade water (solvent A) and acetonitrile (solvent B), at a flow rate of 0.3 mL/min. The elution gradient was an A:B ratio of 20:80 at 0 min, 80:20 at 30 min, 92:8 at 35 min and finally 20:80 at 45 min. The mobile phase for elution of the ACE C4 column consisted of 1 mM acetic acid (A) and 1 mM acetic acid in acetonitrile (B), at a fow rate of 0.4 ml/min. The elution gradient was as follows: A:B ratio 60:40 at 0 min, 0:100 at 30 min, 0:100 at 36 min, 60:40 at 37 min, 60:40 at 41 min.

2.5 The MS run conditions

The nitrogen sheath and auxiliary gas flow rates were maintained at 50 and 17 arbitrary units. The electrospray ionisation (ESI) interface was operated in both positive and negative modes. The spray voltage was 4.5 kV for positive mode and 4.0 kV for negative mode, while the ion transfer capillary temperature was 275°C. Full scan data were obtained in the mass-to-charge ratio (m/z) range of 75 to 1200 for both ionisation modes

on the LC-MS system fully calibrated according to manufacturer's guidelines. The resulting data were acquired using the XCalibur 2.1.0 software package (Thermo Fisher Scientific, Bremen, Germany).

2.6 Data extraction and analysis

Data extraction for each of the samples was carried out using either MZMine or MZMatch software. The extracted ions, with their corresponding m/z values and retention times, were pasted into an Excel macro of the most common metabolites prepared in–house to facilitate identification. The lists of the metabolites obtained from these searches were then carefully evaluated manually by considering the quality of their peaks and their retention time match with the standard metabolite mixtures run in the same sequence. All metabolites were within 3 ppm of their exact masses. Statistical analyses were performed using both univariate with Microsoft Excel and multivariate approaches using SIMCA-P software version 14.1 (Umetrics, Umea, Sweden.).

2.7 Other equipment used

The ultrasonic bath was a Branson 1510 from Branson Ultrasonics (Slough, UK). Automatic pipettes (Gilson) were from Anachem (Luton, UK). All glassware was Fisher Scientific (Loughborough, UK). The centrifuge was a Benchmark MyFuge Mini from Benchmark

Scientific (Edison, NJ, USA). Acrodisc® syringes and filters were purchased from Fisher Scientific (Loughborough, UK).

Chapter Three:

Untargeted Metabolomics Profiling of Ultramarathon Exercise Simulated on a Treadmill in Healthy Adults

3 METABOLOMICS EFFECTS OF ULTRAMARATHON EXERCISE

3.1 Introduction

There has been an upward trend in life expectancy over the past few years in developed countries, but lifestyle risks still pose real challenges to longevity. These risks factors include obesity, unhealthy diet, cigarette smoking, sedentary lifestyle, and alcohol consumption (Harper and Howse, 2008). Regular physical exercise, coupled with proper diet and moderate alcohol consumption, can significantly decrease the impact of these risk factors resulting into an increased life expectancy and wellbeing (Williams, 1997). For instance, it has been reported that regular exercise attenuates sarcopenia and promotes cardiovascular health (Trappe, 2007, Faulkner *et al*., 2008, Sarris *et al*., 2014). It is also prescribed for people with diabetes (Organization, 2009), obesity, and mild to moderate depression (Sarris *et al*., 2014). Moreover, it has been reported that incidences of hypertension, hypercholesterolemia, and diabetes decrease with the frequency of participation in marathons independent of the total distance run annually, but this might be due to longer training runs or genetic and innate differences between marathoners and those who are not (Williams, 2009). However, despite the clear benefits of regular exercise, sedentary behaviour is still widespread. For instance, the indirect cost of physical inactivity is estimated to be 1.5%–3% of total direct healthcare costs in developed countries, including the U.K. where this cost has been estimated at £8.2 billion per annum in England (Scarborough *et al*., 2011).

A clear understanding of the metabolomic effects of physical exercise and how they correlate with physical performance could give insights into comprehensive and individualised healthcare plans for promotion of wellbeing. Currently, data on the metabolomics alterations that occur during exercise are still limited. A previous study of healthy adults subjected to submaximal exercise showed significant increases in a range of purine metabolites and several acyl carnitines (Muhsen Ali *et al*., 2016). High intensity and prolonged exercises such as marathon, which have recently become a worldwide social and fitness phenomenon, can give a better indication of the metabolic changes in the body and enable correlations with other physical performance indicators. Understanding of such metabolic changes could enable the elucidation of individual's ability to maintain peak physical performance and physiological function (Tanaka and Seals, 2008).

Physical performance in a marathon can be affected by gender (Sparling *et al*., 1998, Baker and Tang, 2010, Hunter *et al*., 2011), age (Ransdell *et al*., 2009), lifestyle, and body mass index (BMI) (Knechtle *et al*., 2009), through differences in physiological (e.g., muscle strength, oxygen carrying capacity) and morphological (e.g., percentage of body fat, muscle mass) characteristics of an individual (Lepers and Cattagni, 2012). However, the amount of exercise optimal for a given individual remains unknown due to absence of definitive data on the molecular mechanisms underlying exercise in relation to health.

Thus investigation of the metabolomic effect of exercise on the human metabolome could provide insights into phenotypic responses, permit development of personalised training regimes based on initial metabolic status of an individual (Daskalaki *et al*., 2014),

and yield vital diagnostic and prognostic biomarkers for use by physicians in the management of cardiovascular and other related diseases (Carnethon *et al*., 2005). The maximal rate of oxygen consumption, known as VO_2max, is the most effective indicator of cardiovascular fitness and can be determined by measuring respiratory variables during an incremental exercise test to exhaustion (Muhsen Ali *et al*., 2016).

This investigation was a controlled laboratory study involving 9 healthy male participants in an 80 km marathon simulated on a treadmill. Plasma samples collected at two different points before, and immediately after the marathon were analysed for their metabolomic profiles using both hydrophilic interaction (HILIC) and reversed phase (RP) liquid chromatography-mass spectrometry (LC-MS) methods. Multivariate data analysis was employed with SIMCA by fitting PCA-X, OPLS-DA and OPLS models to determine the metabolic changes due to extreme exercise in order gain some insight into how metabolism is adapted for endurance performance.

3.2 Materials and Methods

3.2.1 Chemicals and Solvents

HPLC grade Acetonitrile (ACN) was purchased from Fisher Scientific (Loughborough, UK) and HPLC grade water was produced by a Direct-Q3 UltrapureWater System (Millipore, Watford, UK). AnalaR-grade formic acid (98%) was obtained from BDH-Merck (Poole, UK). Authentic stock standard metabolites (Sigma-Aldrich, Poole, U.K.) were prepared as

previously described (Zhang *et al*., 2014) and diluted four times with ACN before LC-MS analysis. Ammonium acetate was purchased from Sigma-Aldrich (Poole, UK).

3.2.2 Participants

The healthy participants were 9 males with mean (±SD) age of 33.7±6.7 years, body mass 69.2±7.0 kg, BMI 22.7±2.0 kg/m2. Table 3 summarises all the participants' characteristics.

3.2.3 Plasma samples

Plasma samples were collected from the 9 participants at two time points during an 80 km run on a treadmill, namely: Pre-marathon and Post-80 km. The 'Pre 80 km' samples were collected at rest immediately prior to the participants starting the 80 km treadmill ultramarathon and the 'Post 80 km' samples were taken immediately upon the participants completing the 80 km distance. In addition, baseline samples were taken on the day that the participants came to perform their familiarization, and baseline testing (VO_2 max test) was within a two-day window prior to their 80 km run.

Table 3.1: Participant biographic information and metadata

Participant ID	Gender	Age (years)	Body Mass (kg)	BMI (kg/m^2)	VO_2 max (ml.min.kg)	Total Elapsed Time for 80.5 km (hr:min:sec)
P00	Male	27	65.55	20.5	62.2	08:12:00
P01	Male	32	77.3	23.9	62.5	07:40:48
P03	Male	33	83.55	25.2	55.6	10:37:37
P04	Male	33	61.95	20.5	66.9	07:04:19
P05	Male	34	69.45	22.8	59.5	09:20:37
P07	Male	50	67.47	21.7	57.8	10:02:38
P14	Male	33	72.76	22.7	58.5	10:20:25

P15	Male	29	68.15	20.7	62.4	10:17:14
P16	Male	31	67.5	21.2	68.7	10:00:00

3.2.4 Sample preparation

Exactly 100 µL of the sample was mixed with 400 µL of acetonitrile containing 5 µg/ml of $^{13}C_2$ glycine (Sigma-Aldrich, Poole, U.K.) as an internal standard to ensure retention time stability, and then centrifuged for 10 min before transferring into a vial with an insert. The pooled sample was prepared by pipetting 50 µL from each of the 46 samples and then mixing them together before diluting 0.2 ml of the pooled sample with 0.8 ml of acetonitrile containing 5 µg/mL of $^{13}C_2$ glycine internal standard and centrifuged. Additionally, the prepared mixtures of authentic standard metabolites (Zhang *et al*., 2014) containing 5 µg/mL of $^{13}C_2$ glycine as internal standard were run.

3.2.5 LC-MS conditions

Liquid chromatographic separation was carried out on an Accela HPLC system interfaced to an Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) using both a ZIC-pHILIC column (150 × 4.6 mm, 5 µm, HiChrom, Reading UK) and a reversed phase column (ACE C4, 150 × 3.0 mm, 3 µm, HiChrom Reading UK). The ZICpHILIC column was eluted with a mobile phase consisting of 20 mM ammonium carbonate in HPLC-grade water (solvent A) and acetonitrile (solvent B), at a flow rate of 0.3 mL/min. The elution gradient was an A:B ratio of 20:80 at 0 min, 80:20 at 30 min, 92:8 at 35 min and finally 20:80 at 45 min. The mobile phase for elution of the ACE C4 column consisted of 1 mM acetic acid (A) and 1 mM acetic acid in acetonitrile (B), at a fow rate of 0.4

ml/min. The elution gradient was as follows: A:B ratio 60:40 at 0 min, 0:100 at 30 min, 0:100 at 36 min, 60:40 at 37 min, 60:40 at 41 min. The nitrogen sheath and auxiliary gas flow rates were maintained at 50 and 17 arbitrary units. The electrospray ionisation (ESI) interface was operated in both positive and negative modes. The spray voltage was 4.5 kV for positive mode and 4.0 kV for negative mode, while the ion transfer capillary temperature was 275°C. Full scan data were obtained in the mass-to-charge ratio (m/z) range of 75 to 1200 for both ionisation modes on the LC-MS system fully calibrated according to manufacturer's guidelines. The resulting data were acquired using the XCalibur 2.1.0 software package (Thermo Fisher Scientific, Bremen, Germany).

3.2.6 Data extraction and analysis

Data extraction for each of the samples was carried out by MZMatch software. The extracted ions, with their corresponding m/z values and retention times, were pasted into an Excel macro of the most common metabolites prepared in–house to facilitate identification. The lists of the metabolites obtained from these searches were then carefully evaluated manually, by considering the quality of their peaks and their retention time match, with the standard metabolite mixtures run in the same sequence. All metabolites were within 3 ppm of their exact masses. Statistical analyses were performed using both univariate with Microsoft Excel and multivariate approaches using SIMCA-P software version 14.1 (Umetrics, Umea, Sweden.).

3.3 Results

3.3.1 Physiological response to the marathon

The mean (±SD) values of VO_2max and marathon completion time were 61.56±4.25 ml.min.kg and 9.28 h ±1.3 hours respectively.

3.3.2 Variation of metabolic profile with exercise

The data set of polar metabolites was filtered by excluding 96 metabolites which had RSD values > 20% within the pooled samples. Figure 3.1 shows a clear separation of the pre- and post 80K samples according PCA based on 446 metabolites annotated to MSI level 2 (Sumner *et al*., 2007). The pooled QC samples are clustered in the middle of the plot indicating technical stability throughout the run. There was a technical problem with one of the post-80K samples which was removed from the plot. The model explained 82.6% of the variation in the data in PC1 and PC2. From figure 3.1 it can be seen that ultra-exercise has a strong impact on the levels of polar metabolites in plasma although there is considerable variation between individuals with regard to their response. In addition, there was no separation between two baseline samples one taken prior to the day of the run and the one taken on the day of the run immediately prior to the start (t = 0). This means that the metabolite profiles of the participants at baseline were quite consistent..

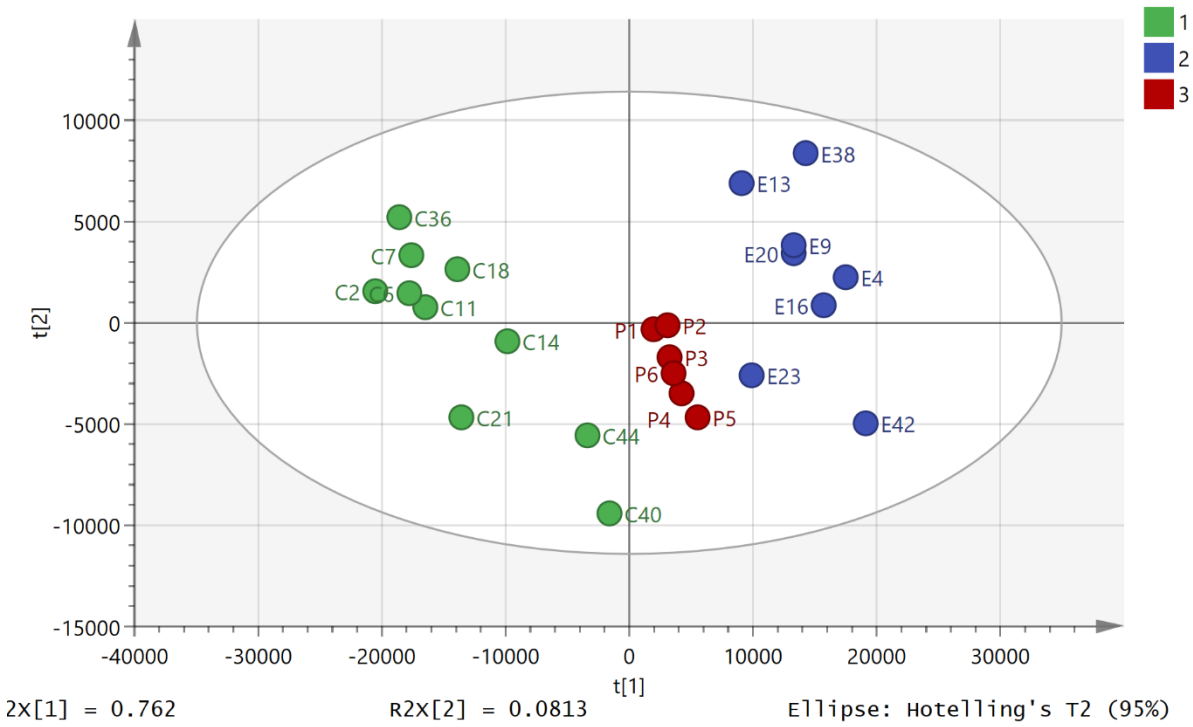

Figure 3.1: PCA separation of pre- 80K samples (C n=9) and post 80K (E n=8) samples based on polar metabolites analysed on a ZICpHILIC column. P= pooled samples. One post sample in the set is missing due to a technical failure. The data was Pareto scaled and log transformed. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

The data for the lipophilic metabolites was filtered by excluding 200 metabolites which had RSD >20% in the pooled samples as previously reported (Muhsen *et al.*, 2016). The PCA model shown in figure 3.2 is based on 220 metabolites annotated to MSI level 2. The model explains 70.2% of the variation in the data in four components. It is thus not as strong as the model based on polar metabolites and pre-80K sample C2 and the corre-

sponding post-80K sample are outliers. Figure S2 shows that there was no clear separation between the two sets of baseline samples although again sample B1 is an outlier for the same individual who produced outliers in figure 3.2.

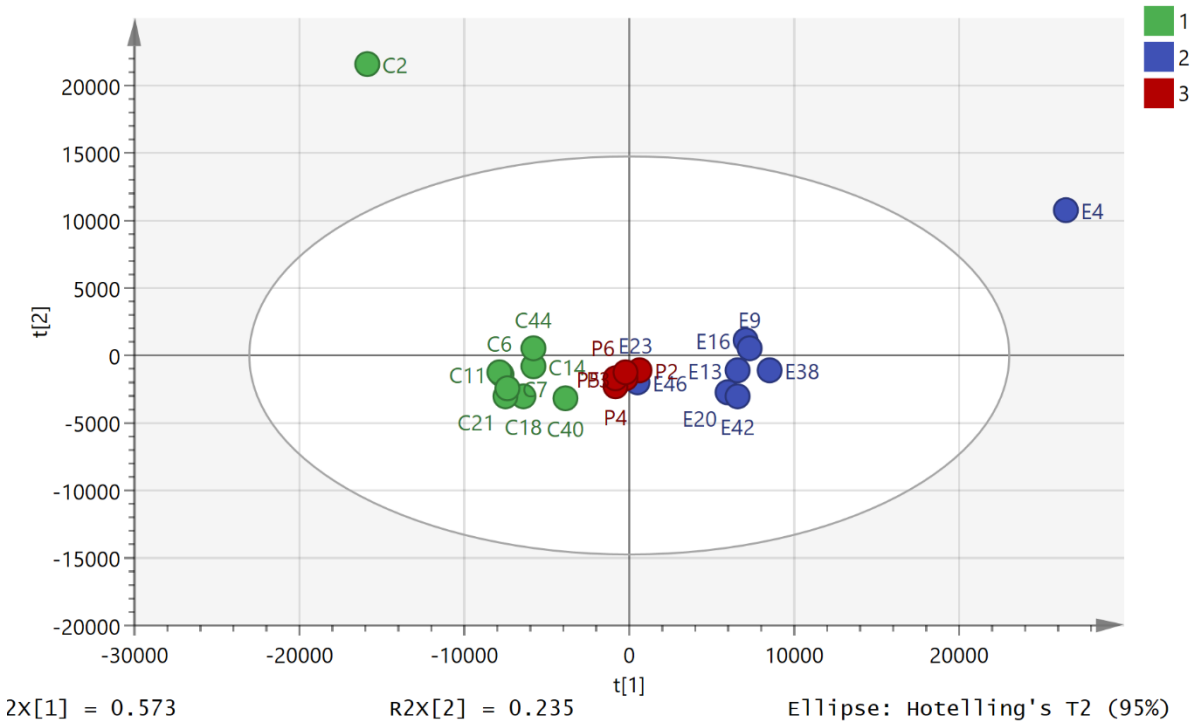

Figure 3.2: PCA separation of pre- 80K samples (C n=9) and post 80K (E n=9) samples based on lipophilic metabolites analysed on an ACE C4 column. P= pooled samples. The data was Pareto scaled. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

3.3.3 Univariate comparisons

As can be seen in table 3.2 there were a very large number of metabolic changes resulting from exercise with many amino acids decreasing in abundance while there were increases in the levels of many acylcarnitines, fatty acids and oxidized fatty acids. The P values from a 2 tailed T test are in many cases are very low for the comparison of the pre- and post-80K samples and a FDR test confirmed the significance to all metabolites with P value <0.05 (Benjamini and Hochberg, 1995). In order to gain a comprehensive overview, analysis was also carried out by RP chromatography which was useful for getting a clearer picture of the lipophilic compounds in plasma including long chain acylcarnitines, fatty acids and oxidized fatty acids. The results from the reversed phase analysis of acylcarnitines, fatty acids and oxidized fatty acids are also shown in table 3.2. The reversed phase mode was better for these classes of compounds since in HILIC mode they all eluted close to the column void volume.

Table 3.2: Table showing all the significant metabolites affected by the marathon exercise.* Matches retention time of standard. ‡Data from runs on ACE C4 column.

Mass	RT (min)	Metabolite	Ratio 80K/pre-80K	P value
Amino acids				
75.0321	15.4	*Glycine	0.510	<0.001
89.0477	14.4	*Alanine	0.603	0.012
103.063	13.4	*3-Amino-isobutanoate	0.392	<0.001
105.043	15.7	*Serine	0.512	<0.001
111.032	9.5	Pyrrole-2-carboxylate	0.413	<0.001
115.063	12.4	*Proline	0.420	<0.001
116.047	1.7	Oxopentanoic acid	0.819	<0.001
117.054	15.5	Guanidinoacetate	0.627	0.001

117.079	12.1	*Valine	0.447	<0.001
117.079	10.8	*Betaine	0.505	<0.001
118.063	1.8	Hydroxypentanoate	1.393	<0.001
119.058	14.4	*Threonine	0.217	<0.001
125.015	15.4	*Taurine	0.565	0.001
130.063	1.8	Oxohexanoic acid‡	0.808	0.725
129.043	14.1	5-Oxoproline	0.352	<0.001
131.058	14.1	*Hydroxyproline	0.361	<0.005
131.095	10.3	*Leucine	0.455	<0.001
131.095	10.8	*Isoleucine	0.430	<0.001
132.079	2.1	Hydroxyhexanoic acid‡	2.237	0.004
132.053	15.2	*Asparagine	0.465	<0.001
132.053	12.5	N-Carbamoylsarcosine	1.430	0.057
132.09	25.5	*Ornithine	0.189	<0.001
138.043	8.8	*Urocanate	0.626	0.019
146.069	14.8	*Glutamine	0.710	<0.001
146.106	23.8	*Lysine	0.369	<0.006
147.053	11.2	*Glutamate	0.528	<0.001
149.051	11.2	*Methionine	0.609	<0.003
154.038	11.7	Imidazol-5-yl-pyruvate	0.469	<0.001
159.068	8.1	Indole-3-acetaldehyde	0.432	0.001
161.069	9.9	O-Acetylhomoserine	0.524	<0.001
165.079	9.6	*Phenylalanine	0.839	0.063
174.112	25.4	*Arginine	0.387	<0.003
175.096	15.6	*Citrulline	0.673	0.047
181.074	12.8	*Tyrosine	0.761	0.016
182.058	9.4	Hydroxyphenyllactate	0.541	0.002
188.116	16.2	*N6-Acetyl-L-lysine	0.233	0.054
189.043	6.4	Kynurenate	2.322	0.001
204.09	11.1	*L-Tryptophan	0.539	<0.001
208.085	10.2	Formylhydroxykynurenamine	0.668	0.004
219.053	4.9	Hydroxyindolepyruvate	5.131	0.010
Acylcarnitines				
204.1227	10.3	*Acetylcarnitine	3.353	<0.001
218.1383	9.1	Propanoylcarnitine	1.420	0.042
232.1539	7.9	Butanoylcarnitine	1.775	0.010
248.1488	3.5	Hydroxybutyrylcarnitine‡	1.109	0.538
258.1695	2.1	Hexenoylcarnitine‡	6.350	0.002
260.1852	2.4	Hexanoylcarnitine‡	9.640	0.011
260.1853	2.9	Hexanoylcarnitine‡	13.091	0.045
274.2008	3.0	Heptanoylcarnitine‡	5.685	0.013

286.2008	3.6	Octenoylcarnitine‡	6.009	0.003
286.2009	3.1	Octenoylcarnitine‡	5.184	0.001
288.2165	4.0	Octanoylcarnitine‡	1.990	0.405
288.2166	4.2	Octanoylcarnitine‡	7.119	0.004
302.2322	4.8	Dimethylheptanoylcarnitine‡	14.587	0.001
312.2165	4.7	Decadienoylcarnitine‡	7.016	0.001
312.2165	3.7	Decadienoylcarnitine‡	16.727	0.102
314.2321	6.2	Decenoylcarnitine‡	7.186	0.039
314.2321	4.8	Decenoylcarnitine‡	1.008	0.994
314.2322	5.8	Decenoylcarnitine‡	6.285	0.004
316.2479	7.1	O-decanoyl-R-carnitine‡	5.017	0.005
330.227	4.4	Keto-decanoylcarnitine‡	13.121	0.000
330.2271	3.0	Keto-decanoylcarnitine‡	7.719	0.003
330.2634	8.3	Dimethylnonanoylcarnitine‡	11.088	0.002
342.2635	9.4	Dodecenoylcarnitine‡	6.439	0.089
342.2635	9.1	Dodecenoylcarnitine‡	8.849	0.004
360.274	6.0	Hydroxylauroylcarnitine‡	4.825	0.003
368.2791	10.8	Tetradecadiencarnitine‡	5.659	0.022
368.2791	9.5	Tetradecadiencarnitine‡	24.743	0.012
368.2792	9.9	Tetradecadiencarnitine‡	19.098	0.055
368.2792	11.9	Tetradecadiencarnitine‡	9.195	0.031
370.2948	13.3	Tetradecenoylcarnitine‡	16.422	0.070
370.2948	12.9	Tetradecenoylcarnitine‡	9.253	0.004
372.3104	15.3	Tetradecanoylcarnitine‡	18.265	0.007
384.2741	6.6	Hydroxytetradecadiencarnitine	11.908	0.001
3862897	8.1	Hydroxytetradecenoylcarnitine‡	6.193	0.007
386.2897	8.9	Hydroxytetradecenoylcarnitine‡	27.813	0.003
388.3053	9.4	Hydroxymyristoylcarnitine‡	4.245	0.006
396.3103	15.2	Hexadecadienoylcarnitine‡	90.958	0.149
396.3105	14.5	Hexadecadienoylcarnitine‡	17.816	0.016
398.3261	17.1	Hexadecenoylcarnitine‡	14.097	0.011
400.3416	19.7	Palmitoylcarnitine‡	4.618	0.089
412.3053	9.3	Hydroxyhexadecadienoylcarnitine‡	6.590	0.003
414.3208	11.1	Hydroxyhexadecenoylcarnitine‡	35.292	0.003
424.3416	18.6	Linoelaidylcarnitine‡	3.955	0.048
424.3416	19.3	Linoelaidylcarnitine‡	6.043	0.121
430.3157	8.1	Hexadecanedioicacid mono-carnitineester‡	114475.436	0.015
Fatty acids and oxidized fatty acids				
172.147	10.0	Decanoic acid‡	1.909	0.034
196.146	10.1	Dodecadienoic acid‡	5.989	0.001

200.178	13.4	Dodecanoic acid	4.342	0.009
202.12	3.2	Decanedioic acid‡	6.045	0.004
210.126	9.3	Hydroxydodecatrienoic acid‡	4.709	0.001
212.178	13.3	Tridecenoic acid‡	13.224	0.006
224.178	13.2	Tetradecadienoic acid‡	10.003	0.013
226.193	14.5	Tetradecenoic acid‡	25.065	0.004
226.193	14.9	Tetradecenoic acid‡	14.409	0.020
230.152	5.0	Dodecanedioic acid‡	9.432	0.014
240.173	8.2	Hydroxytetradecadienoic acid‡	11.109	0.002
240.209	16.5	Pentadecenoic acid‡	3.192	0.007
242.188	11.4	Hydroxytetradecadienoic acid‡	4.066	0.001
244.204	8.2	Hydroxytetradecanoic acid‡	11.109	0.002
244.204	9.3	Hydroxytetradecanoic acid‡	3.581	0.000
252.209	15.8	Hexadecadienoicacid‡	10.174	0.059
252.209	16.3	Hexadecadienoicacid‡	13.108	0.041
254.224	17.5	Hexadecenoic acid‡	38.719	0.006
258.183	7.4	Tetradecanedioic acid‡	7.206	0.006
266.188	13.6	Hydroxyhexatrienoic acid‡	8.977	0.014
268.204	11.3	Hydroxyhexadienoic acid	3.355	0.007
268.24	18.9	Heptadecenoic acid‡	29.923	0.004
270.22	12.1	Hydroxyhexadecenoic acid‡	3.553	0.004
270.22	17.5	Hydroxyhexadecenoic acid‡	8.969	0.001
272.235	11.3	Hydroxyhexadecanoic acid‡	8.969	0.001
276.209	15.6	Octadecatetraenoic acid‡	10.190	0.067
278.225	16.9	Octadecatrienoic acid‡	8.511	0.003
280.24	18.4	Linoleate‡	5.769	0.008
282.256	20.3	Octadecenoic acid‡	6.231	0.000
284.199	9.1	Dihydroxyhexadecadienoic acid‡	2.897	0.001
286.214	10.5	Dihydroxyhexdecenoic acid‡	16.426	0.001
296.235	14.0	Hydroxyoctadecadienoic acid‡	3.145	0.024
300.266	14.3	Hydroxyoctdecanoic acid‡	6.618	0.015
316.261	9.8	Dihydroxyoctadecanoic acid‡	4.040	0.002
327.241	7.0	Nitrooctadecenoic acid‡	10.453	<0.001
328.24	18.3	Docosahexaenoicacid‡	4.266	0.022
330.256	19.0	Docosapentaenoic acid‡	9.179	0.003
332.272	20.5	Docosatetraenoic acid‡	14.588	0.002
Steroids and bile acids				
362.209	4.5	Hydrocortisone	1.787	0.014
364.225	5.0	Urocortisone	3.243	0.003
376.298	3.9	Hydroxycholanate	0.315	0.004
392.293	4.3	Deoxycholanoic acid	0.361	0.026

449.314	4.3	Chenodeoxyglycocholate	0.162	<0.001
465.309	4.9	Glycocholate	0.174	0.003
515.291	4.5	Taurocholate	0.275	0.039
568.324	7.3	Chenodeoxycholic acid glucuronide	0.311	<0.001
612.387	4.5	Cholestane--tetrol-glucuronide	0.443	0.001
Miscellaneous				
136.039	9.8	*Hypoxanthine	1.917	0.003
244.069	9.5	*Uridine	0.420	<0.001
244.07	11.7	Pseudouridine	0.416	<0.001
136.064	23.7	*1-Methylnicotinamide	0.226	0.090
164.069	11.8	Rhamnose	0.348	<0.001
179.079	10.8	Galactosamine	0.181	<0.001
180.064	14.1	Hexose	0.447	<0.001
214.132	9.4	Dethiobiotin	1.517	0.002
416.366	3.4	gamma-Tocopherol	0.529	<0.001
430.381	3.4	Alpha-Tocopherol	0.509	<0.001

3.4 Discussion

The observed clear separation between baseline and samples collected during and immediately after the marathon shows that there are significant metabolic changes induced by physical activity. For some metabolite the changes are very large and are consistent across all the individuals in the trial. The major changes concern fatty acid metabolism. There is a large elevation in acylcarnitine levels in plasma for a wide range of these compounds. The impact of exercise on carnitines has been observed before in a number of studies (Neal *et al*., 2012, Lustgarten *et al*., 2013, Huffman *et al*., 2014, Xu *et al*., 2016, Zhang *et al*., 2017). A possible explanation is that the carnitines reflect the requirement of muscles for glucose as an energy source under the impact of physical activity. The rate

of production of energy from glucose is faster than when fatty acids are used as an energy source (Zajac *et al*., 2014). Most recently it has been hypothesized that acyl carnitines have neuroactive properties that can regulate exertion via interaction with the neurons regulating muscle activity (Zhang *et al*., 2017). Less frequently studied are the products of fatty acid oxidation that accumulate in plasma during exercise (Nieman *et al*., 2014, Nieman *et al*., 2016). Many oxidized fatty acids have potent effects on blood vessels promoting either vasodilation or vasoconstriction. As can be seen in table 3.2 there is a complex mixture of these compounds all of which are greatly elevated in plasma following exercise. The oxidation products of linoleic acid 9-hydroxylinoleic acid and 13-hydroxylinoleic acid have been proposed as markers of oxidative stress following exercise and several isomers of these compounds are elevated particularly in the 80 K samples in comparison to baseline (table 3.2). Figure 3.3 shows extracted ion chromatograms for the pre- and post- levels of oxidized linoleic acid. As can be seen table 3.2 the range of oxidised fatty acids elevated post-exercise is extensive and the increases very marked. Thus the elevation of hydroxylinoleic acids is not exclusive and there are many other hydroxy acids which are elevated post-exercise plus some dioic acids. Whether or not these acids also have biological activities is unknown as is the precise reason for their elevation. When the heat map shown figure 3.4 is considered, it is evident that many of the oxidised fatty acids, although elevated in as shown in table 3.2, are of relatively low abundance. It has been suggested that oxidised acids are a marker of oxidative stress (Nieman *et al*., 2014, Nieman *et al*., 2016) but it might be expected that other readily

oxidised acids present in plasma such as eicosapentaenoic acid (EPA) might also be oxidised in the same but despite EPA being relatively abundant in the plasma no peaks for hydroxy EPAs can be seen. Thus, it is possible that there is some biological mechanism that keeps oxidation products of EPA at low levels since many of these metabolites have potent inflammatory and vasoactive effects. Given the wide range of unsaturated fatty acids and hydroxylated fatty acids shown in table 1 it would seem likely that these compounds are arising from peroxisomal metabolism and this might provide a protective mechanism for ensuring that the levels of oxidised long chain unsaturated acids are kept at low levels. Peroxisomes are known to be responsible for degrading prostaglandins (Wanders and Waterham, 2006). Unlike mitochondrial beta-oxidation of fatty acids peroxisomal beta-oxidation of fatty acids does not necessarily go to completion and acids may only be shortened by 3-4 cycles of 2 carbon chain shortening (Wanders and Waterham, 2006) yielding a molecule of acetyl CoA/acetyl carnitine at each cycle. For instance, it might be significant that hexadecadienoic acid, tetradecadienoic acid and dodecadiencoic acid are all elevated, these are not abundant naturally occurring fatty acids, but they are all products of chain shortening of linoleic acid via beta oxidation (Wanders and Waterham, 2006). Similarly, hexdecatrienoic acid could arise from chain shortening of linolenic acid via one beta-oxidation step. The reason for the metabolism pausing when a double bond is encountered within the fatty acid chain is that at this point further metabolism requires the commitment of NADPH in the reduction of the double bond

before further chain shortening can occur (Wanders and Waterham, 2006). Under conditions of aerobic stress there will be generally a high requirement for NADPH in countering oxidative stress; it is required for instance in the recycling of GSSG back to GSH. The elevated levels of acetylcarnitines are consistent with increased beta-oxidation fatty acids by peroxisomes since they are the major product exported out of peroxisomes resulting from fatty acid beta-oxidation. It has been demonstrated that physical exercise increases peroxisome levels in rat heart (Zipper, 1997). Acetyl carnitine is readily utilised by mitochondria as a source of acetyl CoA which can be metabolised via the Krebs cycle. The major question with regard to carnitines is whether they are waste products or utilisable as substrates for further oxidation. Conversion of acylCoAs to acylcarnitines is necessary in order to preserve free levels of CoA so that further fatty acid metabolism can occur via their conversion of acylCoA esters (Ramsay and Zammit, 2004). The heat map shown in figure 4 indicates in terms of absolute abundance that the common dietary fatty acids are much higher in plasma than the unusual acids which are promoted by exercise observed in the current study. Thus it seems probable that medium chain length unsaturated fatty acids are minor metabolites due to partial metabolism of long chain unsaturated fatty acids by peroxisomes providing an additional source of acetylcarnitine for export to mitochondria. The heat map in figure 3.5 shows the relative abundance of the 40 most abundant acyl carnitines in plasma. Acetyl carnitine is highly abundant while the carnitines corresponding to the medium chain fatty acids are of much lower abundance. Although the levels of some acyl carnitines rise in urine post-exercise they do not

increase to the same extent as the plasma levels in the current study and no increase in urinary acetyl carnitine was observed (Muhsen Ali *et al*., 2016). This suggests that the carnitines may be produced for utilisation as energy substrates. Conversion of free fatty acids to acyl CoAs requires the investment of a molecule of ATP. When the muscles are working hard this is likely to be available in reduced amounts. However, acyl carnitines are an activated form of fatty acid substrate and are convertible into acyl CoAs without the investment of ATP in creating the thioester bond and thus they can be taken up into mitochondria and further metabolised (Ramsay and Zammit, 2004, Reddy and Mannaerts, 1994). Thus, the pattern of fatty acids and carnitines observed in the current study points strongly towards a large increase in peroxisomal metabolism. For example a widely studied substrate of peroxisomal metabolism is phytanic acid which is present in dairy products (Wanders and Waterham, 2006). This compound undergoes α-oxidation in the peroxisomes producing pristanic acid which is then further metabolized by the peroxisomes yielding propanolyl CoA (carnitine) and dimethyl nonanoyl CoA (carnitine) after six cycles of beta oxidation. Both of these carnitines are elevated in post-exercise samples and provide potential substrates for mitochondrial metabolism in the muscles. The increased activity of the peroxisomes is further underlined by elevated levels of some dioic acids (table 3.2) which are also only produced by peroxisomes. The hypothesis that the metabolite patterns are consistent with peroxisomal proliferation is consistent with our earlier observations where it was proposed that exercise increased the proliferation of PPAR-γ ligands in plasma (Thomas *et al*., 2011). From the current case

these ligands might well be long chain unsaturated fatty acids which are substrates for peroxisomal metabolism as discussed above.

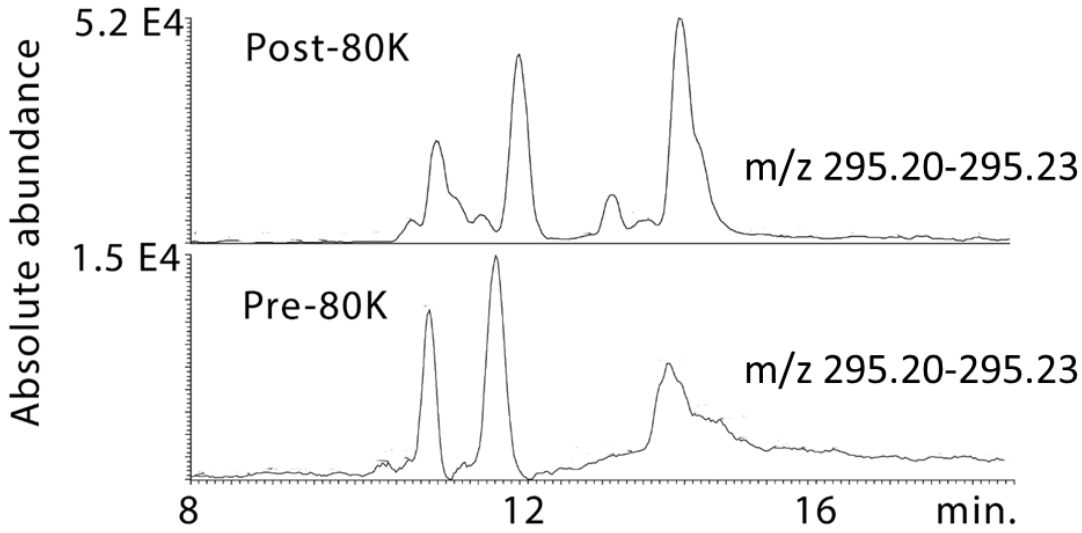

Figure 3.3: Hydroxy linoleic acids in a pre-80 K sample and a post 80K sample run on an ACE C4 column.

Figure 3.4: Heat map showing the relative abundance of the 30 most abundant fatty acids in plasma for the pre- and post-80K samples and two post-exercise samples. Red = highest value (3.93×10^7), Yellow = 1×10^5 and blue = 5×10^3.

Mass	Rt	Fatty acid	Mean pre 80K	Mean post-80K
282.256	20.3	Octadecenoic acid		
280.24	18.4	Octadecadienoic acid		
256.24	19.5	Hexanoic aicd		
254.224	17.5	Hexadecenoic acid		
228.209	16.5	Tetradecanoic acid		
278.225	16.9	Octadecatrienoic acid		
284.271	22.3	Octadecanoic acid		

200.178	13.4	Dodecanoic acid		
226.193	14.5	Tetradecenoic acid		
328.24	18.3	Docosahexaenoicacid		
306.256	19.5	Eicosatrienoic acid		
330.256	19.0	Docosapentaenoic acid		
268.24	18.9	Heptadecenoic acid		
242.225	18.0	Pentadecanoic acid		
332.272	20.5	Docosatetraenoic acid		
172.147	10.1	Decanoic acid		
242.225	17.7	Pentadecanoic acid		
226.193	14.9	Tetradecenoic acid		
298.251	12.5	Hydroxyoctadecenoic acid		
238.157	11.7	Hydroxytetradecatrienoic acid		
224.178	13.2	Tetradecadienoic acid		
314.245	20.3	Dihydroxyoctadecenoic acid		
272.235	11.3	Hydroxypentadecanoic acid		
252.209	15.8	Hexadecadienoic acid		
300.266	14.3	Hydroxyoctadecanoic acid		
258.183	7.4	Tetradecanedioic acid		
316.261	9.8	Dihydroxyoctadecenoic acid		
276.209	15.6	Octadecatetraenoic acid		
230.152	5.0	Dodecanedioic acid		
212.178	13.3	Tridecenoic acid		

The levels of almost all the amino acids in the plasma samples fall significantly. The fall in the amino acids used in protein biosynthesis might be due to an increase in protein

biosynthesis during exercise which has been observed to occur (Harber *et al*., 2009, Walker *et al*., 2011). Hydrocortisone and its metabolite urocortisone are increased during exercise and this has been observed to occur in previous studies (Gatti *et al*., 2005, Dovio *et al*., 2010). Hydrocortisone is responsible for maintaining a homeostasis under stress conditions. The most studied metabolites with regard to the effect of exercise and the determination of fitness are metabolites in the purine pathway such as hypoxanthine and inosine. There is a marked change in levels of hypoxanthine. The re-uptake of hypoxanthine into muscle has been observed to be more efficient in highly trained individuals (Zieliński *et al*., 2013) and the elevation of hypoxanthine in plasma during exercise is less marked than we observed in urine samples taken post-exercise (Muhsen Ali *et al*., 2016). However, since the athletes in the current study were highly trained it might be expected that their metabolism was geared to conserving purines (Stathis *et al*., 2006). Changes in uridine following exercise have been observed previously most often increases have been observed, in the current case there was a marked decrease (Dudzinska *et al*., 2013). Changes in tocopherols have also been observed previously in exercise studies and γ-tocopherol has been correlated to VO_2 max level (Subudhi *et al*., 2001, Lustgarten *et al*., 2013).

Figure 3.5: Changes in the 40 most abundant acylcarnitines in plasma following an ultramarathon analysed by RP method.

row m/z	Rt (min)	Metabolite	Mean Pre-80K	Mean Post-80K
204.1227	1.7	O-Acetylcarnitine		
316.2479	7.1	Decanoyl-R-carnitine		
288.2166	4.2	L-Octanoylcarnitine		
314.2322	5.8	Decenoylcarnitine		
370.2948	12.9	Tetradecenoylcarnitine		
342.2635	9.1	Dodecenoylcarnitine		
286.2009	3.1	Octenoylcarnitine		
288.2165	4.0	L-Octanoylcarnitine		
370.2948	13.3	cis-5-Tetradecenoylcarnitine		
368.2791	10.8	Tetradecadiencarnitine		
342.2635	9.4	trans-2-Dodecenoylcarnitine		
302.2322	4.8	2-6dimethylheptanoylcarnitine		
260.1852	2.4	O-hexanoyl-R-carnitine		
372.3104	15.3	Tetradecanoylcarnitine		
330.2634	8.3	4-8dimethylnonanoylcarnitine		
232.1539	1.8	O-Butanoylcarnitine		
398.3261	17.1	trans-Hexadec-2-enoylcarnitine		
314.2321	6.2	Decenoylcarnitine		
218.1383	1.8	O-Propanoylcarnitine		
312.2165	4.7	Decadienoylcarnitine		
424.3416	18.6	Linoelaidylcarnitine		
386.2897	8.9	Hydroxtetradecenoylcarnitine		
286.2008	3.6	Octenoylcarnitine		
424.3416	19.1	Linoelaidylcarnitine		
400.3416	19.7	O-Palmitoyl-R-carnitine		
424.3416	19.3	Linoelaidylcarnitine		
386.2897	8.1	Hydroxytetradecenoylcarnitine		
312.2165	3.7	Decadienoylcarnitine		
360.274	6.0	Hydroxylauroylcarnitine		
260.1853	2.9	O-hexanoyl-R-carnitine		

396.3105	14.5	Hexadecadienoylcarnitine		
274.2008	3.0	Heptanoylcarnitine		
330.2271	3.0	Keto-decanoylcarnitine		
368.2792	11.9	Tetradecadiencarnitine		
388.3053	9.4	Hydroxymyristoylcarnitine		
330.227	4.4	Keto-decanoylcarnitine		
414.3208	11.1	Hydroxyhexadecenoylcarnitine		
412.3053	9.3	Hydroxyhexadecadienoyl carnitine		
384.2741	6.6	Hydroxytetradecadien carnitine		
248.1488	3.5	Hydroxybutyrylcarnitine		

3.5 Conclusions

The clearest impact of endurance exercise is on fatty acid metabolism but with respect to formation of medium chain unsaturated and partially oxidised fatty acids and conjugates of fatty acids with carnitines. The most likely explanation for the complex pattern of medium chain and oxidised fatty acids formed is that exercise provokes the proliferation of peroxisomes. The peroxisomes may serve two functions one of providing a readily utilisable form of energy in the form of acetyl carnitine and other acyl carnitines for export to mitochondria in the muscles, without the investment of the ATP required to conjugate free fatty acids to CoA. Secondly the peroxisomes may serve to regulate the levels of oxidised metabolites of long chain fatty acids since many of these metabolites can provoke biological responses such as vasoconstriction or have pro-inflammatory activity. It was possible to build a model which was predictive of VO_{2max} based on five metabolites all of which had some potential biological significance with regard to the impact of exercise.

Chapter Four:

Evaluation of the metabolomics effects of *E. coli* incubation in different carbon sources

4 EVALUATION OF THE METABOLOMICS EFFECTS OF *E. COLI* INCUBATION IN DIFFERENT CARBON SOURCES

4.1 Introduction

Among the "Omic" approaches, metabolomics is one of the most reliable techniques currently used in the identification of novel targets for diagnosis and specific markers for diseases, and has been increasingly applied in the characterisation of the link between gut microbiota or host metabolism and pathophysiological alterations in various diseases (De Preter and Verbeke, 2013). Metabolomics, unlike other technologies such as genomics and proteomics which rely on gene expression and protein data, not only indicates the potential for specific metabolic functions, also the effective physiological processes since the influence of several downstream regulatory mechanisms involved are also taken into consideration within the metabolome (De Preter and Verbeke, 2013, Fiehn, 2002). For this reason, metabolomics integrates all the effects of gene regulation, post-transcriptional regulation and pathway interactions, making it a more comprehensive tool in the understanding of a cell's physiological phenotype (Fiehn, 2002, Assfalg *et al*., 2008).

Metabolomics is a powerful exploratory tool for understanding the interactions between nutrients, the intestinal metabolism and the microbiota composition in health and disease and, to gain more insight in metabolic pathways. Although genomic and proteomic

methodologies have, until recently, often been applied to uncover gastrointestinal related pathophysiological processes (Berndt *et al*., 2007, Olsen *et al*., 2009, Arijs *et al*., 2010, Lamendella *et al*., 2012, Yau *et al*., 2013, van Beelen Granlund *et al*., 2013), currently, metabolomics is increasingly used to discover gastrointestinal disease signatures and has been applied for the screening of different pathological conditions that are linked with a metabolic imbalance such as ulcerative colitis, inflammatory bowel disease (IBD), Crohn's disease and irritable bowel syndrome (IBS) (Olsen *et al*., 2009, Arijs *et al*., 2010, Lamendella *et al*., 2012, Yau *et al*., 2013, van Beelen Granlund *et al*., 2013).

The human GIT microbiota, particularly the one associated with large intestines, is considered to be one of the most complex and metabolically active organs of the human body. This microbial ecosystem contains about 500-1000 different species of bacteria, of which, in healthy adults, 80% belong to the phyla of Firmicutes, Bacteroidetes and Actinobacteria (Tuohy *et al*., 2009, Huttenhower *et al*., 2012), with substantial variation between different individuals (Ventura *et al*., 2009). A total of about 10^{14} bacterial cells are present in the adult intestine, which is ten times the number of cells in the human body (Xu and Gordon, 2003). This microbiome outnumbers the host's genetic potential by two orders of magnitude (Methé *et al*., 2012) and provides a diverse range of biochemical and metabolic activities to complement the host's physiology. The presence and metabolic activities of a specific bacterial community play an important role in main-

taining the host's overall health and well-being, and has been shown to respond to metabolic challenges and dietary factors. This complex microbial system varies with the host's age, diet and health status (Claesson *et al*., 2012).

The bacteria in the colon can ferment food nutrients and residues to produce a wide range of chemical compounds which may have an influence on the host's physiological processes, both locally within the colon and systemically, and it has been observed that colonic bacteria exhibit functional redundancy given that several bacteria can ferment the same substrate to produce the same product (Mahowald *et al*., 2009). This implies that both composition and functional capacity of the intestinal microbiota are important in determining the clinical endpoint, although metabolic insights from these processes are still limited due to inability to access intestinal habitat and complexity of the microbiota (Tuohy *et al*., 2009).

It should be noted that, apart from microbial composition and functional capacity of the gut microbiota, various other factors can influence the gut metabolome and these include nutrient availability and its physicochemical properties, age of host, and transit time of the colon. Nutrient availability, particularly the carbohydrate to nitrogen ratio, is believed to be the most important regulator of bacterial metabolism, as it influences preference of saccharolytic vs proteolytic fermentation (De Preter *et al*., 2011). Whereas short-chain fatty acids (SCFA) result from colonic fermentation of carbohydrates, protein fermentation generates phenolic compounds, amines and ammonia, branched chain fatty acids, and S-containing compounds (Cummings, 1981). Protein fermentation gives

rise to a variety of metabolites such as phenolic compounds, branched chain fatty acids, S-containing compounds, amines and ammonia (Smith and Macfarlane, 1996). The former are believed to be beneficial to the host while the latter may have undesirable effects such as toxicity (Cummings, 1981).

In this study, the metabolomics effects of *E. coli* incubation under various carbohydrate/protein sources were investigated. *E. coli* was selected for this study based on its ubiquity as a gut microorganism associated with various disease states. Samples were taken from the cultures at 0, 24 and 48 h. These samples were extracted following the standard metabolomics protocol and injected into a liquid chromatography-mass spectrometry (LC-MS) system based on the Orbitrap and the data obtained were processed through MZMine, identified, and analysed statistically by means of both supervised and unsupervised models in SIMCA-P software. Any observed differences between the carbon sources for each group were compared with those observed with the glucose control group.

4.2 Materials and Methods

4.2.1 Chemicals and Solvents

HPLC grade Acetonitrile (ACN) was purchased from Fisher Scientific (Loughborough, UK) and HPLC grade water was produced by a Direct-Q3 UltrapureWater System (Millipore, Watford, UK). AnalaR-grade formic acid (98%) was obtained from BDH-Merck (Poole, UK). Authentic stock standard metabolites (Sigma-Aldrich, Poole, U.K.) were prepared as

previously described (Zhang *et al*., 2014) and diluted four times with ACN before LC-MS analysis. Ammonium acetate was purchased from Sigma-Aldrich (Poole, UK).

4.2.2 Study samples

Different *E. coli* (DH5a) cultures were incubated in minimal medium with 1% D-glucose (control; G), 1% maize meal (carbohydrates; M), 1% cooked meat medium (protein; B), and 1% olive kernel oil (fat; O) as carbon sources. The incubations were all made in triplicate. Samples were taken from the cultures at 0, 24 and 48 h, resulting to a total of 36 samples. These samples were frozen until the time of analysis. The 36 samples were coded as follows:

Table 4.1: Description summary of the test sample groupings and control

Letter code	Medium (Carbon sources)
G	Control (1% Glucose)
M	Fibre (1% maize carbohydrate meal)
B	Fibre (1% cooked meat medium)
O	Fibre (1% Olive kernel oil as C source)

4.2.3 Sample preparation

Exactly 100 µL of the sample was mixed with 400 µL of acetonitrile and then centrifuged for 10 min before transferring into a HPLC vial with an insert. The pooled sample was prepared by pipetting 50 µL from each of the 126 samples and then mixing them together before diluting 0.2 ml of the pooled sample with 0.8 ml of acetonitrile, centrifug-

ing for 10 min and transferring the supernatant into the HPLC vial for analysis. Additionally, the prepared mixtures of authentic standard metabolites (Zhang *et al*., 2014) were run in the same sequence for later identification of metabolites in the samples.

4.2.4 LC-MS conditions

Liquid chromatographic separation was carried out on an Accela HPLC system interfaced to an Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) using a ZIC-pHILIC column (150 × 4.6 mm, 5 µm, HiChrom, Reading UK). The column was eluted with a mobile phase consisting of 20 mM ammonium carbonate in HPLC-grade water(solvent A) and acetonitrile (solvent B), at a flow rate of 0.3 mL/min. The elution gradient was an A:B ratio of 20:80 at 0 min, 80:20 at 30 min, 92:8 at 35 min and finally 20:80 at 45 min. The nitrogen sheath and auxiliary gas flow rates were maintained at 50 and 17 arbitrary units. The electrospray ionisation (ESI) interface was operated in both positive and negative modes. The spray voltage was 4.5 kV for positive mode and 4.0 kV for negative mode, while the ion transfer capillary temperature was 275°C. Full scan data were obtained in the mass-to-charge ratio (m/z) range of 75 to 1200 for both ionisation modes on the LC-MS system fully calibrated according to manufacturer's guidelines. The resulting data were acquired using the XCalibur 2.1.0 software package (Thermo Fisher Scientific, Bremen, Germany).

4.2.5 Data extraction and analysis

Data extraction for each of the samples was carried out by MZMine software. The extracted ions, with their corresponding m/z values and retention times, were pasted into an Excel macro of the most common metabolites prepared in–house to facilitate identification. The lists of the metabolites obtained from these searches were then carefully evaluated manually by considering the quality of their peaks and their retention time match with the standard metabolite mixtures run in the same sequence. All metabolites were within 3 ppm of their exact masses. Statistical analyses were performed using both univariate with Microsoft Excel and multivariate approaches using SIMCA-P software version 14.1 (Umetrics, Umea, Sweden.).

4.3 Results

4.3.1 Unsupervised analysis

Figure 4.1 shows the PCA-X analysis of all 36 samples and 5 QC samples. PCA-X, an unsupervised model in SIMCA-P, produces a natural scatter of the samples based on their characteristic metabolomics footprints. It can be seen in the figure (samples coloured according to the carbon sources) that, in general, the samples are clustered according to the carbon sources. However, there is some variability in the G and O samples, as reflected in the splitting of their respective clusters; the sub-clusters formed from the split-

tings still contain samples taken at different times, implying that time was not the contributing factor to the sub-clustering. The 5 QC samples are tightly clustered together which confirms reproducibility of different injections throughout the entire experiment.

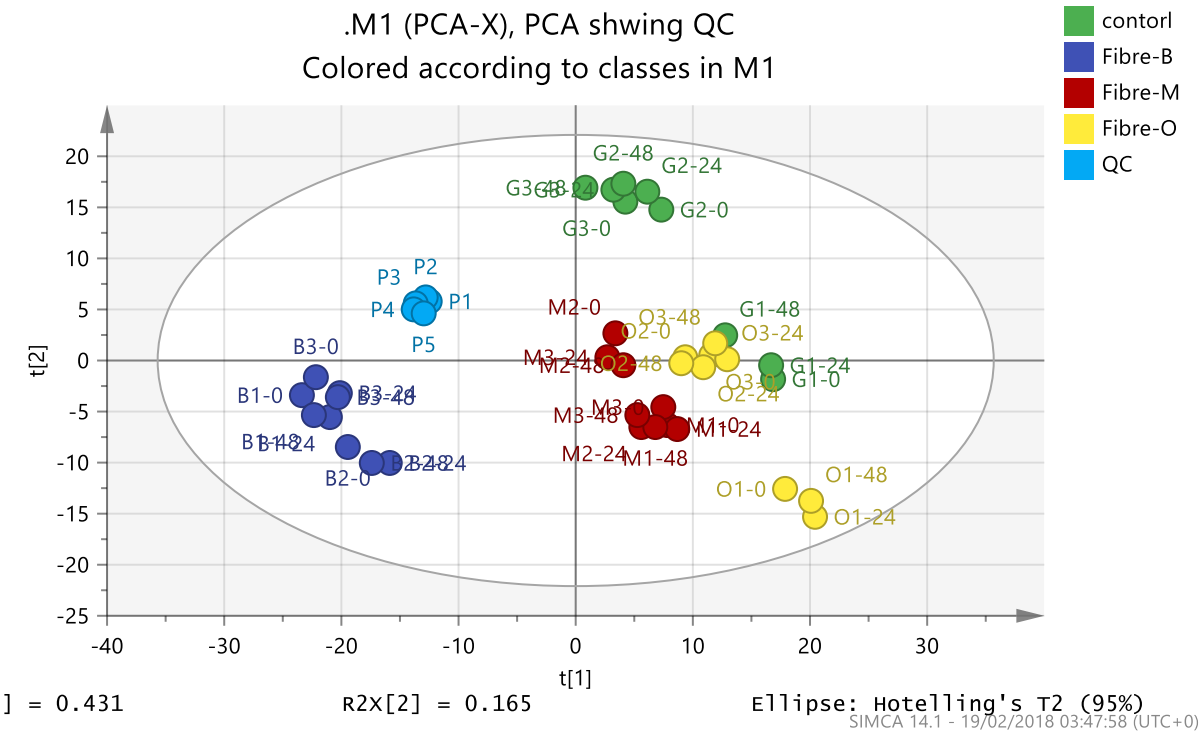

Figure 4.1: PCA-X analysis of the metabolomics footprint of the 36 samples from E. coli cultures in different carbon sources. Circles coloured according to the time of sample collection post diet. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

4.3.2 Supervised analysis

Supervised models enable identification of metabolites that have the most significant contribution to a clustering pattern. In SIMCA-P, supervised analysis can be carried out using OPLS-DA models. To achieve this, three comparisons were made where each of the

samples from a given carbon source was compared with the glucose control. For each OPLS-DA model, permutation tests were carried out to check validity of the model.

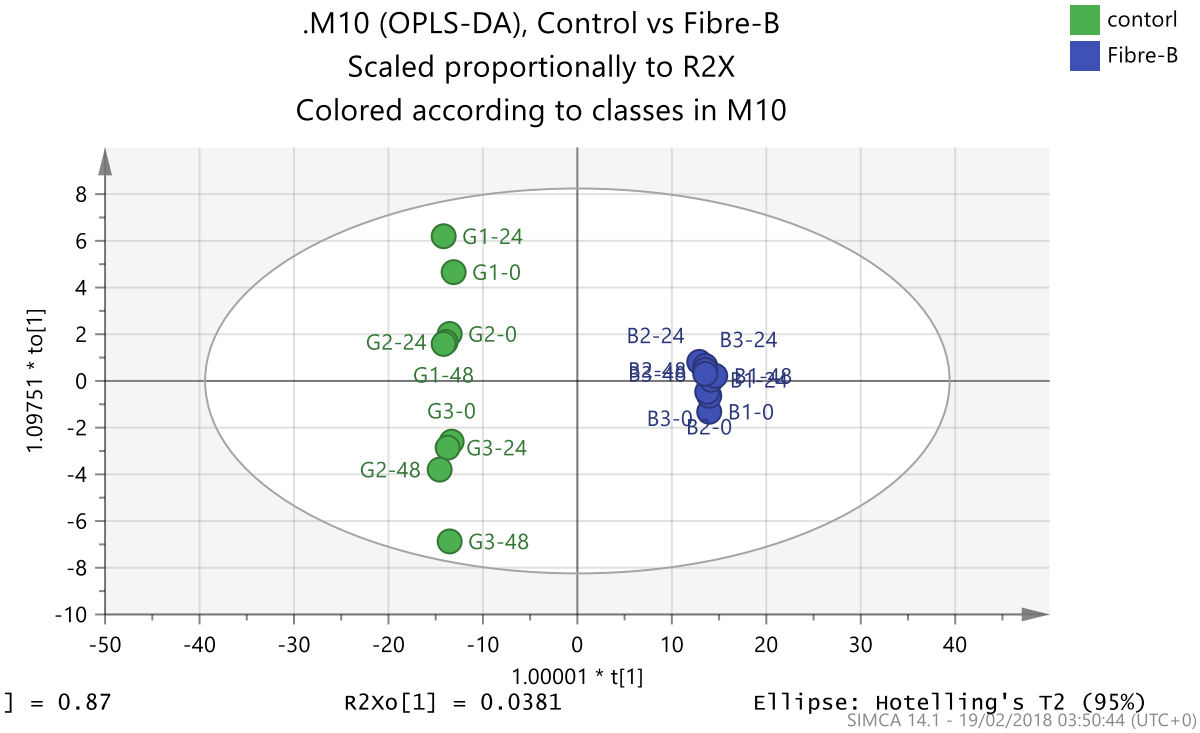

Figure 4.2: OPLS-DA analysis to compare B samples (Fibre = 1% cooked meat medium) with controls (G samples, representing 1% glucose). There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 9.9604e-017. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

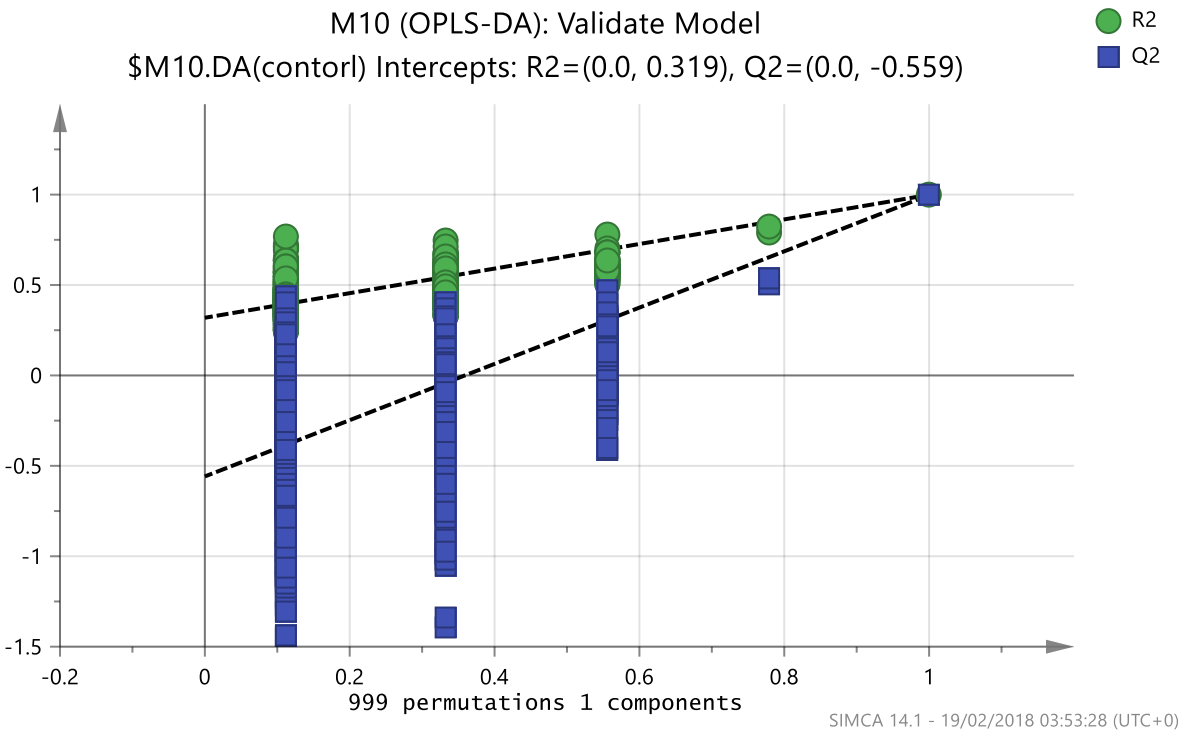

Figure 4.3: Cross validation of the OPLS-DA model comparing B samples (Fibre = 1% cooked meat medium) with controls (G samples, representing 1% glucose).

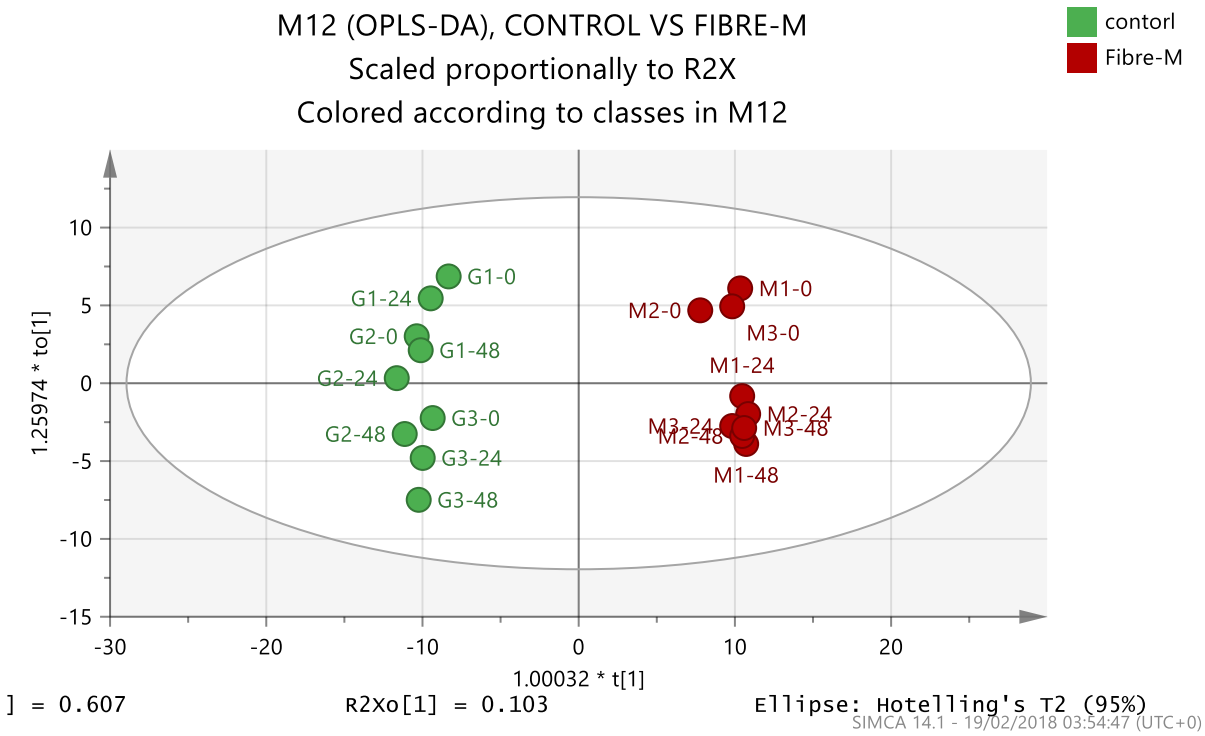

Figure 4.4: OPLS-DA analysis to compare M samples (Fibre = 1% maize carbohydrate meal) with controls (G samples, representing 1% glucose). There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 6.26598e-011. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

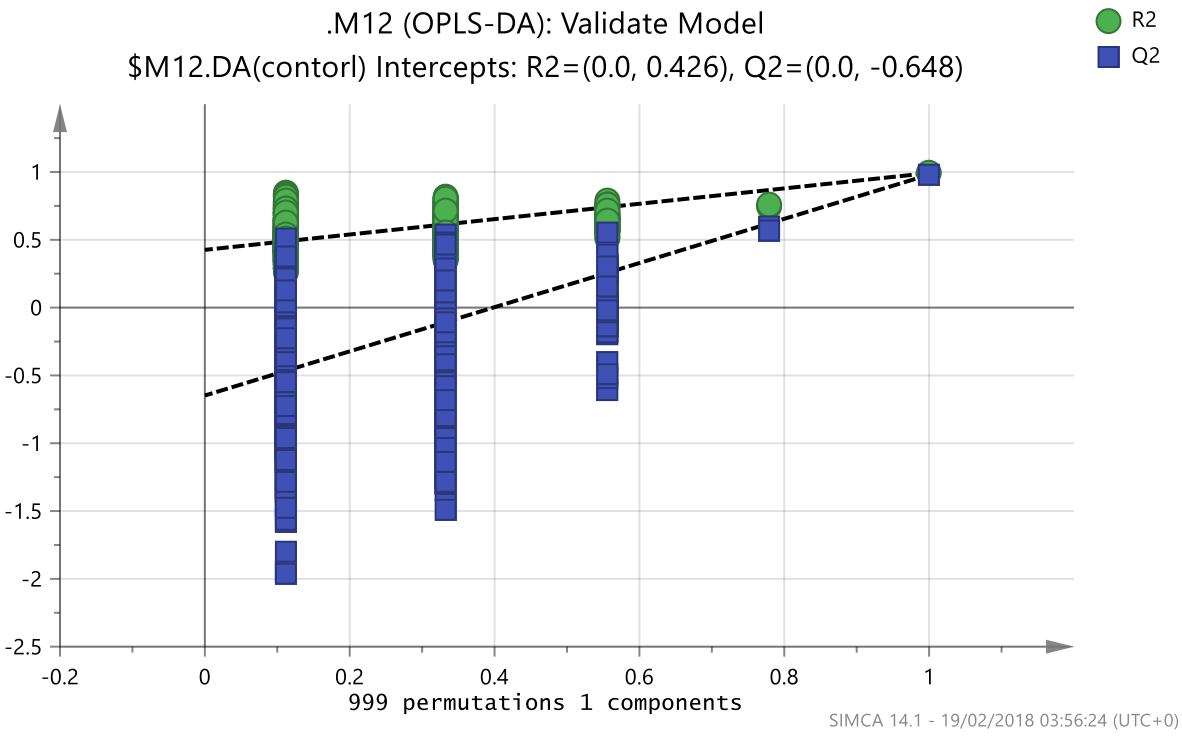

Figure 4.5: Cross validation of the OPLS-DA model comparing M samples (Fibre = 1% maize carbohydrate meal) with controls (G samples, representing 1% glucose).

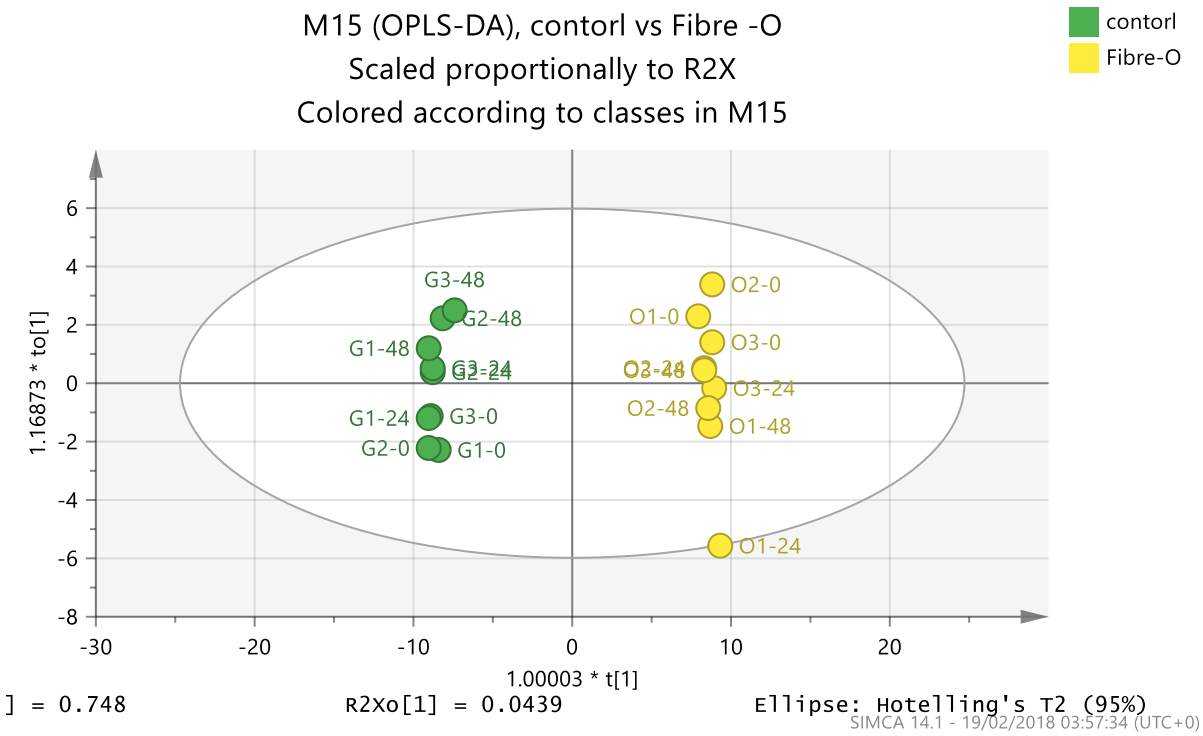

Figure 4.6: OPLS-DA analysis to compare O samples (Fibre = 1% Olive kernel oil as C source) with controls (G samples, representing 1% glucose). There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 4.7657e-012. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

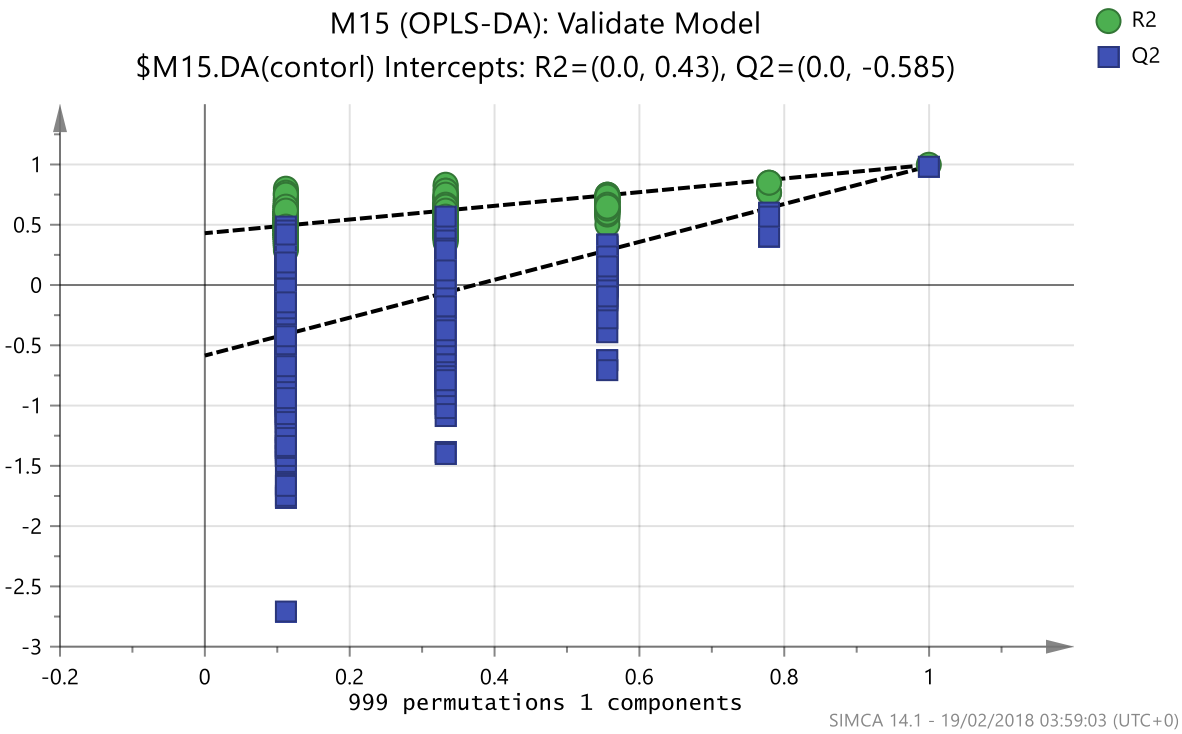

Figure 4.7: Cross validation of the OPLS-DA model comparing O samples (Fibre = 1% Olive kernel oil as C source) with controls (G samples, representing 1% glucose).

4.3.3 Metabolite alterations

Tables 4.2 to 4.4 show the metabolites found to be significantly altered in each carbon source compared with the glucose control. The metabolites have been categorised into the major metabolic pathways of amino acid, protein, carbohydrate, lipid, polyketide, nucleotide metabolism and xenobotics among others. The ratio of each metabolite concentration in the medium relative to its concentration in the control carbon source is also given.

4.3.3.1 Effect of 1% cooked meat fibre on E. coli metabolome

The majority of the metabolites significantly altered by incubation of E. coli in 1% cooked meat medium were those in the amino acid metabolic pathway. Among these were L-Arogenate, L-Glutamyl-5-phosphate, 5-Hydroxy-L-tryptophan, 3-Sulfino-L-alanine, 4-Guanidinobutanamide, (1-Ribosylimidazole)-4-acetate, Guanidinoacetate, Tryptamine, and L-Methionine S-oxide. Others were L-Carnitine, Creatine, O-Acetylcarnitine, and N-Formimino-L-glutamate. On the other hand, the three most significantly downregulated metabolites in this pathway were L-Kynurenine, N-Succinyl-LL-2,6-diaminoheptanedi-oate, and N2-Acetyl-L-aminoadipate.

On the other hand, using carbohydrate as a carbon source produced only a few metab-olites signficantly altered in comparison with 1% cooked meat relative to the control. Moreover, it should be worth noting that all the three significantly altered carbohydrate metabolites were down regulated and these were D-Ribose, D-Glucosamine, and 4-Methylene-L-glutamine.

The incubations with this fibre also produced significant effects on phospholipid precur-sor metabolites whereby there were increased levels of Taurine, sn-Glycerol 3-phos-phate, N-Methylethanolamine phosphate, and Choline phosphate metabolites. In addi-tion, with the exception of lipoate, the rest of the glycerophospholipids such as [FA tri-hydroxy(2:0/2:0)] N,N-dimehyl-9S,11R,15S-trihydroxy-5Z,13E-prostadien-1-amide, [FA

amino(16:0)] 2R-aminohexadecanoic acid, and [FA (11:1)] 2-undecenal were all increased in the E. coli cultures treated with 1% cooked meat compared to minimal glucose treated cultures.

Among the polyketides, only peonidin was down regulated but the rest, including [Fv] Naringenin, Deoxyelephantopin, [PR] Bixindial/ Bixin aldehyde, [SP] 3-dehydrosphinganine, and Cortol were severally increased in cultures incubated with 1% cooked meat.

Additionally, there were increased levels of all the nucleotide metabolites such as xanthine, inosine, guanosine, and guanine, as well as increases in the xenobiotic markers 4-Carboxy-4-hydroxy-2-oxoadipate and 4-Amino-2-hydroxylamino-6-nitrotoluene. However, some of the xenometabolites such as Protoanemonin, Nitrosobenzene, and [FA methyl,oxo(5:0/2:0)] 2-methylene-4-oxo-pentanedioic acid were significantly lowered in the 1% cooked meat incubated *E. coli* cultures. Many of the altered metabolites are xenobiotics and it is difficult tell how genuine these peaks are. However, there are some large changes in glycerol phosphate which can potentially be incorporated into the glycolysis pathway and may be derived from lipid breakdown. Two lyso lipids hexadecanoyl glycerol phosphocholine and octadecenoyl glycerol phosphocholine are greatly increased suggesting that fatty acid chains many also being lost and used as energy substrates and this might be the cause of an increased level of acetyl carnitine which is a product fatty acid oxidation. Amino acids can potentially be used as a nitrogen source. There is no indication of amino acid breakdown, however, xanthine, hypoxanthine and

inosine are all products resulting from the loss of nitrogen from metabolites in the purine pathway and may be evidence if nitrogen scavenging via purine breakdown.

A summary of all significantly altered metabolites has been presented in Table 4.2, showing the fold changes in comparison to the negative control (1% D glucose) and the VIP-pred values reflecting the importance of the given variable in the OPLS-DA model.

Table 4.2: Showing the significant metabolites between Control-G vs meat proteinB.

Var ID	P-value	Var ID	RT	Formula	Metabolite name	Ratio B/G	VIPpred
Amino Acid Metabolism							
1258	9.12E-04	161.107	10.20	C10H12N2	Tryptamine	1.80	0.71
1717	8.28E-13	204.123	9.31	C9H18NO4	O-Acetylcarnitine	1.53	0.74
1409	5.40E-08	291.118	14.94	C11H18N2O7	N-Succinyl-LL-2,6-diaminoheptanedioate	0.07	1.05
1423	2.97E-09	173.057	10.92	C6H10N2O4	N-Formimino-L-glutamate	1.50	0.70
1628	4.91E-16	175.107	13.67	C7H14N2O3	N-Acetylornithine	1.43	0.71
1411	1.00E-10	204.086	12.94	C8H13NO5	N2-Acetyl-L-aminoadipate	0.06	1.20
1088	3.17E-18	118.086	11.39	C5H11NO2	L-Valine	1.30	0.76
1358	7.36E-04	166.053	13.84	C5H11NO3S	L-Methionine S-oxide	1.75	0.69
1272	7.92E-11	209.092	9.11	C10H12N2O3	L-Kynurenine	0.66	0.75
1636	4.30E-07	228.027	15.06	C5H10NO7P	L-Glutamyl-5-phosphate	6.44	1.00
1367	2.46E-22	162.112	13.34	C7H15NO3	L-Carnitine	1.60	0.93
1304	4.24E-09	228.087	15.35	C10H13NO5	L-Arogenate	12.47	1.14
1493	2.43E-11	171.042	13.18	C6H8N2O4	Hydantoin-5-propionate	1.47	0.67
1522	2.02E-15	118.061	16.12	C3H7N3O2	Guanidinoacetate	1.89	0.90
1675	7.11E-08	114.066	9.82	C4H7N3O	Creatinine	1.25	0.67
1524	3.65E-21	132.077	14.86	C4H9N3O2	Creatine	1.57	0.92
1521	6.69E-10	104.107	18.60	C5H13NO	Choline	1.26	0.71
1295	6.66E-06	221.092	10.86	C11H12N2O3	5-Hydroxy-L-tryptophan	4.30	0.97
1585	2.32E-11	174.087	14.53	C6H11N3O3	5-Guanidino-2-oxopentanoate	1.36	0.67
1373	7.30E-13	160.096	6.80	C7H13NO3	5-Acetamidopentanoate	1.32	0.69
1700	1.59E-13	146.092	15.24	C5H11N3O2	4-Guanidinobutanoate	1.39	0.74
1701	1.30E-05	145.108	28.06	C5H12N4O	4-Guanidinobutanamide	2.58	0.97
1590	4.32E-08	152.002	11.26	C3H7NO4S	3-Sulfino-L-alanine	3.45	1.11
1508	9.98E-04	259.092	11.14	C10H14N2O6	(1-Ribosylimidazole)-4-acetate	2.05	0.75

Biosynthesis of Polyketides and Nonribosomal Peptides							
1083	1.37E-09	526.338	9.49	C28H47NO8	Pikromycin	13.65	1.12
1001	1.10E-09	203.102	12.43	C8H14N2O4	Proclavaminic acid	0.76	0.68
1002	9.67E-08	245.124	12.79	C9H16N4O4	Guanidinoproclavaminic acid	4.61	1.11
971	6.56E-08	218.104	9.89	C10H13N5O	cis-Zeatin	7.28	1.08
1008	1.50E-05	193.074	10.44	C8H10N4O2	Caffeine	3.44	0.93
Carbohydrate Metabolism							
794	5.49E-17	149.046	15.24	C5H10O5	D-Ribose	0.68	0.79
967	6.48E-17	180.086	16.59	C6H13NO5	D-Glucosamine	0.69	0.78
906	2.04E-12	159.076	11.17	C6H10N2O3	4-Methylene-L-glutamine	0.61	0.86
Lipid Metabolism							
787	2.53E-08	124.007	15.96	C2H7NO3S	Taurine	2.31	1.02
778	2.59E-09	171.007	15.65	C3H9O6P	Glycerol 3-phosphate	13.89	1.12
776	2.94E-11	156.042	11.40	C3H10NO4P	N-Methylethanolamine phosphate	17.06	1.19
772	8.56E-08	184.073	15.24	C5H14NO4P	Choline phosphate	5.72	1.09
Lipids: Fatty Acyls							
696	2.25E-03	382.294	4.15	C22H39NO4	[FA trihydroxy(2:0/2:0)] N,N-dimehyl-9S,11R,15S-tri-hydroxy-5Z,13E-prostadien-1-amide	1.72	0.67
651	1.87E-04	315.252	4.13	C18H34O4	Dihydroxyoctadecenoic acid	1.95	0.74
743	1.45E-04	272.258	4.98	C16H33NO2	Aminohexadecanoic acid	1.65	0.68
749	3.74E-04	279.231	3.99	C18H30O2	Octadecatrienoic acid	2.15	0.80
550	4.04E-07	324.289	4.14	C20H37NO2	Octadecadienoyl-ethanolamine	6.64	1.09
553	1.26E-04	340.284	4.10	C20H37NO3	Octadecenoyl-glycine	2.61	0.86
676	1.79E-04	265.216	4.10	C17H28O2	Heptadecatrienoic acid	2.52	0.71
565	9.48E-04	167.144	4.29	C11H20O	Undecenal	2.12	0.69
Lipids: Glycerophospholipids							
492	2.42E-06	299.018	17.57	C8H13O10P	Diacyl-sn-glycero-3-phospho-(1'-sn-glycerol)	3.87	1.00
505	1.14E-15	298.143	6.57	C11H26NO6P	Ethyl-2-methyl-sn-glycero-3-phosphocholine	1.42	0.70

508	4.17E-16	312.159	6.34	C12H28NO6P	[PC diethyl(2:0)] 1,2-diethyl-sn-glycero-3-phosphocho-line	1.45	0.72
513	1.41E-07	424.284	4.83	C20H44NO6P	[PC (6:2/6:2)] 1-hexyl-2-hexyl-sn-glycero-3-phospho-choline	5.47	1.09
514	2.22E-17	340.19	5.42	C14H32NO6P	Hexyl-sn-glycero-3-phosphocholine	1.50	0.74
515	9.55E-15	326.174	6.08	C13H30NO6P	[PC (5:2)] 1-pentyl-sn-glycero-3-phosphocholine	1.53	0.72
517	4.13E-05	520.339	4.41	C26H50NO7P	Lyso octadecadienoyl-sn-glycero-3-phosphocholine	2.52	0.92
524	2.77E-08	496.338	4.53	C24H50NO7P	Lysohexadecanoyl-sn-glycero-3-phosphocholine	8.12	1.20
Lipids: Polyketides							
472	6.64E-13	301.07	14.35	C16H12O6	Peonidin	0.64	0.79
476	1.24E-07	273.075	13.34	C15H12O5	[Fv] Naringenin	6.94	1.09
446	1.35E-12	345.133	9.48	C19H20O6	Deoxyelephantopin	1.44	0.75
455	1.81E-07	347.202	4.12	C24H28O2	[PR] Bixindial/ Bixin aldehyde	6.86	1.06
427	9.15E-05	300.289	4.24	C18H37NO2	[SP] 3-dehydrosphinganine	1.71	0.72
414	1.36E-03	369.263	4.24	C21H36O5	Cortol	2.02	0.70
402	7.70E-07	540.335	4.46	C29H51NO6S	[ST hydrox] (25R)-3alpha,7alpha-dihydroxy-5beta-cho-lestan-27-oyl taurine	4.14	1.08
Metabolism of Cofactors and Vitamins							
397	2.65E-10	240.109	10.49	C9H13N5O3	Dihydrobiopterin	14.82	1.18
400	3.84E-13	245.095	6.67	C10H16N2O3S	Biotin	1.40	0.71
327	1.50E-05	144.047	5.60	C6H9NOS	5-(2-Hydroxyethyl)-4-methylthiazole	3.48	1.01
310	1.66E-07	184.06	5.32	C8H9NO4	4-Pyridoxate	5.43	1.08
Nucleotide Metabolism							
278	1.63E-06	151.026	12.49	C5H4N4O2	Xanthine	2.47	1.04
281	7.93E-05	269.087	11.20	C10H12N4O5	Inosine	2.01	0.88
282	1.98E-14	137.046	10.47	C5H4N4O	Hypoxanthine	1.47	0.84
283	4.41E-08	284.098	12.90	C10H13N5O5	Guanosine	1.53	0.76
284	4.12E-10	152.056	12.69	C5H5N5O	Guanine	1.58	0.82

234	1.53E-09	244.092	12.15	C9H13N3O5	Cytidine	1.45	0.71
289	1.80E-09	268.103	8.95	C10H13N5O4	Adenosine	1.33	0.67
Xenobiotics Biodegradation and Metabolism							
213	1.10E-10	95.0136	15.28	C5H4O2	Protoanemonin	0.64	0.75
27	7.42E-10	108.044	7.55	C6H5NO	Nitrosobenzene	0.65	0.69
169	4.21E-05	219.014	10.46	C7H8O8	4-Carboxy-4-hydroxy-2-oxoadipate	2.44	0.88
11	1.85E-09	184.072	12.72	C7H9N3O3	4-Amino-2-hydroxylamino-6-nitrotoluene	9.05	1.16
168	8.19E-07	157.014	16.03	C6H6O5	[FA methyl,oxo(5:0/2:0)] 2-methylene-4-oxo-pentanedi-oic acid	0.10	1.05

4.3.3.2 Effect of 1% maize carbohydrate meal on E. coli metabolome

Unlike in incubation with 1% cooked meat, most of the metabolites in the amino acid metabolic pathway were significantly lowered by treatment with 1% maize carbohydarate meal except L-Glutamyl 5-phosphate, 5-Hydroxy-L-tryptophan, and 4-Guanidinobutanamide. Among those significantly lowered were S-Ribosyl-L-homocysteine, Serotonin, N-Acetylserotonin, N-Acetyl-L-glutamate, N2-Succinyl-L-ornithine, N2-Acetyl-L-aminoadipate, and L-Kynurenine. Others were gamma-Glutamyl-gamma-aminobutyraldehyde, 4-(beta-Acetylaminoethyl)imidazole, and 3-Methyldioxyindole.

In addition, carbohydrate metabolism had all its significantly altered carbohydrate metabolites down regulated and these were D-Ribose, D-Glucose, D-Glucosamine, 4-Methylene-L-glutamine, and 4-Hydroxy-4-methylglutamate. The effect induced by 1% maize carbohydrate meal on carbohydrate metabolism in *E. coli* was similar to the one observed with 1% cooked meat fibre on the same organism.

This fibre also produced significant effects on lipid metabolites whereby there were generally increased levels of both fatty acyls and glycerophospholipids with the exception of [FA (10:0/2:0)] Decanedioic acid and Peonidin Taurine respectively. The fatty acyls [FA hydroxy(18:0)] 9,10-dihydroxy-12Z-octadecenoic acid, [FA (18:3)] 9Z,12Z,15Z-octadecatrienoic acid, [FA (18:2)] N-(9Z,12Z-octadecadienoyl)-ethanolamine, and [FA (18:0)] N-(9Z-octadecenoyl)-glycine were all increased in the *E. coli* cultures treated with 1% maize compared to 1% D-glucose treated cultures. On the other hand, [PC (18:2)] 1-(9Z,12Z-octadecadienoyl)-sn-glycero-3-phosphocholine, [PC (16:0)] 1-hexadecanoyl-sn-glycero-

3-phosphocholine, [PR] Bixindial/ Bixin aldehyde, and [ST hydrox] (25R)-3alpha,7alpha-dihydroxy-5beta-cholestan-27-oyl taurine were some of the glycerophospholipids upregulated by treatment with 1% maize meal.

Among the polyketides, only peonidin was down regulated but the rest, including [Fv] naringenin, deoxyelephantopin, [PR] bixindial/ bixin aldehyde, [SP] 3-dehydrosphinganine, and cortol were severally increased in cultures incubated with 1% cooked meat.

Unlike 1% cooked meat, the 1% maize meal did not induce significant alterations in the levels of nucleotide metabolites such as xanthine, inosine, guanosine, and guanine, except 5,6-dihydrothymine which was decreased. The maize meal however induced several decreases in levels of xenobiotic markers and cofactors/vitamins. Among the latter, only dihydrobiopterin, 5-(2-hydroxyethyl)-4-methylthiazole, and 4-pyridoxate were increased in 1% maize meal treated *E. coli* cultures.

A summary of all significantly altered metabolites has been presented in Table 4.3, showing the fold changes in comparison to 1% D glucose and the VIPpred values which reflect the contribution of the given variable in the OPLS-DA model.

Table 4.3: Significant metabolite based on the comparisons of Control vs Fibre-M

Var ID	P-value	Polarity	RT	Formula	Metabolite name	Ratio B/G	VIPpred
Amino Acid Metabolism							
1357	0.069058	P	15.8	C9H17NO6S	S-Ribosyl-L-homocysteine	0.38	1.03
1269	<0.001	P	5.4	C10H12N2O	Serotonin	0.83	0.98
1270	0.000003	P	5.3	C12H14N2O2	N-Acetylserotonin	0.69	1.31
1632	0.000387	N	9.7	C7H11NO5	N-Acetyl-L-glutamate	0.77	1.10
1634	<0.001	P	13.7	C9H16N2O5	N2-Succinyl-L-ornithine	0.66	1.35
1411	0.000000	P	12.9	C8H13NO5	N2-Acetyl-L-aminoadipate	0.19	2.17
1272	<0.001	P	9.1	C10H12N2O3	L-Kynurenine	0.06	2.52
1636	0.000001	P	15.1	C5H10NO7P	L-Glutamyl 5-phosphate	6.22	1.96
1407	<0.001	P	16.4	C6H11NO4	L-2-Aminoadipate	0.80	1.17
1673	0.000004	P	10.8	C9H16N2O4	gamma-Glutamyl-gamma-aminobutyraldehyde	0.26	2.05
1295	0.003099	P	10.9	C11H12N2O3	5-Hydroxy-L-tryptophan	2.76	1.18
1701	0.018892	P	28.1	C5H12N4O	4-Guanidinobutanamide	1.67	1.03
1506	0.000000	P	7.4	C7H11N3O	4-(beta-Acetylaminoethyl)imidazole	0.82	1.03
1299	0.000000	P/N	6.3	C9H9NO2	3-Methyldioxyindole	0.69	1.42
Biosynthesis of Polyketides and Biosynthesis of Nonribosomal Peptides							
1083	2.5E-08	P	9.5	C28H47NO8	Pikromycin	11.18	1.97
1001	5.8E-09	P/N	12.4	C8H14N2O4	Proclavaminic acid	0.60	1.76
1008	2.1E-03	N	10.4	C8H10N4O2	Caffeine	2.48	1.25
1012	2.4E-09	N	12.7	C8H10N4O4	5-Acetylamino-6-formylamino-3-methyluracil	0.87	0.99
Carbohydrate Metabolism							
794	4.90E-14	N	15.2	C5H10O5	D-Ribose	0.68	1.57
804	3.89E-10	N	12.3	C6H12O6	D-Glucose	0.85	1.15
967	1.78E-06	P	16.6	C6H13NO5	D-Glucosamine	0.41	2.04
906	1.66E-12	P/N	11.2	C6H10N2O3	4-Methylene-L-glutamine	0.58	1.79

913	1.43E-07	N	11.5	C6H11NO5	4-Hydroxy-4-methylglutamate	0.72	1.20
Lipids: Fatty Acyls							
651	0.00678	P	4.1	C18H34O4	Dihydroxy-octadecenoic acid	1.62	1.03
749	0.00112	P	4.0	C18H30O2	Octadecatrienoic acid	2.03	1.46
550	0.00002	P	4.1	C20H37NO2	[FA (18:2)] N-(9Z,12Z-octadecadienoyl)-ethanola-mine	5.20	1.83
553	0.00446	P	4.1	C20H37NO3	[FA (18:0)] N-(9Z-octadecenoyl)-glycine	2.07	1.27
687	0.00376	P	5.3	C10H18O4	[FA (10:0/2:0)] Decanedioic acid	0.41	1.48
Lipids: Glycerophospholipids							
517	1.18E-05	P	4.4	C26H50NO7P	[PC (18:2)] 1-(9Z,12Z-octadecadienoyl)-sn-glycero-3-phosphocholine	2.58	1.85
524	2.77E-08	P	4.5	C24H50NO7P	[PC (16:0)] 1-hexadecanoyl-sn-glycero-3-phospho-choline	7.65	2.34
472	0.00035	P	14.4	C16H12O6	Peonidin	0.60	1.49
476	0.0227	P	13.3	C15H12O5	[Fv] Naringenin	3.31	1.07
446	6.17E-10	P	9.5	C19H20O6	Deoxyelephantopin	1.27	1.15
455	6.92E-06	N	4.1	C24H28O2	[PR] Bixindial/ Bixin aldehyde	5.45	1.80
402	0.000259	N	4.5	C29H51NO6S	[ST hydrox] (25R)-3alpha,7alpha-dihydroxy-5beta-cholestan-27-oyl taurine	2.87	1.52
Metabolism of Cofactors and Vitamins							
302	2.45E-12	P	7.6	C8H11NO3	Pyridoxine	0.71	1.33
303	3.76E-11	P	6.2	C8H12N2O2	Pyridoxamine	0.84	1.01
309	5.43E-14	P	10.8	C8H9NO3	Pyridoxal	0.74	1.27
332	2.38E-07	P	15.3	C11H23N2O7PS	Pantetheine 4'-phosphate	0.19	2.30
397	0.0036	P	10.5	C9H13N5O3	Dihydrobiopterin	6.43	1.25
327	0.0003	P	5.6	C6H9NOS	5-(2-Hydroxyethyl)-4-methylthiazole	2.85	1.68
310	1.56E-06	P/N	5.3	C8H9NO4	4-Pyridoxate	4.71	1.92
376	2.18E-09	N	15.5	C6H8O4	2,3-Dimethylmaleate	0.83	1.20

391	3.12E-13	P	6.5	C7H8N2O	1-Methylnicotinamide	0.81	1.20
Nucleotide Metabolism							
252	4.97E-07	P	10.6	C5H8N2O2	5,6-Dihydrothymine	0.79	1.19
Xenobiotics Biodegradation and Metabolism							
213	4.46E-08	N	15.3	C5H4O2	Protoanemonin	0.72	1.29
27	5.90E-06	P	7.6	C6H5NO	Nitrosobenzene	0.27	1.85
39	6.53E-10	P	6.1	C6H7NO	4-Hydroxyaniline	0.85	1.05
169	0.0118	N	10.5	C7H8O8	4-Carboxy-4-hydroxy-2-oxoadipate	1.73	1.02
208	5.82E-06	P	9.3	C7H9NO3	2-amino-5-methyl-muconate semialdehyde	0.54	1.59
168	2.27E-06	N	16.0	C6H6O5	[FA methyl,oxo(5:0/2:0)] 2-methylene-4-oxo-pentane-dioic acid	0.15	2.07

4.3.3.3 Effect of 1% Olive kernel oil on E. coli metabolome

Unlike in incubation with 1% cooked meat, most of the metabolites in the amino acid metabolic pathway were significantly lowered by treatment with 1% olive kernel oil except 5-Hydroxy-L-tryptophan and L-glutamyl 5-phosphate. These results were almost similar to those obtained with 1% maize meal. Among those significantly lowered were N2-acetyl-L-aminoadipate, gamma-glutamyl-gamma-aminobutyraldehyde, N-acetylserotonin, 3-hydroxyanthranilate, 3-methyldioxyindole, and N2-succinyl-L-ornithine. Others were phenylacetylglycine, N-acetyl-L-phenylalanine, L-2-aminoadipate, and 4-(beta-acetylaminoethyl)imidazole.

In addition, metabolites associated with carbohydrate metabolism were all significantly altered and these were D-Glucosamine, 2-oxoglutarate, 4-methylene-L-glutamine, 4-hydroxy-4-methylglutamate, D-ribose, (S)-Malate, and D-glucose. The effect induced by 1% olive kernel oil on carbohydrate metabolism in *E. coli* was similar to those observed with 1% cooked meat fibre and 1% maize meal on the same organism. This is perhaps not surpising since neither substrate directly provides a source of carbohydrate.

This fibre also produced significant effects on lipid metabolites although the number of lipids altered were lower than those observed with 1% cooked meat or 1% maize meal. With the exception of [PC (16:0)] 1-hexadecanoyl-sn-glycero-3-phosphocholine, the levels of the affected lipids were lowered by the olive kernel oil. Among those lowered were lipoate, [FA trihydroxy(18:0)] 9S,12S,13S-trihydroxy-10E-octadecenoic acid, purpurin, peonidin, and [Fv] matteucinol 7-O-glucoside. These observations on lipids differs from

those observed with both 1% cooked meat and 1% maize meal where there was a general increase in the levels of the lipids.

The other metabolites affected by treatment with 1% olive kernel oil were the metabolites of cofactors and vitamins, nucleotide, and xenobiotics degradants. In general, all these metabolites were significantly lowered in *E. coli* cultures treated with the 1% olive kernel oil compared to those treated with negative control (1% D glucose). The altered cofactors and vitamin metabolites included dethiobiotin, pyridoxal, 1-methylnicotinamide, pyridoxine, pyridoxamine, and 2,3-dimethylmaleate which were all decreased, but 5-(2-Hydroxyethyl)-4-methylthiazole was increased. On the other hand, the altered nucleotide metabolites included uracil, adenosine, 3-oxo-3-ureidopropanoate, and 5,6-dihydrothymine which were all decreased. Finally, the xenobiotic degradants included 2-amino-5-methyl-muconate semialdehyde, protoanemonin, n-acetylisoniazid, 1-7-dimethyluricacid, and 4-hydroxyaniline and were also decreased relative to the negative control treated cultures.

A summary of all significantly altered metabolites has been presented in Table 4.4, showing the fold changes in comparison to 1% D glucose and the VIPpred values which reflect the contribution of the given variable in the OPLS-DA model.

Table 4.4: The significant metabolite for the comparisons of Control vs Fibre-O

Var ID	P-value	Polarity	RT	Formula	Metabolite name	Ratio O/G	VIPpred
Amino Acid Metabolism							
1411	3.20E-09	P	12.94	C8H13NO5	N2-Acetyl-L-aminoadipate	0.11	1.32
1673	1.30E-08	P	10.83	C9H16N2O4	gamma-Glutamyl-gamma-aminobutyraldehyde	0.12	1.31
1270	0.0001	P	5.27	C12H14N2O2	N-Acetylserotonin	0.43	0.95
1266	0.002371	P	11.46	C7H7NO3	3-Hydroxyanthranilate	0.46	0.79
1299	0.000164	P/N	6.33	C9H9NO2	3-Methyldioxyindole	0.52	0.90
1634	0.000339	P	13.69	C9H16N2O5	N2-Succinyl-L-ornithine	0.52	0.85
1335	0.011972	P/N	6.41	C10H11NO3	Phenylacetylglycine	0.57	0.65
1352	5.31E-06	P/N	5.76	C11H13NO3	N-Acetyl-L-phenylalanine	0.74	0.64
1407	5.19E-09	P	16.37	C6H11NO4	L-2-Aminoadipate	0.77	0.72
1506	8.21E-10	P	7.43	C7H11N3O	4-(beta-Acetylaminoethyl)imidazole	0.79	0.64
1295	0.001152	P	10.86	C11H12N2O3	5-Hydroxy-L-tryptophan	2.99	0.81
1636	1.98E-07	P	15.06	C5H10NO7P	L-Glutamyl 5-phosphate	6.77	1.19
Biosynthesis of Secondary Metabolites							
969	0.005398	P/N	5.60	C10H13N5	N6-(delta2-Isopentenyl)-adenine	0.48	0.66
1001	7.47E-12	P/N	12.43	C8H14N2O4	Proclavaminic acid	0.53	1.09
1062	0.001994	P	5.95	C9H12N2	Nornicotine	0.65	0.64
1008	0.002021	N	10.44	C8H10N4O2	Caffeine	2.47	0.76
Carbohydrate Metabolism							
967	5.93E-10	P	16.59	C6H13NO5	D-Glucosamine	0.07	1.52
898	0.0029	N	16.75	C5H6O5	2-Oxoglutarate	0.53	0.80
906	5.37E-13	P/N	11.17	C6H10N2O3	4-Methylene-L-glutamine	0.54	1.07
913	0.000353	N	11.53	C6H11NO5	4-Hydroxy-4-methylglutamate	0.62	0.72
794	1.33E-12	N	15.24	C5H10O5	D-Ribose	0.68	0.90
894	7.10E-10	N	16.65	C4H6O5	(S)-Malate	0.83	0.62

804	7.78E-12	N	12.35	C6H12O6	D-Glucose	0.85	0.64
Lipid Metabolism							
541	6.59E-07	N	43.52	C8H14O2S2	Lipoate	0.16	1.24
534	0.038588	P	4.37	C18H34O5	[FA trihydroxy(18:0)] 9S,12S,13S-trihydroxy-10E-octadecenoic acid	0.42	0.62
524	0.000487	P	4.53	C24H50NO7P	[PC (16:0)] 1-hexadecanoyl-sn-glycero-3-phos-phocholine	5.15	0.92
471	2.67E-09	P	13.71	C14H8O5	Purpurin	0.08	1.15
472	3.54E-08	P	14.35	C16H12O6	Peonidin	0.17	1.35
477	0.003249	P	8.92	C24H28O10	[Fv] Matteucinol 7-O-glucoside	0.69	0.65
Metabolism of Cofactors and Vitamins							
399	0.001159	P	9.55	C10H18N2O3	Dethiobiotin	0.33	0.85
309	3.87E-16	P	10.76	C8H9NO3	Pyridoxal	0.62	0.88
391	4.27E-17	P	6.48	C7H8N2O	1-Methylnicotinamide	0.69	0.88
302	1.18E-13	P	7.60	C8H11NO3	Pyridoxine	0.71	0.75
303	1.12E-12	P	6.23	C8H12N2O2	Pyridoxamine	0.82	0.62
376	4.92E-09	N	15.50	C6H8O4	2,3-Dimethylmaleate	0.83	0.68
327	0.006572	P	5.60	C6H9NOS	5-(2-Hydroxyethyl)-4-methylthiazole	2.28	0.69
Nucleotide Metabolism							
216	1.55E-08	P	13.00	C4H4N2O2	Uracil	0.08	1.15
289	3.09E-05	P	8.95	C10H13N5O4	Adenosine	0.30	1.02
261	0.000623	N	42.56	C4H6N2O4	3-Oxo-3-ureidopropanoate	0.48	0.83
252	1.37E-07	P	10.63	C5H8N2O2	5,6-Dihydrothymine	0.77	0.74
Xenobiotics Biodegradation and Metabolism							
168	3.76E-07	N	16.03	C6H6O5	[FA methyl,oxo(5:0/2:0)] 2-methylene-4-oxo-pen-tanedioic acid	0.09	1.21
208	4.80E-07	P	9.34	C7H9NO3	2-amino-5-methyl-muconate semialdehyde	0.24	1.19
213	0.00044	N	15.28	C5H4O2	Protoanemonin	0.54	0.85

54	0.000518	N	42.53	C8H9N3O2	N-Acetylisoniazid	0.54	0.77
161	1.25E-07	N	14.93	C7H8N4O3	1-7-Dimethyluricacid	0.69	0.77
39	7.78E-11	P	6.06	C6H7NO	4-Hydroxyaniline	0.82	0.65

4.4 Discussion

In this study, the effect on the *E. coli* metabolome following incubation in three different dietary fibres was investigated. These fibres were 1% cooked meat, 1% maize meal and 1% olive kernel oil and each of these was compared to the negative control cultures enriched with 1% D glucose. The key observation is that there were significant effects on various metabolite pathways particularly those associated with amino acid, lipid, carbohydrate, and nucleotide metabolism. In addition, there were effects on intermediates of peptide and polyketide biosynthesis, as well as on xenobiotic breakdown products and vitamin cofactors.

Most importantly, it was revealed that although the pathways affected were the same, the different fibre sources imparted different effects on these pathways in terms of whether the metabolites were down regulated or up regulated. For instance, unlike incubation with 1% cooked meat, most of the metabolites in the amino acid metabolic pathway were significantly lowered by treatment with 1% olive kernel oil or 1% maize meal. However, in the case of carbohydrate metabolism, the overall effect on the pathway was essentially the same irrespective of the dietary fibre. That is, all the three fibres led to significant reduction in the levels of specific metabolites in this pathway.

Regarding effect on lipid metabolism, olive kernel oil produced significant alterations in relatively fewer metabolites compared to those observed with either 1% cooked meat or 1% maize meal fibres. Surprisingly, whereas 1% cooked meat and 1% maize

meal each led to a general increase in the levels of most of the altered lipids, olive kernel oil was associated with lowered lipid content.

Another crucial difference was observed in nucleotide metabolism where unlike 1% cooked meat, the 1% maize meal did not induce significant alterations in the levels of these metabolites. Moreover, whereas nucleotide metabolites were significantly increased in cultures enriched with 1% cooked meat, those enriched with olive kernel oil revealed significant decreases in these metabolites.

Taken together, these findings suggest that *E. coli* metabolome is closely associated with the type of fibre that the microorganism is subjected to. This observation is consistent with many previous studies which suggest that the metabolome of the gut microbiota depends on the microbial composition and functional capacity of the gut microbiota, as well as nutrient availability and its physicochemical properties, age of host, and transit time of the colon. Nutrient availability, particularly the carbohydrate to nitrogen ratio, is believed to be the most important regulator of bacterial metabolism, as it influences preference of saccharolytic vs proteolytic fermentation (De Preter *et al*., 2011).

Chapter Five:

Metabolomic analysis of the effects of different dietary fibres in incubations with fecal samples taken from Crohn's disease patients

5 METABOLOMIC ANALYSIS OF THE EFFECTS OF DIFFERENT DIETARY FIBRES IN CROHN'S DISEASE

5.1 Introduction

The normal flora within the gastrointestinal tract produces a number of compounds through the process of fermentation of nutrients and xenobiotics. The latter are compounds of non-host origin that normally enter the GIT with the diet or are produced by the microbiota. Some of these metabolites produced in the gut are excreted in faeces and others are absorbed through the colonic mucosa and enter the systemic circulation where they can be further modified by human metabolism (De Preter and Verbeke, 2013).

It has been reported, for instance, that the bacterial fermentation of the aromatic amino acid tyrosine in the colon yields the product *p*-cresol which is entirely absorbed into the host's systemic circulation. First pass effect involving conjugation of *p*-cresol in the colon mucosa or within the liver leads to formation of *p*-cresol sulphate or *p*-cresol glucuronide which renders the compound highly water soluble for excretion through the kidney as part of urine (Evenepoel *et al*., 2009).

On the other hand, some metabolites derived from the host's systemic circulation can be returned to the gut through biliary excretion whereupon they are further metabolised by gut microbiota leading to newer metabolites that are consequently reabsorbed back into the host's circulatory system. This latter process occurs for bile

acid metabolites that escape absorption in the terminal ileum which are first deconjugated and the converted to secondary bile acids by microbial metabolism within the colon (Bajor *et al*., 2010).

These host-microbiota metabolic interactions tend to make interpretation of metabolite profiles rather complicated. In addition, the outcome of metabolomic analyses associated with gut microbiota depends on the type of biomatrix selected for analysis. Thus, microbial metabolism is more comprehensively reflected within the faecal metabolome than in the urinary, serum or breath metabolomes (De Preter and Verbeke, 2013). It has been suggested that whereas urinary profiles reflect both human and human-microbial co-metabolites, the serum profiles are normally less enriched with products of microbial metabolism. Nevertheless, the commonest biomatrices used in assessing gut microbial metabolomes are faeces, urine and serum.

In GIT diseases such as inflammatory bowel disease (IBD) and inflammatory bowel syndrome (IBS), it has been suggested that disease initiation and progression are associated with a dysbiosis of the microbiota (De Preter and Verbeke, 2013). In these disease states, studies have identified that there is a disproportion of the predominant bacteria in faecal samples (Mahowald *et al*., 2009, Joossens *et al*., 2011, Krogius-Kurikka *et al*., 2009). For instance, both IBD and IBS are frequently associated with a reduction in the diversity and abundance of Firmicutes. Thus, in order to understand disease pathogenesis in both IBD and IBS, it is necessary to determine the unique differences between metabolic activities of microbiota in these disease states in comparison with healthy controls (McNiven *et al*., 2011, Olivares *et al*., 2013). Such a

measure could also be employed as a diagnostic tool in these conditions (De Preter and Verbeke, 2013).

Inflammatory bowel disease consists of two phenotypes, namely: Crohn's disease (CD) and ulcerative colitis (UC). The two phenotypes share similar features in their pathophysiology and clinical presentation but their therapeutic management and prognosis are quite different. The two disease manifestations are influenced by hereditary factors as well as microbial and environment factors. Current diagnostic criteria for IBD rely on clinical features, endoscopy, radiologic and histological examination all of which require the disease to be in an advanced stage for accurate diagnosis to be possible. An alternative diagnostic tool based on metabolomics would be less invasive as it would rely on convenient biofluids such as urine, serum, or faeces, and it would offer additional advantages over current strategies based on its potential for primary diagnosis, disease surveillance, and early detection of relapses (De Preter and Verbeke, 2013). In addition to discovering new biomarkers, metabolomics is increasingly being considered in improving stratification of IBD patients into the different subtypes.

Among the biomarkers previously reported in IBD and IBS studies and which have been tested in clinical trials include faecal markers (such as lactoferrin, calprotectin, and PMN-elastase), acute phase proteins such as C-reactive protein, and serological markers (antibodies against luminal antigens and anti-glycan antibodies) (Dotan, 2010). Metabolomics has been used to discriminate IBD patients from healthy controls, CD from UC patients, and patients with active disease from those in remission.

These trends suggest that there is a growing acceptance of the role of metabolomics in the understanding of the gut microbiota and its influence on the host's urinary, faecal and serum metabolomes through metabolite profiling studies.

The previous methods used in IBD studies of host-microbiome metabolomes have employed mainly ^{1}HNMR (Marchesi *et al*., 2007, Le Gall *et al*., 2011, Williams *et al*., 2009, Schicho *et al*., 2012, Stephens *et al*., 2013, Bjerrum *et al*., 2009, Balasubramanian *et al*., 2009, Sharma *et al*., 2010, Bajpai *et al*., 1975, Zhang *et al*., 2013) and GC-MS (Walton *et al*., 2013, Ooi *et al*., 2011, Öhman and Simrén, 2013, Ahmed *et al*., 2013) analytical platforms, although techniques such as ICR-FT/MS (Jansson *et al*., 2009) and AA analyser (Hisamatsu *et al*., 2012) have also been reported. Marchesi *et al* (Marchesi *et al*., 2007) was the first to differentiate IBD patients from healthy controls based on 1H-NMR analysis of aqueous extracts of faecal samples and to differentiate CD from UC patients. In that study, dysbiosis—a process of disruption of normal bacterial ecology in the gut—was characterised based on depletion of bacterial metabolites such as short chain fatty acids (SCFA), dimethylamine and trimethylamine, suggesting that alteration of gut microbiota was either a cause or consequence of IBD. A number of other studies (Marchesi *et al*., 2007, Le Gall *et al*., 2011, Williams *et al*., 2009, Schicho *et al*., 2012, Stephens *et al*., 2013, Bjerrum *et al*., 2009, Balasubramanian *et al*., 2009, Sharma *et al*., 2010, Bajpai *et al*., 1975, Zhang *et al*., 2013, Walton *et al*., 2013, Ooi *et al*., 2011, Öhman and Simrén, 2013, Ahmed *et al*., 2013, Jansson *et al*., 2009) have revealed that CD patients can be clearly discriminated from healthy controls based on the metabolite profiles of faecal or urine samples, and that patients with predominantly ileal involvement of the disease can

be separated from those in whom the disease predominantly involves the colon (De Preter and Verbeke, 2013).

It can be seen from the foregoing that despite the successes of the analytical platforms reported in earlier metabolomic studies of IBD, LC-MS has not yet been significantly involved insofar as analysis of gut microbiota is involved. This might be due to the low molecular weights of most metabolites associated with gut microbial fermentation and the fact that such metabolites do not easily ionise which is a key criterion for mass detection. However, such shortcomings have been overcome in other related studies by means of derivatisation prior to analysis (Bawazeer *et al*., 2016).

5.2 Materials and Methods

5.2.1 Chemicals and Solvents

HPLC grade Acetonitrile (ACN) was purchased from Fisher Scientific (Loughborough, UK) and HPLC grade water was produced by a Direct-Q3 UltrapureWater System (Millipore, Watford, UK). AnalaR-grade formic acid (98%) was obtained from BDH-Merck (Poole, UK). Authentic stock standard metabolites (Sigma-Aldrich, Poole, U.K.) were prepared as previously described (Zhang *et al*., 2014) and diluted four times with ACN before LC-MS analysis. Ammonium acetate was purchased from Sigma-Aldrich (Poole, UK).

5.2.2 Participants

Faecal homogenates from three patients with Crohn's disease and 3 normal controls were incubated with seven different dietary fibres (including a blank) and the resulting metabolomics profiles were determined at 0, 24 and 48 hours post resulting in 126 samples. The dietary fibres used in the incubation are shown in table 5.1

Table 5.1 Dietary fibres used in the incubations

Number	Fibre
1	Blank
2	Hi-Miaze
3	Pectin
4	Raftulose
5	Wheat bran
6	Cellulose
7	Mixed fibre

5.2.3 Sample preparation

Exactly 200 µL of the sample was mixed with 800 µL of acetonitrile and then centrifuged for 10 min before transferring into a vial with an insert. The pooled sample was prepared by pipetting 50 µL from each of the 126 samples and then mixing them together before being diluting 0.2 ml of the pooled sample with 0.8 ml of acetonitrile and centrifuging for 10 minutes, then transferring the supernatant into HPLC vials for analysis. Additionally, the prepared mixtures of authentic standard metabolites (Zhang *et al*., 2014) were run in the same sequence for purposes of metabolite identification within the samples.

5.2.4 LC-MS conditions

Liquid chromatographic separation was carried out on an Accela HPLC system interfaced to an Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) using a ZIC-pHILIC column (150 × 4.6 mm, 5 µm, HiChrom, Reading UK). The column was eluted with a mobile phase consisting of 20 mM ammonium carbonate in HPLC-grade water(solvent A) and acetonitrile (solvent B), at a flow rate of 0.3 mL/min. The elution gradient was an A:B ratio of 20:80 at 0 min, 80:20 at 30 min, 92:8 at 35 min and finally 20:80 at 45 min. The nitrogen sheath and auxiliary gas flow rates were maintained at 50 and 17 arbitrary units. The electrospray ionisation (ESI) interface was operated in both positive and negative modes. The spray voltage was 4.5 kV for positive mode and 4.0 kV for negative mode, while the ion transfer capillary temperature was 275°C. Full scan data were obtained in the mass-to-charge ratio (m/z) range of 75 to 1200 for both ionisation modes on the LC-MS system fully calibrated according to manufacturer's guidelines. The resulting data were acquired using the XCalibur 2.1.0 software package (Thermo Fisher Scientific, Bremen, Germany).

5.2.5 Data extraction and analysis

Data extraction for each of the samples was carried out by MZMine software. The extracted ions, with their corresponding m/z values and retention times, were pasted into an Excel macro of the most common metabolites prepared in–house to facilitate identification. The lists of the metabolites obtained from these searches were then carefully evaluated manually by considering the quality of their peaks and their retention time match with the standard metabolite mixtures run in the same sequence.

All metabolites were within 3 ppm of their exact masses. Statistical analyses were performed using both univariate with Microsoft Excel and multivariate approaches using SIMCA-P software version 14.1 (Umetrics, Umea, Sweden.).

5.3 Results

5.3.1 Unsupervised analysis

Figure 5.1 shows the PCA-X analysis of all the samples from the three normal patients with different fibres and at different time points. PCA-X, an unsupervised model in SIMCA-P, produces a natural scatter of the samples based on their characteristic metabolic footprints. It can be seen in the figure (samples coloured according to time of collection for convenience) that, in general, there are two major clusters, namely, at the time = 0 hrs on the one hand, and that of the time = 24 or 48 hours on the other, with just two outliers. This implies that, for these participants at least, time appears to be the most important classifying variable. However, there are no significant differences between individual dietary fibres as shown in Figure 5.2.

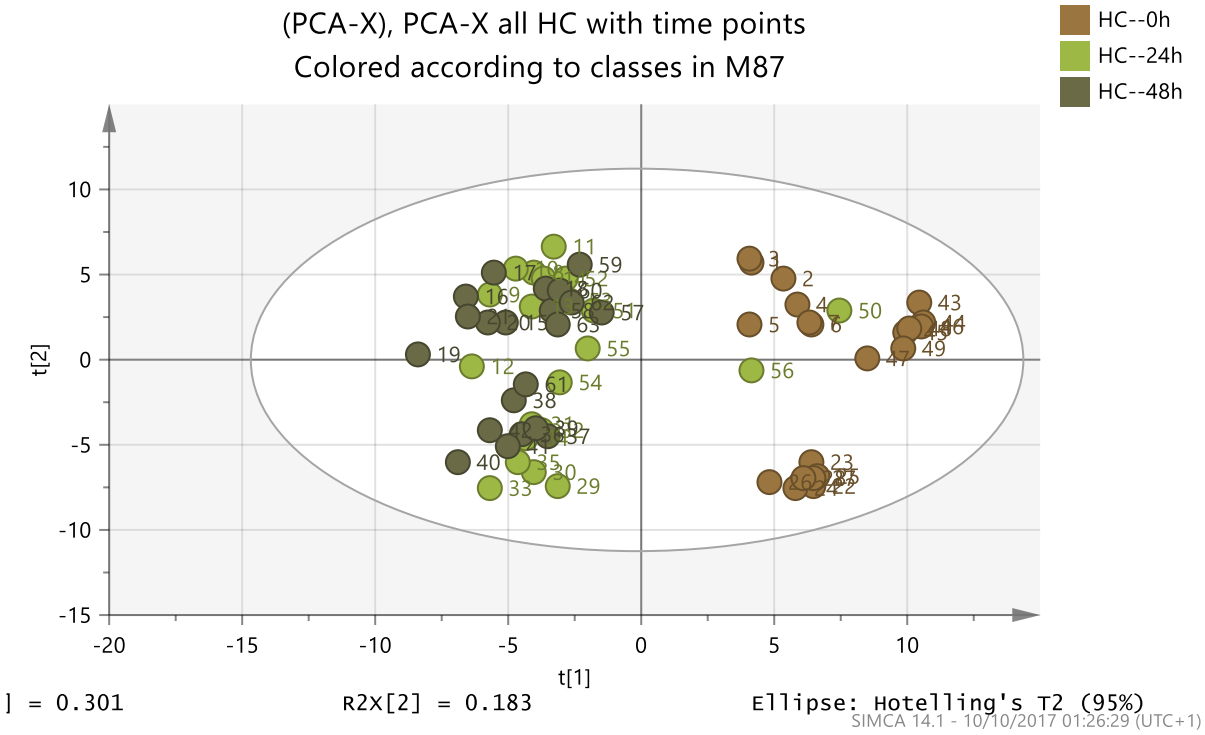

Figure 5.1: PCA-X analysis of the metabolomics footprint of the 63 samples from healthy controls showing that time of sample collection plays a key role in sample classification according to the model. Circles coloured according to the time of sample collection post diet. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

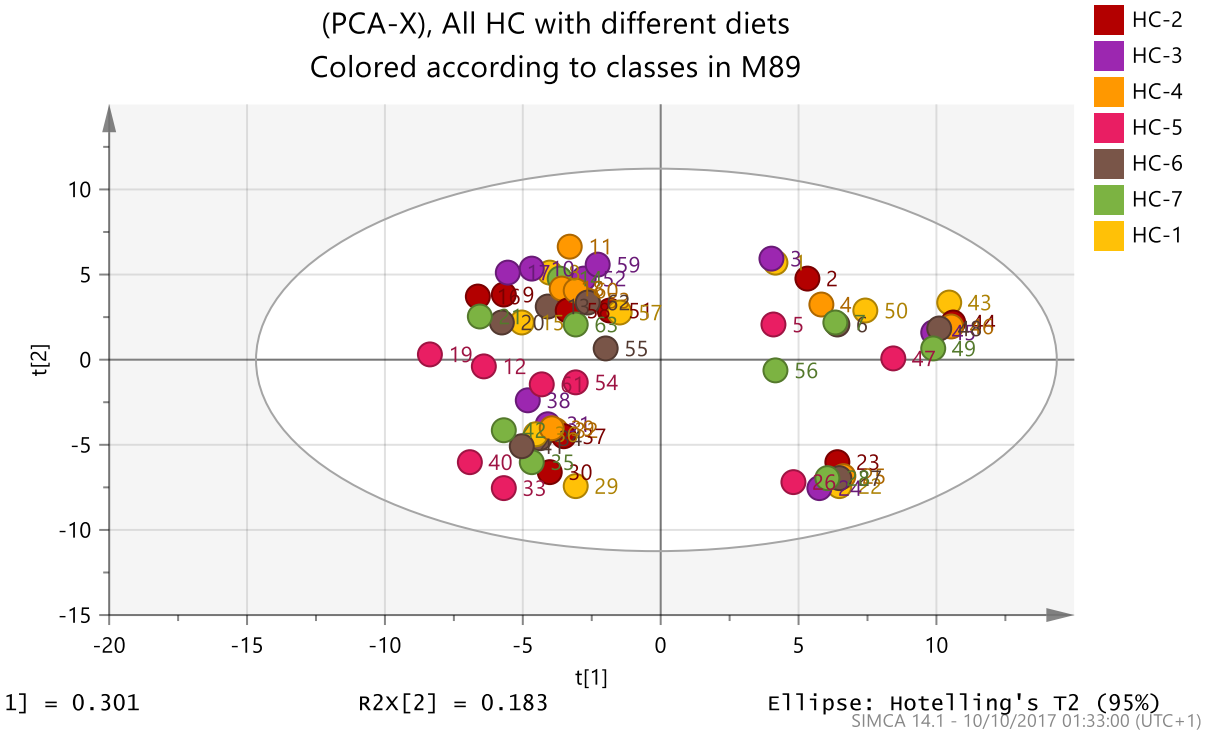

Figure 5.2: PCA-X analysis of the metabolomics footprint of the 63 samples from HC participants showing that diet does not significantly affect sample clustering as each cluster contains all the 7 dietary fibres. Circles coloured according to the dietary fibres. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

On the other hand, samples from patients with Crohn's disease (CD) showed a different pattern of scatter, revealing clusters that could not be wholly accounted for on the basis of time of sample collection. A total of five groupings was formed, two of which consisted entirely of samples at time zero, two comprising both time 24 and 48 samples, and the final one encompassing all three time points, as shown in Figure 5.3. Looking at the samples more closely, it appears that, apart from the time factor, patient differences are also at play. For instance, considering the blue circles (time zero samples), each of the three clusters corresponds to a given patient, as there are also three patients. In addition, the three mixed (24 and 48) clusters correspond to

each of the three CD patients. Thus, in this case, both time and patient individualities have combined to produce a natural clustering in which dietary effects are less noticeable, as can be deduced from Figure 5.4.

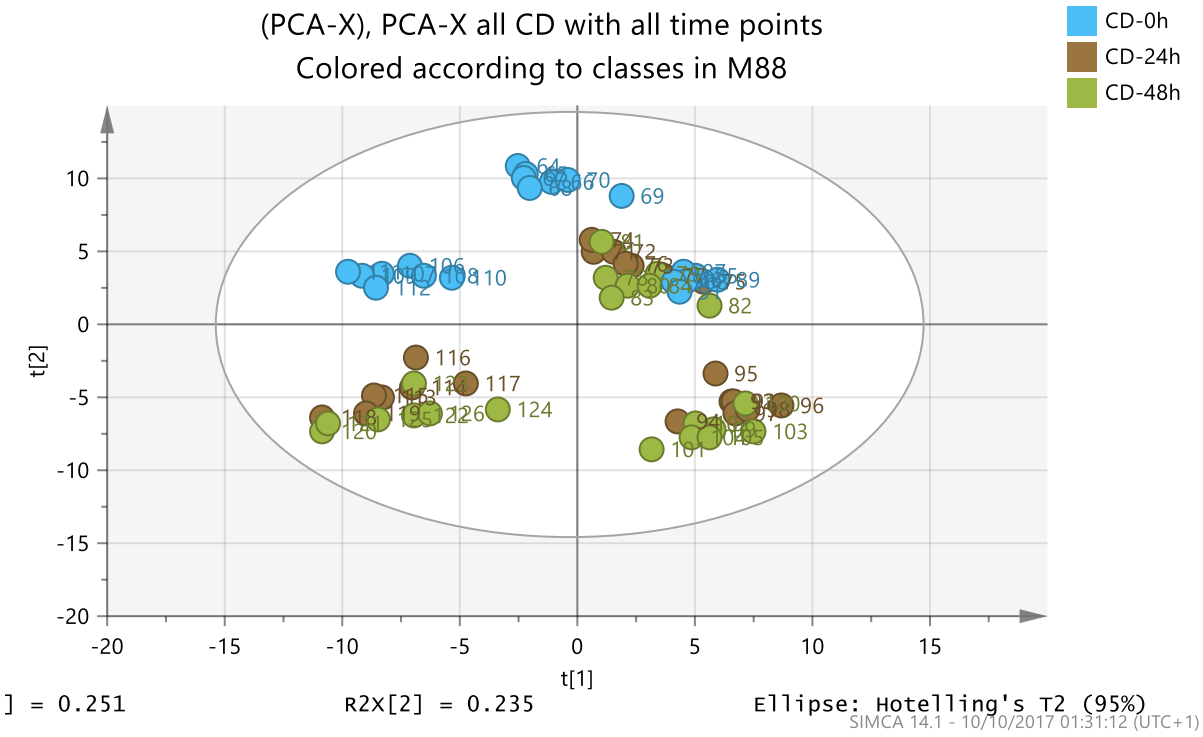

Figure 5.3: PCA-X analysis of the metabolomics footprint of the 63 samples from CD patients in which both time of sample collection and individual patient peculiarities combine to define the model. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

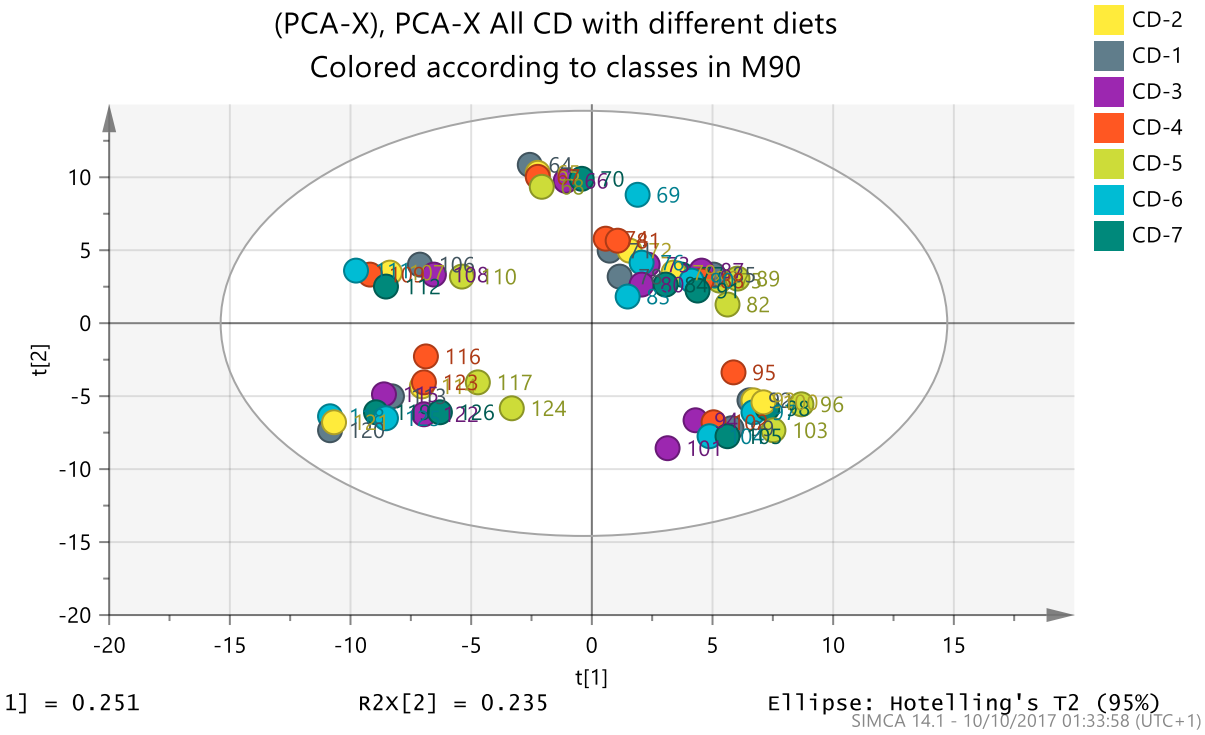

Figure 5.4: PCA-X analysis of the metabolomics footprint of the 63 samples from CD patients. Each of the five natural groupings has representations from all 7 diets, implying that diet did not have much contribution to the scatter. As observed in Figure 3, time and individual patient peculiarities had more significant contribution to the observed clustering pattern. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

Figure 5.5 shows a PCA-X when all the 126 samples were considered altogether. The fact that CD samples were separated from the HC samples (with the exception of two outliers) shows that there are clear metabolomics differences associated with CD.

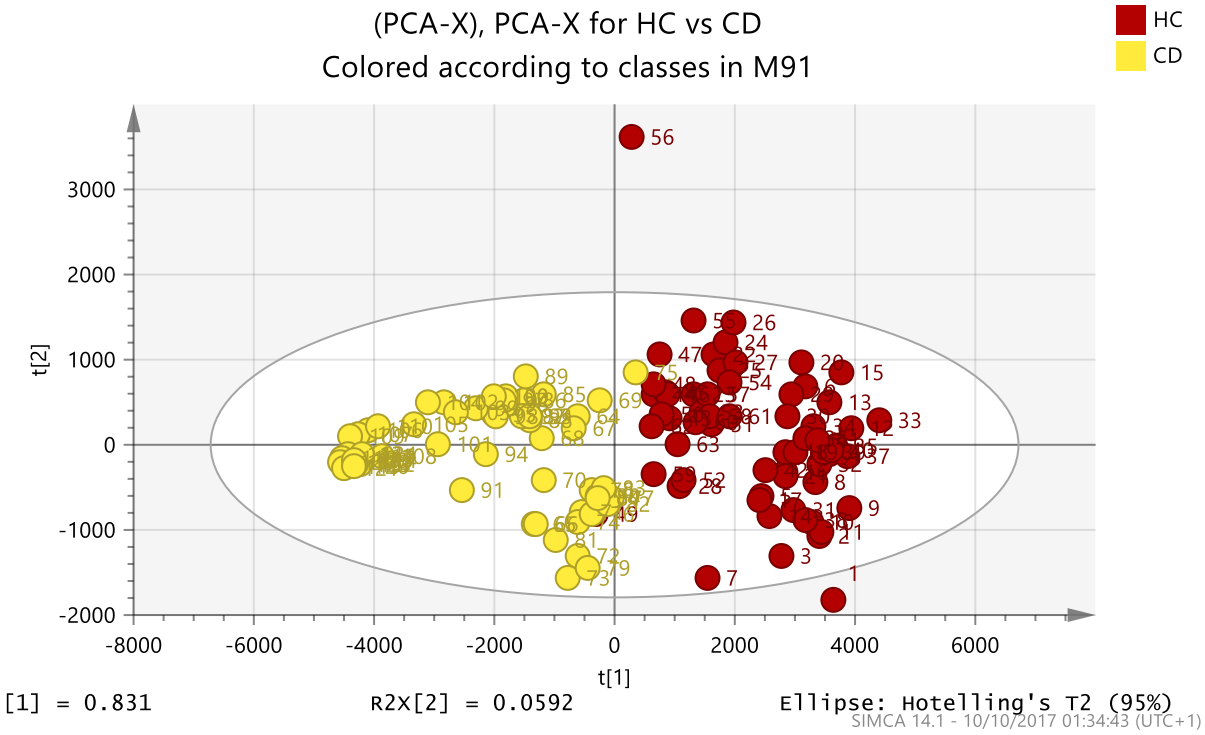

Figure 5.5: PCA-X analysis of all 126 samples showing near complete separation of CD and HC samples. There are two outliers in the model (sample 49 and sample 75) which are misclassified. Samples coloured according to whether they are CD or HC. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

Basing on the findings from unsupervised models described above, it is apparent that the main factors contributing to the metabolomic footprints of the sample which in turn determine their clustering in the scatter plot are the disease status (CD or HC), time of sample collection (0, 24, 48 h) and, especially in the case of CD samples, the individual patient differences.

5.3.2 Supervised analysis

Supervised models enable identification of metabolites that have the most significant contribution to a clustering pattern. Based on the observations from PCA-X described above, it is apparent that metabolomics profiles of samples do not change between 24 and 48 h, and so these have been combined to form one group which was then compared with the group representing samples at 0 h. As seen in Figures 5.6 and 5.7, the OPLS-DA was able to clearly separate the two groups in both healthy controls and CD participants respectively. Both models were found to be valid based on cross validation as shown in Figures 5.8 and 5.9 for HC and CD samples respectively.

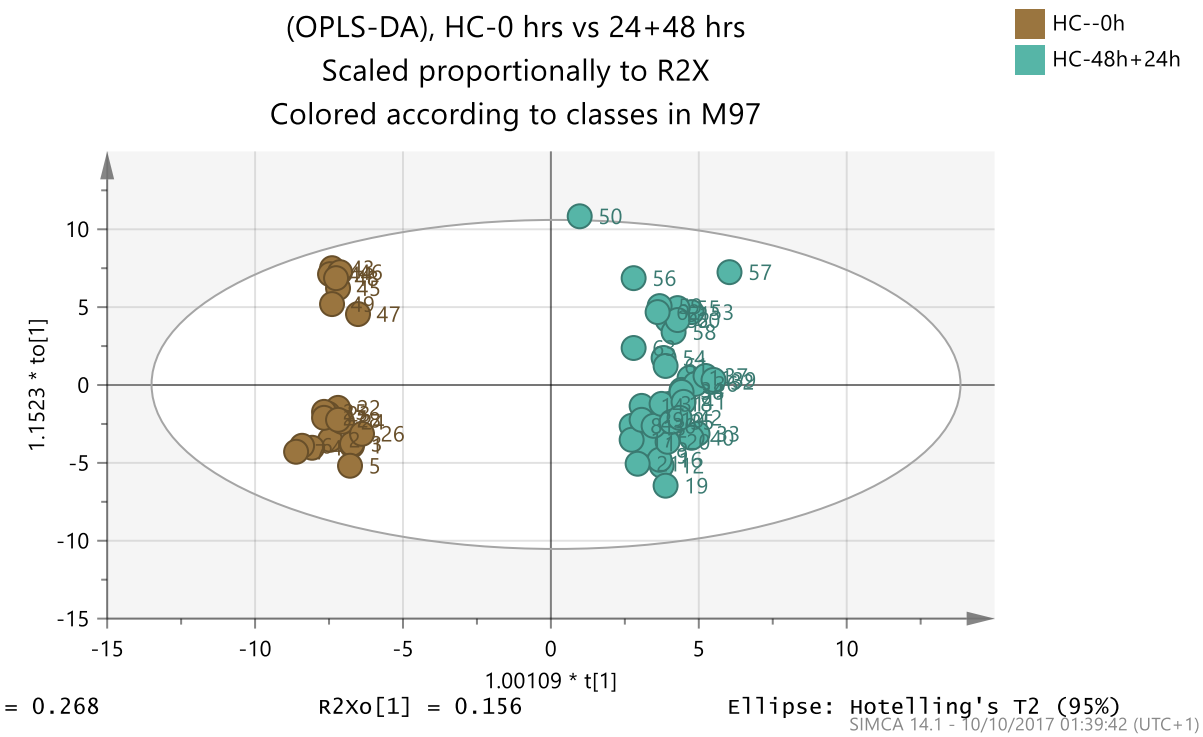

Figure 5.6: OPLS-DA analysis to compare samples at 0h with those at 24 and 48 h in healthy controls. CA-ANOVA =1.9506e-029. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

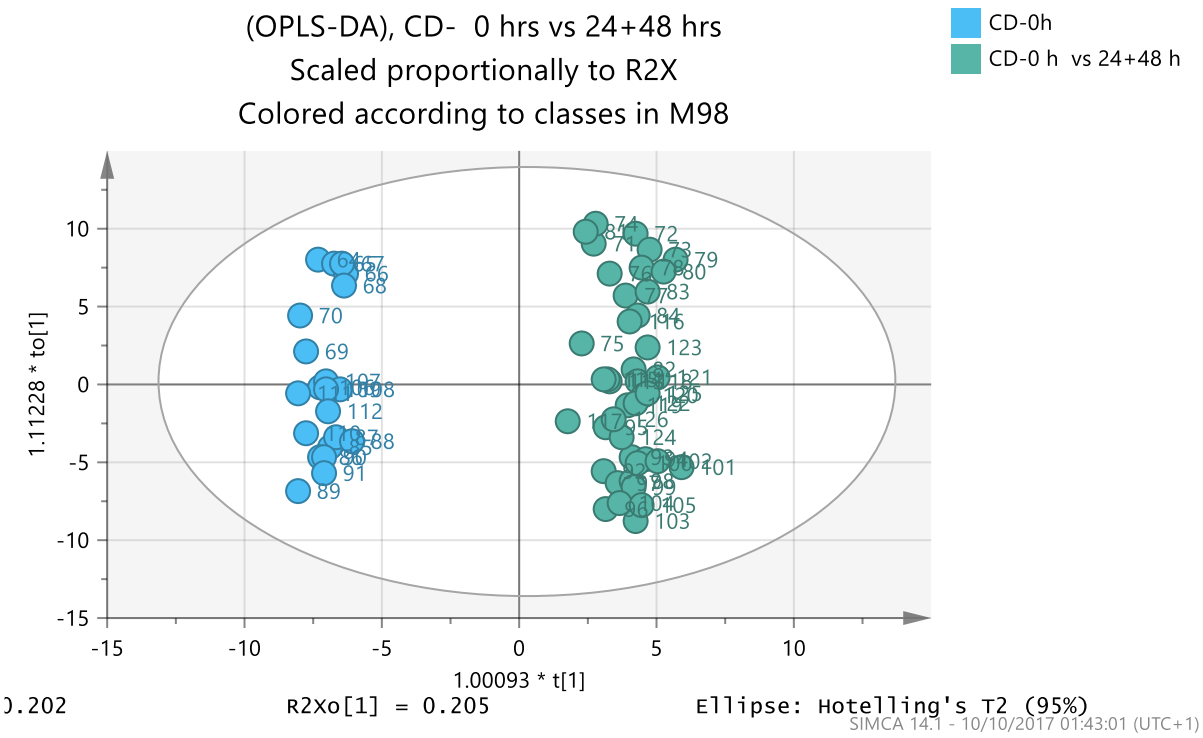

Figure 5.7: OPLS-DA analysis to compare samples at 0h with those at 24 and 48 h in CD patients. CV-ANOVA = 3.08413e-035. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

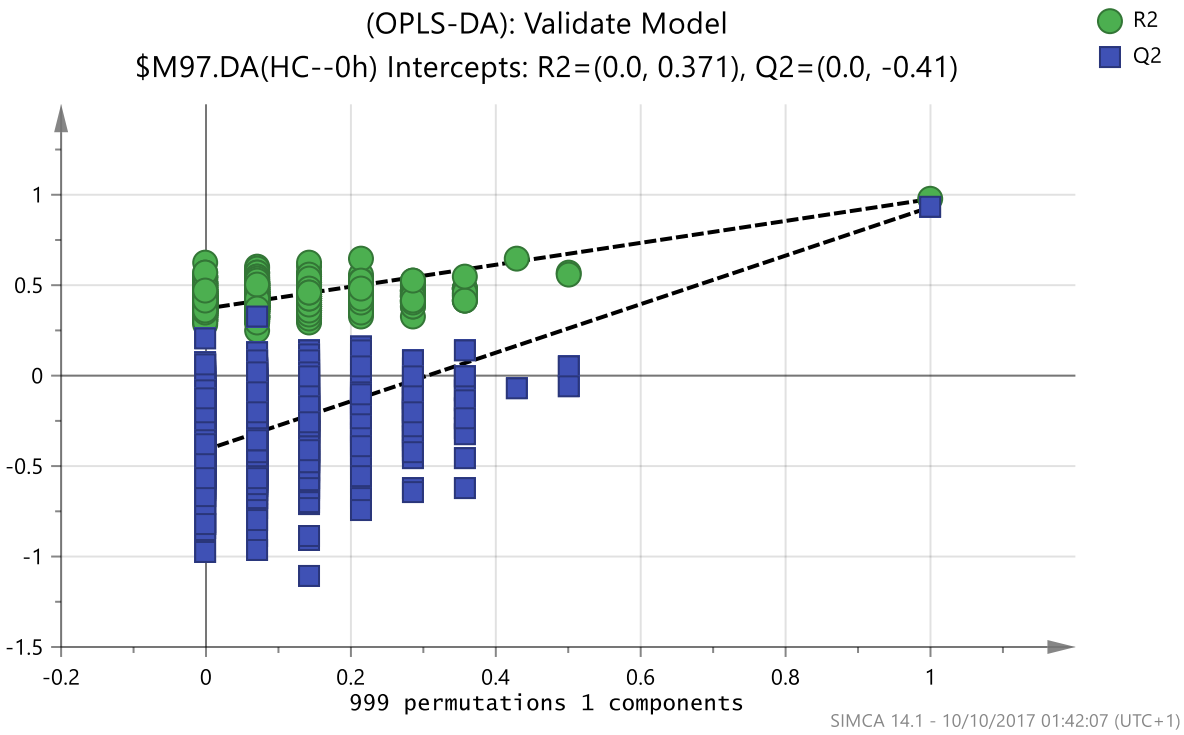

Figure 5.8: Cross-validation of the model for healthy controls.

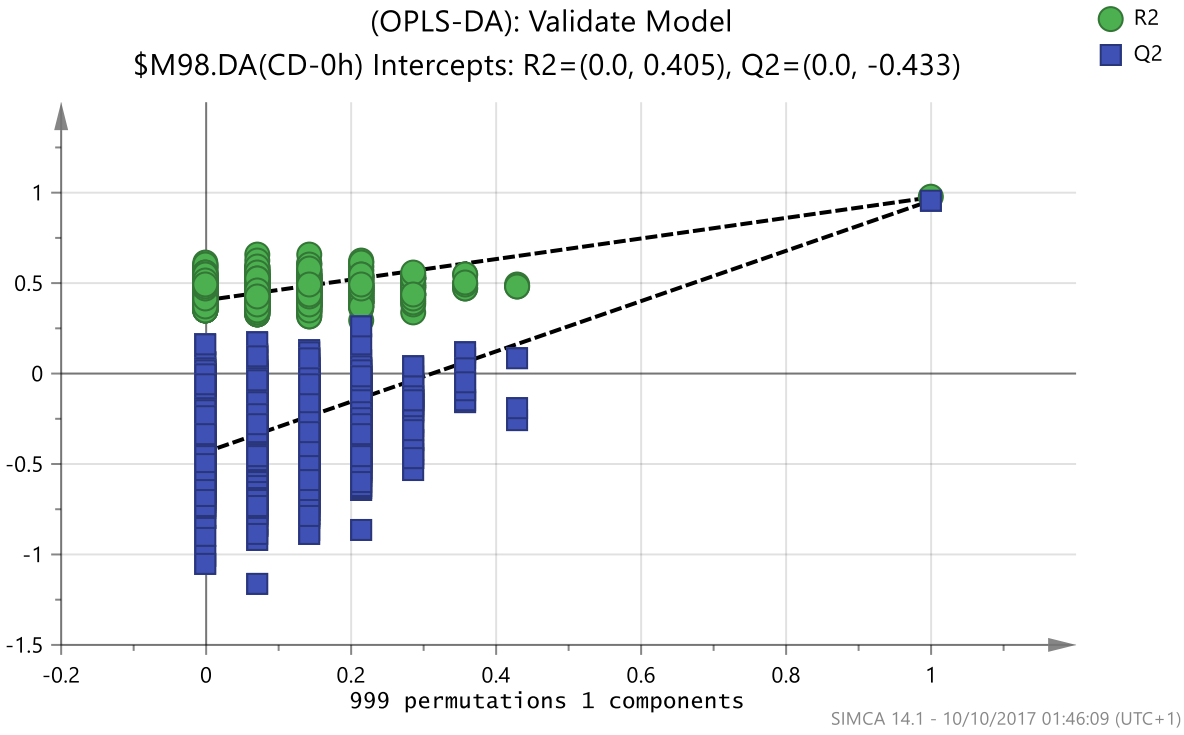

Figure 5.9: Cross-validation of the model for CD patients.

The OPLS-DA analysis of the 126 samples to compare metabolites between CD and HC showed clear separation as shown in Figure 5.10. Besides the between group separation, there were subclusterings based on time of collection which was more pronounced for CD samples. Modelling the samples with respect to time separately it can be seen that both the CD and HC samples separate into 0 h and 24/48 h hour groups (Figures 5.11 and 5.12).

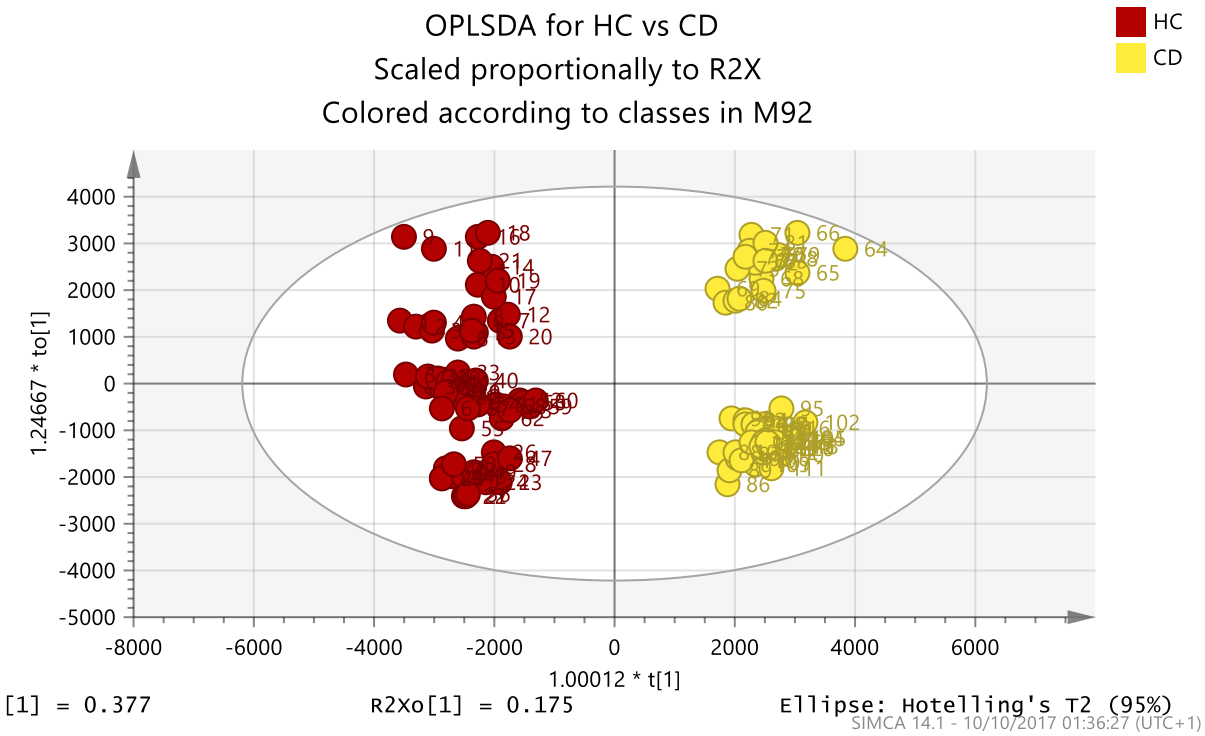

Figure 5.10: OPLS-DA analysis of the 126 samples showed clear separation between CD and HC groupings but with subgroupings based on time of collection (especially in CD samples) and possibly due to patient differences. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

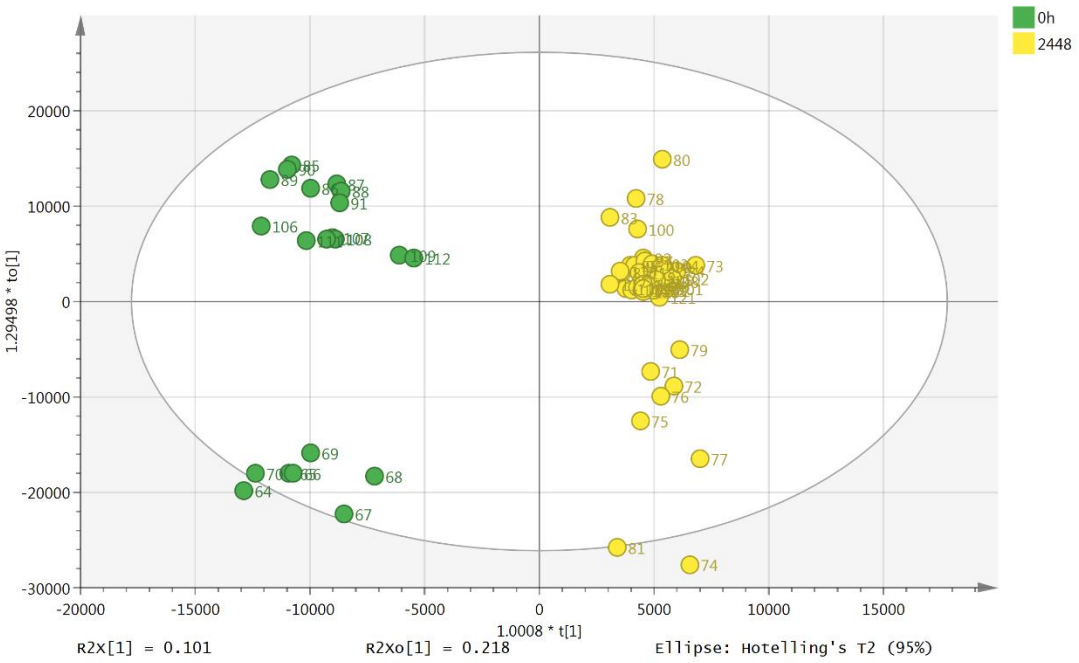

Figure 5.11: OPLS-DA analysis of the 63 samples showing clear separation between CD time 0 and CD 24/48h. The y-axis represents variation within the groups while the x-axis represents variation between the groups. The separation illustrates that the effect of time on the metabolite profiles in the samples from Crohn's disease patients is likely to be significant between 0 and 24 hours and 0 and 48 hours of incubation with the dietary fibres.

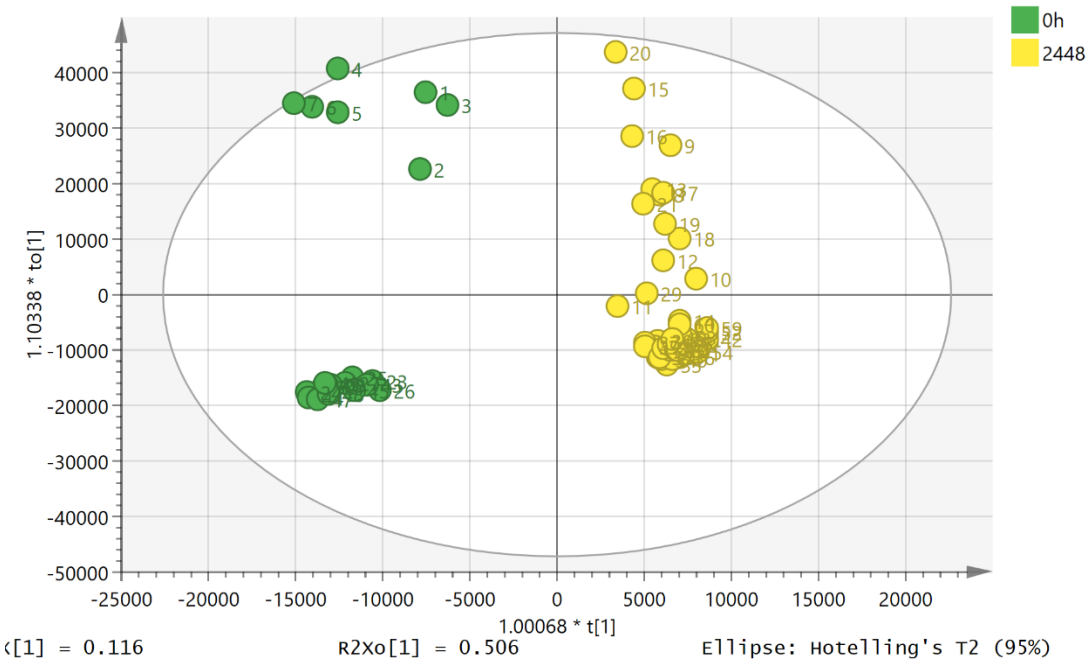

Figure 5.12: OPLS-DA analysis of the 63 samples showing clear separation between HC time 0 and HC 24/48h. The y-axis represents variation within the groups while the x-axis represents

variation between the groups. The separation illustrates that the effect of time on the metabolite profiles of the samples from health controls is likely to be significant between 0 and 24 hours and 0 and 48 hours of incubation with the dietary fibres.

5.3.3 Metabolite alterations

Table 5.1 shows that most of the metabolites altered in CD relative to HC patients belonged to the pathways of amino acids metabolism, biosynthesis of secondary metabolites, carbohydrate and lipids metabolism, metabolism of cofactors and vitamins, and xenobiotics biodegradation and metabolism. Except for L-cystine and aconitine which were raised in CD, all the metabolites were significantly decreased in CD compared to HC.

Table 5.2 shows a comparison of the metabolite levels at 24+48 h on the one hand and those at 0h on the other, within the healthy controls. It can be observed that a number of metabolites significantly differed between the two groups and these belonged to pathways of amino acid metabolism, carbohydrate metabolism, nucleotide metabolism, polyketides, vitamins and cofactors, and xenobiotics. In addition, most of these metabolites are higher at 24+48 h compared to those at 0 h except for a few such as N_2-(D-1-carboxyethyl)-L-arginine, N_2-acetyl-L-aminoadipate and a host of bilirubin intermediates (bilirubin, L-urobilin, D-urobilinogen) which were lowered at 24+48 h compared to zero.

In Table 5.3, a comparison of the metabolites in CD at time 24+48h vs. time 0h revealed varied responses in metabolite levels. Indole-3-acetaldehyde, ethanolamine phosphate, N-acetylornithine and 2-hydroxy-6-oxonona-2,4-diene-1,9-dioate (all

amino acid metabolism) and dihydroceramide and sphinganinewere (lipid metabolites) were significantly raised while the rest were lowered compared to baseline levels at 0h.

Table 5.1: The below table showing the comparisons of healthy patients with Crohn's disease (CD vs HC)

Rt min	MW	Molecular Formula	Metabolite	Ratio (CD/HC)	P value	VIP
Fatty Acids						
4.1	200.178	$C_{12}H_{24}O_2$	Dodecanoic acid	0.060	2.7E-06	3.770
4.0	228.209	$C_{14}H_{28}O_2$	Tetradecanoic acid	0.258	4.8E-09	2.954
4.2	302.225	$C_{20}H_{30}O2$	Eicosapentaenoic acid	0.047	7.9E-07	0.303
4.0	280.24	C18H32O2	Linoleate	0.532	1.9E-05	0.207
3.9	282.256	C18H34O2	Octadecenoic acid	0.496	1.2E-06	0.223
4.2	320.235	C20H32O3	Hydroxyeicosatetraenoic acid	0.261	3.1E-06	0.130
4.1	338.246	C20H34O4	Dihydroxyeicosatrienoic acid	0.392	7.3E-06	0.159
Haemoglobin Metabolism						
10.3	226.095	C10H14N2O4	Porphobilinogen	2.397	3.5E-04	0.327
4.3	562.258	C34H34N4O4	Protoporphyrin	1.875	8.2E-03	0.160
4.0	584.263	C33H36N4O6	Bilirubin	13.793	8.0E-04	0.244
4.8	590.31	C33H42N4O6	D-Urobilinogen	15.911	1.9E-06	1.154
4.6	588.295	C33H40N4O6	D-Urobilin	14.176	2.5E-03	0.363

7.6	594.342	C33H46N4O6	L-Urobilin	0.115	1.7E-11	0.294
Tryptophan Metabolism						
6.4	136.064	C7H8N2O	1-Methylnicotinamide	0.350	2.4E-07	0.129
7.3	152.059	C7H8N2O2	N1-Methyl-2-pyridone-5-carboxamide	2.323	1.1E-05	0.143
5.2	165.043	C8H7NO3	Formylanthranilate	0.312	7.9E-22	0.165
8.2	180.09	C9H12N2O2	5-Hydroxykynurenamine	0.347	0.0E+00	0.161
5.8	218.106	C12H14N2O2	N-Acetylserotonin	0.163	8.5E-20	0.150
6.7	236.08	C11H12N2O4	L-Formylkynurenine	17.791	1.3E-05	0.524
7.0	248.116	C13H16N2O3	6-Hydroxymelatonin	0.151	7.9E-29	0.168
Phenyl Alanine Metabolism						
20.7	137.084	C8H11NO	Tyramine	0.303	1.6E-24	0.146
7.3	165.079	C9H11NO2	L-Phenylalanine	0.436	1.6E-22	0.237
9.2	197.069	C9H11NO4	3,4-Dihydroxy-L-phenylalanine	10.301	1.3E-06	0.545
10.8	207.09	C11H13NO3	N-Acetyl-L-phenylalanine	0.105	2.1E-17	0.233
Purines and Pyrimidines						
8.5	126.043	C5H6N2O2	Thymine	0.463	1.2E-12	0.176

6.3	128.059	C5H8N2O2	5,6-Dihydrothymine	0.141	6.2E-13	0.208
9.5	135.055	C5H5N5	Adenine	0.339	1.2E-08	1.876
8.9	267.097	C10H13N5O4	Adenosine	0.232	1.8E-06	0.478
6.2	278.127	C14H18N2O4	alpha-Ribazole	0.376	1.2E-14	0.159
Creatine Metabolism						
9.8	113.059	C4H7N3O	Creatinine	10.746	3.0E-05	1.979
16.4	117.054	C3H7N3O2	Guanidinoacetate	3.103	5.5E-05	0.149
Bile Acids/Cholesterol Metabolism						
4.1	384.339	C27H44O	Cholestadienol	0.092	5.1E-07	0.496
4.7	392.293	C24H40O4	Cholanoic Acid	2.358	1.3E-02	2.362
4.0	412.371	C29H48O	Dimethyl-cholestadienol	0.242	5.2E-09	0.258
4.0	410.355	C29H46O	Dimethylcholestatrienol	0.395	3.0E-10	0.160
4.2	416.329	C27H44O3	Cholestatrienetriol	0.163	1.0E-10	0.147
Miscellaneous						
6.0	131.095	C6H13NO2	L-Leucine	0.492	6.0E-13	0.314
10.1	140.059	C6H8N2O2	Methylimidazoleacetic acid	0.157	1.8E-18	2.919

12.1	161.105	C7H15NO3	L-Carnitine	0.209	6.0E-07	0.389
9.7	167.058	C8H9NO3	Pyridoxal	0.258	2.0E-30	0.273
13.9	165.046	C5H11NO3S	L-Methionine S-oxide	0.362	7.5E-04	0.462
11.7	169.074	C8H11NO3	Pyridoxine	0.291	2.2E-15	0.334
13.2	169.085	C7H11N3O2	Methyl-L-histidine	1.818	3.2E-05	0.124
11.5	203.079	C8H13NO5	N2-Acetyl-L-aminoadipate	3.545	5.8E-04	0.265
13.8	232.106	C9H16N2O5	N2-Succinyl-L-ornithine	0.158	1.6E-26	0.168
5.4	299.282	C18H37NO2	[SP] 3-dehydrosphinganine	2.292	2.4E-04	0.253
6.4	301.298	C18H39NO2	Sphinganine	0.444	8.2E-07	0.274

Table 5.2: Significant metabolite differences based on time points (0 hrs vs. 24+48 hrs) among the Crohn's disease (CD) **Table 5.3:** Significant metabolite differences based on time points (0 hrs vs. 24+48 hrs) among the Crohn's disease (CD)

Rt min	MW	Molecular Formula	Metabolite	Ratio (CD/HC)	P value	VIP
Fatty Acid Metabolism						
4.1	200.178	C12H24O2	Dodecanoic acid	0.599	2.59E-02	1.299
4.5	399.335	C23H45NO4	Palmitoyl-R-carnitine	0.275	2.80E-03	0.304
Haemoglobin Metabolism						
10.3	226.095	C10H14N2O4	Porphobilinogen	0.512	2.37E-02	0.716
4.3	562.258	C34H34N4O4	Protoporphyrin	0.363	4.17E-06	0.544
4.0	584.263	C33H36N4O6	Bilirubin	0.136	1.94E-04	0.840
4.6	588.295	C33H40N4O6	D-Urobilin	0.083	1.05E-04	1.326
4.4	582.248	C33H34N4O6	Biliverdin	0.201	9.33E-05	0.185
4.4	586.279	C33H38N4O6	(3Z)-Phycocyanobilin	0.137	2.46E-04	0.401
Tryptophan Metabolism						
6.5	139.027	C6H5NO3	6-Hydroxynicotinate	4.510	9.75E-03	0.161
8.0	149.048	C8H7NO2	5,6-Dihydroxyindole	4.935	7.74E-06	0.110

5.3	161.084	C10H11NO	Indole-3-ethanol	26.188	2.61E-08	0.245
6.1	163.063	C9H9NO2	3-Methyldioxyindole	1.673	2.68E-03	0.094
5.2	165.043	C8H7NO3	Formylanthranilate	1.696	4.13E-04	0.136
5.7	159.068	C10H9NO	Indole-3-acetaldehyde	4.338	1.49E-07	0.487
13.5	176.095	C10H12N2O	Serotonin	0.137	1.00E-13	0.259
5.8	218.106	C12H14N2O2	N-Acetylserotonin	6.309	5.21E-05	0.151
7.9	208.085	C10H12N2O3	L-Kynurenine	0.680	2.66E-03	0.155
9.3	220.085	C11H12N2O3	5-Hydroxy-L-tryptophan	0.324	3.03E-03	0.174
8.1	205.074	C11H11NO3	Indolelactate	7.628	9.77E-03	0.320
5.4	232.121	C13H16N2O2	Melatonin	5.360	7.85E-06	0.096
17.3	155.069	C6H9N3O2	L-Histidine	0.040	0.00E+00	2.021
6.7	236.08	C11H12N2O4	L-Formylkynurenine	0.406	2.43E-02	1.111
Purines and Pyridines						
10.2	125.059	C5H7N3O	5-Methylcytosine	4.132	2.22E-03	0.175
9.5	135.055	C5H5N5	Adenine	4.723	6.28E-03	2.917

10.5	136.039	C5H4N4O	Hypoxanthine	0.394	1.49E-04	2.111
12.7	151.049	C5H5N5O	Guanine	0.186	2.21E-11	1.224
12.2	243.086	C9H13N3O5	Cytidine	0.064	1.16E-12	0.601
12.0	251.102	C10H13N5O3	Deoxyadenosine	0.134	3.97E-06	0.242
9.6	252.086	C10H12N4O4	Deoxyinosine	0.045	1.91E-08	0.191
8.9	267.097	C10H13N5O4	Adenosine	1.558	4.22E-02	0.350
6.2	278.127	C14H18N2O4	alpha-Ribazole	0.348	2.70E-15	0.326
13.0	283.092	C10H13N5O5	Guanosine	0.045	5.13E-28	0.614
11.6	323.052	C9H14N3O8P	CMP	0.118	5.42E-06	0.231
9.3	329.053	C10H12N5O6P	3',5'-Cyclic AMP	0.108	1.21E-04	0.250
Phenylalanine Metabolism						
7.3	165.079	C9H11NO2	L-Phenylalanine	0.558	7.38E-07	0.332
7.6	183.09	C9H13NO3	L-Adrenaline	1.397	3.08E-02	0.256
10.8	207.09	C11H13NO3	N-Acetyl-L-phenylalanine	1.867	1.93E-02	0.098
Vitamins and Co-Factors						

10.9	168.09	C8H12N2O2	Pyridoxamine	2.758	8.70E-04	0.107
11.5	214.132	C10H18N2O3	Dethiobiotin	0.129	2.95E-21	0.714
21.2	264.104	C12H17N4OS	Thiamine	26.976	1.88E-03	0.096
8.3	376.138	C17H20N4O6	Riboflavin	3.569	6.02E-03	0.124
8.4	455.155	C20H22N7O6	Methenyltetrahydrofolate	0.268	1.55E-02	0.135
Miscellaneous						
13.2	169.085	C7H11N3O2	N(pi)-Methyl-L-histidine	0.471	6.75E-06	0.496
15.3	160.085	C6H12N2O3	D-Alanyl-D-alanine	0.177	5.15E-26	0.757
15.6	146.069	C5H10N2O3	L-Glutamine	0.052	5.18E-24	3.176
15.6	129.043	C5H7NO3	Pyrroline-3-hydroxy-5-carboxylate	0.094	4.81E-22	1.339
13.9	165.046	C5H11NO3S	Methionine S-oxide	0.068	7.24E-22	1.684
16.5	175.096	C6H13N3O3	L-Citrulline	0.111	1.81E-19	2.580
17.1	240.024	C6H12N2O4S2	L-Cystine	0.015	1.40E-16	1.315
10.0	218.127	C9H18N2O4	Carboxyethyl-L-lysine	0.184	1.69E-14	0.178
14.3	309.106	C11H19NO9	N-Acetylneuraminate	0.041	1.36E-12	0.249

9.4	368.166	C19H28O5S	3beta-Hydroxyandrost-5-en-17-one 3-sulfate	0.020	1.46E-10	0.454
13.8	232.106	C9H16N2O5	N2-Succinyl-L-ornithine	0.330	1.46E-10	0.167
11.7	119.058	C4H9NO3	L-Threonine	0.299	4.98E-10	0.295
8.3	153.09	C7H11N3O	Acetylhistamine	5.316	8.42E-08	0.140
16.0	132.053	C4H8N2O3	L-Asparagine	0.010	2.63E-07	0.391
12.0	221.09	C8H15NO6	N-Acetyl-D-glucosamine	0.475	8.54E-07	0.247
10.1	204.111	C8H16N2O4	N6-Acetyl-N6-hydroxy-L-lysine	0.037	3.53E-06	1.019
15.8	161.069	C6H11NO4	L-2-Aminoadipate	0.464	2.07E-05	0.096
4.5	166.048	C5H10O6	L-Arabinonate	0.301	2.97E-05	0.423
12.4	117.079	C5H11NO2	L-Valine	0.355	6.30E-04	0.932
8.8	111.032	C5H5NO2	Pyrrole-2-carboxylate	5.707	1.08E-03	0.097
4.8	141.019	C2H8NO4P	Ethanolamine phosphate	3.145	1.13E-03	0.131
6.4	301.298	C18H39NO2	Sphinganine	2.427	1.56E-03	0.376
8.3	132.09	C5H12N2O2	L-Ornithine	3.350	1.69E-03	0.197
17.6	246.133	C9H18N4O4	N2-(D-1-Carboxyethyl)-L-arginine	0.466	2.60E-03	0.126

9.6	584.32	C30H48O11	Cholicacidglucuronide	0.000	2.70E-03	0.225
15.6	174.1	C7H14N2O3	N-Acetylornithine	2.974	4.37E-03	0.358
4.2	329.293	C19H39NO3	Dihydroceramide	8.849	7.03E-03	0.186
22.5	188.152	C9H20N2O2	N6,N6,N6-Trimethyl-L-lysine	0.523	1.11E-02	0.248
11.9	174.064	C6H10N2O4	N-Formimino-L-glutamate	0.555	1.94E-02	0.119
25.8	129.09	C5H11N3O	4-Guanidinobutanal	0.764	2.46E-02	0.253
8.9	135.053	C4H9NO4	4-Hydroxy-L-threonine	2.393	2.96E-02	0.157
16.4	117.054	C3H7N3O2	Guanidinoacetate	0.526	3.95E-02	0.306
15.4	147.053	C5H9NO4	L-Glutamate	0.745	4.50E-02	0.198
4.3	376.298	C24H40O3	3alpha-Hydroxy-5beta-cholanate	2.700	4.50E-02	3.510
7.1	204.147	C9H21N2O3	3-Hydroxy-N6,N6,N6-trimethyl-L-lysine	0.457	4.67E-02	0.123
15.1	258.085	C10H14N2O6	Ribosylimidazoleacetate	0.044	2.57E-24	0.585
21.5	124.064	C6H8N2O	Methylimidazole acetaldehyde	0.803	4.90E-02	0.191

Table 5.3: Significant metabolite differences based on time points (0 hrs vs. 24+48 hrs) among the healthy controls (HC)

Rt min	MW	Molecular Formula	Metabolite	24+48h/0h	P value	VIP
Fatty Acid Metabolism						
4.1	200.178	C12H24O2	Dodecanoic acid	0.572	1.4E-01	4.380
4.0	228.209	C14H28O2	Tetradecanoic acid	0.662	6.8E-02	3.078
4.0	280.24	C18H32O2	Linoleate	1.515	3.9E-02	0.231
4.5	354.241	C20H34O5	prostadienoic acid	2.858	3.5E-02	0.224
Haemoglobin Metabolism						
10.3	226.095	C10H14N2O4	Porphobilinogen	0.499	1.3E-04	0.332
4.3	562.258	C34H34N4O4	Protoporphyrin	0.098	5.4E-06	0.388
4.8	590.31	C33H42N4O6	D-Urobilinogen	0.369	4.6E-09	0.442
Tryptophan +Histidine Metabolism						
21.5	124.064	C6H8N2O	Methylimidazole acetaldehyde	0.734	1.1E-02	0.211
10.1	140.059	C6H8N2O2	Methylimidazoleacetic acid	1.203	2.9E-01	1.605

17.3	155.069	C6H9N3O2	L-Histidine	0.048	3.4E-16	0.778
5.2	165.043	C8H7NO3	Formylanthranilate	2.132	1.4E-10	0.220
13.5	176.095	C10H12N2O	Serotonin	0.072	5.3E-20	0.219
5.7	159.068	C10H9NO	Indole-3-acetaldehyde	2.600	1.0E-08	0.301
5.3	161.084	C10H11NO	Indole-3-ethanol	44.636	1.2E-19	0.206
5.8	218.106	C12H14N2O2	N-Acetylserotonin	1.9	7.6 E-5	0.004
5.4	232.121	C13H16N2O2	Melatonin	5.2	1.3 E-7	0.004
Purines and Pyrimidines						
9.5	135.055	C5H5N5	Adenine	2.024	9.2E-04	2.460
10.5	136.039	C5H4N4O	Hypoxanthine	0.352	3.2E-06	1.792
12.7	151.049	C5H5N5O	Guanine	0.166	2.6E-26	1.115
12.2	243.086	C9H13N3O5	Cytidine	0.277	1.1E-11	0.383
12.0	251.102	C10H13N5O3	Deoxyadenosine	0.165	1.2E-15	0.283
15.1	258.085	C10H14N2O6	(1-Ribosylimidazole)-4-acetate	0.002	2.6E-30	0.285
8.9	267.097	C10H13N5O4	Adenosine	1.929	7.8E-02	0.487
6.2	278.127	C14H18N2O4	alpha-Ribazole	0.438	2.7E-17	0.315
13.0	283.092	C10H13N5O5	Guanosine	0.160	5.7E-10	0.358

Vitamins And Cofactors						
11.5	214.132	C10H18N2O3	Dethiobiotin	0.075	3.5E-24	0.429
21.2	264.104	C12H17N4OS	Thiamin	422.058	7.2E-03	0.203
Micellaneous						
9.8	113.059	C4H7N3O	Creatinine	0.428	9.9E-05	0.777
4.7	392.293	C24H40O4	Dihydroxy-5beta-cholan-24-oic Acid	0.539	3.1E-02	1.590
4.3	376.298	C24H40O3	3alpha-Hydroxy-5beta-cholanate	0.798	4.3E-01	2.061
12.4	117.079	C5H11NO2	L-Valine	1.532	3.6E-01	0.385
15.9	103.063	C4H9NO2	4-Aminobutanoate	2.261	1.9E-05	1.140
13.9	165.046	C5H11NO3S	L-Methionine S-oxide	0.012	3.5E-27	1.901
9.7	145.11	C7H15NO2	4-Trimethylammoniobutanoate	2.013	3.5E-02	0.387
25.8	129.09	C5H11N3O	4-Guanidinobutanal	0.648	1.4E-20	0.424
7.6	183.09	C9H13NO3	L-Adrenaline	1.438	1.8E-05	0.276
10.1	204.111	C8H16N2O4	N6-Acetyl-N6-hydroxy-L-lysine	0.019	6.5E-30	1.504

6.4	301.298	C18H39NO2	Sphinganine	1.292	2.1E-01	0.233
7.3	165.079	C9H11NO2	L-Phenylalanine	0.702	3.2E-07	0.264
15.2	131.069	C4H9N3O2	Creatine	0.344	8.1E-06	0.471
15.6	146.069	C5H10N2O3	L-Glutamine	0.034	2.4E-23	1.787
9.0	177.046	C6H11NO3S	N-Formyl-L-methionine	0.284	1.4E-03	0.609
16.5	175.096	C6H13N3O3	L-Citrulline	0.127	1.1E-10	1.267
15.6	129.043	C5H7NO3	L-1-Pyrroline-3-hydroxy-5-carboxylate	0.083	1.1E-21	0.678
12.0	221.09	C8H15NO6	N-Acetyl-D-glucosamine	0.262	1.2E-11	0.372
9.6	584.32	C30H48O11	Cholicacidglucuronide	0.000	4.3E-13	0.418
17.1	240.024	C6H12N2O4S2	L-Cystine	0.000	3.0E-24	0.674
9.4	368.166	C19H28O5S	Hydroxyandrost-5-en-17-one 3-sulfate	0.013	1.3E-32	0.498
16.0	132.053	C4H8N2O3	L-Asparagine	0.015	1.0E-11	0.442
11.7	119.058	C4H9NO3	L-Threonine	0.343	6.3E-12	0.219
4.5	166.048	C5H10O6	L-Arabinonate	0.421	1.0E-06	0.249
14.3	309.106	C11H19NO9	N-Acetylneuraminate	0.003	1.6E-26	0.255

5.4 Discussion

Many gut metabolites are absorbed into the host's circulatory system before being excreted unchanged or after undergoing further metabolism by the human enzymes. Thus excreted metabolite profiles may offer an insight into the composition of gut microbiota associated with certain disease states (Nicholls *et al*., 2003). Several studies have been reported to discriminate samples collected from patients with inflammatory bowel disease from those collected from healthy controls based on profiling of urinary metabolites (Williams *et al*., 2009, Schicho *et al*., 2012, Stephens *et al*., 2013). All these 3 studies showed that hippurate levels were significantly lowered in patients with IBD as compared to healthy controls which suggested that this metabolite may be used as a unique biomarker for IBS.

The best models for discrimination between the 126 samples analysed were those based on time post treatment. None of the comparisons performed between the 7 dietary fibres showed a good discriminatory model in SIMCA. This observation is consistent with a study by Williams *et al*. (2009) who observed that, in urine samples from CD, UC and HC, although discrimination based on patient group was possible, the clustering of samples was independent of the diet and medication (Williams *et al*., 2009). Our study shows that discrimination is possible between CD and HC, and between samples at 0 hours and 24 hours, 0 hours and 48 hours, but not between 24 hours and 48 hours. Due to the similarity in metabolite profiles between samples as 24- and 48-hours, these samples were combined and compared with those at 0 hours in order to determine the metabolites responsible for separation of these groups in the SIMCA models under both supervised and unsupervised approaches.

5.4.1 Metabolites in Crohn's disease vs. healthy controls at all time points

The most important variables discrimating the CD from HC samples were dodecanoic acid and eicosapentaenoic acid. These seemed to be almost absent from the CD samples suggesting possibly and increased rate of fatty acid metabolism in the CD incubations. Previous studies have observed that CD patients exhibit a faster rate of fatty acid oxidation in comparison with controls and have lower levels of medium chain fatty acids (Mingrone *et al*., 1999, Al-Jaouni *et al.*, 2000, De Preter *et al.*, 2015).

A number of degradation products of haemoglobin including bilirubin are elevated in in the CD samples and this is in line in previous studies which report that bilirubin is elevated in the bile of CD patients as a consequence of malbabsorption of bile acids. The malabsorption of bile acids increases their levels in the gut and thus promoting the reuptake of bilirubin into the enterohepatic circulation (Lapidus *et al.*, 2006, Leníček *et al*., 2014). There is no great indication of bile acid elevation in the CD samples apart for chenodeoxy cholate which is elevated 2.3fold in comparison with the HC samples.

The tryptophan degradation product formyl kynurenine is greatly increased in the CD samples and this corresponds with earlier observations where metabolites in the kynurenine pathway were found to be increased in Crohn's disease (Gupta *et al.*, 2011, Kennedy *et al.*, 2017). It was proposed that an increase in tryptophan degradation indicated and aberrant immune response in CD to a sub-population of commensal bacteria in the GI tract. A number of tryptophan metabolites are down regulated in the CD samples including N-acetyl serotonin, formyl anthranilic acid and hydroxyl melotonin. Formyl anthranilate is a direct degradation product of formyl kynurenine and this suggests a downregulation of the enzyme responsible for this conversion in the CD samples.

5.4.2 Metabolites at 0 vs. 24/48 hrs in CD faecal incubations

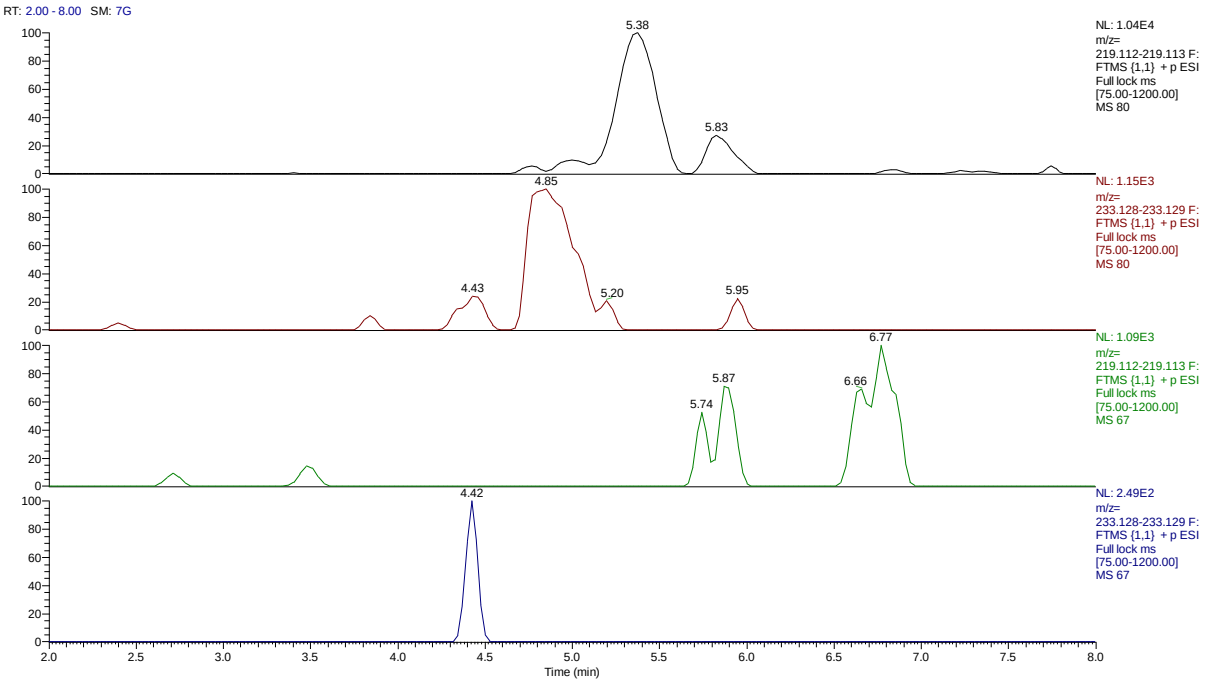

Figure 5.12. Extracted ion trace showing the formation of low levels of melotonin and its precursor N-acetyl serotonin in the CD incubations after 48 h in comparison with time 0 (lower traces).

There is a very strong separation between the time 0 samples for the CD incubations and the 24/48 h samples. This suggests that the microbial activity in the samples is having a marked effect on the metabolome. Dodecanoic acid levels are decreased following inubation but not greatly suggesting that fatty acid metabolism under the conditions of the incubation is not rapid. All of the haemoglobin metabolites decrease markedly following incucation probably because they provide a source of nitrogen. More interesting is the large impact of incubation on tryptophan metabolism. Metabolism of tryptophan appears to be activated via two pathways. The kynurinine pathway in the incubations over time is reduced and tryptophan appears to be catabolised via deamination forming metabolites including indole lactate, indole acetaldehyde and indole ethanol. Possibly of greater significance is the formation of low but detectable levels of melatonin and its precursor N-acetyl serotonin (Figure 5.12). Serotonin is

produced extensively in the GI tract (Kennedy *et al*., 2017) but the formation of melatonin has not been described. In this case it is clear that melotonin is being produced by bacterial activity. While melatonin occurs in blood at very low levels the levels have been found to be up to 100 times higher in the GI tract which is a major source of melatonin in the body and it has been reported the meltonin has an influence on the GI tract immune system (Terry *et al.*, 2008). However, trials using melatonin as a treatment in experimental colitis found that while acute treatment with melatonin reduced the colitis chronic administration provokes it (Marquez *et al*., 2006). Interestingly melotonin production why comesal bacteria in plants has been found to be important in plant function (Tan *et al.*, 2013), it is possible that melatonin production by commensal bacteria might also be important.

There are many other metabolites affected by incubation of the CD samples, however, of greatest interest are those that increase since decreasing metabolites may just reflect break-down of these by the bacteria as nitrogen or carbon sources. There is a marked increase in levels of adenine and adenosine while levels of other purines and pyrimidines largely de-crease. The increase in adenosine might be related to an increased requirement for NAD+ which is required for the production of ATP from the TCA cycle. It would seem likely that oxidative metabolism will increase in the bacterial population in the samples since in the large colon metabolism is anaerobic. This might explain why there is a very large increase in thia-mine during the incubation since thiamine is required for the conversion of pyruvate to acetyl CoA which can then enter the Krebs cycle.

5.4.3 Metabolites at 0 vs. 24/48 hrs in HC faecal incubations

As for the incubations with the CD samples there were not many samples which increased in the HC incubations with time. Dodecanoic acid and and tetradecanoic acid levels were decreased to a similar extent as observed in the CD samples and haemoglobin metabolites were also markedly decreased. In the HC samples the levels of tryptophan metabolites formed via deamination increases as observed for the CD samples. The OPLSDA model separating time 0 and the 24/48 h samples in the HC incubations did not select melatonin and N-acetylserotonin as being important variables. However, as in the case of the CD samples there are marked increases in these low abundance metabolites (Figure 5.13). This reinforces the idea that conversion of N-acetyl serotonin to melatonin is dependent on aerobic metabolism. Melatonin is a powerful antioxidant and this might explain why it is produced when the faecal samples are incubated in air (Tan *et al*, 2013). As in the case of the CD samples there is a marked increase in thiamine indicating a switch to aerobic metabolism. Among the few metabolites which increase in the incubations a marked increase in the levels of GABA is of interest. The production of GABA by commensal bacteria is known to occur (Sharon *et al*, 2014) and the increase in GABA in the CD incubations is smaller and not statistically significant so this could be a marker of a more favourable bacterial profile in the HC samples.

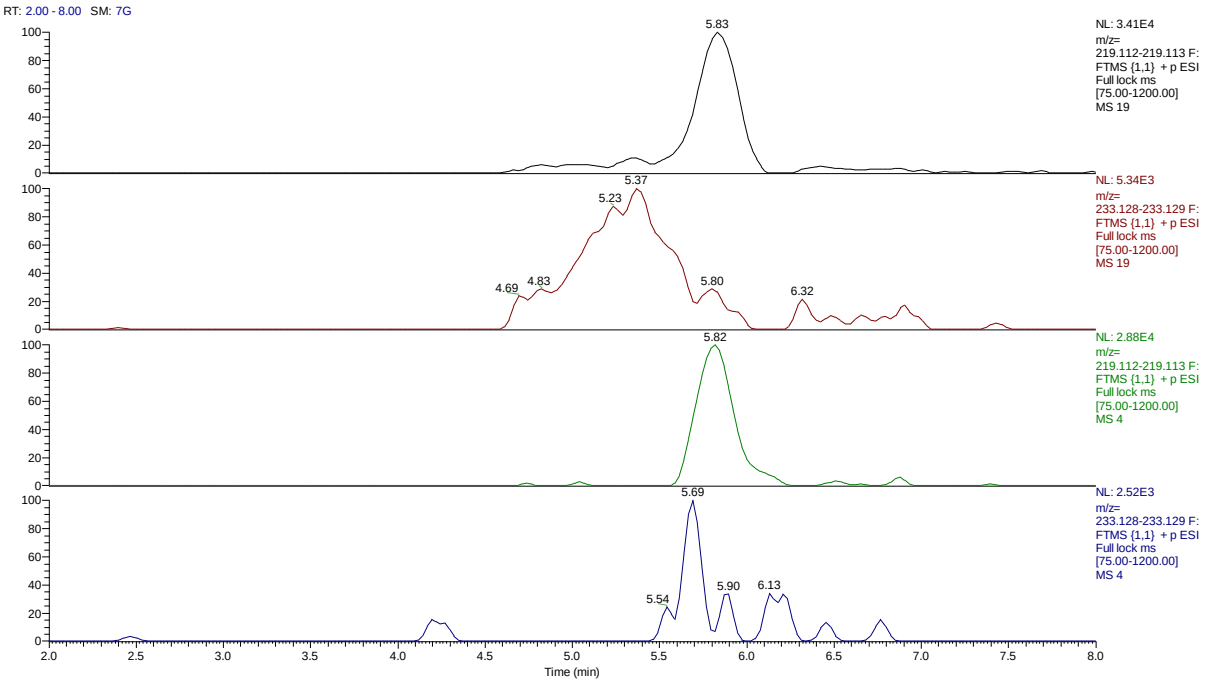

Figure 5.13: Extracted ion current showing increase in N-acetyl serotonin and melatonin in HC incubations in 24/48 h samples in comparison with 0 h samples (lower traces).

5.5 Conclusion

It can be concluded incubation of faecal samples from Crohn's disease patients with 7 dietary fibre options did not yield good discriminatory models in SIMCA, implying that the fibres as carbon sources did not produce significantly different levels of metabolites in each group. It might have been of more interest to observe the effect of the fibres on the production of short chain fatty acids such as butyrate which have a role in microbiome health, however the screen used was not designed to pick up these metabolites. On the other hand, it was found that metabolites differed between the CD samples and HC samples particularly with respect to fatty acid metabolism and haemoglobin metabolism. Most interestingly the time of incubation had a marked effect on the production of melatonin both in the HC and the CD incubations. The ability of the microbiome to influence the immune system and the brain is be-

coming increasingly apparent. The current work provides the first observation of the formation of melatonin as a result of the action of microbes isolated from the gut and indicates that this is in some way linked to aerobic metabolism.

Chapter Six:

Prediction of Cancer Associated Muscle Wasting from Plasma Metabolites

6 PREDICTION OF CANCER ASSOCIATED MUSCLE WASTING FROM HUMAN PLASMA METABOLITES

6.1 Introduction

Cancer cachexia has been defined as "a multifactorial syndrome characterized by an ongoing loss of skeletal muscle mass that cannot be fully reversed by conventional nutritional support that leads to progressive functional impairment" (Fearon *et al*., 2011). The agreed diagnostic criterion for cachexia are either weight loss >5% over 6 months or BMI <20 with any degree of weight loss >2% or appendicular skeletal muscle index consistent with sarcopenia (males <7·26 kg/m^2; females <5·45 kg/m^2) and any degree of weight loss >2% (Fearon *et al.,* 2011). It is characterized by loss of adipose tissue, muscle atrophy, and loss of appetite and impacts negatively on the quality of life of patients with cancer, response to treatment and survival (Fearon *et al*., 2006). Managing cachexia should therefore be considered a central component of cancer patient treatment.

Skeletal muscle mass is maintained by a balance between protein synthesis and degradation, which is principally regulated by physiologic inputs such as nutritional status and physical activity (Fearon *et al.,* 2012). Many studies have shown that muscle loss may occur independently of changes in fat mass and can be an early phenomenon that is difficult to detect against a background of excess body weight (Martin *et al.,* 2013). As cachexia can develop progressively through various stages, from pre-cachexia to cachexia to refractory cachexia, early screening and staging is particularly important to prevent or delay its onset (Fearon *et al.,* 2011). An improved approach for detecting the evolution of muscle and or fat wasting

could possibly help target early intervention and treatment. Imaging methods such as CT and MRI are currently considered the most precise measures of body composition but have several limitations including cost and exposure to radiation (Yip *et al.,* 2015).

Recent progress in high-throughput analytical technologies and bioinformatics now permits simultaneous analysis of hundreds of compounds constituting the metabolome (Mestrangelo *et al.,* 2014). Metabolomic analyses give complex fingerprints that appear to be characteristic of a given metabolic phenotype or diet. There have been very few previous studies attempting to quantify metabolites associated with cachexia. Those that have have identified metabolites that are possibly discriminative of cachexia indicating there maybe scope for metabolomics based studies to identify biomarkers of cachexia (O'Connell *et al.,* 2008, Eisner *et al.,*, 2011, Fujiwara *et al.,* 2014, Yang *et al.,* 2018, Cala *et al.,* 2018).

Although many have suggested that metabolomic analysis has the potential to change how nutrition research is conducted, much of this potential remains unrealised. Building upon the theory that metabolites produced from tissue breakdown are likely to be found in plasma and could potentially be a sensitive indicator of tissue wasting we investigated whether we could detect metabolites associated with cachexia from the plasma of cancer patients. Plasma was selected as the biofluid of choice, since as previously shown several end products of muscle catabolism (e.g. creatinine and methylhistidine) can be easily measured here (Argiles *et al.,* 2014). We therefore attempted to identify patterns of plasma metabolic profiles that discriminate the condition of weight loss.

6.2 Materials and Methods

6.2.1 Chemicals and Solvents

HPLC grade Acetonitrile (ACN) was purchased from Fisher Scientific (Loughborough, UK) and HPLC grade water was produced by a Direct-Q3 UltrapureWater System (Millipore, Watford, UK). AnalaR-grade formic acid (98%) was obtained from BDH-Merck (Poole, UK). Authentic stock standard metabolites (Sigma-Aldrich, Poole, U.K.) were prepared as previously described and diluted four times with ACN before LC-MS analysis. Ammonium acetate was purchased from Sigma-Aldrich (Poole, UK).

6.2.2 Participants

Patients over 18 years of age were recruited to the study from the regional upper GI multidisciplinary team meeting. Written informed consent was obtained from all subjects and ethical approval received from Lothian Research Ethics Committee (UK). Participating patients had a diagnosis of upper gastrointestinal cancer (oesophageal, gastric, pancreatic) and were undergoing surgery with the intent of curative resection of the primary tumour. Skeletal muscle cross-sectional area (CSA) was measured from routine CT scans performed prior to any surgical or oncological intervention. A transverse CT image from the third lumbar vertebrae (L3) was assessed for each scan date and tissue volumes estimated using semi-automated software. Cross-sectional area for muscle was normalized for stature (cm^2/m^2) to calculate the skeletal muscle index (SMI).

6.2.3 Sample collection and storage

Fasting venous blood samples were taken at induction of anaesthesia approximately four to six weeks after the cessation of any neoadjuvant chemotherapy. Samples were allowed to

clot at room temperature. Serum was separated by centrifugation at 1300 RPM for 12 minutes at a temperature of 20 degrees. C-reactive protein (CRP) was measured in Clinical Chemistry, Royal Infirmary, Edinburgh (fully accredited by Clinical Pathology Accreditation Ltd.) using standard automated methods. A CRP ≥5 mg/l was considered consistent with the presence of systemic inflammation. Clinical details, degree of weight loss from self-reported pre-illness stable weight and body mass index (BMI) were recorded. Samples were stored at -80°C until they were transported to the laboratory located at the Strathclyde Institute of Pharmacy and Biomedical sciences in cool bags at -30°C for further analysis. Samples were analysed blindly by the metabolomics laboratory.

6.2.4 Sample preparation

Exactly 200 µL of the sample was mixed with 800 µL of acetonitrile containing 10 µg/ml of 2 ^{13}C glycine (Sigma-Aldrich, Poole, U.K.) as an internal standard to ensure retention time stability, and then centrifuged for 10 min before transferring into a vial with an insert. The pooled sample was prepared by pipetting 50 µL from each of the 18 samples and then mixing them together before diluting 200 µL of the pooled sample with 800 µL of acetonitrile containing 10 µg/mL of 2 ^{13}C glycine internal standard and centrifuging. Additionally, the prepared mixtures of authentic standard metabolites containing 10 µg/mL of 2 ^{13}C glycine as internal standard were run.

6.2.5 LC-MS conditions

Liquid chromatographic separation was carried out on an Accela HPLC system interfaced to an Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) using both a ZIC-pHILIC column (150 × 4.6 mm, 5 µm, HiChrom, Reading UK). The column was eluted with a mobile phase consisting of 20 mM ammonium carbonate in HPLC-grade water (solvent

A) and acetonitrile (solvent B), at a flow rate of 0.3 mL/min. The elution gradient was an A:B ratio of 20:80 at 0 min, 80:20 at 30 min, 92:8 at 35 min and finally 20:80 at 45 min. The nitrogen sheath and auxiliary gas flow rates were maintained at 50 and 17 arbitrary units. The electrospray ionisation (ESI) interface was operated in both positive and negative modes. The spray voltage was 4.5 kV for positive mode and 4.0 kV for negative mode, while the ion transfer capillary temperature was 275°C. Full scan data were obtained in the mass-to-charge ratio (m/z) range of 75 to 1200 for both ionisation modes on the LC-MS system fully calibrated according to manufacturer's guidelines. The resulting data were acquired using the XCalibur 2.1.0 software package (Thermo Fisher Scientific, Bremen, Germany).

6.2.6 Data extraction and analysis

Data extraction for each of the samples was carried out by MZMine software. The extracted ions, with their corresponding m/z values and retention times, were pasted into an Excel macro of the most common metabolites prepared in–house to facilitate identification. The lists of the metabolites obtained from these searches were then carefully evaluated manually by considering the quality of their peaks and their retention time match with the standard metabolite mixtures run in the same sequence. All metabolites were within 3 ppm of their exact masses. Statistical analyses were performed using both univariate with Microsoft Excel and multivariate approaches using SIMCA-P software version 14.1 (Umetrics, Umea, Sweden.).

6.3 Results

Plasma samples from 18 patients were analysed and the metabolomic profiles produced allowed division of the patients into two distinct groups. These profiles corresponded to weight

loss data of the patients with the separation of the groups corresponding with less than or more than 5% weight loss as demonstrated in Table 6.1. There were more males in the <5% weight loss group. The mean % weight loss in the >5% weight loss group was 14.40% and these patients had a higher mean CRP (17.88mg/l compared with 32.56mg/l in the >5% weight loss group). SMI was similar between the two groups (47.17 compared with 45.82 in the >5% weight loss group). Overall a total of 37 metabolites were significantly associated with weight loss. The metabolomes with the highest fold change were species of LysoPE, Lyso PA and LysoPC (P=0.002). The largest affected group were the glycerophospholipids. Fatty acyls and products of lipid and amino acid metabolism were also increased.

Table 6.1: Patient details

	Group 1 **<5% weight loss (n=9)**	**Group 2** **>5% weight loss (n=9)**
Male: Female	8:1	5:4
% weight loss	0.86 (2.47)	14.40 (6.56)
Skeletal Muscle Index	47.17 (6.26)	45.82 (7.72)
BMI (kg^2/m^2)	26.29 (4.64)	24.93 (4.42)
CRP (mg/l)	17.88 (27.06)	32.56 (50.45)
Cancer type	Pancreatic – 1 Oesophageal – 6 Gastric - 2	Pancreatic – 6 Oesophageal – 2 Duodenal - 1

All data are: mean (standard deviation)

Figures 6.1 and 6.2 show the principal components analysis (PCA-X) of all 18 samples and 3 quality control (QC) samples. PCA-X, an unsupervised model in SIMCA-P, produces a natural scatter of the samples based on their characteristic metabolomics footprints. It can be seen in the figure (samples colored according to weight loss) that, in general, there is no separation of samples into groupings based on weight loss. Additionally, one of the three QC samples is

quite separated from the other two which points to some lack of detection stability (Figure 6.1).

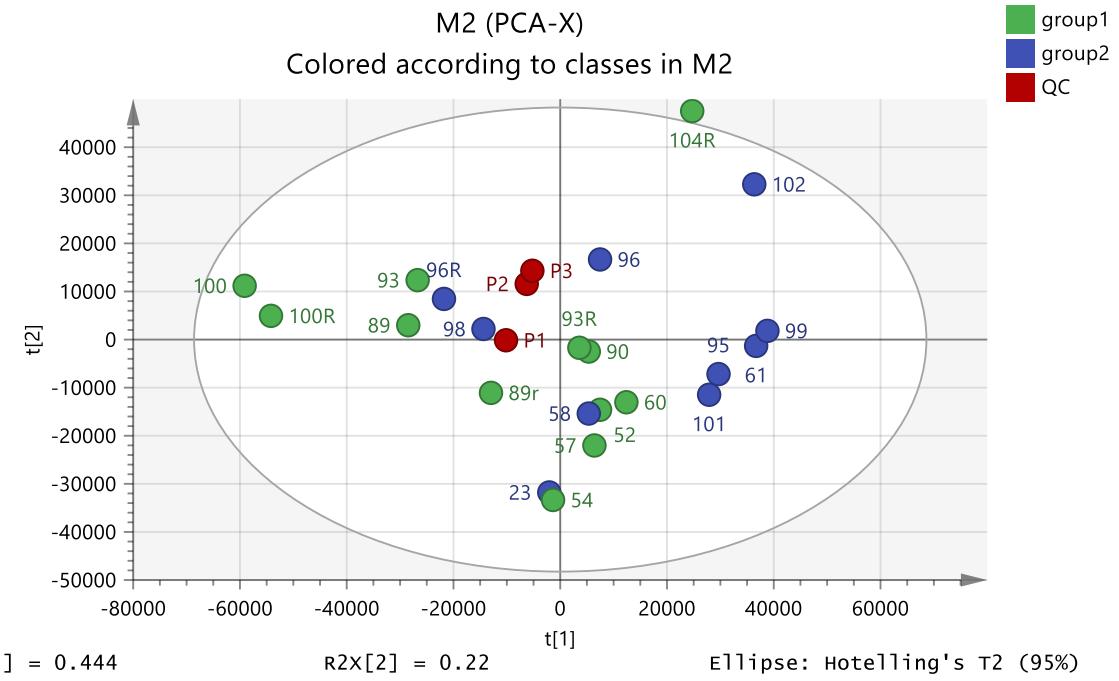

Figure 6.1: PCA-X analysis of the metabolomics footprint of the 18 plasma samples from cancer patients. Circles colored according to the weight loss categorization as weight stable or weight losing. The red circles represent pooled samples. Green circles (group 1) = weight stable. Blue circles (group 2) = weight losing. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

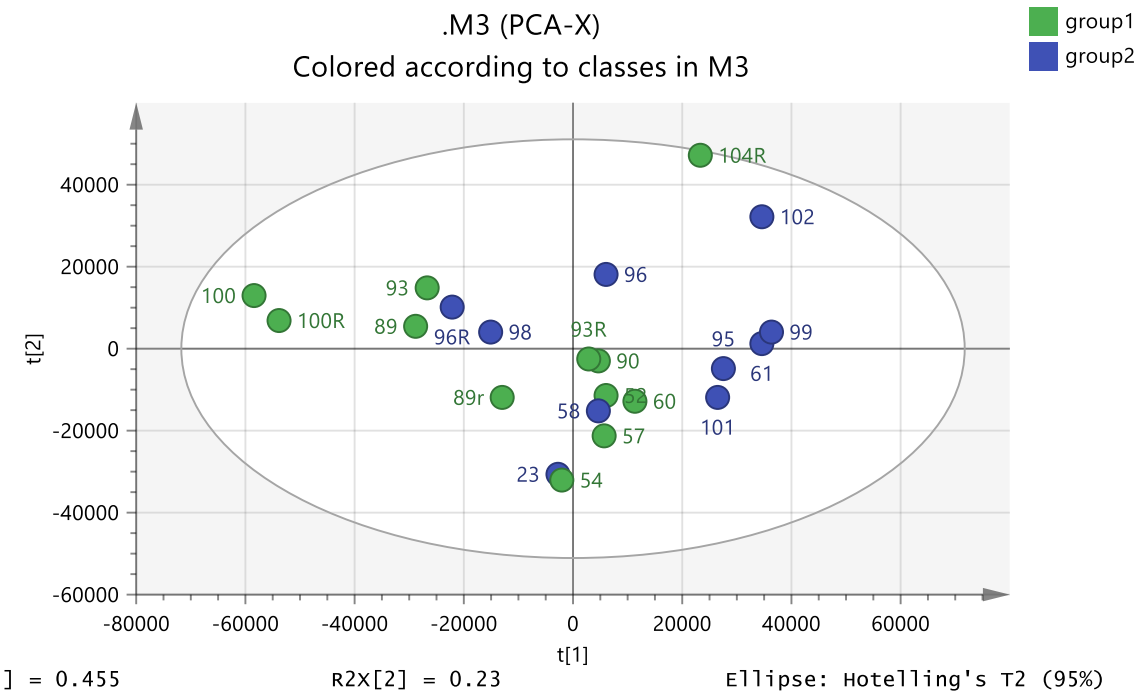

Figure 6.2: PCA-X analysis of the metabolomics footprint of the 18 plasma samples from cancer patients with the pooled samples removed. Green circles (group 1) = weight stable. Blue circles (group 2) = weight losing. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

Supervised models enable identification of metabolites that have the most significant contribution to a given clustering pattern. In SIMCA, supervised analysis can be carried out using OPLS-DA models. The OPLS-DA model works by identifying metabolomics differences in two pre-determined groupings. Based on OPLS-DA, both group 1 (weight stable) and group 2 (weight losing) samples were clearly separated implying that there are clear metabolomics differences between the two groups (Figure 6.3). The permutation test carried out on the model showed validity (Figure 6.4).

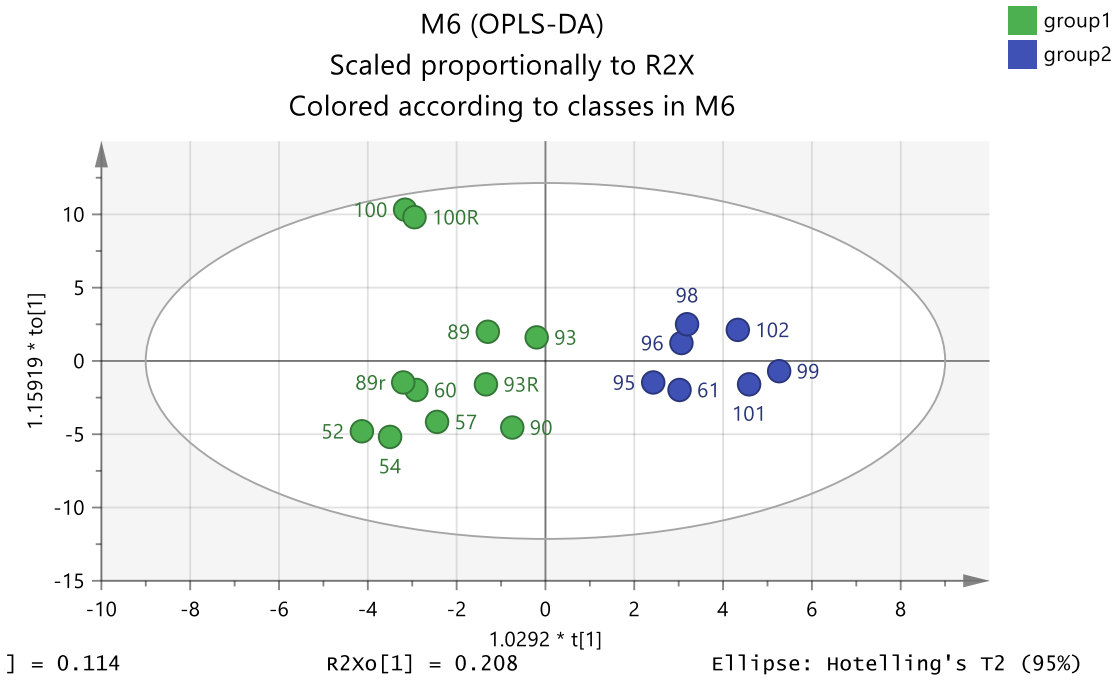

Figure 6.3: OPLS-DA analysis to compare weight losing samples with weight stable samples. There is clear separation of both groups implying significantly different metabolic footprints. The CV-ANOVA = 0.0477247. The y-axis represents variation within the groups while the x-axis represents variation between the groups.

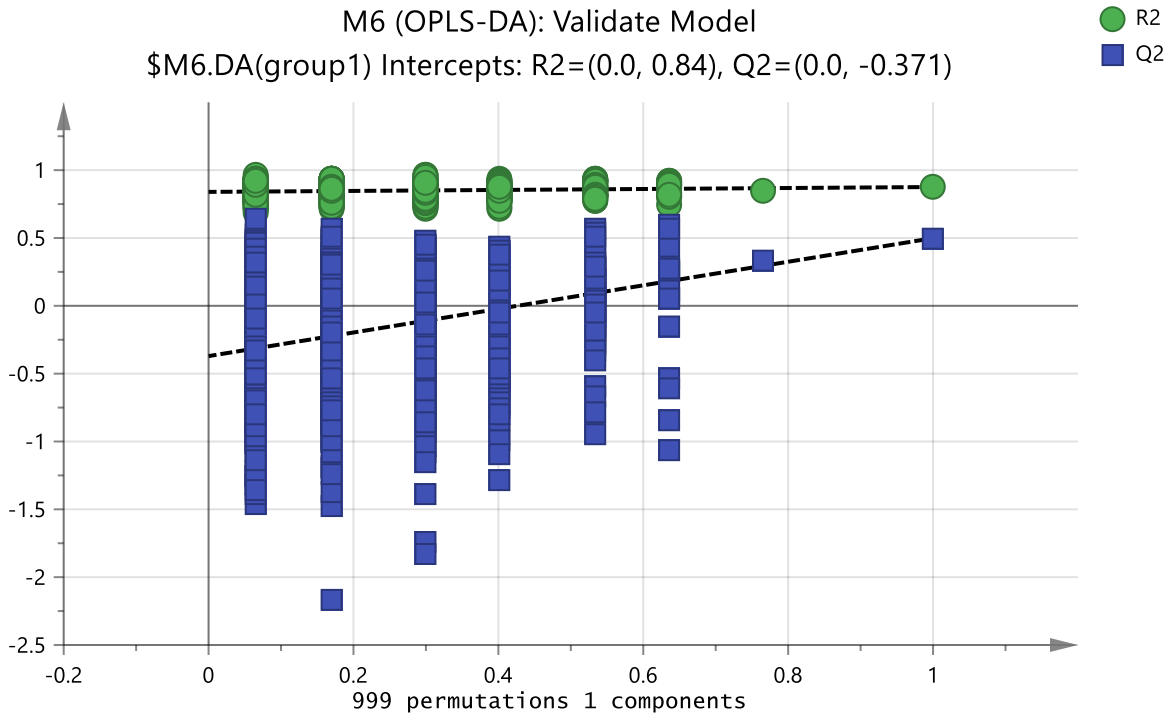

Figure 6.4: Cross validation of the OPLS-DA model comparing the weight losing samples with weight stable samples

Table 6.2 shows the metabolites found to be significantly different between the two weight group categories. The ratio represents the intensity of the metabolites relative to the "weight stable" patient group.

Table 2: Significant metabolites that differ between weight loss vs. weight stable groups

Var ID	M/Z	Rt min	Formula	Metabolite	Ratio 2/1	P-value	VIPpred
Fatty acids							
2585	279.23	4.18	C18H32O2	Linoleate	1.27	0.002	1.2
3144	283.26	4.17	C18H36O2	Octadecanoic acid	1.24	0.015	1
130	303.23	4.12	C20H32O2	Eicosatetraenoic acid	1.32	<0.001	1.3
129	305.25	4.17	C20H34O2	Eicosatrienoic acid	1.37	<0.001	1.4
2223	255.23	4.27	C16H32O2	Hexadecanoic acid	1.2	0.006	1
3629	329.25	4.09	C22H34O2	Docosapentenoic acid	1.38	<0.001	1.5
133	331.26	3.91	C22H36O2	Docosatetraenoic acid	1.66	0.001	1.6
Lipid metabolism/ glycerophospholipids							
24	363.16	4.37	C16H29O7P	LysoPA 16:0	1.55	0.024	1.7
2628	435.25	4.46	C21H41O7P	LysoPA18:1	1.35	0.005	1.4
257	437.27	4.23	C21H43O7P	LysoGP 18:0	1.23	0.01	1.1
296	454.29	4.56	C21H44NO7P	LysoPE 16:0	1.29	0.014	1.3
258	457.24	4.25	C23H39O7P	LysoPA 20:4	1.47	<0.001	1.7
261	464.28	4.42	C22H44NO7P	LysoPC 14:0	1.3	0.007	1.3
2729	478.29	4.39	C23H44NO7P	LysoPE18:2	1.64	0.024	1.9
305	480.31	4.36	C23H46NO7P	LysoPE18:0	1.79	0.012	1.8
277	480.34	4.37	C24H50NO6P	Lyso PC 16:1	1.34	0.026	1.2
3177	485.27	4.27	C25H43O7P	LysoPA20:4	1.61	0.02	2
275	496.34	4.43	C24H50NO7P	LysoPC 16:0	1.3	0.014	1.1
2730	504.31	4.36	C25H48NO7P	LysoPE20:2	1.42	0.021	1.6
281	508.34	4.45	C25H50NO7P	LysoPE 18:0	1.27	0.027	1.2
2722	514.29	4.28	C26H46NO7P	LysoPC18:4	1.79	0.003	1.9
286	520.34	4.36	C26H50NO7P	LysoPC18:2	1.49	0.013	1.6
282	524.37	4.33	C26H54NO7P	LysoPC18:0	1.36	0.031	1.1
2734	526.29	4.29	C27H46NO7P	LysoPE22:5	1.78	0.002	2.2

2733	528.31	4.33	C27H48NO7P	LysoPE22:4	1.82	0.003	1.9
288	540.31	4.35	C28H48NO7P	Lyso PC 20:5	1.38	0.027	1.9
2726	544.34	4.31	C28H50NO7P	LysoPC20:4	1.65	0.002	1.7
2725	546.35	4.29	C28H52NO7P	LysoPC20:3	1.64	0.005	1.8
2727	570.36	4.24	C30H52NO7P	LysoPC22:5	1.61	0.008	2
Bile acids and steroids							
413	367.23	4.16	C24H32O3	Oxocholatrienoic Acid	1.41	0.044	1.8
808	317.25	4.32	C21H34O2	Tetrahydroprogesterone	1.21	0.042	1
Metabolism of cofactors and vitamins							
3059	124.04	7.94	C6H5NO2	Nicotinate	0.58	0.031	1.5

PE = phosphatidyl ethanolamine, PC = phosphatidyl choline, PA =phosphatidic acid.

6.4 Discussion

In this study we performed LC-MS based metabolomics analysis to reveal the metabolic profile of cancer cachexia. We were able to demonstrate distinct profiles associated with more or less than 5% weight loss. It can be seen that most of the metabolites affected fall within the lipid pathways. The clearest effect is that several long chain fatty acids and lysolipids are elevated in the plasma of the patients with higher weight loss. Lysolipids are very abundant in plasma and for instance Lyso PC 16:0 is the most abundant compound by response in this set of sample thus an increase of 1.3 fold between weight stable and weight losing patients represents a major shift in metabolic output, shifts in minor components might not be as significant. Figure 6.5 shows a heat map showing the relative abundance of the lyso lipids in these plasma samples. Lyso PC18:2 is almost as abundant as lyso PC 16:0 and is elevated by 1.49 fold. Beyond these two lysolipids the response is much lower but there are many more minor lipids showing similar or greater fold changes in the weight losing patients. In a previous study it was found that lysolipids in several classes were elevated in blood and ascites from ovarian cancer patients (Xu *et al.,* 2001). The current study therefore demonstrates an association between of lipolysis-promoting activity in the serum of cachectic cancer patients and weight loss. Most current research into cancer cachexia focuses on muscle wasting however, the importance of lipid metabolism is beginning to be recognised. It has been observed that lyso PA lipids stimulate the growth of ovarian and breast cancer cells (Xu *et al.,* 1995).

m/z	Rt min	Chemical formula	Metabolite name	Group 1	Group 2
496.339	4.4	C24H50NO7P	LysoPC 16:0		
520.339	4.4	C26H50NO7P	LysoPC18:2		
544.338	4.3	C28H50NO7P	LysoPC20:4		
435.252	4.5	C21H41O7P	LysoPA18:1		
464.278	4.4	C22H44NO7P	LysoPC 14:0		
546.354	4.3	C28H52NO7P	LysoPC20:3		
437.267	4.2	C21H43O7P	LysoGP 18:0		
480.344	4.4	C24H50NO6P	Lyso PC 16:1		
478.292	4.4	C23H44NO7P	LysoPE18:2		
570.356	4.2	C30H52NO7P	LysoPC22:5		
480.308	4.4	C23H46NO7P	LysoPE18:0		
454.292	4.6	C21H44NO7P	LysoPE 16:0		
526.292	4.3	C27H44NO7P	LysoPE22:5		
457.235	4.3	C23H39O7P	LysoPG 20:4		
524.278	4.3	C27H44NO7P	LysoPC18:0		
508.341	4.3	C25H52NO7P	LysoPE 18:0		
504.31	4.4	C25H48NO7P	LysoPE20:2		
485.267	4.3	C25H43O7P	LysoPA20:4		
363.158	4.4	C16H29O7P	LysoPA 16:0		
514.294	4.3	C26H46NO7P	LysoPC18:4		
528.31	4.3	C27H48NO7P	LysoPE22:4		
540.309	4.4	C28H48NO7P	Lyso PC 20:5		

Figure 6.5: Heat map showing relative levels of lysolipids purple= > 30%, yellow >1% blue >0.1%

Lipids are biochemical intermediates that play an important role in cellular homeostasis, including cell cycle regulation, cell signalling and energy storage (Meer *et al.,* 2008). Adipose loss in cachexia is believed to be mainly caused by an increase in lipolysis rather than a reduction in lipid synthesis (Dahlman *et al.,* 2010). Increased lipolytic activity, evidenced by elevated fasting plasma glycerol and free fatty acids is a driver of fat loss in advanced cancer patients but the underlying causes of elevated lipolysis are not known

(Tsoli *et al.,* 2016). Other mechanisms including decreased lipogenesis impairment in adipogenesis, elevated fat oxidation, and decreased lipid deposition have also been attributed to fat loss in cancer (Zuijdeest-Van *et al.,* 2000). Compared to non-cancer cells, cancer cells demonstrate significant metabolic alterations. Normal cells are able to regulate anabolic and catabolic pathways in response to changes in nutrient availability whereas cancer cells show unregulated growth regardless of nutrient availability (Vander Heiden *et al.,* 2010). Lipid metabolism is highly altered in proliferating cells. Unlike normal cells which rely mostly on the uptake of fatty acids, cancer cells increase adipogenesis which is needed for membrane synthesis and signalling molecules (Belorbi-Diefaflia *et al.,* 2016). Recent studies have demonstrated that activation of growth promoting pathways results in a dependence on unsaturated fatty acids for survival under oxygen deprivation (Ackerman *et al.,* 2015).

Different types of lipids have distinct physiological roles. Triacyglycerols are mainly used as storage within adipose tissue. Following hydrolysis of triacyglycerols, free fatty acids are released into the circulation as an energy source for most cells (Ahmadian *et al.,* 2009). Glycerophospholipids along with cholesterol are the major lipid component of cell membranes. Lysophospholipids (LPL) are usually the result of phospholipase type A enzymatic activity (Raynal *et al.,* 2005). LPLs have been shown to be involved in many physiological and pathological processes such as inflammation and tumorigenesis (Raynal *et al.,* 2005). Over the past decade, it has become clear that medically relevant LPLs activities are mediated by specific G protein-coupled receptor (GPCR), implicating them in the

etiology of a growing number of disorders including cancer (Yan *et al.,* 1996, Sutphen *et al.,* 2004, Zhao *et al.,* 2007). Phosphatidylcholine (PC) is a glycerophospholipid with a polar phosphocholine head group and two non-polar fatty acid hydrocarbon chains (Meer *et al.,* 2008). It is the main membrane-forming phospholipid in mammalian cells. Removal of one of the fatty acids either by an enzyme or spontaneously by hydrolysis results in lyso-phosphatidylcholine (LPC) (Berdel *et al.,* 1986). LPC exerts a lytic action on cell membranes (Taylor *et al.,* 2007). In contrast to the present study several studies have shown that an increased demand for lipids by the tumour is reflected in decreased levels of several lipids in the blood. Decreased levels of lysoPC have been seen in lung and liver cancers (Kriat *et al.,* 1993).

Many different biomarkers have been correlated with cachexia. In particular cancers of the oesophagus and pancreas have been linked with high levels of plasma glycerol and free fatty acids (Shaw *et al.,* 1987, Das *et al.,* 2013). Weight losing cancer patients have been shown to have an increased turnover of both glycerol and fatty acids compared with cancer patients without weight loss (Ebadi *et al.,* 2015). Others have suggested that observed increases in lipolysis and triglyceride-fatty acid cycling in cachectic patients with oesophageal cancer was due to alterations in nutritional status rather than disease presence (Klein *et al.,* 1990). Cachectic ovarian cancer patients have been shown to have increased levels of free fatty acids, monoacylglycerides and diglycerides in their serum

and ascitic fluid (Gercel-Taylor *et al.,* 1996). Free fatty acids may therefore provide energy for the tumour or signaling molecules with the glycerol molecules released during the breakdown of triacylglycerides being used for gluconeogenesis by the liver.

There have been very few previous attempts to profile metabolites associated with cachexia with varying results and important metabolites produced in each study. Metabolomics research in cachexia only began in 2008 in the C26 mouse model of cancer cachexia (O'Connell *et al.,* 2008). This was the first study to demonstrate a distinct metabolic profile associated with the onset of muscle wasting which included increased levels of very low and low density lipoprotein and abberant glycosylation of β-DG. The first human based study in cancer cachexia also found large numbers of glycerophospholipids and metabolites associated with amino acid metabolism, the urea cycle, intermediary metabolism (glycolysis, TCA cycle, 1-carbon metabolism) and creatinine were prominent (Eisener *et al.,* 2011). This group were therefore able to develop a single time-point urine test using concentrations of 63 urinary metabolites to diagnose muscle wasting. This minimally invasive test was rapid, robust, quite accurate (82.2%), and able to detect a small but physiologically relevant rate of muscle loss. One such other quantitative metabolomics study supported the hypothesis that lean and fat mass have distinct metabolic profiles and that broad categories of high and low muscle mass quantity were able to be accurately predicted from metabolite concentrations in urine and serum. They

were therefore able to identify occult sarcopenia in patients with cancer (Stretch *et al.,* 2011).

Recent studies have attempted to separate pre-cachectic, cachectic, weight stable cancer patients and healthy controls using serum metabolomics analysis. They showed a clear separation of the four groups with 45 metabolites and 18 metabolic pathways being associated with cachexia; fifteen of these metabolites being identified as highly discriminative (Yang *et al.,* 2018). Levels of metabolites have been shown to vary throughout the day in patients with pancreatic cancer. Lower levels of paraxanthine (caffeine metabolite) being the only metabolite linked to cachexia not to show diurnal variation. They did not however control for coffee consumption leading to possible bias in their results (Fujiwara *et al.,* 2014). The most recent cachexia-based metabolomics study included 8 cachectic and 7 non cachectic patients and analysed them using three analystical platforms. Contrary to the present study they found a significant reduction in amino acids and glycerophospholipids associated with cachexia – a difference that has not previously been associated with cachexia and a high increase in cortisol levels (Cala *et al.,* 2018). Together these findings suggest that metabolomics based research has potential to profile lipid and amino acid metabolism but large patient numbers are required to account for the multiple variations in metabolites that have already been seen.

All bar one of these studies (Eisener *et al.,* 2011) similarly to ours used 5% weight loss as a cut off to divide patients. Eisener *et al.,* used serial CT scanning and SMI to measure muscle loss or gain over time to compare predictive models. What is particularly interesting about the current study is that the analysis was performed blindly. Patients were not grouped until the metabolic profiles were determined suggesting that metabolomics is able to predict percentage weight loss in these patients. Future studies should also compare cancer patients with proven CT sarcopenia and weight loss to those who are weight stable.

Although many have suggested that metabolomic analysis has the potential to change how nutrition research is conducted, much of this potential remains unrealised. Blood-based metabolomics is a promising method for cachexia research. However, as seen results are often difficult to replicate due to the heterogeneity of the populations and study sizes. One obvious limitation of this study is the number of samples used and the differences in sex between the groups. This was an exploratory study involving patients without refractory cachexia and was not designed to identify sex specific differences. It was designed to give a better understanding of the complex pathophysiology apparent in cachexia. There was a mixture of tumour types in each of the groups. Despite this we were able to demonstrate a metabolic profile consistent with the definition of 5% weight loss. This may have implications when considering the definitions of cachexia for use in clinical trials.

6.5 Conclusion

These results show that metabolomics profiling of plasma from cancer patients is different between cachectic and non-cachectic patients. Differences highlighted in the breakdown of lipids provide an understanding of the mechanisms involved in the pathogenesis of cachexia. A better understanding of these mechanisms and the potential sharing of these datasets between groups might identify novel therapeutic approaches to treat this clinical condition.

Chapter Seven:

General Discussion and Conclusions

7 GENERAL DISCUSSION AND CONCLUSION

From the outset, this study was intended to investigate the application of metabolomics in human health and disease by assessing biomarkers associated with various diseases. All the studies described in this thesis employed a liquid chromatography-mass spectrometry (LC-MS) analytical platform based on the Orbitrap Exactive mass analyser, and using hydrophilic interaction liquid chromatography (HILIC) or/and reversed phase (RP) analytical columns. The LC-MS platform employed XCalibur software through which the system functionality was controlled. This platform has the advantage of accurate mass detection which provides capacity for direct metabolite identification even in the absence of chromatographic resolution. Metabolite identification was based on retention times of the samples relative to authentic reference standards injected at specified intervals into the system in the same sequence. In addition, all the studies employed both unsupervised (PCA-X) and supervised (OPLS-DA) models in SIMCA in order to determine discriminating metabolite biomarkers responsible for the observed natural clustering patterns and supervised separations in OPLS-DA.

The first project under this study (**Chapter 3**) investigated the metabolomic effects of an 80 km ultramarathon exercise. The runners underwent an 80K treadmill and metabolomic profiling was carried out on plasma samples obtained before (pre-80K) and after the trial (post-80K). The samples were analysed by using high resolution mass spectrometry in combination with both HILIC and RP chromatography. The data was extracted and

searched against a metabolite library. The extracted and putatively identified features were modelled using Simca P14.1 software. A principal component analysis (PCA) of the HILIC data showed clear separation between the pre- and post-exercise samples. The pre- and post-samples analysed by RP were also separated by PCA. FDR analysis indicated that all features with a P value <0.05 were significant. In this study, many of the amino acids were lowered in plasma post-exercise but the clearest impact of endurance exercise observed was on fatty acid metabolism but with respect to formation of medium chain unsaturated and partially oxidised fatty acids and conjugates of fatty acids with carnitines. Many of these metabolites were increased several fold. The most likely explanation for the complex pattern of medium chain and oxidised fatty acids formed could be that exercise provokes the proliferation of peroxisomes. The peroxisomes may serve two functions. First, they may provide a readily utilisable form of energy through formation of acetyl carnitine and other acyl carnitines for export to mitochondria in the muscles; which can utilise these substrates without investment of the ATP required to conjugate free fatty acids to CoA. Secondly the peroxisomes may serve to regulate the levels of oxidised metabolites of long chain fatty acids since many of these metabolites can provoke biological responses such as vasoconstriction or have pro-inflammatory activity.

The second study described in **Chapter 4** evaluated the metabolomics effects of *E. coli* incubation in different carbon sources using three types of fibres: 1% cooked meat, 1% maize meal and 1% olive kernel oil. Each of these was compared to the negative control

cultures enriched with 1% D glucose. The key observation from this study was that there were significant effects on various metabolite pathways particularly those associated with amino acid, lipid, carbohydrate, and nucleotide metabolism. In addition, there were effects on intermediates of peptide and polyketide biosynthesis, as well as on xenobiotic breakdown products and vitamins cofactors. Taken together, these findings suggested that the *E. coli* metabolome is closely associated with the type of fibre that the microorganism is subjected to. This observation is consistent with many previous studies which reported that the metabolome of gut microbiota depends on the microbial composition and functional capacity of the microorganisms, as well as nutrient availability and its physicochemical properties, age of host, and transit time of the colon. Nutrient availability, particularly the carbohydrate to nitrogen ratio, is believed to be the most important regulator of bacterial metabolism, as it influences preference of saccharolytic vs proteolytic fermentation.

In **Chapter 5**, the findings of the third study are reported. This study investigated the metabolomic effects of 7 different dietary fibres on incubations of faecal samples taken from Crohn's disease patients and healthy controls. There was little effect on the metabolic profile of the samples resulting from the seven different fibres. The major differences were between the CD samples and the HC samples which had higher levels of fatty acids and lower levels of haemoglobin metabolites. There were also major differences between the 0h incubation samples and the 24/48 h incubation samples for both the HC

and CD samples where the most interesting observation was the formation of the neurologically active compound melatonin with time.

The final study reported in **Chapter 6** investigated whether it was possible to predict cancer associated muscle wasting from plasma metabolites. Participants recruited for this study were men and non-pregnant women over 18 years of age. Written informed consent was obtained from all subjects and ethical approval received from Lothian Research Ethics Committee (UK). Participating patients had a diagnosis of upper gastrointestinal cancer (oesophageal, gastric, pancreatic) and were undergoing surgery with the intent of resection of the primary tumour. The key findings observed in this study were that the levels of significantly altered metabolites were generally higher in patients who had lost so much weight (>7.6 kg weight loss). The discriminating metabolites belonged mainly to the lipid metabolic pathways where long chain fatty acids and lysolipids were affected. Unlike in previous studies on urinary metabolites in cancer cachexia, there were virtually no significantly affected metabolites in the amino acid metabolic pathway, the urea cycle, intermediary metabolism (glycolysis, TCA cycle, 1-carbon metabolism) and creatinine pathway due perhaps to different levels of these metabolites in the two biofluids. It should be remembered that all participants considered in this study had some form of gastrointestinal cancer; the only difference between them was in the extent of the associated weight loss. It is thus not surprising that the differences observed between the two groups of weight loss were not very marked. Nevertheless, the observed

effects on lipid metabolisms in cancer cachexia suggests that there is an increased tendency towards peroxisomal proliferation in patients who had lost significant muscle mass, thus demonstrating an association between lipolysis-promoting activity in the serum of cachectic cancer patients and weight loss.

In conclusion, the various studies presented in this thesis have reiterated the ability of LC-MS based metabolomic profiling in discriminating certain disease states from the physiological state. Although the studies considered vastly differing physical states ranging from healthy participants performing a simulated ultramarathon exercise on a treadmill to diseased participants suffering from either Crohn's disease or gastrointestinal cancer, our findings reveal that LC-MS based metabolomics was capable of determining the metabolic alterations associated with each disease state. This further reinforces the capacity for metabolomics in discovering new biomarkers for various diseases that could be crucial in the diagnosis, monitoring disease progression, therapeutic efficacy evaluation of novel treatment, and detecting relapses following treatment. It is considered that depending on the disease involved, metabolite enrichment in various biofluids (such as plasma or urine) may vary and so it is recommended that careful elaluation of different biofluids should be conducted to ascertain the one with the most unique biomarkers for diagnostic purposes.

FUTURE WORK

With regard to the metabolomics of extreme exercise there is still more information that can be obtained from the the data set since there is rich metadata associated with it such as VO_{2max}, measurments and marathon completion times. Thus, more data modelling could be carried out on the existing data set.

The observations on the incubations of faecal samples with different dietary fibres produced some interesting preliminary observations. It would be interesting to explore the metabolism of tryptophan in detail in these samples by using stable isotope labelled tryptophan as a substrate which both provide strong confirmation of the identities of metabolites such as melatonin and allow the rate of metabolism to be explored.

Extending the work on cachexia would ideally require another cohort of patients to provide samples in order to confirm that the metabolomic changes are reproducible. Discovery of reliable biomarkers could improve treatment through allowing the markers to be monitored to indicate the effectiveness of treatment.

8 REFERENCES

Ackerman D, Simon MC. Hypoxia, lipids and cancer: surviving the harsh tumor microenvironment. Trends Cell Biol. 2015; 24(8):472-478.

Ahmadian M, Duncan RE, Jaworski K, Sarkadi-Nagy E, Sul HS. Triacylglycerol metabolism in adipose tissue. Future Lipidol. 2009;2(2):229-237.

Al-Jaouni, R., Hébuterne, X., Pouget, I. & Rampal, P. Energy metabolism and substrate oxidation in patients with Crohn's disease. Nutrition 16, 173-178 (2000).

Argiles JM, Busquets S, Stemmler B, Lopez-Soriano FJ. Cancer cachexia: understanding the molecular basis. Nat Rev Cancer 2014;14:754–762.

ARIJS, I., QUINTENS, R., VAN LOMMEL, L., VAN STEEN, K., DE HERTOGH, G., LEMAIRE, K., SCHRAENEN, A., PERRIER, C., VAN ASSCHE, G. & VERMEIRE, S. 2010. Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease. Inflammatory bowel diseases, 16, 2090-2098.

ASSFALG, M., BERTINI, I., COLANGIULI, D., LUCHINAT, C., SCHÄFER, H., SCHÜTZ, B. & SPRAUL, M. 2008. Evidence of different metabolic phenotypes in humans. Proceedings of the National Academy of Sciences, 105, 1420-1424.

BACHMANN, J., HEILIGENSETZER, M., KRAKOWSKI-ROOSEN, H., BÜCHLER, M. W., FRIESS, H. & MARTIGNONI, M. E. 2008. Cachexia worsens prognosis in patients with resectable pancreatic cancer. Journal of Gastrointestinal Surgery, 12, 1193.

BAJOR, A., GILLBERG, P.-G. & ABRAHAMSSON, H. 2010. Bile acids: short and long term effects in the intestine. Scandinavian journal of gastroenterology, 45, 645-664.

BAJPAI, J., SINHA, B. & SRIVASTAVA, A. 1975. Clinical study of Volkmann's ischemic contracture of the upper limb. International surgery, 60, 162-164.

BAKER, A. B. & TANG, Y. Q. 2010. Aging performance for masters records in athletics, swimming, rowing, cycling, triathlon, and weightlifting. Experimental aging research, 36, 453-477.

BALASUBRAMANIAN, K., KUMAR, S., SINGH, R. R., SHARMA, U., AHUJA, V., MAKHARIA, G. K. & JAGANNATHAN, N. R. 2009. Metabolism of the colonic mucosa in patients with inflammatory bowel diseases: an in vitro proton magnetic resonance spectroscopy study. Magnetic resonance imaging, 27, 79-86.

BAWAZEER, S., SUTCLIFFE, O. B., EUERBY, M. R., BAWAZEER, S. & WATSON, D. G. 2012. A comparison of the chromatographic properties of silica gel and silicon hydride modified silica gels. Journal of Chromatography A, 1263, 61-67.

BEDAIR, M. & SUMNER, L. W. 2008. Current and emerging mass-spectrometry technologies for metabolomics. TrAC Trends in Analytical Chemistry, 27, 238-250.

Beloribi-Djefaflia S, Vasseur S, Guillaumond F. Lipid metabolic reprogramming in cancer cells. Oncogenesis 2016;5(1): e189.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the False Discovery Rate a Practical and powerful Approach to Multiple Testing. Journal of the Royal Statistical Society, 57, 289-300.

Berdel WE, Von Hoff DD, Unger C, Schick HD, Fink U, Reichert A, Eibl H, Rastetter J. Ether lipid derivatives: antineoplastic activity in vitro and the structure-activity relationship. Lipids 1986;21:301–304.

BERNDT, U., BARTSCH, S., PHILIPSEN, L., DANESE, S., WIEDENMANN, B., DIGNASS, A. U., HÄMMERLE, M. & STURM, A. 2007. Proteomic analysis of the inflamed intestinal mucosa reveals distinctive immune response profiles in Crohn's disease and ulcerative colitis. The Journal of Immunology, 179, 295-304.

BJERRUM, J. T., NIELSEN, O. H., HAO, F., TANG, H., NICHOLSON, J. K., WANG, Y. & OLSEN, J. 2009. Metabonomics in ulcerative colitis: diagnostics, biomarker identification, and insight into the pathophysiology. Journal of proteome research, 9, 954-962.

BLEKHERMAN, G., LAUBENBACHER, R., CORTES, D. F., MENDES, P., TORTI, F. M., AKMAN, S., TORTI, S. V. & SHULAEV, V. 2011. Bioinformatics tools for cancer metabolomics. Metabolomics, 7, 329-343.

BUDCZIES, J., DENKERT, C., MÜLLER, B. M., BROCKMÖLLER, S. F., KLAUSCHEN, F., GYÖRFFY, B., DIETEL, M., RICHTER-EHRENSTEIN, C., MARTEN, U. & SALEK, R. M. 2012. Remodeling of central metabolism in invasive breast cancer compared to normal breast tissue–a GC-TOFMS based metabolomics study. BMC genomics, 13, 334.

Cala MP, Agulló-Ortuño MT, Prieto-García E, González-Riano C, Parrilla-Rubio L, Barbas C, Díaz-García CV, García A, Pernaut C, Adeva J, Riesco MC, Rupérez FJ, and Lopez-Martin JA. Multiplatform plasma fingerprinting in cancer cachexia: a pilot observational and translational study. Journal of Cachexia, Sarcopenia and Muscle 2018; 9: 348–357.

CARNETHON, M. R., GULATI, M. & GREENLAND, P. 2005. Prevalence and cardiovascular disease correlates of low cardiorespiratory fitness in adolescents and adults. Jama, 294, 2981-2988.

CHONG, I.-G. & JUN, C.-H. 2005. Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems, 78, 103-112.

CLAESSON, M. J., JEFFERY, I. B., CONDE, S., POWER, S. E., O'CONNOR, E. M., CUSACK, S., HARRIS, H. M., COAKLEY, M., LAKSHMINARAYANAN, B. & O'SULLIVAN, O. 2012. Gut microbiota composition correlates with diet and health in the elderly. Nature, 488, 178.

CUMMINGS, J. H. 1981. Short chain fatty acids in the human colon. Gut, 22, 763.

Dahlman, I, N Mejhert, K Linder, T Agustsson, D M Mutch, A Kulyte, B Isaksson, et al. 2010. “Adipose Tissue Pathways Involved in Weight Loss of Cancer Cachexia.” British Journal of Cancer 102 (10): 1541–48. doi:10.1038/sj.bjc.6605665.

Das SK, Hoefler G. The role of triglyceride lipases in cancer associated cachexia. Trends Mol Med 2013; 19: 292–301.

DASKALAKI, E., EASTON, C. & G WATSON, D. 2014. The application of metabolomic profiling to the effects of physical activity. Current Metabolomics, 2, 233-263.

DE PRETER, V. & VERBEKE, K. 2013. Metabolomics as a diagnostic tool in gastroenterology. World journal of gastrointestinal pharmacology and therapeutics, 4, 97.

De Preter, V. et al. Faecal metabolite profiling identifies medium-chain fatty acids as discriminating compounds in IBD. Gut 64, 447-458 (2015).

DE PRETER, V., HAMER, H. M., WINDEY, K. & VERBEKE, K. 2011. The impact of pre-and/or probiotics on human colonic metabolism: Does it affect human health? Molecular nutrition & food research, 55, 46-57.

DETTMER, K., ARONOV, P. A. & HAMMOCK, B. D. 2007. Mass spectrometry-based metabolomics. Mass spectrometry reviews, 26, 51-78.

DEWYS, W. D., BEGG, C., LAVIN, P. T., BAND, P. R., BENNETT, J. M., BERTINO, J. R., COHEN, M. H., DOUGLASS JR, H. O., ENGSTROM, P. F. & EZDINLI, E. Z. 1980. Prognostic effect of weight loss prior tochemotherapy in cancer patients. The American journal of medicine, 69, 491-497.

DOTAN, I. 2010. New serologic markers for inflammatory bowel disease diagnosis. Digestive diseases, 28, 418-423.

DOVIO, A., ROVEDA, E., SCIOLLA, C., MONTARULI, A., RAFFAELLI, A., SABA, A., CALOGIURI, G., DE FRANCIA, S., BORRIONE, P. & SALVADORI, P. 2010. Intense physical exercise increases systemic 11β-hydroxysteroid dehydrogenase type 1 activity in healthy adult subjects. European journal of applied physiology, 108, 681-687.

DUDZINSKA, W., LUBKOWSKA, A., JAKUBOWSKA, K., SUSKA, M. & SKOTNICK, E. 2013. Insulin resistance induced by maximal exercise correlates with a post-exercise increase in uridine concentration in the blood of healthy young men. Physiological research, 62.

DUNN, W. B. & ELLIS, D. I. 2005. Metabolomics: current analytical platforms and methodologies. TrAC Trends in Analytical Chemistry, 24, 285-294.

DUNN, W. B., BAILEY, N. J. & JOHNSON, H. E. 2005. Measuring the metabolome: current analytical technologies. Analyst, 130, 606-625.

DUNN, W. B., ERBAN, A., WEBER, R. J. M., CREEK, D. J., BROWN, M., BREITLING, R., HANKEMEIER, T., GOODACRE, R., NEUMANN, S., KOPKA, J. & VIANT, M. R. 2013. Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. Metabolomics, 9, 44-66.

Ebadi M, Mazurak VC. Potential biomarkers of fat loss as a feature of cancer cachexia. Mediators Inflamm. 2015; 820934.

EISNER, R., STRETCH, C., EASTMAN, T., XIA, J., HAU, D., DAMARAJU, S., GREINER, R., WISHART, D. S. & BARACOS, V. E. 2011. Learning to predict cancer-associated skeletal muscle wasting from 1H-NMR profiles of urinary metabolites. Metabolomics, 7, 25-34.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTROM, C. 2013. Multi- and Megavariate Data Analysis: Basic Principles and Application, Sweden, MKS Umetrics AB.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTROM, C. 2013a. Cross validation. Multi- and Megavariate Data Analysis: Basic Principles and Application. 3 ed. Sweden: MKS Umetrics AB.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTROM, C. 2013b. Response permutation and cross-validation. Multi- and Megavariate Data Analysis: Basic Principles and Application. 3 ed. Sweden: MKS Umetrics AB.

ERIKSSON, L., TRYGG, J. & WOLD, S. 2008. CV-ANOVA for significance testing of PLS and OPLS® models. Journal of Chemometrics, 22, 594-600.

EVENEPOEL, P., MEIJERS, B. K., BAMMENS, B. R. & VERBEKE, K. 2009. Uremic toxins originating from colonic microbial metabolism. Kidney International, 76, S12-S19.

FAULKNER, J. A., DAVIS, C. S., MENDIAS, C. L. & BROOKS, S. V. 2008. The aging of elite male athletes: age-related changes in performance and skeletal muscle structure and function. Clinical journal of sport medicine: official journal of the Canadian Academy of Sport Medicine, 18, 501.

Fearon KC, Voss AC, Hustead DS, Cancer Cachexia Study Group. Definition of cancer cachexia: effect of weight loss, reduced food intake and systemic inflammation on functional status and prognosis. American Journal of Clinical Nutrition 2006; 83(6): 1345-1350.

Fearon KCH, Glass DJ, Guttridge DC. Cancer cachexia: Mediators, signaling and metabolic pathways. Cell Metabolism 2012; 16(2):153-166.

FEARON, K. C., VOSS, A. C. & HUSTEAD, D. S. 2006. Definition of cancer cachexia: effect of weight loss, reduced food intake, and systemic inflammation on functional status and prognosis–. The American journal of clinical nutrition, 83, 1345-1350.

FEARON, K., STRASSER, F., ANKER, S. D., BOSAEUS, I., BRUERA, E., FAINSINGER, R. L., JATOI, A., LOPRINZI, C., MACDONALD, N. & MANTOVANI, G. 2011. Definition and classification of cancer cachexia: an international consensus. The lancet oncology, 12, 489-495.

FIEHN, O. 2002. Metabolomics—the link between genotypes and phenotypes. Functional genomics. Springer.

Fujiwara Y, Kobayashi T, Chayahara N, Imamura Y, Toyoda M, Kiyota N, et al. Metabolomics Evaluation of Serum Markers for Cachexia and Their Intra-Day Variation in Patients with Advanced Pancreatic Cancer. PLoS One 2014;9:e113259.

GATES, S. C. & SWEELEY, C. C. 1978. Quantitative metabolic profiling based on gas chromatography. Clinical chemistry, 24, 1663-1673.

GATTI, R., CAPPELLIN, E., ZECCHIN, B., ANTONELLI, G., SPINELLA, P., MANTERO, F. & DE PALO, E. F. 2005. Urinary high performance reverse phase chromatography cortisol and cortisone analyses before and at the end of a race in elite cyclists. Journal of Chromatography B, 824, 51-56.

Gercel-Taylor C. Doering DL, Kraemer FB, Taylor DD. Abberations in normal systemic lipid metabolism in ovarian cancer patients. Gynecol Oncol. 1996;60(1):35-41.

GOODACRE, R. 2007. Metabolomics of a Superorganism1–3. The Journal of Nutrition, 137, 137259S–266S.

GRAY, C., MACGILLIVRAY, T. J., EELEY, C., STEPHENS, N. A., BEGGS, I., FEARON, K. C. & GREIG, C. A. 2011. Magnetic resonance imaging with k-means clustering objectively measures whole muscle volume compartments in sarcopenia/cancer cachexia. Clinical Nutrition, 30, 106-111.

GUAN, W., ZHOU, M., HAMPTON, C. Y., BENIGNO, B. B., WALKER, L. D., GRAY, A., MCDONALD, J. F. & FERNÁNDEZ, F. M. 2009. Ovarian cancer detection from metabolomic

liquid chromatography/mass spectrometry data by support vector machines. BMC bio-informatics, 10, 259.

Gupta, N. K. et al. Serum analysis of tryptophan catabolism pathway: correlation with Crohn's disease activity. Inflammatory bowel diseases 18, 1214-1220 (2011).

HARBER, M. P., CRANE, J. D., DICKINSON, J. M., JEMIOLO, B., RAUE, U., TRAPPE, T. A. & TRAPPE, S. W. 2009. Protein synthesis and the expression of growth-related genes are altered by running in human vastus lateralis and soleus muscles. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 296, R708-R714.

HARPER, S. & HOWSE, K. 2008. An upper limit to human longevity? Journal of Population Ageing, 1, 99-106.

HARRIS, D. C. 2010. Quantitative Chemical Analysis, New York, W. H. Freeman.

HISAMATSU, T., OKAMOTO, S., HASHIMOTO, M., MURAMATSU, T., ANDOU, A., UO, M., KITAZUME, M. T., MATSUOKA, K., YAJIMA, T. & INOUE, N. 2012. Novel, objective, multivariate biomarkers composed of plasma amino acid profiles for the diagnosis and assessment of inflammatory bowel disease. PLoS One, 7, e31131.

HOLMES, E., WILSON, I. D. & NICHOLSON, J. K. 2008. Metabolic phenotyping in health and disease. Cell, 134, 714-717.

HU, Q. Z., NOLL, R. J., LI, H. Y., MAKAROV, A., HARDMAN, M. & COOKS, R. G. 2005. The Orbitrap: a new mass spectrometer. J Mass Spectr, 40, 430-443.

HUFFMAN, K. M., KOVES, T. R., HUBAL, M. J., ABOUASSI, H., BERI, N., BATEMAN, L. A., STEVENS, R. D., ILKAYEVA, O. R., HOFFMAN, E. P. & MUOIO, D. M. 2014. Metabolite signatures of exercise training in human skeletal muscle relate to mitochondrial remodelling and cardiometabolic fitness. Diabetologia, 57, 2282-2295.

HUNTER, S. K., STEVENS, A. A., MAGENNIS, K., SKELTON, K. W. & FAUTH, M. 2011. Is there a sex difference in the age of elite marathon runners? Medicine & Science in Sports & Exercise.

HUTTENHOWER, C., GEVERS, D., KNIGHT, R., ABUBUCKER, S., BADGER, J. H., CHINWALLA, A. T., CREASY, H. H., EARL, A. M., FITZGERALD, M. G. & FULTON, R. S. 2012. Structure, function and diversity of the healthy human microbiome. Nature, 486, 207.

JANSSON, J., WILLING, B., LUCIO, M., FEKETE, A., DICKSVED, J., HALFVARSON, J., TYSK, C. & SCHMITT-KOPPLIN, P. 2009. Metabolomics reveals metabolic biomarkers of Crohn's disease. PloS one, 4, e6386.

JOHNSON, H. E., BROADHURST, D., GOODACRE, R. & SMITH, A. R. 2003. Metabolic fingerprinting of salt-stressed tomatoes. Phytochemistry, 62, 919-928.

JOOSSENS, M., HUYS, G., CNOCKAERT, M., DE PRETER, V., VERBEKE, K., RUTGEERTS, P., VANDAMME, P. & VERMEIRE, S. 2011. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut, gut. 2010.223263.

KADERBHAI, N. N., BROADHURST, D. I., ELLIS, D. I., GOODACRE, R. & KELL, D. B. 2003. Functional genomics via metabolic footprinting: monitoring metabolite secretion by Escherichia coli tryptophan metabolism mutants using FT–IR and direct injection electrospray mass spectrometry. Comparative and functional genomics, 4, 376-391.

KATAJAMAA, M. & OREŠIČ, M. 2007. Data processing for mass spectrometry-based metabolomics. Journal of Chromatography A, 1158, 318-328.

Kennedy, P. J., Cryan, J. F., Dinan, T. G. & Clarke, G. Kynurenine pathway metabolism and the microbiota-gut-brain axis. Neuropharmacology 112, 399-412 (2017).

KIM, D.-H., JARVIS, R. M., XU, Y., OLIVER, A. W., ALLWOOD, J. W., HAMPSON, L., HAMPSON, I. N. & GOODACRE, R. 2010. Combining metabolic fingerprinting and footprinting

to understand the phenotypic response of HPV16 E6 expressing cervical carcinoma cells exposed to the HIV anti-viral drug lopinavir. Analyst, 135, 1235-1244.

KIM, K., ARONOV, P., ZAKHARKIN, S. O., ANDERSON, D., PERROUD, B., THOMPSON, I. M. & WEISS, R. H. 2009. Urine metabolomics analysis for kidney cancer detection and biomarker discovery. Molecular & cellular proteomics, 8, 558-570.

KIRWAN, G. M., JOHANSSON, E., KLEEMANN, R., VERHEIJ, E. R., WHEELOCK, A. M., GOTO, S., TRYGG, J. & WHEELOCK, C. E. 2012. Building multivariate systems biology models. Anal Chem, 84, 7064-71.

Klein S, Wolfe RR. Whole-body lipolysis and triglyceride-fatty acid cycling in cachectic patients with oesophageal cancer. J Clin Invest. 1990; 86(5): 1403-1408.

KNECHTLE, B., WIRTH, A., KNECHTLE, P., ZIMMERMANN, K. & KOHLER, G. 2009. Personal best marathon performance is associated with performance in a 24-h run and not anthropometry or training volume. British Journal of Sports Medicine, 43, 836-839.

KOPKA, J. 2006. Current challenges and developments in GC–MS based metabolite profiling technology. Journal of biotechnology, 124, 312-322.

KRAJ, A., DESIDERIO, D. M. & NIBBERING, N. M. 2008. Mass spectrometry: instrumentation, interpretation, and applications, John Wiley & Sons.

Kriat M, Vion-Dury J, Confort-Gouny S, Favre R, Viout P, Sciaky M, Sari H, Cozzone PJ. Analysis of plasma lipids by NMR spectroscopy: application to modifications induced by malignant tumors. J Lipid Res. 1993 Jun; 34(6):1009-19.

KROGIUS-KURIKKA, L., LYRA, A., MALINEN, E., AARNIKUNNAS, J., TUIMALA, J., PAULIN, L., MÄKIVUOKKO, H., KAJANDER, K. & PALVA, A. 2009. Microbial community analysis reveals high level phylogenetic alterations in the overall gastrointestinal microbiota of diarrhoea-predominant irritable bowel syndrome sufferers. BMC gastroenterology, 9, 95.

LAMENDELLA, R., VERBERKMOES, N. & JANSSON, J. K. 2012. 'Omics' of the mammalian gut–new insights into function. Current opinion in biotechnology, 23, 491-500.

Lapidus, A., Åkerlund, J.-E. & Einarsson, C. Gallbladder bile composition in patients with Crohn's disease. World journal of gastroenterology: WJG 12, 70 (2006).

LE GALL, G., NOOR, S. O., RIDGWAY, K., SCOVELL, L., JAMIESON, C., JOHNSON, I. T., COLQUHOUN, I. J., KEMSLEY, E. K. & NARBAD, A. 2011. Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. Journal of proteome research, 10, 4208-4218.

Leníček, M. et al. The relationship between serum bilirubin and Crohn's disease. Inflammatory bowel diseases 20, 481-487 (2014).

LEPERS, R. & CATTAGNI, T. 2012. Do older athletes reach limits in their performance during marathon running? Age, 34, 773-781.

LOZANO, N. B., OLIVEIRA, R. F., WEBER, K. C., HONORIO, K. M., GUIDO, R. V., ANDRICOPULO, A. D., DE SOUSA, A. G. & DA SILVA, A. B. 2014. Pattern recognition techniques applied to the study of leishmanial glyceraldehyde-3-phosphate dehydrogenase inhibition. Int J Mol Sci, 15, 3186-203.

LUSTGARTEN, M. S., PRICE, L. L., LOGVINENKO, T., HATZIS, C., PADUKONE, N., REO, N. V., PHILLIPS, E. M., KIRN, D., MILLS, J. & FIELDING, R. A. 2013. Identification of serum analytes and metabolites associated with aerobic capacity. European journal of applied physiology, 113, 1311-1320.

MACINTYRE, L., ZHANG, T., VIEGELMANN, C., MARTINEZ, I. J., CHENG, C., DOWDELLS, C., ABDELMOHSEN, U. R., GERNERT, C., HENTSCHEL, U. & EDRADA-EBEL, R. 2014. Metabolomic tools for secondary metabolite discovery from marine microbial symbionts. Marine drugs, 12, 3416-3448.

MAEDA, J., HIGASHIYAMA, M., IMAIZUMI, A., NAKAYAMA, T., YAMAMOTO, H., DAIMON, T., YAMAKADO, M., IMAMURA, F. & KODAMA, K. 2010. Possibility of multivariate function composed of plasma amino acid profiles as a novel screening index for non-small cell lung cancer: a case control study. BMC Cancer, 10, 690.

MAHOWALD, M. A., REY, F. E., SEEDORF, H., TURNBAUGH, P. J., FULTON, R. S., WOLLAM, A., SHAH, N., WANG, C., MAGRINI, V. & WILSON, R. K. 2009. Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. Proceedings of the National Academy of Sciences, 106, 5859-5864.

MAKAROV, A. & SCIGELOVA, M. 2010. Coupling liquid chromatography to Orbitrap mass spectrometry. J Chromatogr A, 1217, 3938-45.

MARCHESI, J. R., HOLMES, E., KHAN, F., KOCHHAR, S., SCANLAN, P., SHANAHAN, F., WILSON, I. D. & WANG, Y. 2007. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. Journal of proteome research, 6, 546-551.

Marquez, E., Sánchez-Fidalgo, S., Calvo, J. R., Lastra, C. A. l. d. & Motilva, V. Acutely administered melatonin is beneficial while chronic melatonin treatment aggravates the evolution of TNBS-induced colitis. Journal of pineal research 40, 48-55 (2006).

Martin L, Birdsell L, MacDonald N, Reiman T et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. J Clin Oncol. 2013; 31(12):1539-47.

Mastrangelo A, Armitage EG, Garcia A, Barbas C. Metabolomics as a tool for drug discovery and personalised medicine. A review Curr Top Med Chem 2014;14:2627–2636.

MCNIVEN, E. M., GERMAN, J. B. & SLUPSKY, C. M. 2011. Analytical metabolomics: nutritional opportunities for personalized health. The Journal of nutritional biochemistry, 22, 995-1002.

Meer GV, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. Nat Rev Mol Cell Biol. 2008;9(2):112-124

METHÉ, B. A., NELSON, K. E., POP, M., CREASY, H. H., GIGLIO, M. G., HUTTENHOWER, C., GEVERS, D., PETROSINO, J. F., ABUBUCKER, S. & BADGER, J. H. 2012. A framework for human microbiome research. Nature, 486, 215.

MICHALSKI, A., DAMOC, E., HAUSCHILD, J.-P., LANGE, O., WIEGHAUS, A., MAKAROV, A., NAGARAJ, N., COX, J., MANN, M. & HORNING, S. 2011. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. Mol Cell Proteomics, 10, 1-11.

MICHALSKI, A., DAMOC, E., LANGE, O., DENISOV, E., NOLTING, D., MULLER, M., VINER, R., SCHWARTZ, J., REMES, P., BELFORD, M., DUNYACH, J. J., COX, J., HORNING, S., MANN, M. & MAKAROV, A. 2012. Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. Mol Cell Proteomics, 11, 1-11.

Minderhoud, I. M., Oldenburg, B., Schipper, M. E., Ter Linde, J. J. & Samsom, M. Serotonin synthesis and uptake in symptomatic patients with Crohn's disease in remission. Clinical Gastroenterology and Hepatology 5, 714-720 (2007).

Mingrone, G. et al. Elevated diet-induced thermogenesis and lipid oxidation rate in Crohn disease. The American journal of clinical nutrition 69, 325-330 (1999).

MOSES, A., SLATER, C., PRESTON, T., BARBER, M. & FEARON, K. 2004. Reduced total energy expenditure and physical activity in cachectic patients with pancreatic cancer can be modulated by an energy and protein dense oral supplement enriched with n-3 fatty acids. British journal of cancer, 90, 996.

MUHSEN ALI, A., BURLEIGH, M., DASKALAKI, E., ZHANG, T., EASTON, C. & WATSON, D. G. 2016. Metabolomic Profiling of Submaximal Exercise at a Standardised Relative Intensity in Healthy Adults. Metabolites, 6.

NEAL, C. M., HUNTER, A. M., BRENNAN, L., O'SULLIVAN, A., HAMILTON, D. L., DEVITO, G. & GALLOWAY, S. D. 2012. Six weeks of a polarized training-intensity distribution leads to greater physiological and performance adaptations than a threshold model in trained cyclists. Journal of applied physiology, 114, 461-471.

NICHOLLS, A. W., MORTISHIRE-SMITH, R. J. & NICHOLSON, J. K. 2003. NMR spectroscopic-based metabonomic studies of urinary metabolite variation in acclimatizing germ-free rats. Chemical research in toxicology, 16, 1395-1404.

NICHOLSON, J. K. & LINDON, J. C. 2008. Systems biology: metabonomics. Nature, 455, 1054.

NICHOLSON, J. K. 2006. Global systems biology, personalized medicine and molecular epidemiology. Molecular systems biology, 2, 52.

NIEMAN, D. C., A, S. R., B, L., MEANEY, M. P., A, D. D. & L, P. K. 2014. Metabolomics approach to assessing plasma 13-and 9-hydroxy-octadecadienoic acid and linoleic acid metabolite responses to 75-km cycling. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 307, R68-R74.

NIEMAN, D. C., MEANEY, M. P., JOHN, C. S., KNAGGE, K. J. & CHEN, H. 2016. 9-and 13-hydroxy-octadecadienoic acids (9+ 13 hode) are inversely related to granulocyte colony stimulating factor and il-6 in runners after 2 h running. Brain, behavior, and immunity, 56, 246-252.

O'Connell TM, Ardeshirpour F, Asher SA, Winnike JH, Yin X, George J, et al. Metabolomic analysis of cancer cachexia reveals distinct lipid and glucose alterations. Metabolomics 2008;4:216

ÖHMAN, L. & SIMRÉN, M. 2013. Intestinal microbiota and its role in irritable bowel syndrome (IBS). Current gastroenterology reports, 15, 323.

OLIVARES, M., LAPARRA, J. M. & SANZ, Y. 2013. Host genotype, intestinal microbiota and inflammatory disorders. British Journal of Nutrition, 109, S76-S80.

OLSEN, J., GERDS, T. A., SEIDELIN, J. B., CSILLAG, C., BJERRUM, J. T., TROELSEN, J. T. & NIELSEN, O. H. 2009. Diagnosis of ulcerative colitis before onset of inflammation by multivariate modeling of genome-wide gene expression data. Inflammatory bowel diseases, 15, 1032-1038.

OOI, M., NISHIUMI, S., YOSHIE, T., SHIOMI, Y., KOHASHI, M., FUKUNAGA, K., NAKAMURA, S., MATSUMOTO, T., HATANO, N. & SHINOHARA, M. 2011. GC/MS-based profiling of amino acids and TCA cycle-related molecules in ulcerative colitis. Inflammation Research, 60, 831-840.

ORGANIZATION, W. H. 2009. Unhealthy diets and physical inactivity. World Health Organization: Geneva, Switzerland, 1-2.

PRADO, C. M., BIRDSELL, L. A. & BARACOS, V. E. 2009. The emerging role of computerized tomography in assessing cancer cachexia. Current opinion in supportive and palliative care, 3, 269-275.

PRELORENDJOS, A. 2014. Multivariate Analysis of Metabonomic Data. PhD, Strathclyde.

RAMSAY, R. R. & ZAMMIT, V. A. 2004. Carnitine acyltransferases and their influence on CoA pools in health and disease. Molecular aspects of medicine, 25, 475-493.

RANSDELL, L. B., VENER, J. & HUBERTY, J. 2009. Masters athletes: an analysis of running, swimming and cycling performance by age and gender. Journal of Exercise Science & Fitness, 7, S61-S73.

RASMUSSEN, L. G., SAVORANI, F., LARSEN, T. M., DRAGSTED, L. O., ASTRUP, A. & ENGELSEN, S. B. 2011. Standardization of factors that influence human urine metabolomics. Metabolomics, 7, 71-83.

Raynal, P.; Montagner, A.; Dance, M.; Yart, A. Lysophospholipids and cancer: Current status and perpespectives. Pathol. Biol. 2005, 53, 57–62.

REDDY, J. K. & MANNAERTS, G. P. 1994. Peroxisomal lipid metabolism. Annual review of nutrition, 14, 343-370.

ROESSNER, U., WAGNER, C., KOPKA, J., TRETHEWEY, R. N. & WILLMITZER, L. 2000. Simultaneous analysis of metabolites in potato tuber by gas chromatography–mass spectrometry. The Plant Journal, 23, 131-142.

SANTALI, E. Y., EDWARDS, D., SUTCLIFFE, O. B., BAILES, S., EUERBY, M. R. & WATSON, D. G. 2014. A Comparison of Silica C and Silica Gel in HILIC Mode: The Effect of Stationary Phase Surface Area. Chromatographia, 77, 873-881.

SARRIS, J., O'NEIL, A., COULSON, C. E., SCHWEITZER, I. & BERK, M. 2014. Lifestyle medicine for depression. BMC psychiatry, 14, 107.

SCARBOROUGH, P., BHATNAGAR, P., WICKRAMASINGHE, K. K., ALLENDER, S., FOSTER, C. & RAYNER, M. 2011. The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006–07 NHS costs. Journal of public health, 33, 527-535.

SCHAUER, N., STEINHAUSER, D., STRELKOV, S., SCHOMBURG, D., ALLISON, G., MORITZ, T., LUNDGREN, K., ROESSNER-TUNALI, U., FORBES, M. G. & WILLMITZER, L. 2005. GC–MS libraries for the rapid identification of metabolites in complex biological samples. FEBS letters, 579, 1332-1337.

SCHICHO, R., SHAYKHUTDINOV, R., NGO, J., NAZYROVA, A., SCHNEIDER, C., PANACCIONE, R., KAPLAN, G. G., VOGEL, H. J. & STORR, M. 2012. Quantitative metabolomic profiling of serum, plasma, and urine by 1H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. Journal of proteome research, 11, 3344-3357.

SHARMA, U., SINGH, R. R., AHUJA, V., MAKHARIA, G. K. & JAGANNATHAN, N. R. 2010. Similarity in the metabolic profile in macroscopically involved and un-involved colonic mucosa in patients with inflammatory bowel disease: an in vitro proton (1H) MR spectroscopy study. Magnetic resonance imaging, 28, 1022-1029.

Sharon, G. et al. Specialized metabolites from the microbiome in health and disease. Cell metabolism 20, 719-730 (2014).AHMED, I., GREENWOOD, R., DE LACY COSTELLO, B., RATCLIFFE, N. M. & PROBERT, C. S. 2013. An investigation of fecal volatile organic metabolites in irritable bowel syndrome. PloS one, 8, e58204.

Shaw JH, Wolfe RR. Fatty acid and glycerol kinetics in septic patients and in patients with gastrointestinal cancer. The response to glucose infusion and parenteral feeding. Ann Surg 1987; 205: 368–376.

SHULAEV, V. 2006. Metabolomics technology and bioinformatics. Briefings in bioinformatics, 7, 128-139.

SMITH, E. A. & MACFARLANE, G. T. 1996. Enumeration of human colonic bacteria producing phenolic and indolic compounds: effects of pH, carbohydrate availability and retention time on dissimilatory aromatic amino acid metabolism. Journal of Applied Bacteriology, 81, 288-302.

SPARLING, P. B., O'DONNELL, E. M. & SNOW, T. K. 1998. The gender difference in distance running performance has plateaued: an analysis of world rankings from 1980 to 1996. Medicine and Science in Sports and Exercise, 30, 1725-1729.

STATHIS, C. G., CAREY, M. F., HAYES, A., GARNHAM, A. P. & SNOW, R. J. 2006. Sprint training reduces urinary purine loss following intense exercise in humans. Applied physiology, nutrition, and metabolism, 31, 702-708.

STEPHENS, N. S., SIFFLEDEEN, J., SU, X., MURDOCH, T. B., FEDORAK, R. N. & SLUPSKY, C. M. 2013. Urinary NMR metabolomic profiles discriminate inflammatory bowel disease from healthy. Journal of Crohn's and Colitis, 7, e42-e48.

Stretch C, Eastman T, Mandal R, Eisner R, Wishart DS, Moutzakis M, Prado CMM, Damaraju S, Ball RO, Greiner R, Baracos VE. Prediction of skeletal muscle and fat mass in patients with advanced cancer using a metabolomic approach. The Journal of Nutrition 2011; 1(1):14-21

SUBUDHI, A. W., DAVIS, S. L., KIPP, R. W. & ASKEW, E. W. 2001. Antioxidant status and oxidative stress in elite alpine ski racers. International journal of sport nutrition and exercise metabolism, 11, 32-41.

SUGIMOTO, M., KAWAKAMI, M., ROBERT, M., SOGA, T. & TOMITA, M. 2012. Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis. Current bioinformatics, 7, 96-108.

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. 2007. Proposed minimum reporting standards for chemical analysis. Metabolomics, 3, 211-221.

Sutphen, R.; Xu, Y.; Wilbanks, G.D.; Fiorica, J.; Grendys, E.C.; LaPolla, J.P.J.; Arango, H.; Hoffman, M.S.; Martino, M.; Wakeley, K.; Griffin, D.; Blanco, R.W.; Cantor, A.B.; Xiao, Y.J.; Krischer, J.P. Lysophospholipids are potential biomarkers of ovarian cancer. Cancer Epidemiol.Biomarkers Prev. 2004, 13, 1185–1191.

TAKAHASHI, H., MORIMOTO, T., OGASAWARA, N. & KANAYA, S. 2011. AMDORAP: Non-targeted metabolic profiling based on high-resolution LC-MS. BMC Bioinformatics, 12, 259.

Tan, D. X. et al. Mitochondria and chloroplasts as the original sites of melatonin synthesis: a hypothesis related to melatonin's primary function and evolution in eukaryotes. Journal of pineal research 54, 127-138 (2013).

TANAKA, H. & SEALS, D. R. 2008. Endurance exercise performance in Masters athletes: age-associated changes and underlying physiological mechanisms. The Journal of physiology, 586, 55-63.

Taylor LA, Arends J, Hodina AK, Unger C, Massing U. Plasma lyso-phosphatidylcholine concentration is decreased in cancer patients with weight loss and activated inflammatory status. Lipids Health Dis. 2007;6:17.

Terry, P. D., Villinger, F., Bubenik, G. A. & Sitaraman, S. V. Melatonin and ulcerative colitis: evidence, biological mechanisms, and future research. Inflammatory bowel diseases 15, 134-140 (2008).

THOMAS, A. W., DAVIES, N. A., MOIR, H., WATKEYS, L., RUFFINO, J. S., ISA, S. A., BUTCHER, L. R., HUGHES, M. G., MORRIS, K. & WEBB, R. 2011. Exercise-associated generation of PPARγ ligands activates PPARγ signaling events and upregulates genes related to lipid metabolism. Journal of applied physiology, 112, 806-815.

TRAPPE, S. 2007. Marathon runners. Sports Medicine, 37, 302-305.

TRIBA, M. N., LE MOYEC, L., AMATHIEU, R., GOOSSENS, C., BOUCHEMAL, N., NAHON, P., RUTLEDGE, D. N. & SAVARIN, P. 2015. PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. Mol Biosyst, 11, 13-9.

TRYGG, J., HOLMES, E. & LUNDSTEDT, T. 2007. Chemometrics in metabonomics. J Proteome Res, 6, 469-79.

Tsoli, Maria, Michael M. Swarbrick, and Graham R. Robertson 2016. “Lipolytic and Thermogenic Depletion of Adipose Tissue in Cancer Cachexia.” Seminars in Cell & Developmental Biology, Mechanisms of cancer cachexia. Meiosis and Recombination, 54 (June): 68–81. doi:10.1016/j.semcdb.2015.10.039.

TUCA, A., JIMENEZ-FONSECA, P. & GASCÓN, P. 2013. Clinical evaluation and optimal management of cancer cachexia. Critical reviews in oncology/hematology, 88, 625-636.

TUOHY, K. M., GOUGOULIAS, C., SHEN, Q., WALTON, G., FAVA, F. & RAMNANI, P. 2009. Studying the human gut microbiota in the trans-omics era-focus on metagenomics and metabonomics. Current pharmaceutical design, 15, 1415-1427.

VAN BEELEN GRANLUND, A., FLATBERG, A., ØSTVIK, A. E., DROZDOV, I., GUSTAFSSON, B. I., KIDD, M., BEISVAG, V., TORP, S. H., WALDUM, H. L. & MARTINSEN, T. C. 2013. Whole genome gene expression meta-analysis of inflammatory bowel disease colon mucosa demonstrates lack of major differences between Crohn's disease and ulcerative colitis. PLoS One, 8, e56818.

VAN DER GREEF, J. & SMILDE, A. K. 2005. Symbiosis of chemometrics and metabolomics: past, present, and future. Journal of Chemometrics: A Journal of the Chemometrics Society, 19, 376-386.

Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: The metabolic requirements of cell proliferation. 2010;324(5930): 1029-1033

VENTURA, M., TURRONI, F., CANCHAYA, C., VAUGHAN, E. E., O'TOOLE, P. W. & VAN SINDEREN, D. 2009. Microbial diversity in the human intestine and novel insights from metagenomics.

WALKER, D. K., DICKINSON, J. M., TIMMERMAN, K. L., DRUMMOND, M. J., REIDY, P. T., FRY, C. S., GUNDERMANN, D. M. & RASMUSSEN, B. B. 2011. Exercise, amino acids and aging in the control of human muscle protein synthesis. Medicine and science in sports and exercise, 43, 2249.

WALTON, C., FOWLER, D. P., TURNER, C., JIA, W., WHITEHEAD, R. N., GRIFFITHS, L., DAWSON, C., WARING, R. H., RAMSDEN, D. B. & COLE, J. A. 2013. Analysis of volatile organic compounds of bacterial origin in chronic gastrointestinal diseases. Inflammatory bowel diseases, 19, 2069-2078.

WANDERS, R. J. & WATERHAM, H. R. 2006. Biochemistry of mammalian peroxisomes revisited. Annu. Rev. Biochem., 75, 295-332.

WANG, X., ZHANG, A., HAN, Y., WANG, P., SUN, H., SONG, G., DONG, T., YUAN, Y., YUAN, X. & ZHANG, M. 2012. Urine metabolomics analysis for biomarker discovery and detection of jaundice syndrome in patients with liver disease. Molecular & Cellular Proteomics, mcp. M111. 016006.

WATSON, D. G. 2012. Pharmaceutical Analysis: a Textbook for Pharmacy Students and Pharmaceutical Chemists, London, Elsevier Health Sciences.

WATSON, J. T. & SPARKMAN, O. D. 2007. Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation, John Wiley & Sons.

WESTERHUIS, J. A., HOEFSLOOT, H. C. J., SMIT, S., VIS, D. J., SMILDE, A. K., VAN VELZEN, E. J. J., VAN DUIJNHOVEN, J. P. M. & VAN DORSTEN, F. A. 2008. Assessment of PLSDA cross validation. Metabolomics, 4, 81-89.

WHEELOCK, A. M. & WHEELOCK, C. E. 2013. Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. Mol Biosyst, 9, 2589-96.

WILLIAMS, H. R., COX, I. J., WALKER, D. G., NORTH, B. V., PATEL, V. M., MARSHALL, S. E., JEWELL, D. P., GHOSH, S., THOMAS, H. J. & TEARE, J. P. 2009. Characterization of inflammatory bowel disease with urinary metabolic profiling. The American journal of gastroenterology, 104, 1435.

WILLIAMS, P. T. 1997. Relationship of distance run per week to coronary heart disease risk factors in 8283 male runners The National Runners' Health Study. Archives of Internal Medicine, 157, 191.

WILLIAMS, P. T. 2009. Lower prevalence of hypertension, hypercholesterolemia, and diabetes in marathoners. Medicine and science in sports and exercise, 41, 523.

Wishart, D.S., 2008. Applications of metabolomics in drug discovery and development. Drugs in R & D, 9(5), pp.307-322.

XU, G., HANSEN, J., ZHAO, X., CHEN, S., HOENE, M., WANG, X.-L., CLEMMESEN, J. O., SECHER, N. H., HÄRING, H. & PEDERSEN, B. 2016. Liver and muscle contribute differently to the plasma acylcarnitine pool during fasting and exercise in humans. The Journal of Clinical Endocrinology & Metabolism, 101, 5044-5052.

XU, J. & GORDON, J. I. 2003. Honor thy symbionts. Proceedings of the National Academy of Sciences, 100, 10452-10459.

Xu, Y.; Fang, X.; Casey, G.; Mills, G., Lysophospholipids activate ovarian and breast cancer cells. Biochemical Journal 1995, 309, (3), 933-940.

Xu, Y.; Xiao, Y.-j.; Baudhuin, L. M.; Schwartz, B. M., The role and clinical applications of bioactive lysolipids in ovarian cancer. Journal of the Society for Gynecologic Investigation 2001, 8, (1), 1-13.

YAMAMOTO, H., YAMAJI, H., ABE, Y., HARADA, K., WALUYO, D., FUKUSAKI, E., KONDO, A., OHNO, H. & FUKUDA, H. 2009. Dimensionality reduction for metabolome data using

PCA, PLS, OPLS, and RFDA with differential penalties to latent variables. Chemometrics and Intelligent Laboratory Systems, 98, 136-142.

Yan, Y.; Schoenwaelder, S.M.; Salem, H.H., Jackson, S.P. The Bioactive phospholipid, lysophosphatidylcholine, induces cellular effects via G-protein-dependent activation of adenylylcyclase. J. Biol. Chem. 1996, 271, 27090–27098.

Yang QJ, Zhao JR, Hao J, Li B, Huo Y, Han YL, Wan LL, Li J, Huang J, Lu J, Yang GJ, Guo C. Serum and urine metabolomics study reveals a distinct diagnostic model for cancer cachexia. J Cachexia Sarcopenia Muscle 2018;9(1):71-85.

YAU, Y., LEONG, R. W., ZENG, M. & WASINGER, V. C. 2013. Proteomics and metabolomics in inflammatory bowel disease. Journal of gastroenterology and hepatology, 28, 1076-1086.

Yip C, Dinkel C, Mahajan A, Siddique M, Cook GJR, Goh V. Imaging body composition in cancer patients: visceral obesity, sarcopenia and sarcopenic obesity may impact on clinical outcome. Insights imaging 2015; 6(4):489-497.

ZAJAC, A., POPRZECKI, S., MASZCZYK, A., CZUBA, M., MICHALCZYK, M. & ZYDEK, G. 2014. The effects of a ketogenic diet on exercise metabolism and physical performance in off-road cyclists. Nutrients, 6, 2493-2508.

ZHANG, A., SUN, H., WU, X. & WANG, X. 2012. Urine metabolomics. Clinica Chimica Acta, 414, 65-69.

ZHANG, J., LIGHT, A. R., HOPPEL, C. L., CAMPBELL, C., CHANDLER, C. J., BURNETT, D. J., SOUZA, E. C., CASAZZA, G. A., HUGHEN, R. W. & KEIM, N. L. 2017. Acylcarnitines as markers of exercise-associated fuel partitioning, xenometabolism, and potential signals to muscle afferent neurons. Experimental physiology, 102, 48-69.

ZHANG, R., WATSON, D. G., WANG, L., WESTROP, G. D., COOMBS, G. H. & ZHANG, T. 2014. Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based untargeted metabolite profiling of Leishmania parasites. J Chromatogr A, 1362, 168-179.

ZHANG, R., ZHANG, T., ALI, A. M., AL WASHIH, M., PICKARD, B. & WATSON, D. G. 2016. Metabolomic Profiling of Post-Mortem Brain Reveals Changes in Amino Acid and Glucose Metabolism in Mental Illness Compared with Controls. Comput Struct Biotechnol J, 14, 106-16.

ZHANG, Y., LIN, L., XU, Y., LIN, Y., JIN, Y. & ZHENG, C. 2013. 1H NMR-based spectroscopy detects metabolic alterations in serum of patients with early-stage ulcerative colitis. Biochemical and biophysical research communications, 433, 547-551.

Zhao, Z.; Xiao, Y.; Elson, P.; Tan, H.; Plummer, S.J.; Berk, M.; Aung, P.P.; Lavery, I.C.; Achkar, J.P.; Li, L.; Casey, G.; Xu. Y. Plasma lysophosphatidylcholine levels: Potential biomarkers for colorectal cancer. J. Clin. Oncol. 2007, 25, 2696–2701.

ZIELIŃSKI, J., KRASIŃSKA, B. & KUSY, K. 2013. Hypoxanthine as a predictor of performance in highly trained athletes. International journal of sports medicine, 34, 1079-1086.

ZIPPER, J. 1997. Proliferation of myocardial peroxisomes caused by several agents and conditions. Journal of molecular and cellular cardiology, 29, 149-161.

Zuijdgeest-van Leeuwen, Sonja D., J. Willem O. van den Berg, J. L. Darcos Wattimena, Ate van der Gaast, G. Roelof Swart, J. H. Paul Wilson, and Pieter C. Dagnelie. 2000. “Lipolysis and Lipid Oxidation in Weight-Losing Cancer Patients and Healthy Subjects.” Metabolism 49 (7): 931–36. doi:10.1053/meta.2000.6740.