The University of Nottingham

Mass Spectrometry for High-Throughput Metabolomics Analysis of Urine

Salah M.A. Abdelrazig, B. Pharm. (Hons.), M. Pharm., MSc.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy in Pharmacy

December 2015

Abstract

Direct electrospray ionisation-mass spectrometry (direct ESI-MS), by omitting the chromatographic step, has great potential for application as a high-throughput approach for untargeted urine metabolomics analysis compared to liquid chromatography-mass spectrometry (LC-MS). The rapid development and technical innovations revealed in the field of ambient ionisation MS such as nanoelectrospray ionisation (nanoESI) chip-based infusion and liquid extraction surface analysis mass spectrometry (LESA-MS) suggest that they might be suitable for high-throughput metabolomics analysis.

In this thesis, LC-MS and high-throughput direct ESI-MS methods using high resolution orbital trap mass spectrometer were developed and validated for untargeted metabolomics of human urine. Three different direct ESI-MS techniques were explored and compared with LC-MS: flow injection electrospray ionisation-MS (FIE-MS), chip-based infusion and LESA-MS of dried urine spots on a cell culture slide. A high-throughput sample preparation protocol was optimised using in-house artificial urine. Urine samples after consumption of green tea and healthy controls were used as a model to explore the performance and classification ability of the direct ESI-MS. High-throughput data pre-processing and multivariate analysis protocols were established for each method. The developed methods were finally applied for the analysis of clinical urine samples for biomarker discovery and to investigate the metabolic changes in osteoarthritis and malaria. Also, the methods were applied to study the effect of oligofructose diet on the gut microbial community of healthy subjects. The analytical performance of the methods for urine metabolomics was validated using quality control (QC) and principal component analysis (PCA) approaches. Rigorous validation including cross-validation, permutation test, prediction models and area under receiver operating characteristic (ROC) curve (AUC) was performed across the generated datasets using the developed methods.

Analysis of green tea urine samples generated 4128, 748, 1064 and 1035 ions from LC-MS, FIE-MS, chip-based infusion and LESA-MS analysis, respectively. A selected set of known green tea metabolites in urine were used to evaluate each method for detection sensitivity. 15 metabolites were found with LC-MS compared to 8, 5 and 6 with FIE-MS, chip-based infusion and LESA, respectively.

The developed methods successfully differentiated between the metabolic profiles of osteoarthritis active patients and healthy controls (Q^2 0.465 (LC-MS), 0.562 (FIE-MS), 0.472 (chip-based infusion) and 0.493 (LESA-MS)). The altered level of metabolites detected in osteoarthritis patients showed a perturbed activity in TCA cycle, pyruvate metabolism, β-oxidation pathway, amino acids and glycerophospholipids metabolism, which may provide evidence of mitochondrial dysfunction, inflammation, oxidative stress, collagen destruction and use of lipolysis as an alternative energy source in the cartilage cells of osteoarthritis patients. FIE-MS, chip-based infusion and LESA-MS increased the analysis throughput and yet they were able to provide 33%, 44% and 44%, respectively, of the LC-MS information, indicating their great potential for diagnostic application in osteoarthritis.

Malaria samples datasets generated 9,744 and 576 ions from LC-MS and FIE-MS, respectively. Supervised multivariate analysis using OPLS-DA showed clear separation and clustering of malaria patients from controls in both LC-MS and FIE-MS methods. Cross-validation R^2Y and Q^2 values obtained by FIE-MS were 0.810 and 0.538, respectively, which are comparable to the values of 0.993 and 0.583 achieved by LC-MS. The sensitivity and specificity were 80% and 77% for LC-MS and FIE-MS, respectively, indicating valid, reliable and comparable results of both methods. With regards to biomarker discovery, altered level of 30 and 17 metabolites were found by LC-MS and FIE-MS, respectively, in the urine of malaria patients compared to healthy controls. Among these metabolites, pipecolic acid, taurine, 1,3-diacetylpropane, N-acetylspermidine and N-acetylputrescine may have the potential of being used as biomarkers of malaria.

LC-MS and FIE-MS were able to separate urine samples of healthy subjects on oligofructose diet from controls (specificity/sensitivity 80%/88% (LC-MS) and 71%/64% (FIE-MS)). An altered level of short chain fatty acids (SCFAs), fatty acids and amino acids were observed in urine as a result of oligofructose intake, suggesting an increased population of the health-promoting *Bifidobacterium* and a decreased *Lactobacillus* and *Enterococcus* genera in the colon.

In conclusion, the developed direct ESI-MS methods demonstrated the ability to differentiate between inherent types of urine samples in disease and health state. Therefore they are recommended to be used as fast diagnostic tools for clinical urine samples. The developed LC-MS method is necessary when comprehensive biomarker screening is required.

Dedication

To

Kamal Eldin Eltayeb Ibrahim

The scholar, the friend, the father and the role model

Acknowledgements

I wish to convey special acknowledgments to many people for their help in this research. Firstly, I would like to record my gratitude to Prof David Barrett for his supervision, help, advice, guidance, patience and enthusiasm throughout the work. I would also like to express my deep and sincere gratitude to my second supervisor Dr Catherine Ortori for her advice, supervision, and invaluable practical assistance, in particular with the mass spectrometers, data analysis and sample preparation. I acknowledge Paul Cooling for his technical support, practical assistance and HPLC instruments training. My sincere gratitude to Prof Gail Davey, Brighton and Sussex Medical School, UK, for her help in the transport of malaria samples from Ethiopia to the UK. I would also like to thank Dr Erik Johansson, Umetrics AB, Sweden, for his kind help, knowledge and support in implementing the multivariate design and SUS plots in osteoarthritis study.

I wish to thank all members of the Centre for Analytical Biosciences for their help, support, knowledge and warm welcome. My friends who volunteered in the intervention study also require special acknowledgments; I'm really grateful for their kind support in a short notice. I would also like to thank Dr Toby Gould for his help, guidance and suggestions throughout the writing-up of this thesis. I convey my sincere acknowledgement to Dr Magdi Awadalla for his kind help and support with all administrative and financial issues at the University of Khartoum, Sudan. Also, I would like to express my gratitude to the IDB for funding and giving me the support to complete this thesis.

On a personal note, I would also like to thank Dr Muhammed Abdulrahman for his help and support during the last year of this research, I am grateful in every possible way, more than he knows. Collective and individual acknowledgments are also owed to all my friends in Nottingham for being of great support and making life in Nottingham somehow perpetually refreshing, easy and memorable.

Lastly, and definitely not least, I would like to thank my parents, siblings, Dr Mohammed Hassan and Dr Fatima Eltinay for their continuous encouragement and invaluable support through all the ups and downs of this work.

Contents

List of Abbreviations

ACTs	Artemisinin-based combination therapies
APCI	Atmospheric pressure chemical ionisation
API	Atmospheric pressure ionisation
APPI	Atmospheric pressure photoionisation
ATS	Amphetamine-type stimulants
AUC	area under receiver operating characteristic (ROC) curve
BEH	Ethylene Bridged Hybrid
BMI	Body mass index
BPC	Base peak chromatogram
C	Control urine sample
CAWG	Chemical analysis working group
CDF	Common data format
CE	Capillary electrophoresis
CE-MS	Capillary electrophoresis-mass spectrometry
COMP	Cartilage oligomeric protein
COX-2	Cyclo-oxygenase-2
CRM	Charged residue mechanism
CV	Cross-validation
DAPCI	Desorption atmospheric pressure chemical ionisation
DAPPI	Desorption atmospheric pressure photoionisation
DART	Direct analysis in real time
D-COMP	Deaminated cartilage oligomeric protein
DESI	Desorption electrospray ionisation
DIMS	Direct infusion mass spectrometry
Direct ESI-MS	Direct electrospray ionisation-mass spectrometry
DKK1	Dickkopf-related protein 1
DMOAD	Disease-modifying OA drugs
1D-NMR	One dimensional nuclear magnetic resonance
2D-NMR	Two dimensional nuclear magnetic resonance
DOPAC	Dihydroxyphenylacetic acid
EASI	Easy ambient sonic ionisation
EESI	Extractive electrospray ionisation
EMP	Embden–Meyerhof–Parnas pathway
ESI	Electrospray ionisation
ESI-MS	Electrospray ionisation-mass spectrometry
ESR	Erythrocyte sedimentation rate
FAB-MS	Fast atom bombardment-mass spectrometry
FDA	Food and Drug Administration
FDR	False discovery rate
Fib	Fibulin

FIE-MS	Flow injection electrospray ionisation-mass spectrometry
FIE-MS/MS	Flow injection electrospray ionisation-tandem mass spectrometry
FODMAP	Fermentable oligo-, di-, mono-saccharides and Polyols
FOG study	A trial to study the effect of low **F**ODMAP diet and dietary **o**ligofructose on **g**astrointestinal form, function and microbiota in healthy volunteers
FPR	False positive rate
FT-ICR-MS	Fourier transform ion cyclotron resonance mass spectrometer
FWER	Family wise error rate
FWHM	Full width at half maximum
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
GI	Gastrointestinal
Glc-Gal-PYD	Glucosyl–galactosyl–pyridinoline
^{1}H NMR	Proton nuclear magnetic resonance
HCA	Hierarchical cluster analysis
HILIC	Hydrophilic interaction liquid chromatography
HMDB	Human Metabolome Database
HPLC	High performance liquid chromatography
HPLC-MS	High performance liquid chromatography-mass spectrometry
HPLC-MS/MS	High performance liquid chromatography-tandem mass spectrometry
HRMS	High resolution mass spectrometer
HRP II	Histidine-rich protein II
hs-CRP	High sensitivity C reactive protein
IBS	Irritable bowel syndrome
ICTs	Immunochromatographic diagnostic tests
IEM	Ion evaporation mechanism
IL	Interleukin
IS	Internal standard
IT-ToF	Ion trap-time of flight mass spectrometer
LAESI	Laser ablation electrospray ionisation
LC	Liquid chromatography
LC/MS^{n}	Liquid chromatography-tandem mass spectrometry
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LC-NMR	Liquid chromatography-nuclear magnetic resonance
L-DOPA	L-3,4-Dihydroxyphenylalanine
LESA-MS	Liquid extraction surface analysis mass spectrometry
LIT	Linear ion trap mass spectrometer
LIT-FT-ICR-MS	Linear ion trap-fourier transform ion cyclotron resonance mass spectrometer

LLE	Liquid-liquid extraction
LTP	Low temperature plasma
m/z	Mass-to-charge ratio
mDa	Millidalton
MHPG sulphate	3-Methoxy-4-hydroxyphenylglycol sulphate
MIIS	Matrix-induced ion suppression
MMP	Matrix metalloproteinase
MRI	Magnetic resonance imaging
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS2	Tandem mass spectrometry
MSI	Metabolomics Standards Initiative
MSTUS	Mass spectrometry total useful signal
MVA	Multivariate analysis
MW	Molecular weight
NanoESI	Nanoelectrospray ionisation
NDDC	Nottingham Digestive Diseases Centre (NDDC), Nottingham, UK
NMR	Nuclear magnetic resonance
N-PCA	Nonlinear-principal component analysis
NSAIDs	Non-steroidal anti-inflammatory drugs
OA	Osteoarthritis
OPLS	Orthogonal-partial least square
OPLS-DA	Orthogonal-partial least square-discriminant analysis
PCA	Principal component analysis
PCR	Polymerase chain reaction
PF	*Plasmodium falciparum*
PH3	Primary hyperoxaluria type 3
PIIANP	N-propeptide IIA of type II collagen
PIIBNP	N-propeptide IIB of type II collagen
PIICP	C-propeptide of collagen type II
PIINP	N-propeptide II of type II collagen
pLDH	Plasmodial lactate dehydrogenase
PLS	Partial least square
PLS-DA	Partial least square-discriminant analysis
ppm	Part per million
Q^2	Predictive ability
QC	Quality control
QhQ	Quadrupole-hexapole-quadrupole mass spectrometer
QMC	Queen's Medical Centre, Nottingham, UK
qPCR	Quantitative real-time polymerase chain reaction
QqQ	Triple quadrupole mass spectrometer

Q-ToF	Quadrupole-time of flight mass spectrometer
Q-Trap	Quadrupole-ion trap mass spectrometer
R^2Y	Fitness of model
RBCs	Red blood cells
RDTs	Rapid diagnostic tests
ROC curve	Area under receiver operating characteristic (ROC) curve
RP	Reversed-phase
RSD	Relative standard deviation
RT	Retention time
SCFAs	Short chain fatty acids
SIM	Selected ion monitoring
sOB-Rb	Soluble leptin receptor
SOST	Sclerostin
SPE	Solid-phase extraction
SPMMRC	Sir Peter Mansfield Magnetic Resonance Centre, Nottingham, UK
SRM	Selected reaction monitoring
TCA	Tricarboxylic acid cycle
TIC	Total ion current
TIMP	Tissue inhibitor of matrix metalloproteinase
TLC	Thin layer chromatography
TNR	True negative rate
ToF	Time of flight mass spectrometer
TPR	True positive rate
UHPLC-MS	Ultra-high performance liquid chromatography-mass spectrometry
UHPLC-Q-ToF-MS	Ultra-high performance liquid chromatography-time of flight-mass spectrometry
UMS	Universal Metabolome-Standard
UV	Unit variance
UVA	Univariate analysis
VIP	Variable importance for the projection
WHO	World Health Organisation
WSR	Wide-scan range

CHAPTER ONE

General Introduction

1. General Introduction

1.1 Urine chemistry and composition

Human urine is a sterile, transparent, amber coloured fluid produced by the kidneys as a result of extraction of soluble waste, excess water and small molecular weight metabolic byproducts from bloodstream in a process called "urination". Urination is the primary route of elimination of water-soluble waste products from the human body, which generates 1,500-2,000 mL of urine/day. Urine normally consists of 95% water, urea, ammonia, inorganic salts, organic substances and water-soluble metabolic products of endogenous compounds, toxins and xenobiotics (Table 1-1). Urea, a waste product produced during proteins and amino acids breakdown in the liver, accounts for almost 50% of solutes in urine. Urobilin, a haemoglobin breakdown pigmented product, is also excreted in urine and accounts for the characteristic colour of urine. However, the concentrations of these substances may vary throughout the day under the influence of many factors such as physical activity, dietary intake, body metabolism and endocrine functions. Urine volume is determined by the amount of water excreted by the kidneys as a response to state of hydration of the body. The major factors that have a direct influence on the amount of urine excreted are fluid intake/fluid loss, amount of antidiuretic hormone secreted and the urge to eliminate increased amount of dissolved waste (Bouatra et al., 2013, Strasinger and Di-Lorenzo, 2008).

Table 1-1 Human urine main composition

Component	Process involved/comments
1. Water	Kidney excretion (95%)
2. Organic	
Urea	Amino acids breakdown
Creatinine	Derived from creatine, a nitrogen containing substance in muscles
Uric acid	Derived during metabolism of nucleic acid from food and cell lysis
Hippuric acid	Benzoic acid metabolite
3. Inorganic	
Sodium chloride	Principal salt (varies with intake)
Potassium	Occurs as salt of chloride, sulphate, and phosphate
Sulphate	Derived from amino acids metabolism
Phosphate	Occurs mainly as sodium salt (blood buffer)
Ammonium	Protein metabolism
Magnesium	Occurs as salt of chloride, sulphate, and phosphate
Calcium	Occurs as salt of chloride, sulphate, and phosphate
4. Other substances	Carbohydrates, pigments, fatty acids, mucin, enzymes, hormones may present in small amounts depending on diet and health

1.2 Metabolomics analysis and its application to urine

Urine generally viewed as a human waste but it is also considered as a valuable diagnostic biofluid. The analysis of urine (also known as urinalysis) is routinely performed to diagnose various types of diseases and disorders. Simple tests such as urine dipsticks are used to measure urinary glucose, bilirubin, urobilinogen, ketone bodies, nitrates, leukocyte esterase, specific gravity, haemoglobin and protein (Bouatra et al., 2013). While sophisticated analytical tests are used to determine a variety of urinary metabolites which help in the study of metabolic changes during normal or disease states, e.g. bladder (Issaq et al., 2008) and kidney diseases (Kim et al., 2009). Different tables of normal reference concentration ranges of 100s of urine solutes have been published and used for such purposes (Kim et al., 2008, Luo et al., 1997, Bollard et al., 2005, Bouatra et al., 2013). In addition to these detailed referential ranges, urine as readily available biofluid has been an attractive choice for more comprehensive investigations using metabolomics approaches. Urine is non-invasive, simple and is less likely to be volume-limited which provides enough aliquots for metabolite identification, quantification and subsequent data analysis. It can be easily sampled at different time intervals allowing temporal metabolic changes to be investigated. Also, urine being a waste product pool of the biological system and it is not under homeostatic regulation, can provide insights into metabolic dis-regulation associated with physiological changes during normal or disease processes (Chen et al., 2013a).

Metabolomics deals with quantitative and qualitative investigations of small molecules, known as metabolites within a biological system framework. Typically, metabolites are defined as a class of organic molecules of molecular mass less than 1.5 kDa (Lokhov and Archakov, 2009). Metabolomics mainly delivers its target goals by analysing metabolite levels in biological samples followed by extensive data analysis and interpretation. Although metabolomics is considered as a well-established science now, there is still no known standalone single-instrument platform capable of performing a complete analysis of a metabolome, an organism's complete set of metabolites. The metabolome represents a huge set of components that belong to different compound classes, such as lipids, nucleotides, amino acids ... etc. This diversity is accompanied not only by differences in their chemical and physical properties but also their presence in wide concentration ranges in the organism. Therefore, a successful metabolomics

study using experimental design is crucial, desirable and challenging (Andrade et al., 2010).

Metabolomics analyses are mainly carried out by two different approaches, targeted and untargeted (global) metabolomics. Targeted metabolomics is a directed approach that focuses on the quantitative analysis of a pre-defined set of metabolites with similar chemical properties from a class of molecules (e.g. carbohydrates, lipids or amino acids), members of a specific metabolic pathway (e.g. histidine metabolism) or belonging to a certain group of metabolites. The interest here is on a certain group of compounds in the metabolome and usually selective sample preparation steps are used to improve the data quality and optimize the method. This manipulation of sample preparation leads to the higher sensitivity and precision desired for the quantification of the specific set of metabolites (metabolite pool size). Analysing a specific set of metabolites with the ignorance of the rest in the metabolome, is clearly a targeted approach and cannot be considered a truly 'global' metabolomics approach. However, establishing a suit of assembled quantitative measurements targeting key metabolites can give rise to the concept of metabolomics. For instance, it promises to provide an effective investigative quantitative tool for studying microbial metabolism, in which the good knowledge in the intracellular metabolite concentrations changes provides a direct way to study kinetics of enzymes of the underlying metabolic pathway (Buchholz et al., 2002).

Untargeted metabolomics, also known as global metabolomics, metabolic profiling or metabolic fingerprinting, ideally provides an unbiased overview of the whole metabolic activity in a biological system by detecting all the metabolites from a representative set of samples (Figure 1-1). It involves the analysis of treated (altered) and control sets of samples and then by comparing these sets using multivariate analysis (MVA), the tentative identity of metabolites of interest can be extracted. The identity of those metabolites could be further confirmed using a targeted method (Dettmer et al., 2007). In 2003, another variant of fingerprinting called "metabolic footprinting" was introduced. It differs in analysing extracellular metabolites instead of intracellular ones (Allen et al., 2003). A few years later, newer terms, endo- and exo-metabolome, were coined by Nielsen and Oliver to distinguish between intracellular and extracellular metabolites, respectively (Nielsen and Oliver, 2005).

Many studies have applied metabolomics approaches to urine, mainly in the area of nutrition (Khymenets et al., 2015), metabolic disorders (Luan et al., 2015) and disease biomarker discovery (Peng et al., 2014). Recently, an extensive effort has been made to gather most of the metabolites found in urine using metabolomics approaches; as a result a comprehensive list of 449 metabolites (identified) and 378 metabolites (quantified) was compiled (Bouatra et al., 2013). In addition, 3100 urine metabolites were annotated along with 3900 concentration values in the Human Urine Metabolome Database, http://www.urinemetabolome.ca/, as part of the Human Metabolome Database (HMDB) project (Wishart et al., 2013).

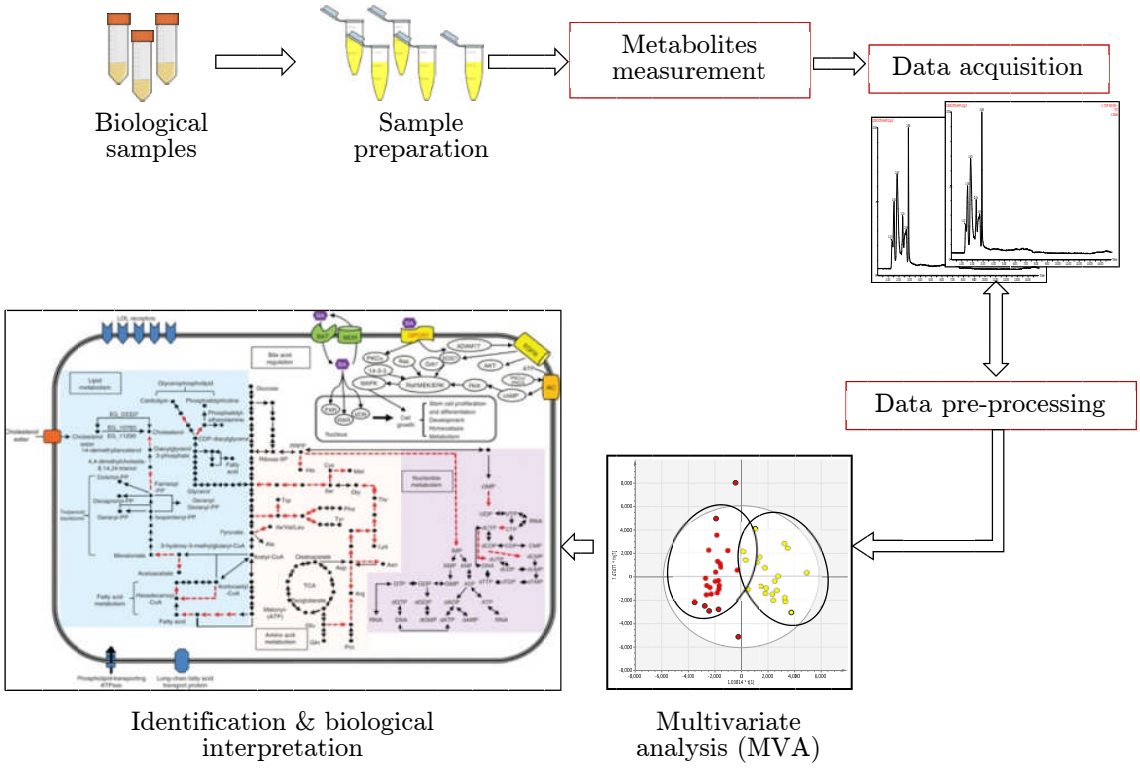


Figure 1-1 Typical untargeted metabolomics workflow

1.3 Analytical techniques used in metabolomics

Different traditional and modern analytical techniques have been used for metabolite analysis, such as thin layer chromatography (TLC) (Tweeddale et al., 1998), enzymatic assays (Bergmeyer, 1984), nuclear magnetic resonance (NMR) (De Graaf et al., 1999), gas chromatography (GC) (Womersley, 1981), liquid chromatography (LC) including high performance liquid chromatography (HPLC) (Bhattacharya et al., 1995) and capillary electrophoresis (CE) (Edwards

et al., 2006) and mass spectrometry (MS) (Castrillo et al., 2003, Fuzfai et al., 2004, Dubbelboer et al., 2012). Metabolomics analysis requires highly sensitive and selective methods, capable of handling small sample volumes and yet providing reliable measurement of a large spectrum of compounds in the metabolome of interest. The metabolome constitutes a very small percentage of the total biomass of a cell. Hence, the volume of the metabolome constituents in a biological sample used for the analysis is limited. The volume of the biological sample can be increased to overcome this problem but the routine way of sample dilution decreases metabolite concentrations for the analysis. Also, the detector used in the measurement should measure the compound of interest and reduce the sample matrix background signals to the minimum (Oldiges et al., 2007). Therefore, in metabolomics, mass spectrometry (MS) and nuclear magnetic resonance (NMR) are the most commonly used techniques; especially when coupled with chromatography and capillary electrophoresis as pre-detection separation techniques, e.g. GC-MS, LC-MS, CE-MS and LC-NMR.

Different NMR and MS techniques have been applied for urine metabolomics such as high resolution NMR (Yang et al., 2008), HPLC-MS (Wilson, 2011), high performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS) (Navarro-Munoz et al., 2012), ultra-high performance liquid chromatography-mass spectrometry (UHPLC-MS) (Jimenez Giron et al., 2012), fast atom bombardment-mass spectrometry (FAB-MS) (Luo et al., 1997), two-dimensional-GC-MS (Zhang et al., 2011) and nano-flow LC-MS/MS (Kim et al., 2008). However, the selection of an analytical technique is often based on the speed of the runs, sensitivity and selectivity. As it is a difficult task to combine all of them in a method and the selection of the most suitable technique is generally a compromise between them. The acquired data from MS and NMR is usually complementary and therefore, ideally it is preferable to use both techniques on a sample set, but in practice this is rarely achieved. For example, Law et al., applied GC-MS, LC-MS and ^{1}H NMR for metabolomics investigation of human urine after consumption of green tea to help improve metabolite coverage (Law et al., 2008).

NMR is a non-destructive technique that allows the same samples to be further analysed by other methods which gives the technique a unique advantage where the sample volume is a crucial issue. However, the low sensitivity of ^{1}H NMR compared to MS (which is several orders of magnitude more sensitive), limits its use in the field. In addition, difficulty associated with the identification of large

groups of metabolites due to overlapping peaks in NMR favours other techniques such as MS. The use of 2D-NMR methods can overcome identification problems to some extent but it is usually restricted to the characterization of unidentified analytes from the 1D-NMR spectrum (Ward et al., 2007). On the other hand, there are many applications in the literature describing the use of MS routinely coupled to GC and LC for metabolomics analysis to increase sensitivity, selectivity and to provide metabolite identification by measurement of mass. (Bouatra et al., 2013). MS analysis reduces the background signals and peaks overlapping due to the matrix interferences which are usually associated with biological samples. In comparison with NMR, MS is less robust, lower in reproducibility and is limited to ionisable metabolites. However, MS has the ability to detect the "invisible NMR" moieties, e.g. sulphates, and it provides higher sensitivity and therefore, it is now established as the cornerstone technique in metabolomics (Xiayan and Legido-Quigley, 2008).

Hyphenated MS methods such as LC-MS and GC-MS have been used for clinical analysis of both polar and non-polar metabolites. GC-MS enables both quantification and identification of wide ranges of metabolites, but its use is limited to thermally stable volatile compounds and derivatisation step is required to include non-volatile ones (Rauha et al., 2001). LC-MS is now considered as the primary method in MS-based metabolomics as it delivers adequate selectivity and sensitivity of a wider range of metabolites in the sample including thermally labile and non-volatile metabolites compared to GC-MS without the need for chemical derivatisation, the essential pre-analysis step for non-volatile metabolites in GC-MS analysis (Andrade et al., 2010).

1.4 LC-MS based urine metabolomics

The rapid development and inventions in the field of mass spectrometry have played an important role in the adoption of LC-MS as one of the most useful technique for metabolomics. The introduction of atmospheric pressure ionisation (API) interfaces overcame most of the problems associated with coupling of LC with MS such as relatively high flow rate of effluent from the LC column as well as coupling of high pressure technique into a vacuum working environment of MS. Many interfaces or soft ionisation methods usually provide molecular ions with fewer fragmentations and are now used routinely such as electrospray ionisation (ESI), atmospheric pressure chemical ionisation (APCI) and atmospheric pressure photoionisation (APPI) (Gross, 2011). A literature review

of most recent LC-MS urine metabolomics publications (Table 1-2) shows that most studies used ESI as the ionisation technique with either HPLC-MS or UHPLC-MS. However, the combination of different ionisation interfaces may increase the detectable range of masses in metabolic profiling and footprinting. For instance, ESI and APCI were effectively combined for untargeted metabolomics study in which more metabolites have been detected (Nordstrom et al., 2007). In addition, the use of nanoESI with nanoUHPLC has been reported to improve MS sensitivity of low abundance compounds in urine metabolomics (Chetwynd et al., 2014, Chetwynd et al., 2015).

The introduction of UHPLC-MS with its improved chromatographic resolution, efficiency, shorter analysis time and lower sample volume, has resulted in the detection of more metabolites than conventional HPLC-MS (Han et al., 2009). Many urine metabolomics studies have used UHPLC-MS for untargeted approach with either reversed-phase (RP) (Llorach et al., 2014, Jin et al., 2014, Peng et al., 2014) or hydrophilic interaction liquid chromatography (HILIC) columns (Peretz et al., 2012). The choice of the column for LC-MS analysis mainly depends on the matrix and the analytes of interest. In urine, the metabolites of interest are fairly polar or semi-polar (i.e. hydrophilic) in nature. Hence, the use of RP columns with good retentive power is an adequate standard choice for urine metabolomics. However, highly polar metabolites will not be retained on a RP column, thus limiting the method detection coverage especially for untargeted metabolomics. For these applications, the use of a HILIC column, which is more suited for the analysis of polar compounds, provides complementary information to those obtained with RP columns. Combined LC-MS analysis with HILIC and RP columns has already been applied for urine analysis to enhance metabolite coverage (Huang et al., 2013, Zhang et al., 2013c).

In targeted metabolomics, tandem LC-MS (LC/MS^n) is a widely used platform for urine analysis such as tandem triple quadrupole MS (Struck et al., 2013) and tandem Q-Trap (Blydt-Hansen et al., 2014), as it provides good sensitivity and accuracy in quantification of metabolites of interest (Sawada et al., 2009). The application of internal or external calibration is essential when accurate measurement is required, e.g. screening of biomarkers in clinical metabolomics (Ceglarek et al., 2009). Several ion modes of MS and tandem MS can be used such as selected ion monitoring (SIM), selected reaction monitoring (SRM) and multiple reaction monitoring (MRM). However, in untargeted metabolomics the use of low resolution mass analysers such as quadrupole and ion trap has

significant disadvantages. Such features in a mass spectrometer cause severe overlapping of metabolite masses which makes their identification a difficult task to attain (Han et al., 2009). The use of high resolution mass analysers such as Q-ToF and orbital trap MS for metabolomics in general and for urine analysis is desirable. Therefore, most of untargeted urine analysis (Table 1-2) has been carried out with high resolution mass spectrometers (HRMS) such as hybrid Quadrupole-ToF (Q-ToF) (Llorach et al., 2014, Mattarucchi et al., 2012) (Jin et al., 2014), orbital trap MS (May et al., 2013, Ridder et al., 2014) (Peretz et al., 2012) and hybrid ion trap ToF (IT-ToF) (van Wietmarschen et al., 2012).

For a successful application of LC-MS method for urine analysis, careful control of chromatography is essential to ensure stability and reproducibility of the analysis. In addition, the chromatographic separation step adds some degree of complexity to the metabolome datasets, therefore, significant time, efforts and expertise are required for data processing in order to extract, align, deconvolute and annotate peaks correctly. LC runs take a relatively long time for each sample, thus becoming a bottle-neck in the analysis and processing of large batches of samples for high-throughput metabolomics studies. During this long analysis time or upon repetitive use, chromatographic columns may gradually deteriorate, resulting in the need to address problems with a significant drop in the quality of the acquired datasets across the analytical run. Moreover, LC-MS analytical run time is limited to the stability of urine sample during the analysis. Gika et al., demonstrated that urine sample can be stable for a maximum of 48 hours at 4 °C for LC-MS analysis and beyond that significant changes in urine composition may occur (Gika et al., 2008c). Thus, for urine metabolomics studies involving analysis of large batches of samples, an alternative high-throughput approach is essentially needed. The development of MS profiling methods such as direct ESI-MS without recourse to any chromatographic separation might be a suitable approach for high-throughput urine metabolomics. The omission of the separation step increases sample throughput. Also, the absence of the need to undertake chromatographic peak alignment and reduce dataset dimensionality significantly simplifies data processing (Fuhrer et al., 2011, Southam et al., 2007). The direct ESI-MS method meeds to provide the necessary information related to the total pool of metabolites in urine samples for metabolomics analysis.

Table 1-2 A selected list of recent LC-MS urinary metabolomics publications

Application	MS method	Data processing software	Normalisation method	Statistical approach	Reference
Aronia-citrus juice	HPLC-Q-ToF (RP)	MarkerView (AB Sciex)	Not applied	UVA and OPLS-DA	(Llorach et al., 2014)
Asthma	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	MSTUS	OPLS-DA	(Mattarucchi et al., 2012)
Bladder cancer	HPLC-Q-ToF (RP)	MZMine 2	MSTUS	UVA, OPLS-DA,and ROC	(Jin et al., 2014)
Bladder cancer	HPLC-Q-ToF (RP)	XCMS	^{13}C-labeled UMS	PCA, OPLS-DA,and ROC	(Peng et al., 2014)
Bladder cancer	HPLC-Q-ToF (RP) and CE-Q-ToF	MassHunter (Agilent)	Creatinine	UVA, PCA and OPLS-DA	(Alberice et al., 2013)
Bladder, kidney and prostate cancer	HPLC-triple quardrupole MS (RP)	Targeted method	Creatinine	UVA, PCA and PLS-DA	(Struck et al., 2013)
Cervical cancer	HPLC-Q-ToF (RP)	MarkerVeiw (AB Sciex), XCMS and MZMine 2	Creatinine	UVA, PCA, OPLS-DA and ROC	(Chen et al., 2013b)
Cervical cancer	HPLC-Q-ToF (RP)	MassHunter (Agilent)	TIC	UVA and PLS	(Liang et al., 2014)
Chinese herbal medicine	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	MSTUS	UVA, PCA and OPLS-DA	(Liu et al., 2013)
Chinese herbal medicine	HPLC-Orbital trap MS (RP)	SIEVE (Thermo)	Not applied	UVA and OPLS-DA	(Lu et al., 2014)
Chinese herbal medicine	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	TIC	UVA, PCA and OPLS-DA	(Su et al., 2013)
Citrus fruit intervention	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	Not applied	UVA, PCA, PLS-DA and HCA	(Pujos-Guillot et al., 2013)
Citrus intervention	HPLC-Q-ToF (RP)	Profile Analysis (Bruker)	Highest intensity	PCA	(Medina et al., 2013)
Cardiovascular disease	HPLC-Q-ToF (RP)	MarkerView (AB Sciex)	Not applied	UVA and OPLS-DA	(Llorach et al., 2013)
Coffee metabolites	UHPLC-Q-ToF (RP)	XCMS	Not applied	UVA, PLS-DA and ROC	(Rothwell et al., 2014)
Dietary pattern	UHPLC-Q-ToF (RP)	MZMine 2	MSTUS	UVA, PCA and PLS-DA	(Andersen et al., 2014)
Dioxin exposure	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	MSTUS	UVA and OPLS-DA	(Jeanneret et al., 2014)
Doping control	UHPLC-Q-ToF (RP)	Profile Analysis (Bruker)	Not applied	UVA, PCA and OPLS-DA	(Kiss et al., 2013)
Dietary intervention	HPLC-Orbital trap MS (RP)	MSInspect	Not applied	UVA and PCA	(May et al., 2013)
Goji tea intervention	HPLC-Q-ToF (RP)	Profile Analysis (Bruker)	Not applied	UVA, PCA and PLS-DA	(Tseng and Li, 2014)
Green tea intervention	HPLC-Orbital trap MS (RP)	MAGMa	Not applied	Not applied	(Ridder et al., 2014)
Gastric Cancer	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	Not applied	UVA, PCA, OPLS-DA, HCA and ROC	(Fan et al., 2012)
Idiopathic nephrotic syndrome	UHPLC-Q-ToF (RP)	Progenesis Comet (Nonliner Dynamics)	TIC	OPLS-DA	(Sedic et al., 2014)
Jaundice syndrome	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	TIC	OPLS-DA and HCA	(Wang et al., 2012a)
Kidney and Bladder cancer	HPLC-Q-ToF (RP and HILIC)	Profile analysis (Bruker)	MSTUS	UVA, PCA, OPLS-DA,and ROC	(Huang et al., 2013)
Kidney cancer	UHPLC-IT (RP)	Not mentioned	MSTUS	UVA	(Ganti et al., 2012)

Application	MS method	Data processing software	Normalisation method	Statistical approach	Reference
Liver cancer	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	Not applied	OPLS-DA	(Zhang et al., 2013b)
Liver cancer	UHPLC-Q-ToF (RP)	XCMS	Not applied	UVA and PCA	(Zhang et al., 2013a)
Liver cirrhosis	UHPLC-Orbital trap MS (RP)	MS finder	MSTUS	UVA (Non parametric *t*-test) and PCA	(Dai et al., 2014)
Lung carcinoma	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	MSTUS and creatinine	UVA, OPLS-DA,and ROC	(Wu et al., 2014)
Male infertility	HPLC-Q-ToF (RP)	Profile analysis (Bruker)	MSTUS	UVA, OPLS-DA,and ROC	(Zhang et al., 2014a)
Metabolic syndrome	UHPLC-Q-ToF (RP)	Profile analysis (Bruker)	IS and creatine	UVA, PCA and OPLS-DA	(Yu et al., 2014)
Neuroleptic-Naïve Schizophrenia	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	TIC	UVA and OPLS-DA	(Cai et al., 2012)
Method protocol	UHPLC-Q-ToF (RP and HILIC)	MarkerLynx (Waters)	Not applicable	UVA and OPLS-DA	(Want et al., 2010)
Method development	NanoUHPLC-nanoESI-Q-TOF (RP)	MarkerLynx (Waters)	Total spectral area	UVA and OPLS-DA	(Chetwynd et al., 2015)
Ovarian cancer	UHPLC-Q-ToF (RP)	XCMS	MSTUS	UNA, PCA and PLS-DA	(Chen et al., 2012)
ThioTEPA metabolism	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	Not applied	UVA and OPLS-DA	(Li et al., 2011)
Prostate cancer	HPLC-Orbital trap MS (RP and HILIC)	MZMine 2	Creatinine, osmolality and MSTUS	UVA, PCA, OPLS-DA,and ROC	(Zhang et al., 2013c)
Radiation effect	UHPLC-ToF (RP)	MarkerLynx (Waters)	Creatinine	UVA and PCA	(Laiakis et al., 2014)
Renal clearance rate	UHPLC-Orbital trap MS (RP)	MZMine 2	Not applied	UVA	(Sirich et al., 2013)
Rheumatoid arthritis subtypes	UHPLC-IT-ToF (RP)	Profiling Solution (Shimadzu) CATPCA (SPSS)	Not applied	NPCA and PLS-DA	(van Wietmarschen et al., 2012)
Rejection in kidney transplantation	HPLC-Q-Trap (RP)	Targeted analysis	Creatinine	UVA (Non parametric *t*-test), PLS-DA and ROC	(Blydt-Hansen et al., 2014)
Stomach cancer	HPLC-Q-ToF (RP)	MassHunter (Agilent)	Not applied	PCA	(Zhang et al., 2013d)
Tuberculosis treatment signature	HPLC-Q-ToF (RP)	MassHunter (Agilent)	Creatinine	UVA and PCA	(Mahapatra et al., 2014)
Tea polyphenols profiling	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	Not applied	UVA, PCA and OPLS-DA	(Xie et al., 2012)
α-Tocopherol metabolism	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	IS	UVA, PCA and OPLS-DA	(Johnson et al., 2012)
Urinary signature of tetrahydrocannabinol	UHPLC-Q-ToF (RP)	MarkerLynx (Waters)	Not applied	UVA (ANOVA), PLS-DA and OPLS	(Kiss et al., 2014)
Xanthinuria biomarkers	HPLC-Orbital trap MS (ZIC-HILIC)	SIEVE (Thermo)	Not applied	UVA	(Peretz et al., 2012)

RP: reverse phase, IS: internal standard, HILIC: hydrophilic interaction liquid chromatography, MSTUS: MS total useful signal, TIC: total ion current, UMS: Universal Metabolome-Standard, UVA: univariate analysis, PCA: principal component analysis, N-PCA: nonlinear-principal component analysis, PLS: partial least square, PLS-DA: partial least square-discriminant analysis, OPLS: orthogonal-partial least square, OPLS-DA: orthogonal-partial least square-discriminant analysis, ROC: receiver operator characteristic and HCA: hierarchical cluster analysis.

1.5 Direct ESI-MS: a potential high-throughput approach for urine metabolomics

Direct electrospray ionisation-mass spectrometry (direct ESI-MS) is an introduction technique of samples into MS source omitting the chromatographic separation step for both targeted and untargeted metabolomics. It is considered as a high-throughput approach as it provides short time sample analysis (Kayser and Warzecha, 2012). Among different types of ionisation sources in mass spectrometry, electrospray ionisation (ESI) can be considered as standard, and is probably the most commonly used ion source in MS.

1.5.1 Electrospray ionisation: the standard ion source in MS

Electrospray ionisation (ESI) is a soft ionisation technique, it provides minimal fragmentations of molecular ions and less complex mass spectra could be obtained (Kayser and Warzecha, 2012). It was first described in the late 60s after Malcolm Dole and his colleagues had performed a provocative ingenious series of experiments in which they attempted to form in vacuo beams of macro ions in a gas-phase from a dilute polymer (macromolecules) solution (Dole et al., 1968). In mid 1980s, a breakthrough in the field was reported by two different research groups as they were able to couple ESI to a mass spectrometer and generating intact ions from large and complex chemical species in solution (Yamashita and Fenn, 1984). Few years later, Fenn group was able to obtain MS spectra of biopolymers including oligonucleotides and proteins such as insulin using ESI coupled to Fourier transform MS (Fenn et al., 1989). Twenty years later, John B. Fenn and Koichi Tanaka were shared half of the 2002 Nobel Prize in Chemistry "for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules" (Nobel Media AB, 2014). The advent of coupling ESI to MS has had great impact on reshaping the future use of mass spectrometry and makes it one of the most important techniques for biomolecule analysis.

The principle of ion generation in ESI involves the use of electrical energy to assist the transfer of ionic species from solution into the gas phase. Its mechanism involves formation of charged droplets mist followed by evaporation of solvent and subsequent ions formation from the droplets. In ESI source, a sample solution is pumped through a capillary tube maintained at high voltage, e.g. 1.5 – 6.0 kV. This leads to formation of a fine spray cone of charged droplets at the

tip end of the electrospray capillary known as 'Taylor cone'. The use of a nebulising gas, e.g. nitrogen, improves handling higher sample flow rates. These droplets pass down towards the mass spectrometer region by pressure and potential gradients. With the aid of a high ESI source temperature and the use of a drying gas, the solvent starts to evaporate from the droplets and charge density on the surface increases. At this point, there are two different competing mechanisms and controversial scenarios of how ions in the charged droplets transferred into the gaseous phase, namely, single ion or charged residue mechanism (CRM) and ion evaporation mechanism (IEM).

In the CRM, when Rayleigh limit is reached the droplets begin to subdivide into smaller ones as the coulomb repulsion forces start to overcome the surface tension holding the droplet together. The fission process continues until each droplet contains only one ion at the end. The evaporation of the remaining solvent leads to formation of single ions in the gas phase (Yamashita and Fenn, 1984). In the IEM, when the electric field strength limit is reached within the charged droplets, it makes the ejection of the ions at the surface more energetically and kinetically favourable (Thomson and Iribarne, 1979); shown below is a combination model of CRM and IEM (Figure 1-2). At the end, the generated ions are sampled by a skimmer cone and then accelerated into the mass analyser for subsequent measurement of molecular mass and intensities (Ho et al., 2003).

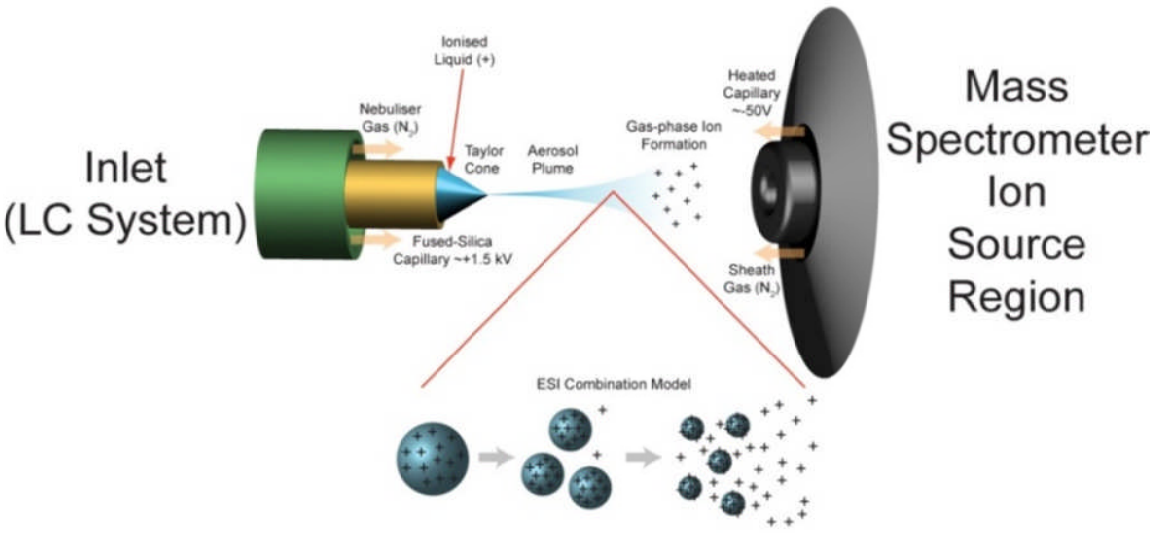


Figure 1-2 A schematic of ions formation process in electrospray ionisation (Lamond-Lab, 2014).

1.5.2 Direct ESI-MS techniques

A literature review of publications reporting the use of direct ESI-MS analysis (Table 1-3) highlights that the method has been applied with different types of introduction techniques, ranging from classical approaches using syringe pumps to advanced and newly emerging robotic techniques.

1.5.2.1 Direct infusion mass spectrometry (DIMS)

Direct ESI-MS is performed, classically, by using a syringe pump or an infusion device to deliver samples directly into MS which is known as direct infusion mass spectrometry (DIMS). The use of DIMS has found many applications in metabolomics. Lin et al., proposed a DIMS method for the fast diagnosis of kidney cancer. They reported a high-throughput method capable of detecting 23 potential biomarkers in a 0.5 min sample run compared to 30 min LC-MS method (Lin et al., 2010). Zhou et al. proposed simple and rapid DIMS method for the identification of Vinca alkaloids from a crude extract of *Catharanthus roseus*. The method not only revealed the molecular weight information of 5 major Vinca alkaloids but also provides useful structural information for the identification. In comparison, with an HPLC-MS method, the proposed method reduces the run time from 1 h to a few min. They described the technique as sensitive, rapid and convenient for targeted metabolomics in complex and mixed plant extracts. Moreover, they suggested the method as a valuable addition for metabolic profiling and quantitative analyses that may be necessary in quality control (Zhou et al., 2005). Despite that DIMS provides a rapid MS analysis, but the use of the syringe pumps limits the throughput in which each sample must be infused manually. Therefore, some direct ESI-MS approaches use an LC as a flow injection system to improve the throughput by automating sample injection and it is known as flow injection ESI–MS (Beckmann et al., 2008).

1.5.2.2 Flow injection ESI mass spectrometry (FIE-MS)

In FIE-MS, the sample is injected using the LC autosampler into a stream of solution entering the ESI source at steady rate. A typical FIE-MS profile takes about 1 min, in which a sharp increase in the signal to the peak's apex is followed by a gradual tailing-off as sample slightly delayed due to week diffusion and interaction processes within the solvent lines enter the ESI source (Figure 1-3). An injection of a blank or a few minute delays before the next sample injection

can minimise any carry-over effect between consecutive runs and also provides a signal noise region for background subtraction (Beckmann et al., 2008). FIE-MS has been used in different applications in metabolomics focusing to date mainly in the analysis of microbes (Higgs et al., 2001, Parker et al., 2008, Kaderbhai et al., 2003, Smedsgaard and Frisvad, 1996) and plants extracts (Chen et al., 2010b, Enot et al., 2006, Mas et al., 2007, Beckmann et al., 2007). FIE-MS also has been reported for targeted and untargeted analysis of body fluids such as blood and urine in animal (Beckmann et al., 2010, Yang et al., 2009) and nutritional (Fave et al., 2011, Lloyd et al., 2011a, Lloyd et al., 2011b) studies.

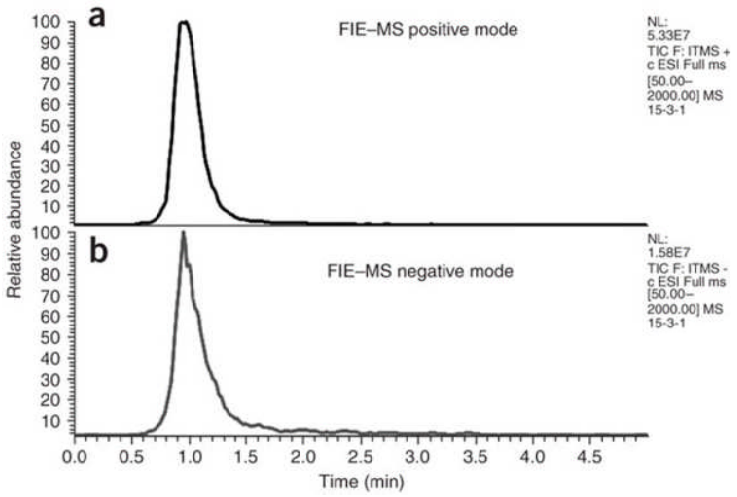


Figure 1-3 Typical FIE-MS chromatogram of a plant extract acquired using a linear ion trap mass spectrometer in (a) ESI+ and (b) ESI- modes (Beckmann et al., 2008).

1.5.2.3 Chip-based infusion and liquid extraction surface analysis-mass spectrometry (LESA-MS)

Ambient ionisation MS implies the use of an ionisation source that operates at room temperature and atmospheric pressure. The technique promises high specificity in studying and investigating in situ biochemical and cellular changes in organisms with minimal or no sample pre-treatment. In 2004, desorption electrospray ionisation (DESI) was introduced by Cooks group and represented a breakthrough in direct high-throughput MS analysis (Takats et al., 2004). Following the invention of DESI, more than 20 Ambient ionisation MS techniques were introduced, including extractive electrospray ionisation (EESI), laser ablation electrospray ionisation (LAESI), desorption atmospheric pressure photoionisation (DAPPI), direct analysis in real time (DART) (Han et al., 2009), nanoESI chip-based infusion and liquid extraction surface analysis-mass spectrometry (LESA-MS) (Li et al., 2012, Le Gac and van den Berg, 2009). Some

least popular ambient ionisation MS methods also have been reported in the literature, including surface desorption atmospheric pressure chemical ionisation (DAPCI) (Pi et al., 2011), low temperature plasma (LTP) (Harper et al., 2008), a plasma-based active capillary ionisation (Nudnova et al., 2012) and easy ambient sonic ionisation (EASI) (Amaral et al., 2011).

Ambient ionisation MS uses ESI such as nanoESI chip-based infusion and liquid extraction surface analysis-mass spectrometry (LESA-MS), can be considered a form of an advanced DIMS, suggesting that it might be suitable for high-throughput metabolomics. The minimal sample pre-treatment and fast analysis times offering potential advantages (Li et al., 2012, Le Gac and van den Berg, 2009). NanoESI chip-based devices such as NanoMate (Advion Bioscience, USA), allow direct infusion, fraction collection and LESA-MS. The system is used in conjunction with a nanoESI chip when interfaced with mass a spectrometer. The technique provides more informative data from complex samples much better in comparison with those obtained from standalone systems with ordinary interfaces (Li et al., 2012). A silicon-based integrated nanoESI microchip device is used which contains 400 nanoESI nozzles with different internal diameters that capable of delivering flow rate down to 30-50 nL/min (Almeida et al., 2008). A designed robotic arm inside the device picks up a pipette tip from a rack and aspirates the sample from 96/384 multi-well samples plate to be engaged into the back of the ESI chip. Applying high voltages through the tip while spraying produce ions that can be introduced directly into the mass spectrometer (Figure 1-4). This method aids intact protein analysis, non-covalent interactions and metabolite/lipid analysis (Le Gac and van den Berg, 2009, Southam et al., 2007).

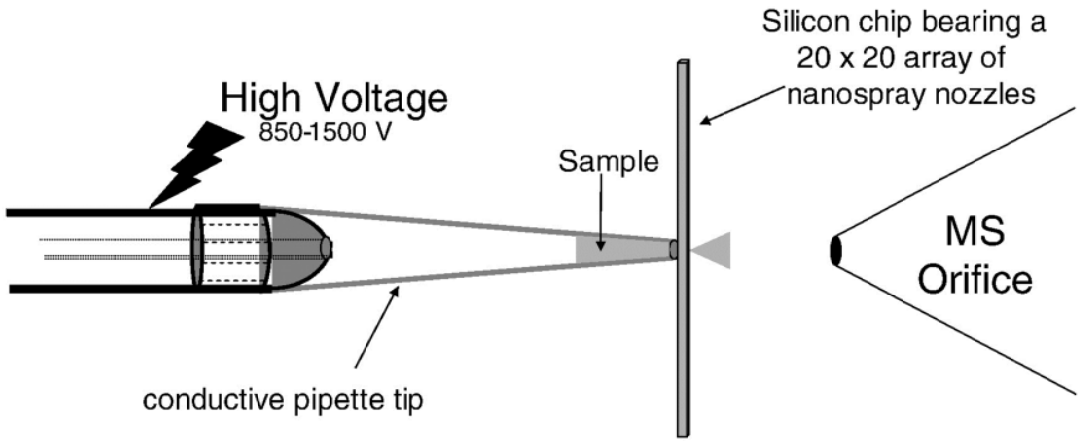


Figure 1-4 Sample infusion into MS by nanoESI using the silicon chip technology integrated in the NanoMate robot (Flangea et al., 2011).

Different applications of chip-based infusion were reported in the literature for DIMS analysis (Table 1-3) and they were mainly used for targeted analysis of metabolites in urine (Yang et al., 2004) or plasma (Han et al., 2008a). While LESA-MS has been extensively used for quantification of metabolites from different types of clinical samples, e.g. dried blood spots (Kertesz and Van Berkel, 2010) and tissues (Tomlinson et al., 2014). Flangea et al. reviewed the use of NanoMate for DIMS analysis focusing on strategies for brain gangliosides mapping and sequencing. In comparison with classical capillary based infusion ESI-MS, the review concluded that chip-based ESI-MS provides efficient ionisation, sensitive (detection limits up to fmol range) and high-throughput analyses. Furthermore, the use of the technique in conjunction with ultra-high resolution mass spectrometers enables multistage precursor ions fragmentations and subsequently, provides structural information of minor gangliosides which are often represented as invaluable biomarkers (Figure 1-5) (Flangea et al., 2011).

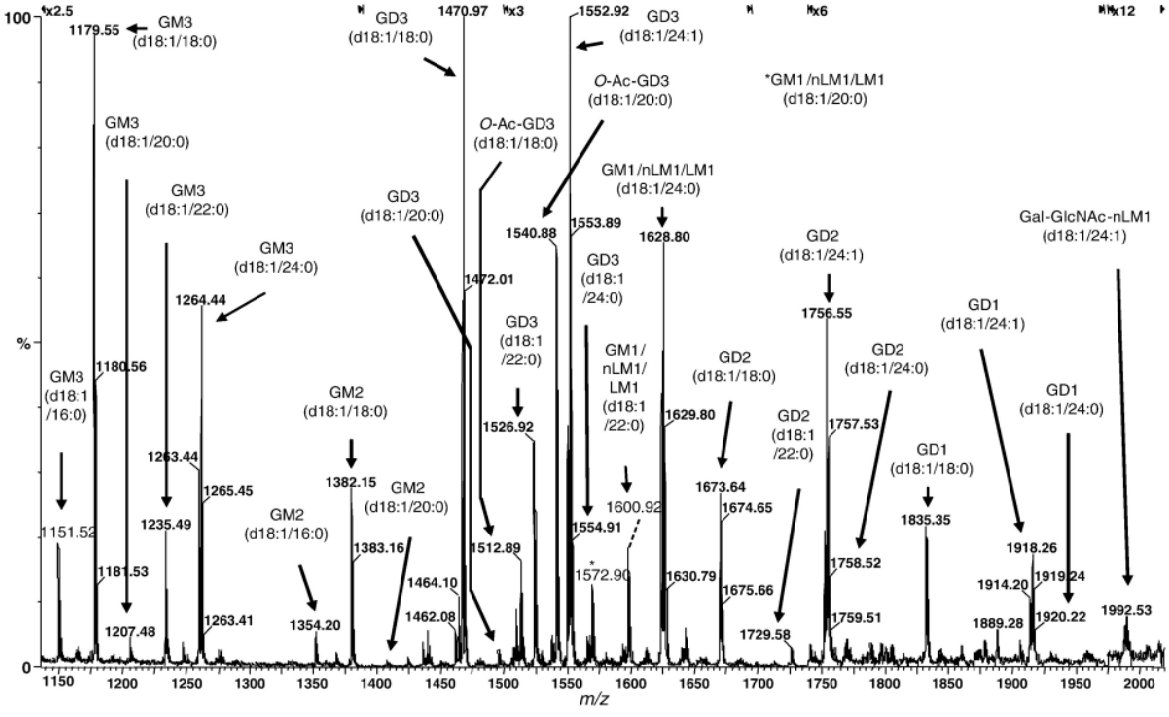


Figure 1-5 A chip based nanoESI-Q-ToF mass spectrum of the native gliosarcoma gangliosides mixture showing different types of metabolites detected in the negative mode (Flangea et al., 2011).

Table 1-3 A selected list of direct ESI-MS based publications for metabolomics analysis

Method	Sample/Application	Type/purpose of the study	Mass spectrometer	Reference
1. Low resolution MS/nominal mass direct ESI-MS				
DIMS	Olive oil, adulteration	Classification	LIT	(Alves et al., 2010)
DIMS	Botanical dietary supplements	Classification	LIT	(Mattoli et al., 2011)
DIMS	Berry fruits, polyphenol content	Classification	LIT	(McDougall et al., 2008)
DIMS	Catharanthus roseus anthocyanins	Targeted metabolomics	LIT-MS/MS	(Piovan et al., 1998)
DIMS	Hypericins in Hypericum elodes	Targeted metabolomics	LIT-MS/MS	(Piovan et al., 2004)
DIMS	Biodiesel typification	Quality control	Q-ToF	(Catharino et al., 2007)
DIMS	Yeast, footprinting	Classification	Q-ToF	(Pope et al., 2007)
DIMS	Wine aging	Classification	Q-ToF	(Sawaya et al., 2011)
DIMS	Crude cell extracts, bacterial identification	Classification	Q-ToF	(Vaidyanathan et al., 2002)
DIMS	Humic Substances	Classification	Q-Trap (MS/MS)	(Baigorri et al., 2008)
DIMS	Yeast intracellular metabolites	Specific protocol	QhQ	(Castrillo et al., 2003)
DIMS	Mouse tissue, lipidomics	Methodology	QqQ	(Han et al., 2008b)
DIMS	Broccoli fingerprinting	Classification	Single quadruple MS	(Luthria et al., 2008)
DIMS	Ginkgo biloba extract	Method development	Single quadruple MS	(Mauri et al., 1999)
DIMS	Medicinal plant extracts	Screening	Single quadruple MS	(Mauri and Pietta, 2000a)
DIMS	Plant extract	Screening	Single quadruple MS	(Mauri and Pietta, 2000b)
DIMS	Intact bacteria	Classification	ToF	(Goodacre et al., 1999)
DIMS	Olive oil characterisation	Classification	ToF and Q-ToF	(Goodacre et al., 2002)
DIMS	Plant extract	Classification	ToF	(Goodacre et al., 2003)
DIMS	Microbial cells suspensions	Classification	ToF	(Vaidyanathan et al., 2001)
FIE-MS	herbal products adulteration	Quality control	LIT	(Chen et al., 2010b)
FIE-MS	Urine, dog breeds	Classification	LIT	(Beckmann et al., 2010)
FIE-MS	Arabidopsis and potato profiling	screening strategy	LIT	(Enot et al., 2006)
FIE-MS	Grape anthocyanins	Classification	LIT	(Favretto and Flamini, 2000)
FIE-MS	Catharanthus roseus extracts	Targeted analysis	LIT	(Favretto et al., 2001)
FIE-MS	Raw plant extracts	Method development	LIT	(Koulman et al., 2007)
FIE-MS	Rice blast disease	Protocol	LIT	(Parker et al., 2008)
FIE-MS	Plant extracts	Protocol	LIT	(Beckmann et al., 2008)
FIE-MS	Yeast mutants	Classification	LIT	(Mas et al., 2007)
FIE-MS	Dietary supplement	Quality control	LIT	(Sun and Chen, 2011)
FIE-MS	Grapefruit fingerprinting	Quality assurance	LIT and ToF	(Chen et al., 2010a)
FIE-MS	Microbial Extracts	Method development	Single quadruple MS	(Higgs et al., 2001)
FIE-MS	Crude fungal extracts	Classification	Single quadruple MS	(Smedsgaard and Frisvad, 1996)
FIE-MS	Penicillium extracts, screening	Classification	Single quadruple MS	(Smedsgaard, 1997)
FIE-MS	Potato fingerprinting	Classification	ToF	(Beckmann et al., 2007)
FIE-MS	Genetically modified Potato	Classification	ToF	(Catchpole et al., 2005)

Method	Sample/Application	Type/purpose of the study	Mass spectrometer	Reference
FIE-MS	Yeast, mutation	Classification	ToF	(Allen et al., 2003)
FIE-MS	Antifungal modes of action	Classification	ToF	(Allen et al., 2004)
FIE-MS	Plant leaf extracts	Classification	ToF	(Grata et al., 2007)
FIE-MS	Plant extracts	Classification	ToF	(Johnson et al., 2007)
FIE-MS	Escherichia coli tryptophan metabolism	Classification	ToF	(Kaderbhai et al., 2003)
FIE-MS	Enhance Plant Resistance to necrotrophic fungus	Classification	ToF	(Lloyd et al., 2011c)
FIE-MS	Plant extracts	Classification	ToF	(Scott et al., 2010)
FIE-MS	Corn Stover Prehydrolysates	Conditioning Processes assessment	QqQ	(Helm et al., 2010)
2. Low plus ultra-high resolution MS: hierarchical direct ESI-MS				
FIE-MS	Human urine, diet assessment	Classification	LIT/LIT-FT-ICR-MS	(Fave et al., 2011)
FIE-MS	Human urine, dietary biomarkers	Biomarkers	LIT/LIT-FT-ICR-MS	(Lloyd et al., 2011a)
FIE-MS	Human urine, dietary biomarkers	Biomarkers	LIT/LIT-FT-ICR-MS	(Lloyd et al., 2011b)
FIE-MS	Rice blast disease	Classification	LIT/LIT-FT-ICR-MS	(Parker et al., 2009)
FIE-MS	Bacterial infection in plants	Classification	LIT/LIT-FT-ICR-MS	(Ward et al., 2010)
3. High resolution MS: accurate mass direct ESI-MS				
DIMS	Urine, drug metabolites	Targeted	LIT-Orbital trap MS	(Rathahao-Paris et al., 2014)
DIMS	Tomato profiling	Classification	ToF	(Overy et al., 2005)
DIMS	Plant populations profiling	Classification	ToF	(Davey et al., 2008)
DIMS	Plant extract	Methodology evaluation	ToF	(Dunn et al., 2005)
DIMS	Rice blast disease	Classification	ToF and Q-ToF	(William Allwood et al., 2006)
DIMS	Serum, kidney cancer	Biomarkers	Q-ToF	(Lin et al., 2010)
DIMS	Plasma, lung cancer	Biomarkers	Q-ToF	(Lokhov et al., 2012)
DIMS	Crude plant extracts	Classification	FT-ICR-MS	(Aharoni et al., 2002)
DIMS	Phospholipids role for salmonella growth in bile	untargeted metabolomics	FT-ICR-MS	(Antunes et al., 2011a)
DIMS	Salmonella infection on host metabolism	untargeted metabolomics	FT-ICR-MS	(Antunes et al., 2011b)
DIMS	Alcohol liver toxicity (animal model)	Classification	FT-ICR-MS	(Bradford et al., 2008)
DIMS	wine discrimination	Quality control	FT-ICR-MS	(Cooper and Marshall, 2001)
DIMS	Nut oil analysis	Identification	FT-ICR-MS	(Fard et al., 2003)
DIMS	Wine diversity	Classification	FT-ICR-MS	(Gougeon et al., 2009)
DIMS	plant physiology	Classification and feature	FT-ICR-MS	(Gray and Heath, 2005)
DIMS	Plasma/mouse serum	Targeted approach	FT-ICR-MS	(Han et al., 2008a)
DIMS	Rat urine, phospholipidosis	Biomarkers	FT-ICR-MS	(Hasegawa et al., 2007)
DIMS	Rat urine, drug-induced toxicity	Biomarkers	FT-ICR-MS	(Hasegawa et al., 2010)
DIMS	Faecal water, Crohn's disease	Biomarkers	FT-ICR-MS	(Jansson et al., 2009)
DIMS	Plant metabolism	Identification	FT-ICR-MS	(Kai et al., 2009)
DIMS	Champagne Wine	profiling	FT-ICR-MS	(Liger-Belair et al., 2009)

Method	Sample/Application	Type/purpose of the study	Mass spectrometer	Reference
DIMS	Mycobacterium extract	Identification	FT-ICR-MS	(Mougous et al., 2002)
DIMS	Tobacco plant metabolism	Identification	FT-ICR-MS	(Mungur et al., 2005)
DIMS	Cell culture, metabolic activity	Identification	FT-ICR-MS	(Nakamura et al., 2007)
DIMS	Plant metabolic disorders	Methodology	FT-ICR-MS	(Ohta et al., 2007)
DIMS	Plant extracts	Methodology	FT-ICR-MS	(Oikawa et al., 2006)
DIMS	Nut oil, aging	Quality assessment	FT-ICR-MS	(Proschogo et al., 2012)
DIMS	Bacterial extract, geographic isolation	Classification	FT-ICR-MS	(Rossello-Mora et al., 2008)
DIMS	Poppy alkaloids extract	Screening	FT-ICR-MS	(Schmidt et al., 2007)
DIMS	Accurate mass and isotope pattern	Methodology	FT-ICR-MS	(Stoll et al., 2006)
DIMS	E. coli, growth conditions	Identification	FT-ICR-MS	(Takahashi et al., 2008)
DIMS	Plasma, insulin sensitivity	Biomarkers	FT-ICR-MS	(Lucio et al., 2010)
DIMS	Bacterial extract	Classification	FT-ICR-MS	(Nam et al., 2008)
DIMS	Fish liver extract	Methodology	FT-ICR-MS	(Southam et al., 2007)
DIMS	Water flea extract, Cu-toxicity	Classification	FT-ICR-MS	(Taylor et al., 2009)
DIMS	E. coli CYP450	Substrate screening	LIT-FT-ICR-MS	(Furuya et al., 2008)
DIMS	Plant extracts, isotope labeling	Methodology	LIT-FT-ICR-MS	(Giavalisco et al., 2008)
DIMS	Liver transplantation	Methodology	LIT-FT-ICR-MS	(Hrydziuszko et al., 2010)
DIMS	Wine discrimination	Classification	LIT-FT-ICR-MS	(Villagra et al., 2012)
FIE-MS	Yeast, Cd treatment	Classification	LIT-Orbital trap MS	(Madalinski et al., 2008)
FIE-MS	Plasma (mice), hypercholesterolemia	phosphatidylcholine biomarkers	ToF	(Yang et al., 2009)
FIE-MS	Escherichia coli extract, profiling	Method development	Q-ToF	(Fuhrer et al., 2011)
FIE-MS	Yeast extract, footprinting	Classification	Q-ToF	(Hojer-Pedersen et al., 2008)
FIE-MS	Penicillium extracts	Classification	Q-ToF	(Smedsgaard et al., 2004)
FIE-MS	Coumarins extracts	Targeted metabolomics	Q-ToF (MS/MS)	(Yue et al., 2011)

DIMS: direct infusion mass spectrometry using chip-based infusion device (NanoMate) or a syringe pump, FIE-MS: flow injection mass spectrometry, LIT: linear ion trap MS, QqQ: triple quadrupole MS, QhQ: quadrupole-hexapole-quadrupole MS, ToF: time of flight MS, Q-ToF: quadrupole-time of flight MS, FT-ICR-MS: fourier transform ion cyclotron resonance MS and Q-Trap: quadrupole ion trap MS.

1.5.3 Mass spectrometers used for direct ESI-MS

Different types of mass spectrometers and data processing techniques have been used for direct ESI-MS (Table 1-3). Direct ESI-MS experiments were originally performed with nominal mass analysers such as single-stage quadrupole (Luthria et al., 2008, Mauri et al., 1999, Mauri and Pietta, 2000a, Mauri and Pietta, 2000b), linear ion trap (LIT) (Alves et al., 2010, Mattoli et al., 2011, Piovan et al., 1998, Piovan et al., 2004), quadrupole ion trap (Q-Trap) (Baigorri et al., 2008) and triple quadrupole (QqQ) (Helm et al., 2010) instruments. The analysis with low resolution mass spectrometers is highly robust and offers fast scanning over a wide *m/z* range (Gross, 2011), hence, nominal mass ESI-MS datasets are quite reproducible and fast to generate. Metabolite profiling using nominal mass ESI-MS is usually carried out when there is a need for a high-throughput screening approach with wide coverage of metabolites. This method usually provides classification and discrimination between samples according to their biological relevance without the need to identify or quantify metabolites in the sample (Beckmann et al., 2008). Most of nominal mass ESI-MS applications have been concentrated in the area of plant metabolomics (Grata et al., 2007, Johnson et al., 2007), food quality assessment (Alves et al., 2010, Chen et al., 2010a) and microbial profiling/footprinting (Allen et al., 2003, Kaderbhai et al., 2003). However, in order to resolve isobaric compounds (compounds with the same nominal masses) as in the case of more complex matrices such as plasma and urine, high-resolution mass spectrometers (HRMS) are the instruments of choice. Nevertheless, the reproducibility of the analysis with HRMS can decrease as mass resolution increases. It is therefore, in some direct ESI-MS studies, HRMS such as ToF (Beckmann et al., 2007, Catchpole et al., 2005) and Q-ToF (Pope et al., 2007, Sawaya et al., 2011) has been used to generate nominal mass ESI-MS datasets. It is quite understandable they sacrificed resolution with the promise of improving method reproducibility. Alternatively, for the purpose of identification using accurate mass and retaining reproducibility, direct ESI-MS has been employed with hierarchical approach in which nominal mass ESI-MS profiling is followed by a targeted analysis of specific ions using HRMS. Hierarchical nominal mass FIE-MS has found applications to urine in the area of nutrition using ion trap MS followed by FT-ICR-MS analysis (Fave et al., 2011, Lloyd et al., 2011a, Lloyd et al., 2011b). However, this approach increases the capital cost associated with the method as it requires at least two MS instruments to perform nominal and accurate mass analyses.

In contrast to nominal mass ESI-MS, the use of HRMS instruments in direct ESI-MS provides higher mass resolution and therefore, offers a means to annotate metabolites based on their accurate mass. The desire behind using exact mass methodologies is that the measured *m/z* values of specific metabolites can be directly used to interrogate publically available metabolite databases. Thus, accurate mass information obtained from the analysis could tentatively identify those metabolites. Furthermore, HRMS instruments, depending on their resolving power, are capable of resolving most of isobaric compounds that would normally be binned into a single *m/z* in nominal mass ESI-MS and therefore, more spectral information could be obtained from the analysis. However, with the increased resolution, a valid analysis protocol followed by a powerful data processing technique is needed to ensure method reproducibility. Direct ESI-MS untargeted metabolomics have been reported with different types of HRMS instruments such as FT-ICR-MS (Hasegawa et al., 2010), ToF-MS (Davey et al., 2008), Q-ToF (Smedsgaard et al., 2004) and orbital trap MS (Koulman et al., 2009). Q-ToF MS has a resolution of 10,000-17,500 at full width at half maximum (FWHM) depending on the geometry of the ion optics, the flight-tube (e.g. use of a reflectron) and instrument tuning with a typical mass accuracy of 5-10 ppm (Kayser and Warzecha, 2012, Waters, 2005, Draper et al., 2013). A reported FIE-MS method using Q-ToF for microbial profiling was able to detect around 1500 reproducible ions at a resolution of 10,000 (FWHM) with mass error within 1 mDa in the range of *m/z* 100-1,000 (Fuhrer et al., 2011). However, recent types of Q-ToF can achieve a mass resolution up to 50,000 (FWHM) when internal calibration is used (Ranasinghe et al., 2012). The increase in the resolution of MS is usually associated with a loss of sensitivity and dynamic range (highest to lowest ratio of analyte concentration detected) and could only be achieved upon expense of mass accuracy. With the advent of new ultra-high resolution mass spectrometers such as FT-ICR-MS and orbital trap MS, such problems are reduced significantly. FT-ICR-MS has a very high resolution of 100,000–1,000,000 (FWHM) with sub-ppm accuracy while orbital trap MS resolution in the range of 10,000-130,000 (FWHM) and 1-2 ppm accuracy (Gross, 2011). Rathahao-Paris et al., reported detection of more than 3400 ions with average mass error of 2.4 ppm in the mass range of *m/z* 125-1,000 from rat urine analysed by FIE-MS using orbital trap MS (Rathahao-Paris et al., 2014).

Different techniques such as selected ion monitoring (SIM) stitching have been reported to address the dynamic range limitation of direct ESI-MS for both nominal mass and accurate mass profiling. In SIM stitching, datasets are

generated as a series of overlapping spectra with narrow mass windows and then those spectra are stitched (combined) together after acquisition. Southam et al., described a SIM stitching strategy to widen the dynamic range of FT-ICR-MS for DIMS (chip-based infusion) and therefore, low and high concentrations metabolites could be detected with high accuracy (Southam et al., 2007). Datasets were generated in SIM mode in the range of *m/z* 70-500 with 30 *m/z* window overlapped by 10 *m/z* and then the acquired spectra were stitched using custom-written MATLAB algorithms (Payne et al., 2009). They used fish liver samples to demonstrate that, when a wide mass range is used the SIM stitching method increased the detection of peaks by a factor of 5.3 with mass error less than 0.5 ppm compared to standard wide-scan range (WSR). Similarly, SIM stitching methods for two mass ranges (low: *m/z* 15-200 and high: *m/z* 110-2,000) have been reported with nominal mass ESI-MS for the analysis of dog urine (Beckmann et al., 2010) and infected plant extracts (Parker et al., 2008, Parker et al., 2009). This technique was further extended to acquire four scans in a mass range of *m/z* 15-1200 and has been applied to great effect for urine metabolomics (Fave et al., 2011, Lloyd et al., 2011a, Lloyd et al., 2011b). The collection of datasets in narrow mass windows in nominal mass ESI-MS generally improved classification characteristics of multivariate models. However, the use of SIM stitching technique for direct ESI-MS requires an instrument with MS2 capability such as LIT-orbital trap and LIT-FT-ICR-MS which increases the capital cost of the analysis. In addition a sophisticated algorithm is used for SIM stitching with intensive data pre-processing steps, therefore significant time, effort and expertise is required to handle such data.

In summary, direct ESI-MS has been extensively used for both targeted and untargeted metabolomics using different approaches. However, few studies were reported for urine analysis (Table 1-3) and most of direct ESI-MS have not been validated properly; a point will be discussed further in the next chapter. Regarding instrumentation, the high resolution and accuracy of FT-ICR-MS makes it the ideal instrument for direct ESI-MS untargeted urine metabolomics. However, the high capital cost of the instrument limited its use for large and well-funded laboratories. While the advent of relatively affordable benchtop mass spectrometers such as Orbital trap MS offers the opportunity to perform direct ESI-MS analysis more routinely at higher mass resolutions (Kayser and Warzecha, 2012).

1.6 Data processing in LC-MS and direct ESI-MS based metabolomics

Untargeted MS-based metabolomics involves the analysis of treated and control sets of samples, generating large, complex datasets containing the analytical measurements. Difficulties in extracting useful analytical information from these datasets may arise due to (1) the high matrix MS background noise and impurities associated with the sample, (2) poor separation of isomers, (3) multiple analytes co-elution, (4) protonation/deprotonation, ion fragments and cluster/adduct formation associated with the single analyte leading to multiple features in the data sets, and (5) the wide range of analytes in the sample. The numerous features and the complex datasets remain a hurdle in the development of reliable and unbiased data processing approach in high-throughput metabolomics. Therefore, a standardised workflow of metabolomics experiments is essential in order to be analytically informative and useful (Han et al., 2009). In 2007, the chemical analysis working group (CAWG) Metabolomics Standards Initiative (MSI) recommended minimum requirements that should be reported in any metabolomics experiments to maximise the utility of the metadata generated. These requirements include reporting of: collection/sample preparation protocol, experimental analysis, quality control, data processing and identification of metabolites detected in the experiment (Sumner et al., 2007).

The pre-processing workflow of the LC-MS datasets for metabolomics includes, in general, (1) spectral filtering (uses baseline subtraction to reduce spectral noise), (2) peak detection and deconvolution (deconvolution is an algorithm-based process utilised to reverse the recorded data convolution effect) by finding *m/z* values of a single ion in all scans to reconstruct ion chromatograms through peak deconvolution, (3) Peak alignment (matches peaks across replicates) and (4) normalisation (to reduce systematic errors). Each step has it is own parameters to setup with no standard way to do so. Therefore, researchers primarily setup parameters from their own experience which potentially affect their final output. In conclusion, although many data processing packages are available, the gap between accurate data acquisition and accurate data processing still exists and remains an obstacle in high-throughput metabolomics using MS approaches (Han et al., 2009).

In targeted metabolomics, platform-dependent software provided with the instrument is quite satisfactory in processing the limited number of metabolites

of interest. Nevertheless, open source programmes for automatic quantification of floods of data can be used, e.g. MetaQuant (Bunk et al., 2006). On the other hand, in metabolic profiling and fingerprinting, thousands of mass spectral sets including metabolite features such as *m/z*, retention time, and peak intensity as well as different ion fragments can be generated from a single metabolome (Borner et al., 2007). Therefore, an important data processing step is to use unbiased mass spectral deconvolution software such as Progenesis QI. Progenesis QI is platform-independent mass spectrometry software, and it has the ability to analyse different file format directly from different instrument vendors. It provides identification of each individual component in the metabolomics profile. It extracts the analyte peaks and intensities from the raw data and aligns them with a reference spectrum. Progenesis QI performs normalisation with different methods across the dataset of samples (Nonlinear-Dynamics, 2014).

Several commercial instrument vendors and open-source software tools are now available with more advanced features which improve, to some extent, data processing for MS-based metabolomics. The software packages from MS instrument vendors are restricted to certain format that should be generated from a certain platform and they are not applicable, in general, for cross platforms data processing. They include MarkerLynx from Waters, MarkerView from MDS/Applied Biosystems, SIEVE from ThermoFisher Scientific, Profile Analysis form Bruker Daltonics, and Mass Profiler Professional from Agilent Technologies (Han et al., 2009). Open source software tools such as MetaboliteDetector (Hiller et al., 2009), MET-IDEA(Broeckling et al., 2006), TagFinder (Luedemann et al., 2008), MetaAlign (De Vos et al., 2007), MZmine (Katajamaa and Oresic, 2005), MZmine2 (Katajamaa et al., 2006), XCMS (Smith et al., 2006), XCMS2 (Benton et al., 2008) and MathDAMP (Baran R Fau - Kochi et al., 2006) are also available. However, the process results should be exported into common data format files (CDF), e.g. netCDF, mzData, mzML, mzXML and ASCII, first before processing. The analysis of LC-MS based metabolomics datasets may require relatively long processing time and therefore, limiting the overall throughput of the analysis (Han et al., 2009).

In contrast to LC-MS datasets, direct ESI-MS datasets are lower in dimensionality in which the retention time (RT) is not a variable. LC-MS software are not designed, in general, for processing direct ESI-MS datasets, therefore, many research groups had developed their own data processing algorithms in R or MATLAB environment in combination with sophisticated

signal acquisition procedures (Southam et al., 2007, Han et al., 2008a). However, these data processing protocols require extensive output validation, efforts, time and expertise to ensure the quality of the extracted data matrices and therefore, increasing the debates of considering them as a high-throughput data analysis approach. Alternatively, some dedicated software for direct ESI-MS data pre-processing are available such as MetaboAnalyst (Xia et al., 2012) and SpecAlign (Wong et al., 2005). However, the raw datasets should be converted into common data format files, e.g. netCDF or comma-separated values (.csv), first before pre-processing. On the other hand, some LC-MS data processing software such as Progenesis QI, has the capability to handle direct ESI-MS raw datasets without conversion (Nonlinear-Dynamics, 2014). This makes data analysis a simple process with minimum intervention, hence high-throughput analysis could be achieved.

In conclusion, the wide availability of data analysis software and bioinformatics tools promises to facilitate interpretation and data processing of many metabolomics studies worldwide. However, such tools still require further refinement. Finding an efficient metabolic identification strategy is still a bottleneck hurdle in pursuing high-throughput in metabolomics and a challenging future task.

1.7 Multivariate analysis (MVA) in MS based metabolomics

The complexity of data generated by metabolomics datasets result in 2 dimensional multivariate data matrices (n x m) where n denotes metabolite features (variable) and m denotes samples (observation). To obtain useful information from these matrices, multivariate analysis (MVA) techniques are needed for data modelling and exploratory analysis. These MVA approaches include the use of unsupervised approach with no need for prior sample information such as principal component analysis (PCA) which reduces the number of variable in the matrix into specific small number of projections named as “principal components”. Then, the differences or similarities of each metabolic profile of the sample are displayed as a plot consists of 2 or 3 principal components called a “score plot”. Unfortunately the score plot does not give any information about variable trends; instead, a correlated loading plot can reveal such trends. Partial least square combined with discriminant analysis (PLS-DA), is an alternative supervised extension of PCA. It has the advantage of maximising the separation between groups by providing classes’ information to

the model. However, the modelling of the data using PLS-DA should be followed by proper model validation as data overfitting is a possibility, which could deviate the model (Li et al., 2009). In Figure 1-6, PCA and PLS-DA score plots are compared for human plasma metabolic profiles of groups of 10 healthy volunteers (control) against 10 metabolic syndrome patients (study) analysed by UHPLC-Q-ToF-MS. It is obvious that supervised PLS-DA plot shows better groups separation than PCA plot. Another extension of PLS is orthogonal PLS-DA (OPLS-DA) which is additionally featured orthogonal signal correction. In OPLS-DA, the datasets are divided into two parts; one part contains the systematic variations of the variables (n) which are linearly related to the observations (m) and the second part is orthogonal. This facilitates and provides better model fitting and interpretation of the data compared to PLS-DA model (Trygg and Wold, 2002).

It is worth mentioning that, the data pre-processing steps, e.g. normalisation, scaling and possibly transformation before MVA have a large impact on the subsequent models generated and can significantly affect the metabolites extracted and identified by MS-based metabolomics. Van den Berg, et al., demonstrated (using LC-MS datasets) the effect of centering, scaling, and transformations on the quality of biological information that can be gained from a metabolomics study (Van den Berg et al., 2006).

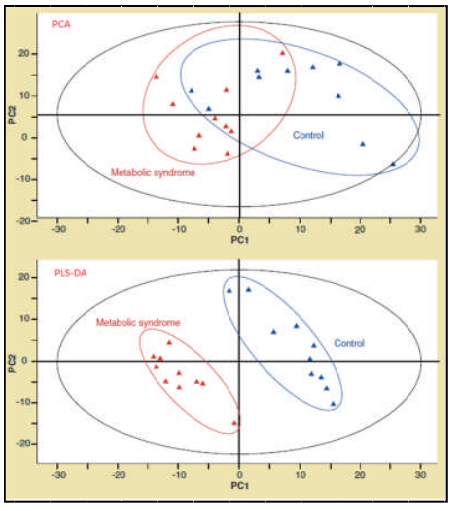


Figure 1-6 Principal component analysis (PCA) and partial least square-discriminant analysis (PLS-DA) score plots of human metabolic syndrome study. Comparison between the score plots of the 1st VS 2nd principal components from PCA (upper) and PLS-DA (lower) models of datasets from metabolic syndrome patients (red triangles) and healthy controls (blue triangles) analysed by C8 UHPLC-Q-ToF-MS (Van den Berg et al., 2006).

1.8 Metabolite identification in MS-based metabolomics

Extracted metabolites from HRMS-based metabolomics are putatively identified by searching local or web-based databases against their measured accurate mass. Then, metabolites identities can be confirmed further by performing MS/MS analysis in which the obtained fragmentation pattern is compared with reference spectral libraries. Whereas absolute identification is achieved only by comparing metabolites retention times and MS/MS spectra with reference standards (Gika et al., 2014). The large amount of data that can be obtained from the metabolome drives the development of many bioinformatics software and metabolome databases. For complexity and purpose of interest, the metabolome open source databases can be categorized into (1) compound-specific, e.g. PubChem (Wang et al., 2012b), (2) species-specific, e.g. HMDB (human specific metabolites database) (Kouskoumvekaki and Panagiotou, 2011), KNApSAcK (plant and microorganisms specific database) (Afendi et al., 2012) (3) pathway specific, e.g. KEGG (general pathways) (Krishnappa, 2011) and (4) reference spectral database, e.g. MassBank (MS/MS standard chemical substances spectral database) (Horai et al., 2010).

1.9 General considerations in the development of urine MS-based metabolomics methods

1.9.1 Analytical considerations

In MS-based metabolomics studies there are different sources of variability associated with the analysis of biological samples. They could be genuine due to the nature of the biological sample (biological variability) or artificial (introduced variability) due to pre-analytical or analytical manipulations in the study. These sources of variability are mainly of biological, pre-analytical (e.g. sample collection and preparation) and analytical origins. In urine analysis the biological variability is quite high due to extreme diversity of metabolite classes present in the sample which exist in wide dynamic concentration ranges and they are continuously subjected to huge inter-individual variability (Da Silva et al., 2013). Therefore pre-analytical and analytical variability should be kept at the minimum in order to estimate the biological one (Gika et al., 2012).

1.9.1.1 Biological variability

It is advantageous during the stage of designing the experiment for urine metabolomics to minimise the possible factors that could increase the biological variability and they are not relevant to the aim of the study such as differences in gender and age of subjects. Un-matching gender and ages of the subjects in urine metabolomics have been demonstrated to have a substantial impact on metabolic profiles (Holmes et al., 2008). It likewise has been shown that the urine metabolic profiles can easily detect differences associated with samples collected from different cities and countries which is possibly due to different environment, cultural habits and diet from one place to another (Lenz et al., 2004). Therefore, during subject recruitment for urine metabolomics, inclusion and exclusion criteria should be in favour of controlling such differences. This will be discussed in details in chapter 2, section 2.1.2.2.

1.9.1.2 Pre-analytical variability

When clinical samples are collected, the collection procedure should be highly standardised, e.g. specify samples time of collection during the day (Fave et al., 2011). Diurnal variation has been demonstrated to have a great influence on the urine metabolic profiling (Lenz et al., 2003), a point will be discussed in detail in the next chapter. In addition, the handling steps of urine samples during collection and sample preparation may introduce a considerable bias in the study, these steps include: sample collection containers, time before freezing, storage temperature, thaw/freeze cycles and sample preparation procedure (SMRS, 2005). Therefore, many studies developed different protocols to standardise urine collection and sample preparation procedure and they demonstrated significant improvement in the analytical outcomes (Fave et al., 2011, Araki et al., 1990, Carrieri et al., 2000, Saude et al., 2007).

1.9.1.3 Analytical variability

In MS-based metabolomics, especially LC-MS, analysis of large batches of samples are required and potentially takes 10s to 100s hours of continuous analysis. Such analysis is only practical if good analytical stability is achieved. The repeatability/reproducibility of the MS-based analytical methods is more prone to variability than other methods such as NMR-based metabolomics. This is not just because of the nature of the MS instrument, but also due to the extensive samples pre-treatment required prior analysis. Therefore, the

instrument performance should be checked and optimised to reduce the variability to the minimum. The use of pooled quality control (QC) samples and/or a set of analytical standards interspaced between samples are the most common techniques to assess the performance of the MS instrument. Different protocols in the literature suggest following a roadmap for a "good practice" in LC-MS based metabolomics (Want et al., 2010, Zelena et al., 2009).

1.9.2 Data analysis considerations

The datasets generated by MS-metabolomics are multivariate by nature; hence, the majority of metabolomics studies use multivariate analysis (MVA) approaches, e.g. PCA and OPLS-DA to report their findings. Most of instrument vendors, as mentioned earlier, provide different software packages for multivariate analysis. Despite the elegance of MVA models, their predictive power might be compromised due to model overfitting. Therefore, a substantial MVA model validation is needed such as cross-validation (CV), permutation and training/prediction sets can be handled using the classical univariate statistical approaches which consider only one variable at a time (Saccenti et al., 2014). However, univariate approaches overlook correlations within *m/z* and/or retention time features and subsequently information related to correlation trends in the data are not retained. Therefore, it is also recommended to use both multivariate and univariate analysis in parallel to maximise the confidence in the extracted sets of features from metabolomics studies (Goodacre et al., 2007). Moreover in univariate analysis, normal distribution is assumed when parametric statistical tests such as ANOVA or Student *t*-test is used and in many MS-derived datasets normality might be violated (Vinaixa et al., 2012). If the data is not normally distributed (i.e. skewed distribution), transformations such as log transformation (Broadhurst and Kell, 2006), power transformation (Van den Berg et al., 2006) or ArcSinh transformation (Jones, 2008) can be effective ways to restore normality to the data.

When univariate analysis is used for MS-based metabolomics, the number of tests equates the number of the MS features detected. Subsequently, type I error (false positive) may occur due to the multiple testing problem. Different approaches have been used for multiple testing corrections such as family wise error rate (FWER) and false discovery rate (FDR). FWER such as Bonferroni correction remains the most commonly used parameter for ascribing significance levels to statistical test in metabolomics. However, its conservative power to keep a strict

control on avoiding Type I error (false positive) increases the probability of Type II error (false negative). Other methods such as FDR strike a balance between the need to avoid false discoveries (type I error) and retaining true discoveries from being missed (type II error). FDR calculates q-value (adjusted p-value) for each tested feature by estimating the rate of significant features being false. In another words, the number of false positives is computed out of the significantly varied metabolic features, for example, a p-value = 0.05 means that 5% of all computed tests might be false positive while a q-value = 0.05 implies that only 5% from the significant tests might result in false positive (Vinaixa et al., 2012). Some commercial LC-MS software such as Progenesis QI calculates both ANOVA p-values and FDR q-values after applying ArcSinh transformation to restore normality (Nonlinear-Dynamics, 2014).

All these practical and data analysis aspects should be considered carefully when designing and executing an MS-based metabolomics study. The choice of following a certain workflow above another depends mainly on the biological context of the study, instrumentation and data processing approaches.

1.10 General aims and objectives of the study

In untargeted metabolomics studies, when a large batch of samples must be analysed, one of the main concerns in the analysis is the stability of the samples during the run especially for unstable metabolites (Han et al., 2008a). The use of the common chromatography MS with its long separation time even make the situation worse in term of sample stability. A truly effective analytical technique must be capable not only of sensitive and selective detection of diverse molecules present in the metabolome, but also of accurate quantitation over a wide range of concentrations, in a high-throughput fashion (Gamache et al., 2004).

This PhD research aims to develop and compare several high-throughput direct ESI-MS approaches for urine metabolomics. The specific objectives required to fulfil this aim are as follows:

- Optimise an LC-MS method for urine metabolomics.
- Develop three different direct ESI-MS methods for urine metabolomics.
- Compare and validate the developed direct ESI-MS methods with the LC-MS.

Apply the developed direct ESI-MS methods and the LC-MS for clinical urine metabolomics to study:

- Osteoarthritis (OA) urinary signature.
- Malaria urinary biomarkers.
- The dietary effect of oligofructose on gastrointestinal microbiota in healthy volunteers.

CHAPTER TWO

Development of High-Throughput Mass Spectrometry Approaches for Clinical Urine Metabolomics

2. Development of High-Throughput Mass Spectrometry Approaches for Clinical Urine Metabolomics

2.1 Introduction

2.1.1 Current use of direct mass spectrometry for urine-based metabolomics

Urine composes of aqueous media with a high salt content and many small molecules, e.g. creatinine, urea, and amino acids but also different organic acids and xenobiotics (Strasinger and Di-Lorenzo, 2008). The interest of using urine in metabolomics has its advantages as it involves non-invasive sample collection, well suited for clinical or epidemiological applications. However, there are relatively few studies reported in the literature related to the use of flow injection electrospray-mass spectrometry (FIE-MS) for the analysis of urine samples, mainly in toxicological and nutritional areas as summarised in Table 2-1. High-throughput FIE-MS targeted screening for inborn errors of metabolism in urine samples under controlled pH conditions has confirmed that 4-hydroxyglutamate is a biomarker for primary hyperoxaluria type 3 (PH3) (Pitt et al., 2015). A further study has used FIE-MS/MS as an alternative tool to immunoassays for screening for mephedrone, six amphetamine-type stimulants (ATS), ketamine and its metabolites in human urine (Lua et al., 2012). Similarly, FIE-MS has found its use for rapid quantification of creatinine using stable isotope dilution tandem mass spectrometry in urine (Niesser et al., 2012). The use of a liquid handling robotic system (e.g. NanoMate) as a recent automated direct infusion mass spectrometry (DIMS) introduction system with chip-based nanoelectrospray (nanoESI) as an alternative to manual syringe infusion has also getting more intension for clinical studies. NanoMate has been applied for structural elucidation of drug metabolites in urine using Q-ToF in MS/MS mode (Trunzer et al., 2007). Studies involving complex carbohydrate analysis have demonstrated the high-throughput power of the technique for glycoconjugates detection in urine (Froesch et al., 2004) or structure assignment of carbohydrate components in complex mixtures (Zamfir et al., 2004).

In many cases FIE-MS has been employed with low resolution MS analysers such as quadrupole and ion trap (Castrillo et al., 2003, Cao et al., 2008), but it is accepted that high resolution accurate mass analysers, e.g. time of flight (ToF),

FT-ICR-MS and orbital trap MS, are more suited for untargeted metabolomics (Giardi et al., 2010). FT-ICR-MS has been reported for FIE-MS human urine metabolomics for dietary biomarkers (Lloyd et al., 2011c, Lloyd et al., 2011a, Fave et al., 2011). A recent study described an FIE-MS method with an orbital trap MS instrument for the rapid detection of vinclozolin metabolites in the urine of rats (Rathahao-Paris et al., 2014). In direct ESI-MS the quality of the data, performance and outcomes obtained employing electrospray ionisation (ESI) is quite promising when it is used in combination with chemometric data analysis tool such as cluster analysis and principal component analysis. Generally, its use, in comparison with other detection techniques, is considered competitive in terms of speed and simplicity especially when large samples sets are analysed. Nevertheless, it must be dealt with caution as its potential gains lie in compromise between speed and information (Giardi et al., 2010).

Table 2-1 An overview of direct ESI-MS based publications for urine metabolomics

Technique	Type of the study	Aim	Mass spectrometer	Reference
FIE-MS	Untargeted metabolomics (Dietary biomarkers)	Biomarkers related to consumption habitual citrus fruit	LIT, LIT-FT-ICR-MS	(Lloyd et al., 2011a)
FIE-MS	Untargeted metabolomics	Food biomarkers	LIT, LIT-FT-ICR-MS	(Lloyd et al., 2011b)
FIE-MS	Untargeted metabolomics (Diet assessment)	Feature selection and classification	LIT, LIT-FT-ICR-MS	(Fave et al., 2011)
DIMS	Untargeted metabolomics (Rat urine, toxicology)	Phospholipidosis biomarkers	FT-ICR-MS (syringe pump)	(Hasegawa et al., 2007)
DIMS	Untargeted metabolomics (Rat urine, toxicology)	Drug induced toxicity	FT-ICR-MS (syringe pump)	(Hasegawa et al., 2010)
FIE-MS	Targeted metabolomics (confirmation)	hyperoxaluria type 3 (PH3) biomarker	Triple quadrupole	(Pitt et al., 2015)
FIE-MS	Targeted metabolomics	Drug screening	Triple quadrupole	(Lua et al., 2012)
FIE-MS	Targeted metabolomics	Creatinine quantification	Triple quadrupole	(Niesser et al., 2012)
DIMS	Targeted metabolomics	Structural elucidation of urinary metabolites	Q-ToF (NanoMate)	(Trunzer et al., 2007)
DIMS	Targeted metabolomics	Carbohydrate screening (Schindler's disease)	FT-ICR-MS (NanoMate)	(Froesch et al., 2004)
DIMS	Targeted metabolomics	Carbohydrate analysis	Q-ToF (NanoMate)	(Zamfir et al., 2004)
FIE-MS	Targeted metabolomics	Identification of xenobiotics metabolites	LIT-Orbital trap MS	(Rathahao-Paris et al., 2014)
FIE-MS	Untargeted metabolomics	Breast cancer biomarkers	Triple quadrupole	(Chen et al., 2015a)
DIMS	Untargeted metabolomics (Mice urine, development)	Alzheimer's disease biomarkers	Q-ToF (syringe pump)	(Gonzalez-Dominguez et al., 2014)
DIMS	Targeted metabolomics	Quantification of Methylphenidate	Triple quadrupole (NanoMate)	(Yang et al., 2004)
FIE-MS	Untargeted metabolomics (Dogs urine, Diet)	Dietary metabolism in domestic dogs	FT-ICR-MS	(Beckmann et al., 2010)

DIMS: direct infusion mass spectrometry using chip-based infusion device (NanoMate) or a syringe pump, FIE-MS flow injection mass spectrometry, Q-ToF: quadrupole-time of flight MS, LIT: linear ion trap MS, and FT-ICR-MS: fourier transform ion cyclotron resonance MS.

2.1.2 Analytical challenges in the development of direct ESI-MS methods for urine metabolomics

2.1.2.1 Ion suppression

Ion suppression during the ionisation phase is one of the most important factors that can affect the performance of the mass spectrometer for direct ESI-MS analysis of complex biological samples. Sample related issues such as co-eluting analytes and sample matrix (e.g. salts in urine) may have a high impact on this effect. Although it affects both electrospray ionisation (ESI) and atmospheric pressure chemical ionisation (APCI), a comparative study using biological extracts has shown that ESI is more prone to ion suppression than APCI. Further investigations in the same article have indicated that the change in the droplet properties due to presence of non- or less volatile materials is the main cause of ion suppression in the ESI (King et al., 2000). Samples containing non-volatile solutes such as salts, xenobiotics/metabolites, endogenous compounds or ion pairing agents have a direct effect on the efficiency of the formation and evaporation of droplets during the ESI process. This in turn has a direct effect on the population of ions generated in the gas phase and subsequently, deteriorates the detection performance of the mass spectrometer. Other factors such as mass and charge of the individual analytes may also contribute as a source of ion suppression for the analyte itself or for other ions. High mass molecules suppress the ionisation of smaller ones (Sterner et al., 2000) and polar compounds are more prone to ion suppression than non- or less polar compounds (Bonfiglio et al., 1999).

2.1.2.2 Urine variability

In metabolomics studies, when sample preparation and the analytical procedure are adequately standardised, the remaining source of variation is mainly that of biological origin. This variability is a main concern in the case of urine as it is continuously subjected to huge inter-individual variability in terms of ionic strength, pH and osmolarity. The unpredictable dilution throughout the day causes urine volume to vary, it is not uncommon in a metabolomics study to encounter up to 15-fold differences in urine volumes (Tsuchiya et al., 2003). As a result, the amount of metabolites in urine also varies. All these sources of variability might question the suitability of urine for metabolomics analysis. However, there are many advantages associated with urine samples make it an attractable choice. Urine collection is non-invasive, simple and is less likely to be

volume-limited which provides enough aliquots for metabolite identification, quantification and subsequent data analysis. Furthermore, it can be easily sampled at different time intervals allowing temporal metabolic changes to be investigated. In addition, urine being a waste product pool of the biological system and it is not under homeostatic regulation, it can provide insights into metabolic dis-regulation associated with physiological changes during normal or disease processes (Chen et al., 2013a). Therefore, different techniques and approaches such as post-analysis normalisation, sample-intervened normalisation and standardised urine protocols have been reported to compensate such variability for metabolomics studies. Post-analysis normalisation mainly include normalisation to creatinine (Alberice et al., 2013, Wagner et al., 2010, Cone et al., 2009, Heavner et al., 2006), urine volume normalisation (Warrack et al., 2009), osmolality normalisation (Chadha et al., 2001) and total area normalisation (Sen et al., 2013, Chetwynd et al., 2015). Creatinine excretion may vary due to gender, lean body mass and kidney function (Miller et al., 2004, Waikar et al., 2010). Therefore normalisation to creatinine should be limited to subjects with normal kidney function. Normalisation to osmolality is used to correct the concentration of all metabolites in urine samples based on the premise that the concentration of osmolite is a direct measure of total metabolites in the sample (Gyamlani et al., 2003). However, osmolality is often affected by the presence of insoluble substances in urine (Chadha et al., 2001) and therefore, might not reflect the true amount of metabolites in urine. Total area normalisation is also used in MS-based analysis, however, the background noise and ion suppression of urine matrix may significantly interfere with the total signal in ESI-MS. In addition, metabolites response in ESI is a compound specific, therefore normalisation to total peak area might yield biased results as it may not reflect the total concentrations of metabolites in urine (Chen et al., 2015a).

Sample-intervened normalisation such as normalisation to multiple internal standards (IS) (Bijlsma et al., 2006) or chemically labelled metabolites (Wu and Li, 2012) has also been reported for MS-based metabolomics and might be applicable to different clinical samples including urine. Recently, Chen et al., described an FIE-MS with a matrix-induced ion suppression (MIIS) normalisation technique for urinary metabolomics (Chen et al., 2015a). In this approach, the intensity of an ion suppression indicator spiked into different aliquots of diluted urine samples was measured, and then a correction factor was calculated by subtracting the intensity of the indicator in the samples from a

blank. This correction factor was then used to estimate the extent of suppression in the signal by the urine matrix. Subsequently, a regression equation was generated to estimate the relative concentration of unknown urine metabolites. A good reproducibility with this method was obtained with relative standard deviations (RSDs) below 8.95% for the intensity of the indicator across 50 samples. However, this type of sample-intervened normalisation techniques require each sample to be chemically labelled or spiked with internal standard which will increase the analytical cost and decrease the throughput of the experiment. A different normalisation concept called "MS total useful signal" (MSTUS) has been introduced by Warrack et al. (Warrack et al., 2009). MSTUS normalisation uses the total intensity of peaks that are common in all samples to generate a normalisation factor which is then used to normalise individual peaks and, therefore, avoids the possible xenobiotics and artefact interferences and subsequently, more realistic measure of urinary metabolites profiles could be achieved. This method has been proven to reduce variations between biological replicates and detect significant differences in the metabolic profiles of urine samples (Chen et al., 2013a). However, it does not provide a control measure to estimate the suitable amount of urine sample to be introduced into mass spectrometer, a point that will be discussed further in the next section.

Although much efforts have been focused on normalisation, but different studies also have considered managing other sources contributed to urine variability. Dietary intake habit was reported to have a remarkable impact on urinary composition and significant changes in endogenous urinary metabolites were observed (German et al., 2003, Phipps et al., 1998, Gavaghan et al., 2001). However, pre-intervention diet intake (e.g. fasting) can be controlled, and this has proven to reduce both inter and intra-individual variability in urinary metabolomics (Lenz et al., 2003, Walsh et al., 2006, Winnike et al., 2009, Wallner-Liebmann et al., 2014). Moreover, the use of a standardised urine collection procedure has shown to minimise such variability (Fernandez-Peralbo and Luque de Castro, 2012). Urine is normally collected as random spot samples (Rybi–Szuminska et al., 2014) or 24 h pooled samples (Le et al., 1999), sometimes longer sampling times were used in pharmacokinetics studies (Blum et al., 1994). Collection of random-spot samples, including first void in the morning or late-void samples, has the potential of being convenient, manageable and improves subject compliance. However, diurnal variation is not taken into consideration and therefore different metabolite profiles may be observed as a result of random and short time collection of urine samples (Araki et al., 1990,

Carrieri et al., 2000). Saude et al., calculated the variability associated with quantification of 24 metabolites in spot urine samples (without controlling dietary intake) and as expected, they observed high variance between metabolites in the samples (Saude et al., 2007). Fave et al., developed a standardised protocol for urine collection and demonstrated that spot urine collection at different times of the day gave distinctly different metabolic profiles (Fave et al., 2011). Therefore, when overall individual metabolic status is investigated, 24 h pooled urine samples are preferred, if possible, to minimise the variability in the urinary profiles obtained with shorter collection periods (Fernandez-Peralbo and Luque de Castro, 2012).

2.1.3 Justification of the study

Different sample preparation protocols have been used for urine metabolomics such as solid-phase extraction (SPE) (Chetwynd et al., 2015, Idborg-Bjorkman et al., 2003), liquid-liquid extraction (LLE) and simple urine dilution (Fernandez-Peralbo and Luque de Castro, 2012). The amount of salts in urine has a direct effect on the extent of ion suppression experienced in ESI (van Hout et al., 2003). Therefore, extensive sample clean-up may reduce the amount of salts in urine and subsequently the ion suppression effect on ESI. However, for high-throughput FIE-MS analysis, urine dilution was reported to be a quick and reasonable approach compared to SPE and LLE in terms of simplicity, reproducibility and no metabolites loss during sample preparation which leads to unbiased urinary metabolomics investigations (Gonzalez-Dominguez et al., 2014). Nevertheless, an adequate volume of urine in the sample for the analysis should be carefully estimated. The amount of the sample injected in MS analysis is one of the important factors for the analyte detection. If a small amount is used, the concentration of some compounds might be below the MS limit of detection and subsequently those compounds will not be detected. On the other hand, if a large amount of sample is used, ESI source and MS detector can be easily saturated and high abundance ions might obscure small ones and make them undetectable (Mattarucchi and Guillou, 2012). Most of the reported urine metabolomics studies (Lloyd et al., 2011b, Want et al., 2010) used different percentages of urine in the sample without addressing the need to estimate the proper urine amount for the analysis.

In FIE-MS samples are infused directly into MS analyser, which may reduce sensitivity and deteriorate metabolite identification. Regardless the type of MS used; matrix and ion suppression effect on ESI is a major inevitable issue notably

with urine samples. Therefore, for FIE-MS urine metabolomics it is essential to optimise MS source parameters to improve ionisation with minimum ion suppression. The use of standards mixtures of chemicals naturally found in urine has been reported to enhance detection in MS (Fitzgerald et al., 2012). However, most of reported direct ESI-MS urinary studies (Table 2-1) optimised the MS source parameters without considering the diversity of chemical classes present in urine and the need to improve instrument performance in order to maximise metabolites coverage, e.g. by using a mixture of standards. Thus there is a need to establish a urine analysis protocol which will help enhance the detection of urinary metabolites, reduce analytical variability and minimise ion suppression effect of urine salts on ESI.

Variability associated with urine samples has been thoroughly investigated with different techniques in order to maximise reproducibility in urine metabolomics, as detailed in the previous section. However, most of the reported normalisation methods were either flawed, time consuming or require high cost and efforts (Wu and Li, 2012). Nevertheless, some reported methods, despite the drawbacks, such as MS total useful signal (MSTUS) normalisation has proven to reduce the possible xenobiotics and artefact interferences associated with urine samples (Warrack et al., 2009). However, this method lacks the ability to estimate the optimum sample amount for MS-based analysis. Chen et al., proposed a strategy of combining MSTUS normalisation with a pre-analysis optimisation of urine volume for LC-MS and better results were obtained (Chen et al., 2013a). Yet this approach has not been fully investigated for direct ESI-MS and might be an adequate approach for reducing urine variability in order to attain the essential prerequisites for clinical urine metabolomics.

Several attempts in the literature have been reported to improve the performance of direct ESI-MS and LC-MS for urine metabolomics using advanced ESI sources such as nanoelectrospray ionisation (nanoESI) (Trunzer et al., 2007, Froesch et al., 2004, Zamfir et al., 2004, Chetwynd et al., 2014, Chetwynd et al., 2015). NanoESI using a chip-based infusion device, NanoMate, for DIMS or LESA-MS analysis has demonstrated better ionisation efficiency, no carry over effects and higher throughput than the conventional FIE-MS using standard ESI (Yang et al., 2004). Yet, this technique, with all its merits, has not been fully evaluated for clinical urine untargeted metabolomics. Most of the published urinary studies using chip-based infusion were concentrated mainly in the area of toxicology (animal models) or targted analysis (Table 2-1). LESA-MS has been used for metabolite quantification in different body fluids other than urine, e.g. dried

blood spots (Kertesz and Van Berkel, 2010) and tissues (Tomlinson et al., 2014). This technique showed to produce different MS spectral profiles compared to LC-MS with standard ESI (Trunzer et al., 2007). This indicates that detection is improved for the analytes that could not be seen by standard ESI or it forms different types of ions, adducts or fragments. Thus, it is essential to investigate and compare the performance of this technique with standard ESI interface when untargeted metabolomics is planned.

In conclusion, direct ESI-MS for urine metabolomics has been used with different sample preparation protocols, normalisation methods and ESI interfaces. There is no general consensus of a standard preparation protocol and most of the reported normalisation methods were either time consuming or flawed. There is no reported study evaluate these conditions to improve the performance of direct ESI-MS for urine untargeted metabolomics and most reported studies lack proper validation (Hasegawa et al., 2010, Hasegawa et al., 2007) or evaluation against conventional methods such as LC-MS. The use of nanoESI interface in either infusion or LESA-MS mode, with all its anticipated advantages, has not been fully evaluated for urine untargeted metabolomics. Thus there is a need for a rigorous comparative study to determine the optimum conditions for the analysis and compare different methodologies to evaluate the most suitable one for urine untargeted metabolomics.

2.1.4 Aims and objectives

- Adapt and optimise a HILIC LC-MS method for urine metabolomics.
- Develop and test three high-throughput direct ESI-MS approaches for human urine untargeted metabolomics (Figure 2-1). These methods are:
    1. Flow injection ESI-MS (FIE-MS).
    2. Chip-based infusion MS.
    3. Liquid extraction surface analysis-MS (LESA-MS).
- Establish a high-throughput data analysis protocol for direct ESI-MS methods.
- Compare and validate the performance of direct ESI-MS methods with LC-MS for urine untargeted metabolomics.

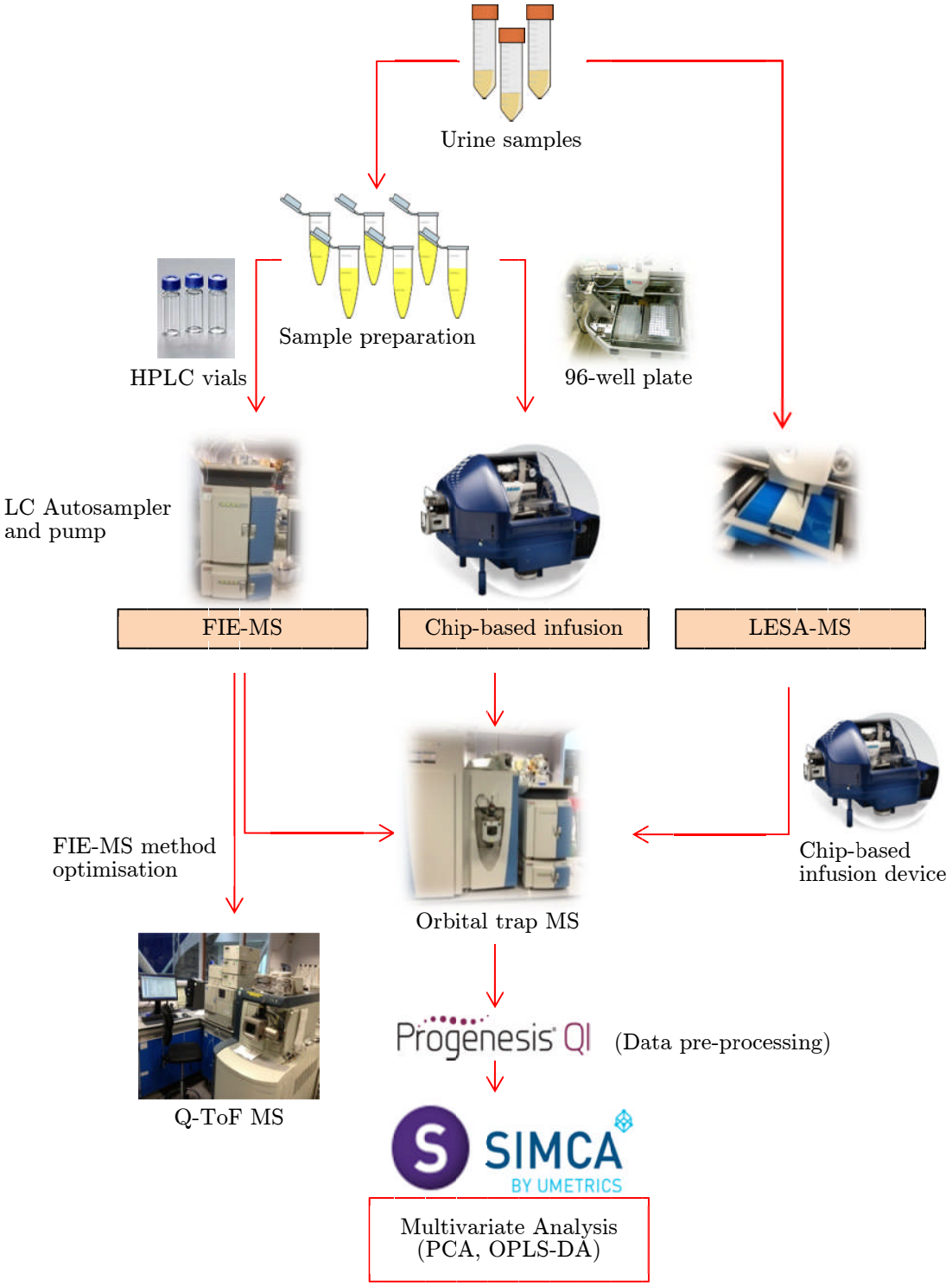


Figure 2-1 A simplified workflow of the direct ESI-MS methods for urine metabolomics

2.2 Materials and methods

2.2.1 Materials and reagents

All chemicals and reagents used are analytical grade, HPLC grade or MS grade and were used without any further purification but minimum handlings were performed to minimise any possible contamination. The materials and reagents and their sources are summarized alphabetically in Table 2-2.

Table 2-2 List of materials and reagents

Standards and Reagents	Description	Supplier
Acetonitrile	Hypersolv Chromanorm acetonitrile 99.9%	VWR international, EU
Adrenaline	L- Adrenaline 98+%	Alfa Aesar, UK
Alanine	L-Alanine	Amresco, USA
4-Aminohippuric acid	4-Aminohippuric acid 99%	Acros Organics, USA
Ammonium acetate	LC-MS ultra-ammonium acetate 99%	Fluka, Sigma-Aldrich, Netherland
Aspartic acid	L-Aspartic acid monosodium salt	Sigma Chemical Co, USA
Citric acid	Citric acid-anhydrous	Sigma-Aldrich, Germany
Creatine	Creatine monohydrate	Acros Organics, USA
Creatinine	Creatinine 98%	Alfa Aesar, UK
Cytidine	Cytidine Ultra-Pure Grade	Amresco, USA
3,4-Dihydroxymandelic acid	DL-3,4-Dihydroxymandelic acid 98%	Sigma-Aldrich, Germany
3,4-Dihydroxyphenylacetic acid (DOPAC)	3,4-Dihydroxyphenylacetic acid 98%	Acros Organics, USA
3,4-Dihydroxyphenyl alanine (L-DOPA)	3,4-Dihydroxyphenylalanine 97%	Sigma-Aldrich, Germany
Formic acid	LC-MS formic acid 98%	Sigma-Aldrich, Germany
0.1% Formic acid	LC-MS 0.1% formic acid in water	Sigma-Aldrich, Germany
Glutamic acid	DL-Glutamic acid monohydrate ≥ 98%	Sigma-Aldrich, Germany
Hippuric acid	Hippuric acid Sodium salt 99%	Acros Organics, USA
Histidine	DL-Histidine 98%	Sigma-Aldrich, Germany
Homovanillic acid	Homovanillic acid 98%	Acros Organics, USA
5-Hydroxyindole-3-acetic acid	5-Hydroxyindole-3-acetic acid ≥ 98%	Sigma-Aldrich, Germany
2-Hydroxyisobutyric acid	2-Hydroxyisobutyric acid	Alfa Aesar, UK
4-Hydroxy-3-methoxymandelic acid	DL-4-Hydroxy-3-methoxymandelic acid ≥ 98%	Sigma-Aldrich, Germany
4-Hydroxyphenylacetic acid	4-Hydroxyphenylacetic acid 98%	Sigma-Aldrich, Germany
Hypoxanthine	Hypoxanthine 99%	Alfa Aesar, UK
Lactic acid	DL-Lactic acid lithium salt 98%	Sigma Chemical Co, USA
Leucine enkephalin	LC-MS Leucine enkephalin	Sigma-Aldrich, Germany
Metanephrine	DL-Metanephrine HCl	Sigma-Aldrich, Germany
Methanol	LC-MS Methanol	Fisher Scientific, UK
MHPG sulphate	3-Methoxy-4-hydroxyphenylglycol (MHPG) sulphate potassium salt	Sigma-Aldrich, Germany

Standards and Reagents	Description	Supplier
3-Methyl-L-histidine	3-Methyl-L-histidine	Sigma-Aldrich, Germany
4-Methyl-2-oxovaleric acid	4-Methyl-2-oxovaleric acid ≥ 98.0%, liquid	Sigma-Aldrich, Germany
Normetanephrine	Normetanephrine HCl 98%	Aldrich Chemical, Germany
Phenylalanine	L-Phenylalanine 99%	Aldrich-Chemie, Germany
2-Propanol	LC-MS 2-propanol 99.96%	Fisher Scientific, UK
Serotonin	Serotonin HCl	Sigma-Aldrich, Germany
Sodium Hydroxide	LC-MS Sodium Hydroxide 97%	Sigma-Aldrich, Germany
trans-Aconitic acid	trans-Aconitic acid 98%	Alfa Aesar, UK
Threonine	L-Threonine 98%	Acros Organics, USA
Tryptophan	L-Tryptophan Pure Ph.	AppliChem GmbH, Germany
Tyrosine	L-Tyrosine	Sigma Chemical Co, USA
Urea	Urea Sigma Ultra	Sigma Chemical Co, USA
Uric acid	Uric acid 99%	Alfa Aesar, UK
Uridine	Uridine ≥ 99%	Sigma-Aldrich, Germany
Water	LC-MS Chromasolv water	Fluka, Sigma-Aldrich, Switzerland

2.2.2 Mass spectrometers

Q-ToF MS (Micromass Q-ToF Premier, Waters, USA) coupled with an HPLC system (SIL-HTc autosampler, LC10ADvp pumps, Shimadzu, UK) was extensively used for FIE-MS method development, optimisation and analysis in both positive and negative ion modes. Initial setting of MS conditions were: capillary voltage (kV) 2.56 (ESI+), 2.75 (ESI-), sampling cone voltage (V) 38 (ESI+), -48 (ESI-), desolvation gas flow (L/h) 250, cone gas flow (L/hr) 70, source temperature 120 °C and desolvation temperature 250 °C. The data acquired in continuum (profile) mode (*m/z* 30-1000) employing 1 µg/mL leucine enkephalin as a lock mass with 1 Hz (1 scan/s) scan rates for 1.0 min. The optimised source parameters were: capillary voltage (kV) 1.5 (ESI+), 1.8 (ESI-), sampling cone voltage (V) 40 (ESI+), 65 (ESI-), desolvation gas flow (L/h) 250 (ESI+), 400 (ESI-), cone gas flow (L/h) 70.0, source temperature 120°C and desolvation temperature 250 °C. The data acquired in continuum mode (*m/z* 30-1000) with 1 Hz scan rate for 1.0 min.

Orbital trap mass spectrometer (Exactive-Orbitrap, Thermo Fisher Scientific, USA) was used in both ESI+ and ESI- modes for all types of analyses. The optimised source parameters were: spray voltage (kV) 1.5 (ESI+), 1.8 (ESI-), capillary voltage (V) 25 (ESI+), -40 (ESI-), sheath, auxiliary and sweep gas flow rate (arbitrary unit) were: 50, 5 and 5 (ESI+/-), respectively. Capillary and

heater temperature were maintained at 250 and 120 °C, respectively, in both modes. Data were acquired in full scan mode with resolution 25,000 from *m/z* 50-1000 with 4 Hz scan rate. A syringe pump (Razel, Connecticut, USA) was used at a flow rate of 50 µL/min to infuse samples from an analytical 5.0 mL syringe (SGE analytical Science, Australia) with a flow of 300 µL/min of 0.1% formic acid/water for MS source parameters optimisation.

2.2.3 Chromatography

Chromatography was performed using an Accela UHPLC system (Thermo Fisher, USA) on BEH HILIC UHPLC column (2.1 x 100 mm, 1.7 µm particle size, Waters, USA) coupled to the orbital trap MS. The column was maintained at 40 °C and a flow rate of 400 µL/min. Mobile phases used were: (A) 50:50 acetonitrile:ammonium acetate (10 mM final concentration) and (B) 95:5 acetonitrile:ammonium acetate (10 mM final concentration). The gradient started with 1% (A) and increased to 100% (A) over 12 min then the composition was returned to its initial conditions and maintained for the second run (15 min total). The injection volume was 5 µL and samples were maintained at 4 °C during the analysis.

2.2.4 Direct ESI-MS analyses

In FIE-MS analysis, the method employed the LC as a flow injection system to inject samples from the pre-chilled autosampler (4 °C) directly without using a column into the MS. 10 µL sample volume was injected in a flow of 300 µL/min of 0.1% formic acid for 0.5 min run. Triversa NanoMate (Advion, USA) coupled to orbital trap MS was used for chip-based infusion and LESA-MS. In chip-based infusion, 10 µL sample volume was infused from a 96-well plate into MS. While in LESA-MS, 10 µL of an extraction solvent of 0.1% formic acid in 1:9 (ESI+) and 7:3 (ESI-) methanol:water in positive and negative ion modes, respectively, was used to extract analytes from 1 µL dried urine spots on Permanox cell culture slide (Thermo Scientific, USA). NanoESI spray conditions comprised a nitrogen gas head pressure of 0.7 psi and 1.5 kV (ESI+), 1.8 kV (ESI-) spray voltage controlled by Chipsoft Manager software (version 8.3.3, Advion BioSciences, USA). Urine dried spots positions on the cell culture slide were mapped for LESA-MS analysis using LESA-Points software (version 1.1, Advion BioSciences, USA).

2.2.5 Human urine samples collection and storage

Two sets of human urine samples; control and after ingestion of green tea were used as a study model to evaluate and compare the suitability of the proposed direct ESI-MS methodologies against a conventional LC-MS method for untargeted metabolomics. Control urine was collected without the use of preservatives from 6 healthy male volunteers (28-35 years old of 21.8-29.8 BMI) in the morning after an overnight 12 hour fasting period. Then, 200 mL of brewed 2 green tea bags with water, *Camellia sinensis* (Twinings, R. Twinings & company, UK) was ingested, followed by 350 mL water to enhance diuresis. The urine samples were collected in a period of 0-4 h after consumption of the green tea (Figure 2-2). Urine samples collected from each subject at different collection intervals were pooled in pre-labelled 50 mL sample collection containers. Eight aliquots (1 mL) from each sample were transferred into 1.5 mL cryogenic tubes and stored immediately in -80 °C freezer.

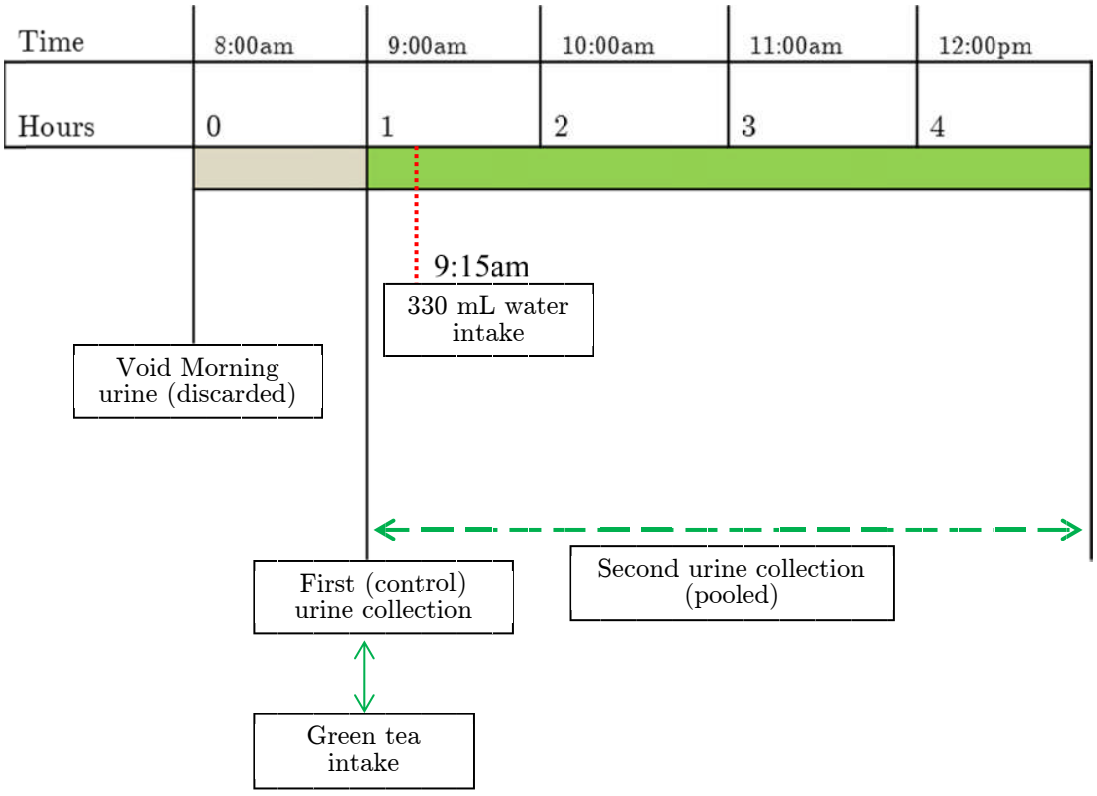


Figure 2-2 Urine samples collection protocol

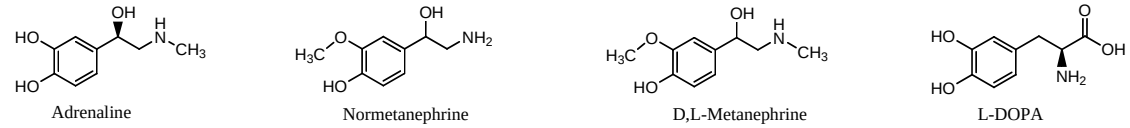
2.2.6 Artificial urine preparation

35 metabolites were selected to represent a range of metabolite structures (Figure 2-3) naturally found or expected in healthy human urine (Saude et al., 2007, Bouatra et al., 2013, Brooks and Keevil, 1997, Rane et al., 2013). A standard solution of each metabolite was prepared in water and was used to generate calibration curves and fingerprint mass spectra with FIE-MS using Q-ToF. A suitable concentration of each metabolite, representative of a typical concentration found in urine was selected to prepare a standards mixture at a concentration range of 2.7-67.8 μg/mL as "artificial urine" (Appendix A). They were stored in Falcon tubes at -20 °C between analyses. Artificial urine was diluted 10 times with 0.1% formic acid in water and used for the optimisation of MS source conditions.

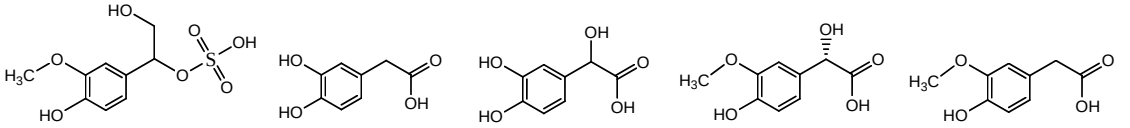
2.2.7 Preparation of urine samples for LC-MS and direct ESI-MS analysis

Thawed urine/artificial urine samples for metabolomics analysis were prepared in three different ways: (1) LC-MS: 60 μL were centrifuged at 10,000 g for 10 min to remove particulate. 50 μL of the supernatant was added to 100 μL water in HPLC amber glass vials containing 200 μL micro glass inserts. (2) FIE-MS and chip-based infusion: 50 μL aliquots of thawed artificial urine/urine were diluted with 450 μL pre-chilled 0.1% formic acid in 1:8 methanol:water spiked with leucine enkephalin to give 0.2 μg/mL final solution. Samples were shaken for 15 min and centrifuged (13,000 g, 4 °C) for 5 min. 60 μL of the supernatant was transferred to HPLC amber glass vials containing 200 μL micro glass inserts. Blank was prepared following the same protocol without including artificial urine/urine. (3) LESA-MS: 1 μL of urine was placed on Permanox cell culture slide and left to dry prior analysis.

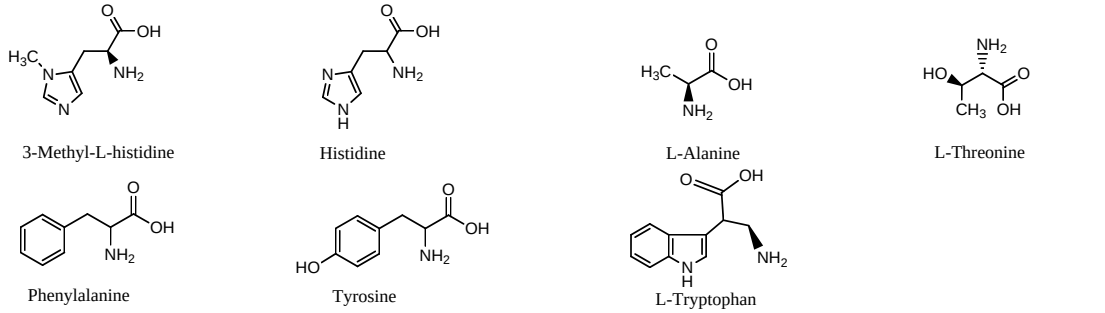
For metabolomics analysis, a pooled QC sample was prepared by mixing 20 μL aliquots taken from each urine sample in the study. Pooled control urine sample was prepared the same but only from the control urine samples. The pooled urine samples were treated the same as described for the samples in this section.



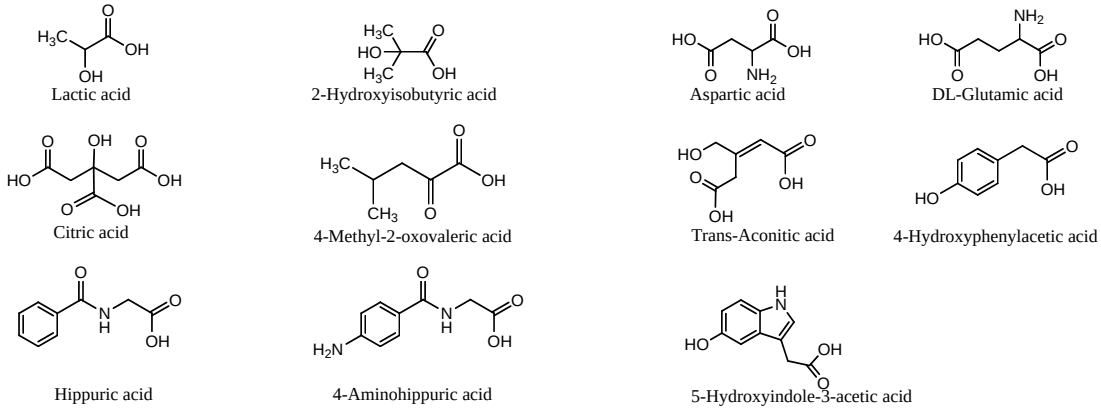
Catecholamines: different side chanis: amines (primary, secondery), alcohols and acids



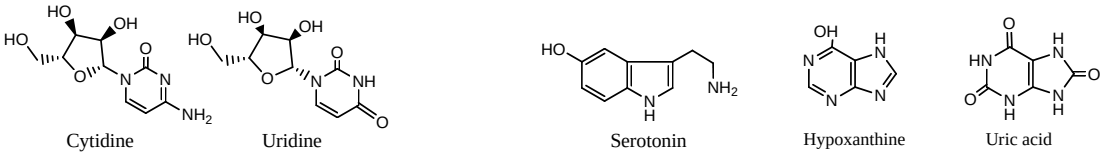
Catechol acids: alcohols, organic and inorganic acids side chains



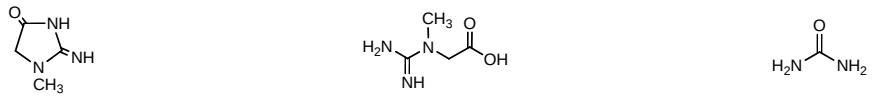
Amino acids and derivatives



Carboxylic acids: aliphatic and aromatic side chains with differnt functional groups



Pyrimidine nucleosides **Heterocyclic compounds**



Nitrogen containing compounds (protein breakdown end metabolites)

Figure 2-3 Structures and chemical groups of metabolites selected for artificial urine preparation

2.2.8 Optimisation of urine dilution for direct ESI-MS analysis

Aliquots (50 µL) of control urine were diluted with different volumes (0 - 450 µL) of methanol. The total ion current (TIC) was then measured for the samples against their methanol blanks to assess the quality of the analysis and find out the optimum volume of methanol that gives the best signal from urine samples.

2.2.9 Assessing the ion suppression effect of urine on standard metabolites

The ion suppression effect in ESI-MS experiments has been assessed using different approaches (Matuszewski et al., 1998, King et al., 2000, van Hout et al., 2003, Fitzgerald et al., 2012). The most direct approach involves comparison of the MS response of a mixture of standards (100% response) with the same amount of standards spiked into pre-extracted samples (the effect of sample matrix on MS response, the ion suppression). The addition of the same amount of standards before extraction highlights whether the loss of the signal is attributed to the ion suppression or extraction process (Matuszewski et al., 1998). For urine dilution optimisation, the above principle was used to study the effect of ion suppression of urine salts and subsequently determine the adequate dilution of urine that will give minimum ion suppression. Different aliquots of control urine (0 - 350 µL) were added to 50 µL of artificial urine and 50 µL methanol (i.e. urine spiked with artificial urine). All the volumes were kept constant at 500 µL with MS grade water. A second set of samples were treated the same; omitting artificial urine (i.e. urine). Triplicates of these samples were analysed by FIE-MS in positive and negative modes. Maximum peak counts of artificial urine analytes were used to compare differences in signal before and after addition of urine. As some of these analytes already found in human urine, the following steps were used to estimate the % ion suppression effect of urine salts on their signals:

(a) Expected analyte level in the samples of (urine + artificial urine) is:

$$\text{Peak count}_{(expected)} = \text{Peak count}_{(artificial\ urine)} + \text{Peak count}_{(urine)} \quad (1)$$

(b) Actual analyte level in the samples of (urine +artificial urine) is:

$$\text{Peak count}_{(actual)} = \text{Peak count}_{(urine+artifical\ urine)} \quad (2)$$

(c) From (1) and (2), the % ion suppression on analyte signal is:

$$\%\ \text{ion suppression} = \frac{\text{Peak count}_{(expected)} - \text{Peak count}_{(actual)}}{\text{Peak count}_{(expected)}} \times 100\% \quad (3)$$

2.2.10 Experimental design for untargeted metabolomics

Control urine (n = 6) and urine samples after consumption of green tea (n = 6) were randomised and analysed in both positive and negative ion modes with the LC-MS, FIE-MS, chip-based infusion and LESA-MS. The whole analysis time for direct ESI-MS methods including operation time was about 1.0 min for each sample. MS signal was sufficiently stable within the injection time for chip-based infusion and LESA-MS and the injection peak was eluted in about 0.3 min in FIE-MS. Therefore, longer infusion time than 0.5 min was not used. Artificial urine triplicates were injected in the beginning and at the end of the run as a reference test mix to check the stability and performance of the instrument for accurate mass measurements. To monitor mass accuracy within each run a reference standard of leucine enkephalin, *m/z* 556.2771 (ESI+), 554.2615 (ESI-) was spiked in each sample to give final concentration of 0.2 µg/mL. Pooled quality control (QC) samples were interspaced with samples for the purpose of monitoring the stability, robustness, reproducibility and performance of the proposed analytical platforms. Blank samples were injected after each sample in FIE-MS to minimise the carryover effect, if any.

2.2.11 Multivariate analysis and metabolite identification

The raw data were acquired and visualised with Xcalibur v2.1 software (Thermo Scientific, USA). The performance of the analytical methods were validated by monitoring a representative set of urine metabolites in the pooled QC sample for retention time (RT) shifts (LC-MS), mass accuracy, relative standard deviations (RSD%) of peak areas/counts. Xcalibur Quan Browser was used to integrate and extract peaks areas of (RT, *m/z*) metabolites pairs from the LC-MS raw data. While, ToxID v2.12.57 software (Thermo Scientific, USA) was used to extract intensities of *m/z* values of metabolites of interest from the FIE-MS, chip-based infusion and LESA-MS raw data. The obtained results were compared with the acceptable limits for bioanalysis (FDA, 2013). Furthermore, the quality of the datasets obtained from the LC-MS analysis was assessed by determining the variability (RSD%) in the mean peak areas of all peaks present in at least 80% of the QC samples using a metabolomics approach proposed by Want et al., (Want et al., 2010). The results were evaluated against the suggested limit for a reliable metabolomics analysis, in which the RSD% across at least 70% of the mean peak areas should be less than 30%. For direct ESI-MS, the quality of the data was assessed using principal component analysis (PCA) as proposed by Beckmann et al., (Beckmann et al., 2008).

For metabolomics analysis, the full datasets of control urine samples and urine samples after ingestion of green tea were imported and pre-processed by Progenesis QI software (Nonlinear-Dynamics, Waters, USA). LC-MS main parameters were set as follows: high resolution-accurate mass MS, profile mode, mass range *m/z* 60-1000, automatic chromatographic peak alignment, automatic peak picking and 5 ppm mass window. For direct ESI-MS, the acquired raw data of urine samples were imported directly into Progenesis QI; the main parameters were set the same as LC-MS except: Adduct, M+H (positive) and M-H (negative) with no chromatographic peak alignment. Pre-processing steps of Progenesis QI include: peak picking, peak alignments and normalisation to all compounds which is a modified type of MS total useful signal (MSTUS) (Warrack et al., 2009). Automatic deconvolution of the extracted *m/z* peaks was also carried out using Progenesis QI to remove isotopes, adducts, and other confounding peaks resulting from MS detection.

The selected metabolites (RT, *m/z*) pairs and *m/z* values of LC-MS and direct ESI-MS methods, respectively, were then exported with their normalised abundances for multivariate analysis (MVA) using Simca P+14 (Umetrics AB, Sweden). As retention times (t_R) were not relevant in the infusion data sets, the separation and clustering of metabolites will be only a function of (*m/z*, intensity) pairs of the variables. Imported datasets were mean-centred and scaled to unit variance (UV). The mean-centering technique subtracts the mean of the variables intensity and hence, shifts the data towards the mean. UV scaling procedure gives the weight of each variable by its standard deviation, which shrinks the weight of intense features and stretches the weight of smaller ones so that all features rest with equal weights (Eriksson et al., 2006b). Principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) were used for modelling the differences between samples and controls. Cross-validation using leave-one out method (1 out of 7) was used to evaluate the robustness of the models by monitoring the fitness of model (R^2Y) and predictive ability (Q^2) values. Prediction method based on randomly selected training (50%) and test sets (50%) using OPLS-DA model were also performed.

The ions responsible for the class separation in OPLS-DA models between samples and controls were selected by means of Variable Importance for the Projection (VIP) and variables loadings plots. To increase the confidence in the selection of those ions, univariate analysis using Student's *t*-test was computed in parallel to test the significant difference of the selected ions between the two groups. Prior to univariate analysis, ArcSinh transformation (Jones, 2008) was

performed for the normalised data to restore normality using Progenesis QI. Accounting for multiple testing problems, p-values were adjusted using false discovery rate (FDR) technique, which was computed also using Progenesis QI. A set of metabolites reported in urine after consumption of green tea was used to generate a local database of green tea metabolites (Ridder et al., 2014, Spencer, 2003, van der Hooft et al., 2012). The exact mass of the significant ions generated from the OPLS-DA models was then used to interrogate the green tea metabolite database for possible identification based on accurate mass measurements within 3 mDa error range.

2.3 Results and discussion

2.3.1 Generation of artificial urine database of structures and standard ESI spectra

The use of standards mixtures of chemicals naturally found in biofluids have been reported to enhance detection in MS (Fitzgerald et al., 2012, Zamboni and Fisher, 2012). Hence, the preparation of artificial urine containing most of metabolite classes to improve MS detection of different urine components will be beneficial. Creatinine standard solution (22.6 µg/mL) was infused, initially, to optimise Q-ToF source parameters (see initial Q-ToF settings in section 2.2.2). The optimised settings were applied to construct calibration curves of artificial urine standards. FIE-MS mass spectra of individual artificial urine compounds were used to build an "artificial urine database" in access format (Microsoft access, Microsoft, USA). Figure 2-4 shows the interface of the artificial urine database, the database includes ESI+ and ESI- spectra, expected *m/z* values and actual *m/z* values detected for each compound with Q-ToF alongside its basic chemical properties. A softcopy of artificial urine database is attached in Appendix B. The MS information of each artificial urine compound was used to select a suitable concentration, which gave good signal in Q-ToF for the preparation of artificial urine as detailed in section 2.2.6.

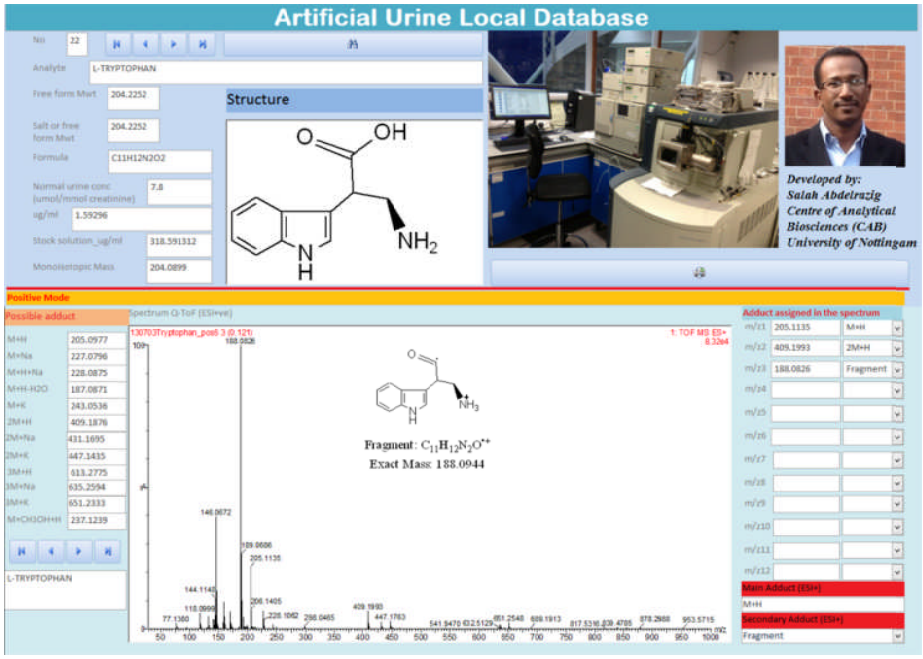


Figure 2-4 The main interface of artificial urine database.

2.3.2 Optimisation of FIE-MS using artificial urine

The effect of different cone and capillary voltages of the Q-ToF ESI source on total ion current (TIC) signal from direct infusion of artificial urine is shown in Figure 2-5. The highest intensities observed were at 40 and 65 V cone voltages for positive and negative modes, respectively. These new cone voltages were used to optimise the capillary voltage over a range of 0.9-3.0 kV with 0.1 kV increment intervals. The optimum voltages were 1.5 kV (ESI+) and 1.8 kV (ESI-). The optimum desolvation temperature was found to be 250˚C for both modes with desolvation gas flow rates of 250 and 400 L/h for the positive and negative modes, respectively. Different composition of mobile phases (water, methanol and acetonitrile), flow rates, injection volumes, mass ranges and scan rates were explored and optimised for FIE-MS using artificial urine (data not shown). The acquired data of artificial urine samples (n = 3) was used to evaluate and compare the signal quality before and after MS optimisation with artificial urine. Figure 2-6 presents a direct comparison of TIC and combined spectral counts of the infused artificial urine samples before and after MS source optimisation. TICs and summed peak counts of artificial urine improved by a factor of 3 in both modes. This indicates better sensitivity of the system with the new settings.

This is the first time that artificial urine has been for this type of optimisation. All reported in-house or commercial artificial urine mixtures have been intended for other purposes such as improving in-vitro growth of urinary pathogens (Brooks and Keevil, 1997) or to prevent formation of certain pathogens biofilms (Rane et al., 2013). Nevertheless, similar optimisation strategies of using chemical standards to enhance sensitivity of direct ESI-MS have been reported for Q-ToF analysis. Zamboni and Fisher, used a mixture of 80 standards to improve the MS coverage of *Escherichia coli* extract analysed with FIE-MS using 6550 Q-ToF (Zamboni and Fisher, 2012). They reported an increase in the number of detected peaks by a factor of 11 and one order of magnitude in ion counts. However, the improvement in sensitivity was evaluated against another study performed with an old version of the instrument, the 6520 Q-ToF (Fuhrer et al., 2011), making the comparison a function of the improved capability of the new instrument rather than demonstrating the improvement in sensitivity of the same instrument.

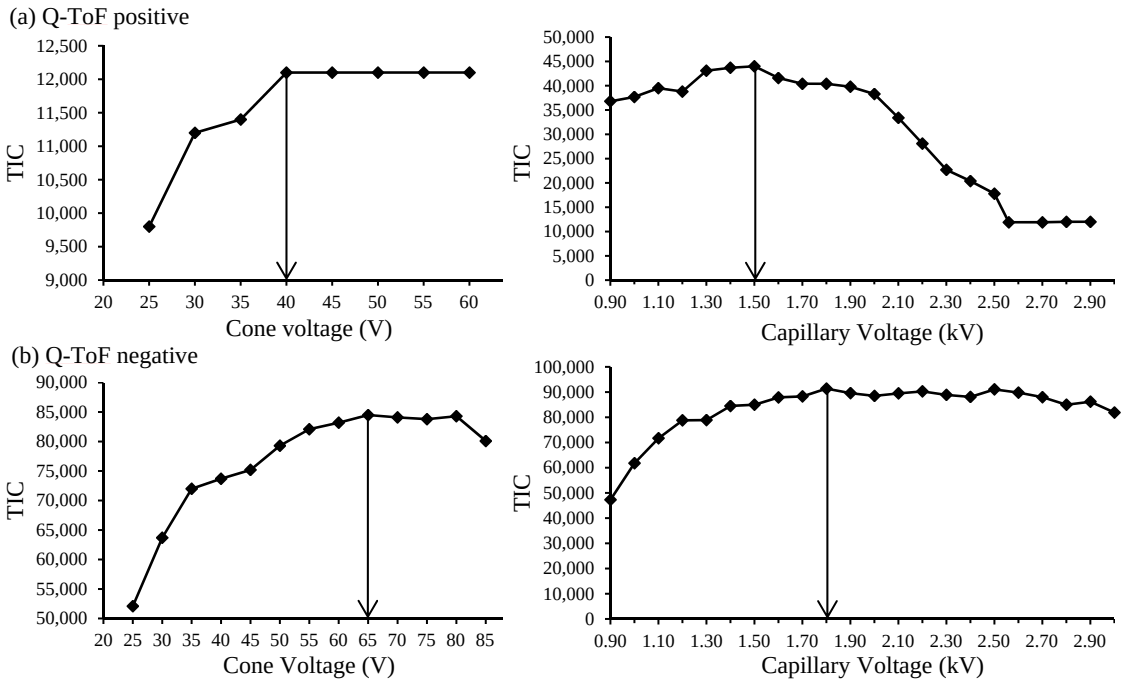


Figure 2-5 The effect of cone and capillary voltages on Q-ToF performance with artificial urine. TIC of infused artificial urine with Q-ToF cone and capillary voltages ramping between 25-90 V and 0.9-3.0 kV in both (a) positive and (b) negative ESI modes, respectively. The vertical arrows represent the selected (optimum) cone and capillary voltages, 40V (ESI+), 65V (ESI-) and 1.5 kV (ESI+), 1.8 kV (ESI-), respectively.

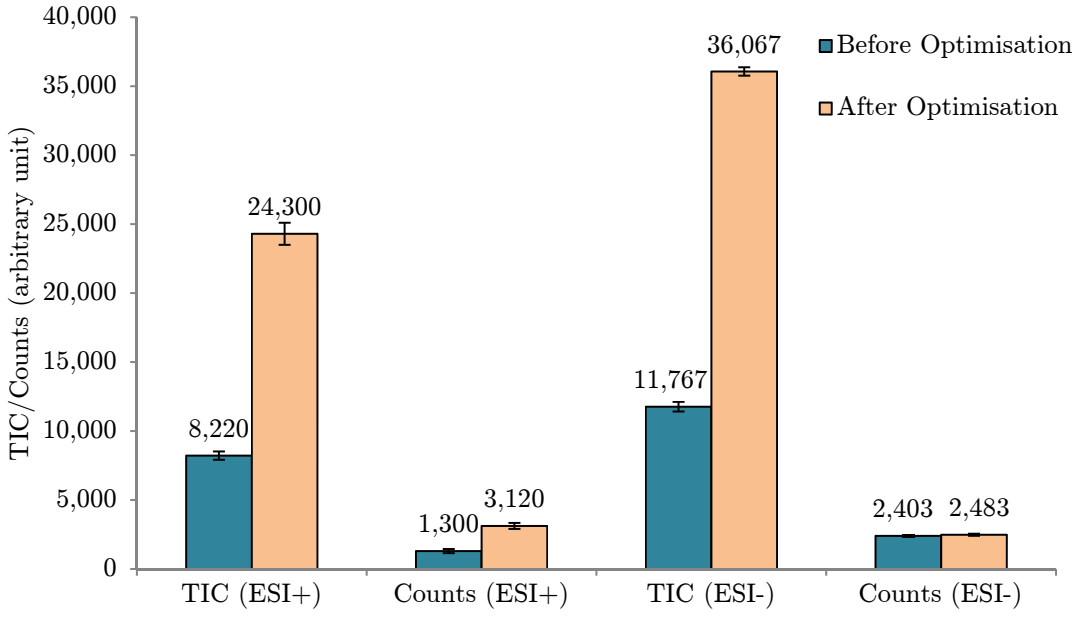


Figure 2-6 The performance of FIE-MS before and after optimisation with artificial urine. The histogram presents a direct comparison before and after optimisation (n = 3). The TIC and peak counts of combined mass spectra in positive and negative ESI modes using Q-ToF were used for the comparison.

2.3.3 Optimisation of orbital trap MS for artificial urine analysis

Artificial urine was then used to transfer and optimise the flow injection and ESI source parameters of the developed FIE-MS (Q-ToF) to the orbital trap mass spectrometer. To give an overview of the capability of the optimised FIE-MS (orbital trap-MS) on detection of different compounds found in artificial urine, typical summed mass spectra in both ESI+ and ESI- modes are presented in Figure 2-7. Some compounds such as histidine gave a good response in both ionisation modes, whereas others such as nitrogen containing compounds with basic functional groups, e.g. 3-methyl-L-histidine gave a response in ESI+ only. Compounds with carboxylic acid or phenolic groups such as lactic acid, 2-hydroxyisobutyric acid, hippuric acid and uridine ionised better in the negative mode. Therefore, it is essential to carry out analysis in both modes to maximise metabolite coverage of the MS methods. It is worth mentioning that some of the ions in the positive ion mode have a *m/z* difference of 21.9819 Da corresponding to the mass difference of Na and H. This indicates that compounds were forming sodium adducts as well as M+H, e.g. cytidine, and this is a well-known phenomenon in ESI+ MS when sodium ions are present in the injected solution (Nielsen et al., 2003). This is quite consistent with the common ions or adducts usually expected with nominal mass FIE-MS analysis (Table 2-3) (Overy et al., 2008, Beckmann et al., 2008). Some of the expected ions/adducts were also observed in artificial urine but at low intensities (highlighted in bold face in the table).

Table 2-3 Nominal mass ion adducts commonly expected in FIE–MS analysis.

Cation	Mass	Cation	Mass	Anion	Mass
$[M+H]^+$	**M+1**	**$[2M+H]^+$**	**2M+1**	**$[M-H]^-$**	**M - 1**
$[M+Li]^+$	**M+7**	$[2M+H+Na]^+$	M+20	**$[M-H2O-H]^-$**	**M - 19**
$[M+NH4]^+$	**M + 18**	**$[2M+Na]^+$**	**2M+23**	$[M+Na-2H]^-$	M + 21
$[M+Na]^+$	**M + 23**	$[2M+2H+3H2O]^{2+}$	2M+28	**$[M+Cl]^-$**	**M + 35**
$[M+H+CH3OH]^+$	**M + 33**	**$[2M+K]^+$**	**2M+39**	$[M+K-2H]^-$	M + 37
$[M+K]^+$	**M + 39**	$[2M+2Na-H]^+$	2M+45	**$[2M-H]^-$**	**2M - 1**
$[M+2Na-H]^+$	M + 45	$[M+2H]^{2+}$	M/2 + 1	$[2M+Na-2H]^-$	2M + 21
$[M+2K-H]^+$	M + 77	$[M+H+Na]^{2+}$	M/2 + 12	**$[2M+Cl]^-$**	**2M + 35**
$[M+H-HCOOH]^+$	M-45	$[M+H+K]^{2+}$	M/2 + 20	$[2M+K-2H]^-$	2M + 37
$[M+H-H2O]^+$	**M-17**	$[M+2Na]^{2+}$	M/2 + 23	$[M-3H]^{3-}$	M/3 - 1
$[M+H-NH3]^+$	**M-16**			$[M-2H]^{2-}$	M/2 - 1

Ions detected with flow injection electrospray–mass spectrometry (FIE–MS) of artificial urine analysis are highlighted in bold font. However, their occurrence and abundance are analyte specific and matrix dependent (Overy et al., 2008, Beckmann et al., 2008).

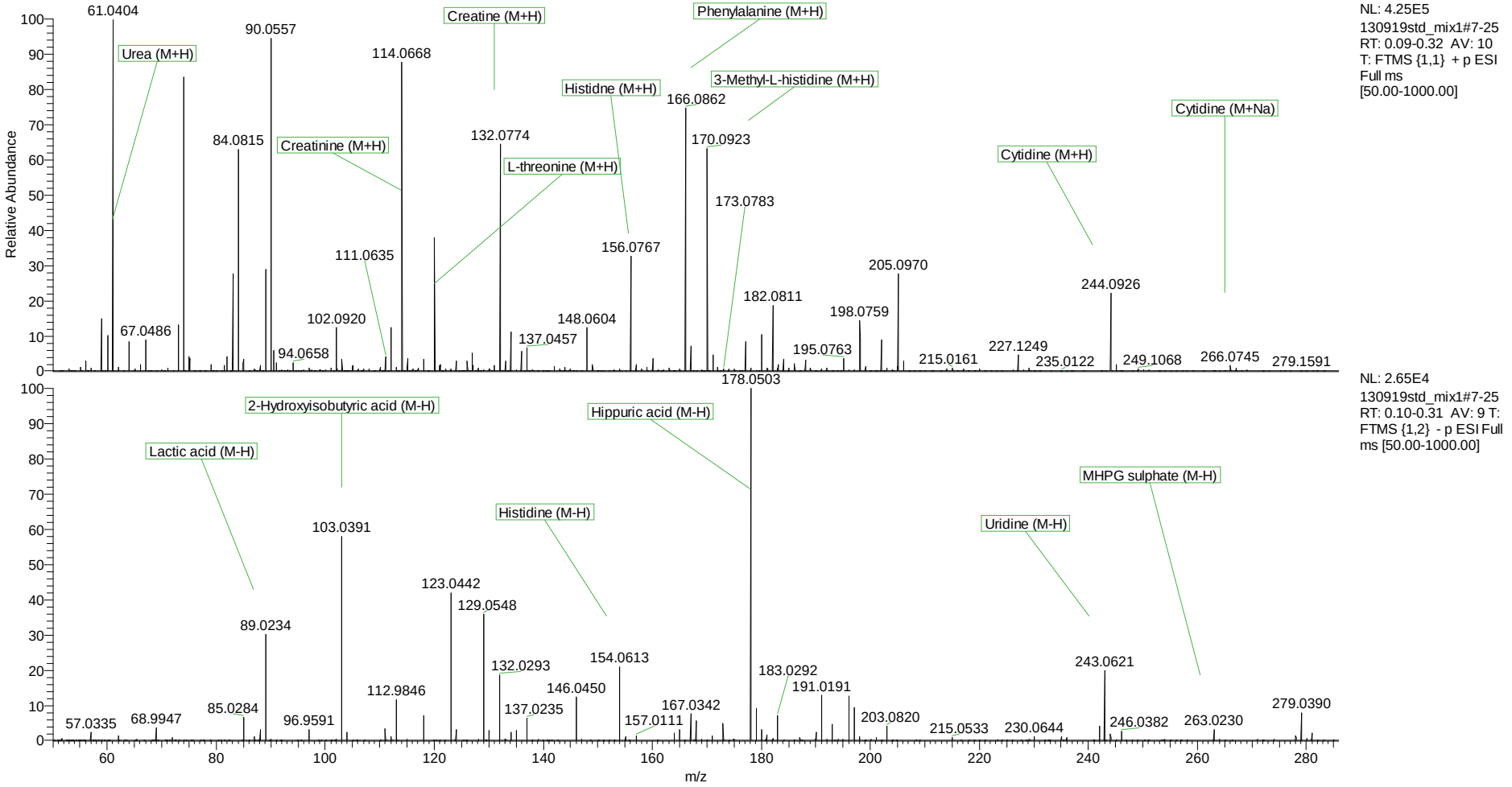


Figure 2-7 MS fingerprint of artificial urine analysed with FIE-MS. Typical summed spectra of artificial urine analysed by FIE-MS (orbital trap MS) in positive (upper) and (b) negative (lower) ion modes.

2.3.4 Optimisation of urine dilution to maximise direct ESI-MS response

A urine dilution method was considered in this study as it is compatible with a simple high-throughput sample preparation protocol, minimal sample volume usage and unbiased metabolites coverage could be attained with no sample loss (i.e. 100% recovery of metabolites).

2.3.4.1 Optimisation of sample dilution solvent for direct ESI-MS

Urine contains a small amount of protein. It is advantageous to remove any proteins to minimise matrix effect in the ESI and to reduce any interfering peaks resulting from multiply-charged protein ions. Methanol has the ability to cause proteins to precipitate but a high percentage of methanol in the sample can enhance ionisation of some ions and the signal of low abundance ions maybe overlaid leading to a decrease in the number of the detected peaks (Iavarone et al., 2000). Therefore, it is essential to find an optimum point at which methanol precipitates proteins in the sample without affecting total number of ions detected. A urine dilution-protocol (Fave et al., 2011) for FIE-MS was adapted as a starting point for the optimisation of urine sample preparation. In this protocol, a 10-fold dilution with 3.5:1 methanol:water was used for urine samples preparation for FIE-MS.

Optimisation of methanol volume for the dilution protocol was studied using orbital trap MS. Different loads of methanol were incorporated with a constant control urine volumes (50 μL). Hence, any increment in the total ion current (TIC) is either due to the enhanced ionisation of urine metabolite or increased background ions signal by methanol as illustrated in Figure 2-8. The lower graphs present TIC of urine:methanol ratio in the sample against its volume ratio. The increase in the signal ratio corresponds to the methanol effect on ESI. Little or no effect was observed above 1:1 urine:methanol and therefore, such ratio was used for the proposed method protocol. Although methanol has been extensively used as organic modifier for urine dilution or as flow injection solvent (Beckmann et al., 2008). However, it is proportion in the dilution solvent was quite different from one protocol to another, e.g. 1:1 methanol:water (Gonzalez-Dominguez et al., 2014), 2:1 methanol:water (Chen et al., 2015a) and 3.5:1 methanol:water (Fave et al., 2011, Lloyd et al., 2011b). Some studies had used water only for dilution without considering other organic modifiers (Want et al.,

2010, Fitzgerald et al., 2012). None of the reported studies considered the need to optimise methanol volume in the sample which makes the results obtained with this experiment quite valuable. It minimised the increase in the background signal due to unnecessary large volume of methanol and at the same time precipitates possibly existed proteins in the sample. This optimisation enhanced the signal of urine metabolite detected by FIE-MS and minimised the background signals, hence more peaks can be observed.

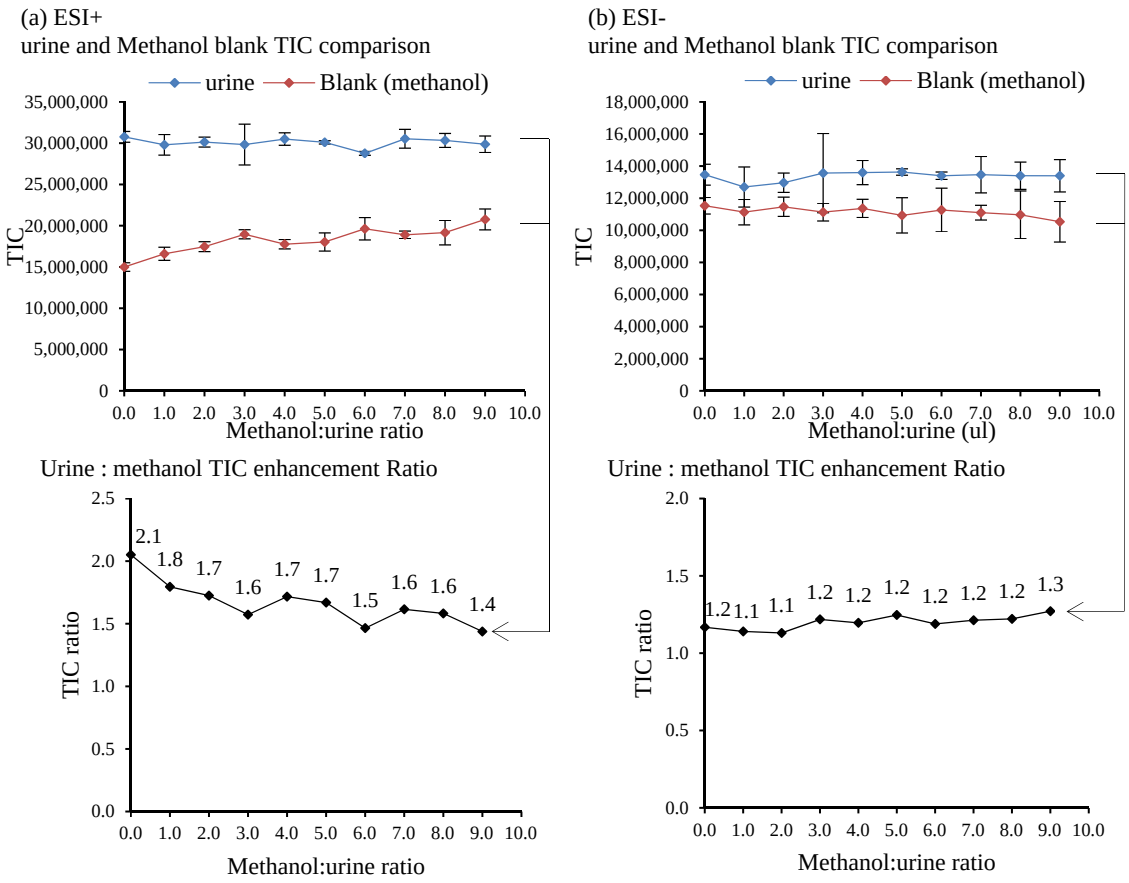


Figure 2-8 Optimisation of methanol load for urine dilution. (a) Positive and (b) negative TIC orbital trap MS of urine samples (blue line) and methanol (blank) (brown line) VS different loads of methanol in the samples analysed by FIE-MS. The lower graphs represent the ratio of TIC of urine:methanol VS methanol loads in the samples (the highest ratio represents the corresponding optimum methanol load in the sample).

2.3.4.2 Ion suppression effect of urine salts on ESI

Selection of a suitable dilution of urine sample for direct-ESI-MS was investigated using artificial and control urine. Different dilutions (2 to 10-fold) of human urine were reported in the literature for MS analysis (Gonzalez-Dominguez et al., 2014, Want et al., 2010, Fitzgerald et al., 2012). Among these, the most suitable dilution of urine for direct ESI-MS was assessed by investigating the effect of ion suppression of urine salts on ESI. Different loads of urine were added to a constant volume of artificial urine and analysed with FIE-MS using Q-ToF and orbital trap MS as described in the method section. Creatinine, *m/z* 114.0667 $[M+H]^+$ and uridine, *m/z* 243.0617 $[M-H]^-$ were selected as an example to display the ion suppression effect on ESI signal as a result of changing urine load in the sample (i.e. different dilution folds) as illustrated in Figure 2-9. Different levels of ion suppression on the signal of all analytes in artificial urine were observed in both modes; the degree of ion suppression increased with increasing urine load in the sample. King et al., illustrated that the main event that induces ion suppression is the efficiency of droplet formation and evaporation process during ESI (King et al., 2000). This theory explains the different degree of ion suppression observed with urine samples. The efficiency of ionisation decreases with the increase of non-volatile substances in the sample (urine salts in this case) and subsequently suppresses the signal of the analytes. Therefore, any suppression on the signals of artificial urine compounds is a function of the increased load of salts from urine.

Figure 2-10 gives an overview of % ion suppression of the detected artificial urine analytes in the samples in either positive or negative ion modes of orbital trap MS. The appropriate urine dilution that gave a reasonable signal and detection of most artificial urine analytes (89%) with minimum ion suppression on ESI was 10-fold dilution (0.1 urine ratio in the sample). These results were quite consistent with the literature as 10-fold dilution of urine was the most used dilution for FIE-MS (Fave et al., 2011, Lloyd et al., 2011a, Lloyd et al., 2011b). However, these methods did not evaluate the ion suppression effect on ESI. Van Hout et al., highlighted the benefits of increasing analyte/urine matrix ratio by extensive sample clean-up for improving MS sensitivity (van Hout et al., 2003). They demonstrated the importance of performing ion suppression validation by analysing urine samples spiked with different concentrations of clenbuterol (45-93 µg/L). The ion suppression on clenbuterol signal at the lowest concentration was 69% and decreased to 37% at the highest concentration, indicating that ion

suppression effect is more prominent at low level of analyte in the samples. Similarly, metabolites in urine were found to experience different levels of ion suppression based on their chemical structures. Fitzgerald et al., used a 10-fold dilution method for the measurement of drug of abuse in urine from different chemical classes (Fitzgerald et al., 2012). They reported that ion suppression less than 31% was observed with benzoylecgonine, amphetamine, methamphetamine, hydromorphone, and hydrocodone. Whereas a considerable ion suppression of urine salts was observed with morphine (81%) and codeine (89%) at the same concentration range. The obtained outcome from the above two articles highlighted the importance of using a mixture of diverse chemical classes in a wide concentration range to investigate the ion suppression effect in urine and hence, maximise metabolite coverage.

A recent DIMS study compared the use of SPE, LLE and simple dilution for decreasing the ion suppression effect of urine salts on ESI using Q-ToF (Gonzalez-Dominguez et al., 2014). In this article, assessment of ion suppression was carried out by measuring the ionic conductivity of urine extracts as it provides a mean to estimate the concentration of salts in the sample after treatment. They concluded that, 10-fold dilution of urine samples with 50% methanol/water gave a better result than LLE and was comparable to SPE for unbiased DIMS metabolomics fingerprinting. Moreover, they noticed that dilution higher than 10-folds provoked a notable decrease in the sensitivity of the instrument and therefore, should be avoided. Although, this supports the findings in this study, but their results showed a high degree of variability in the analysis with RSDs of 39% giving the developed protocol in this study, with 10% methanol/water dilution better reproducibility with maximum RSD of 23% across the detected artificial urine analytes.

In conclusion, 10-fold urine dilution gave good urine metabolites coverage (89% of artificial urine analytes were detected) with minimum ion suppression and better reproducibility compared to literature.

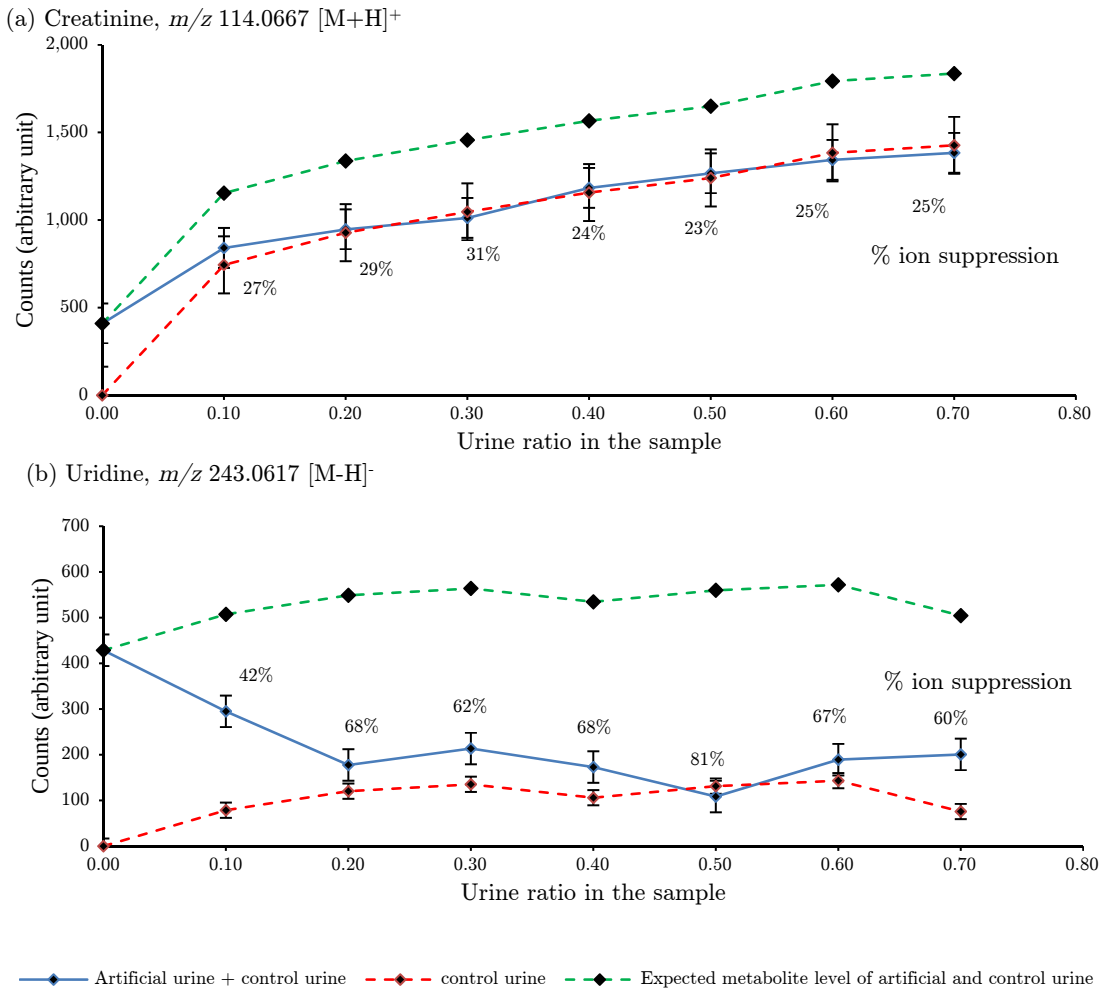


Figure 2-9 The effect of urine salts on ionisation of ESI-Q-ToF. (a) and (b) present creatinine (ESI+) and uridine (ESI-) levels measured by Q-ToF, respectively. The dotted red line represents the metabolite levels in the urine samples; the green dotted line gives the expected levels of metabolites in the samples of (urine spiked with artificial urine). The solid blue line represents the measured metabolite levels and their % ion suppression (n = 3).

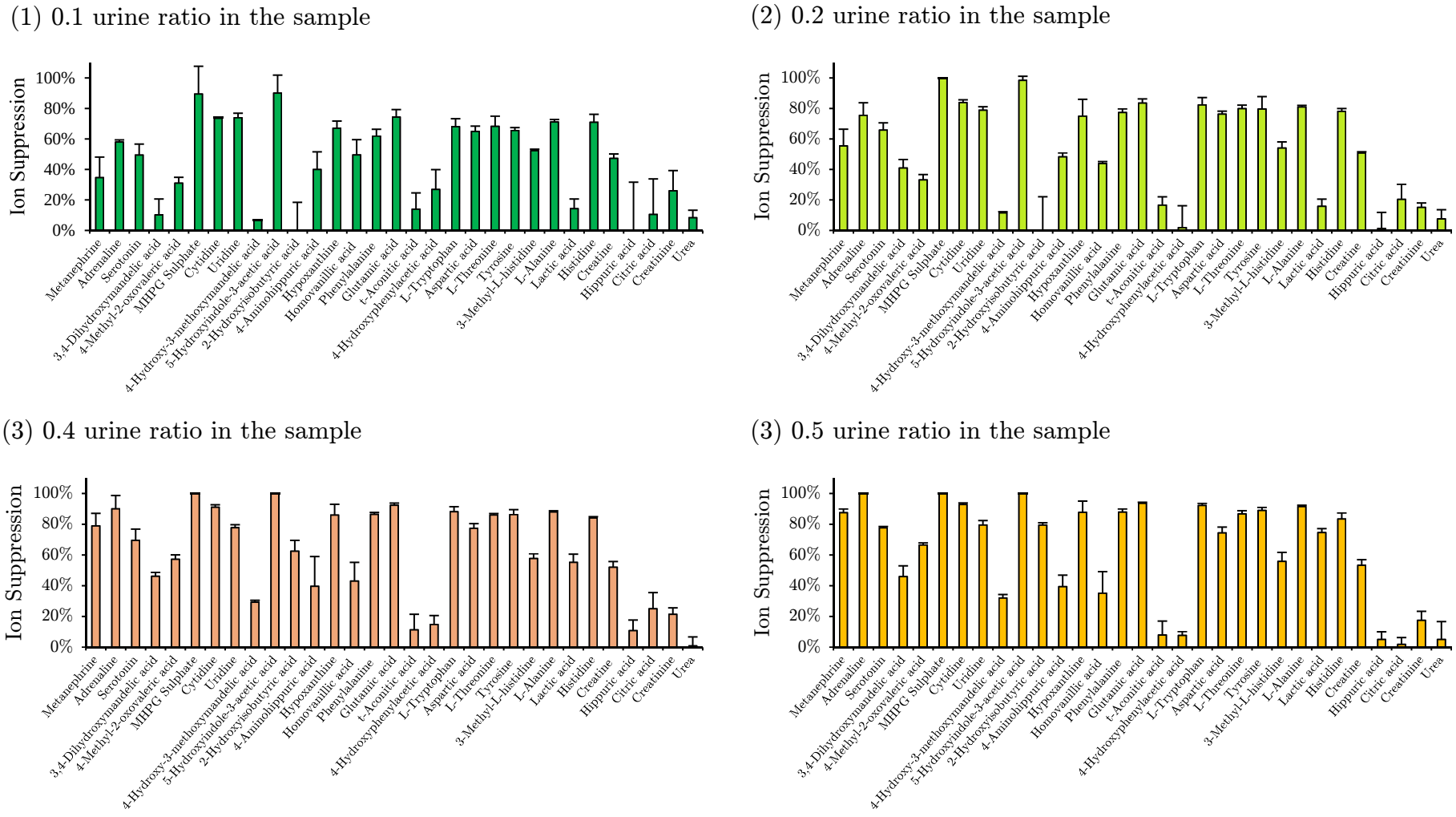


Figure 2-10 The effect of ion suppression of urine salts on the FIE-MS performance. Histograms (1-4) present an estimate of ions suppression (%) of artificial urine compounds with 0.1, 0.2, 0.4 and 0.5 control urine ratio in the sample analysed with orbital trap MS, respectively. Minimum ion suppression experienced by most of the ions corresponds with the 0.1 urine ratio in the sample (10-fold dilution) (n = 3).

2.3.5 Cross-platform comparison of direct ESI-MS performance with artificial urine

Artificial urine was used to check and compare the performance of the optimised sample dilution protocol, MS source and flow injection parameters on Q-ToF, orbital trap MS and chip-based infusion. Artificial urine was injected 6 times to evaluate the analytical stability of the systems across the analysis. TIC, reproducibility, S/N ratio, number of peaks detected and mass accuracy of the direct ESI-MS methods were compared and listed in Table 2-4. FIE-MS analysis with orbital trap MS detected more features with higher peak counts, accuracy and repeatability in comparison with Q-ToF analysis. This suggested different sensitivities between the two instruments, which is a logical reflection of the instrument capabilities. The high resolution (10,000 - 130,000 at FWMH) of Exactive Orbitrap (orbital trap MS) (Zubarev and Makarov, 2013) compared to a maximum of 17,500 FWMH resolution of Q-ToF Premier (Kayser and Warzecha, 2012) enables resolving more mass features from compounds with the same nominal masses, hence, more peaks were detected. Exactive improved ions focusing optics increase its sensitivity over Q-ToF. Glauser et al., compared the performance of Q-ToF (Synapt G2 Q-TOF, Waters, USA) with orbital trap MS (Exactive Plus Orbitrap, Thermo Fisher Scientific, USA) for untargeted metabolomics (Glauser et al., 2013). They reported that sensitivity of Exactive was higher for certain compounds compared to Q-ToF, while higher accuracy was observed with Q-ToF. However, the overall sensitivity, number of detected features and repeatability were comparable between the two instruments. These results contradicted with the findings in this experiment as better performance was observed with Exactive compared to Q-ToF Premier. However, the Q-ToF used in this study was not the most sensitive on the market and more recent Q-ToF insurtuments might provide comparable sensitivity to the orbital trap MS.

The best direct ESI-MS performance in terms of number of features detected was obtained with chip-based infusion. Nevertheless, higher variability was observed compared to FIE-MS with orbital trap MS and Q-ToF. The performance of Q-ToF compared to orbital trap MS, as explained above, is entirely a result of the very different instrumental design and capabilities. On the contrary, the varied performance between chip-based infusion and FIE-MS (orbital trap MS) is not due to mass analyser capabilities as they were both performed on the same

instrument. This might be a function of the improved ionisation with nanoESI-chip over standard ESI, appoint already has been discussed in general introduction chapter. Also, the concentration of the analytes in the sample was relatively higher in chip-based infusion compared to FIE-MS as the dilution effect of flow analysis solvent was opted out in chip-based infusion.

MS fingerprints of artificial urine obtained with the developed methods are presented in Figure 2-11. Different base peaks were observed, e.g. 3-methyl-L-histidine, *m/z* 170.0944 $[M+H]^+$ is the most abundant peak in FIE-MS (Q-ToF), while it was creatinine, *m/z* 114.0668 $[M+H]^+$ in both FIE-MS (orbital trap MS) and chip-based infusion. This indicates that direct ESI-MS ionisation strictly platform dependent. The detection of different ions with different sensitivity has been highlighted before in the literature. Gika et al., performed simultaneous urine analysis with two mass spectrometers, Q-ToF and Q-Trap, for untargeted metabolomics (Gika et al., 2010). They reported that urine sample analysed at the same time with the two instruments gave different MS profiles. This, as observed in this study, may be a result of the very different ESI configurations and design between the two instruments. Hence, for the purpose of comparison, evaluation and validation of the developed methods; orbital trap MS were used for further FIE-MS, chip-based infusion and LESA-MS analyses. In addition, the mass shift observed with the Q-ToF is quite high and not suitable for untargeted metabolomics.

Table 2-4 A direct comparison of the performance of direct ESI-MS using artificial urine

MS Measured parameters	FIE-MS				Chip-based infusion	
	Q-ToF	RSD%	Orbital trap MS	RSD%		RSD%
TIC (arbitrary unit)						
ESI+	62,400	11%	9,448,333	5%	130,666,667	2%
ESI-	48,567	10%	301,333	4%	11,153,333	16%
Sample:blank TIC ratio						
ESI+	4.46	-	1.83	-	2.23	-
ESI-	4.34	-	5.58	-	1.09	-
S/N ratio (TIC)						
ESI+	248	-	555	-	2646	-
ESI-	133	-	620	-	538	-
Number of peaks detected						
ESI+	1407	3%	9237	4%	78514	4%
ESI-	658	11%	1650	6%	29873	18%
mass accuracy (Δppm) (Leucine enkephalin)						
ESI+ (*m/z* 556.2771)	31	-	4	-	4	-
ESI- (*m/z* 554.2615)	16	-	3	-	3	-

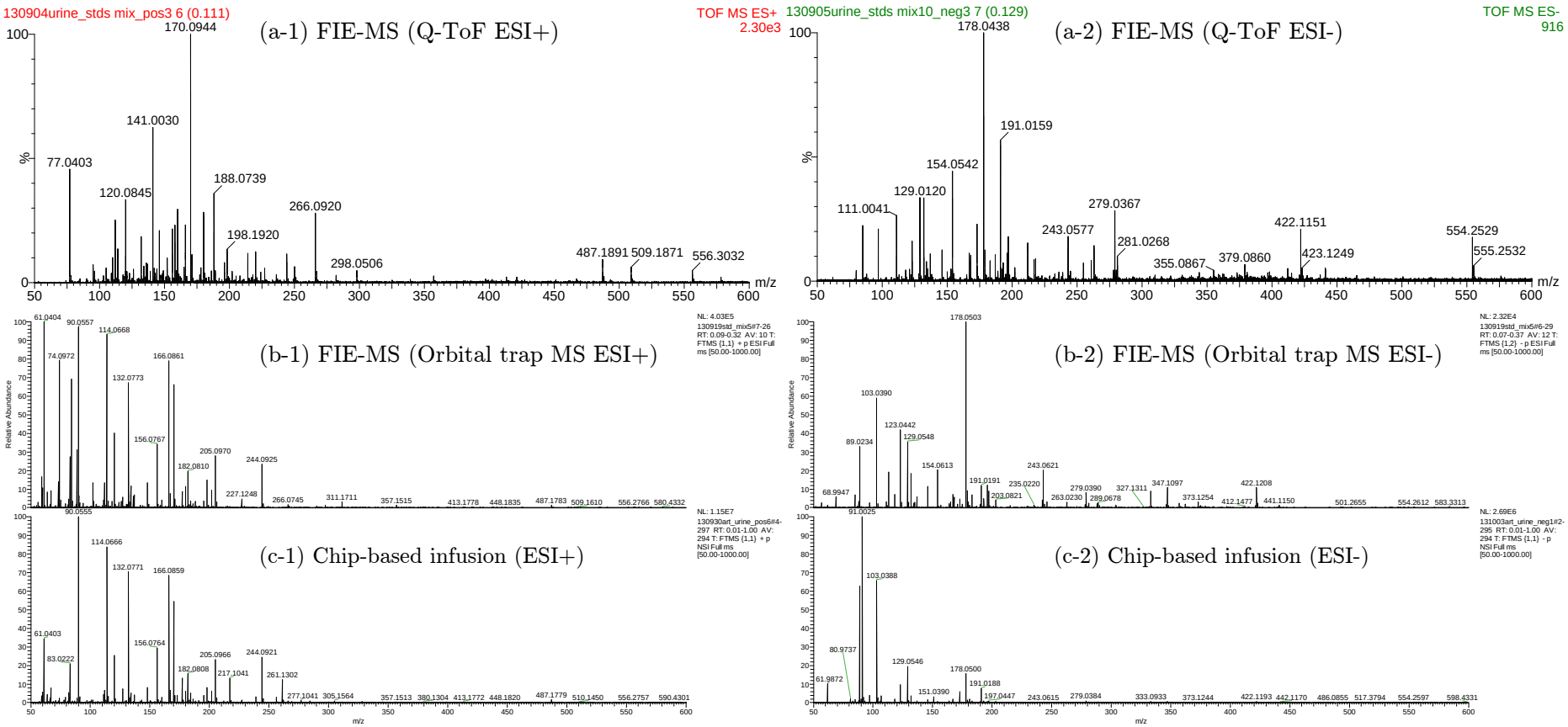


Figure 2-11 Cross platform evaluation of direct ESI-MS performance with artificial urine. Combined mass spectra of artificial urine analysed by direct ESI-MS using Q-ToF (FIE-MS) (a-1, a-2), Orbital trap MS (FIE-MS) (b-1, b-2) and chip-based infusion (c-1, c-2) in both positive and negative ESI modes, respectively.

2.3.6 Direct ESI-MS performance with urine samples

Control urine samples spiked with artificial urine [1:1] (n =6) were analysed with FIE-MS, chip-based infusion and LESA-MS. The analysis was carried out with orbital trap MS in both ESI modes to check the performance and coverage of the direct ESI-MS methods for urine analysis. In addition, a standard HILIC UHPLC-MS urine profiling protocol (Want et al., 2010) has been adapted for the analysis of samples as a reference method. Detailed tabulated results of the detected peaks were provided in Appendix C. Most of artificial urine compounds have been detected as $[M+H]^+$ or $[M-H]^-$ ions within 5.0 ppm mass accuracy. 89%, 86%, 86% and 94% of artificial urine compounds were detected by LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively. The RSDs of the peak areas/intensities were calculated across the samples; LC-MS, FIE-MS, chip-based infusion and LESA-MS had %RSD within 8%, 21%, 25% and 30%, respectively. The retention time (RT) shift RSD% of the LC-MS method was within 1%. The stability of the analytical platform is one of the major issues in obtaining valid metabolomics results. There is no general consensus on accepting a preferred analytical technique as suitable for biomarker discovery. However, for conventional bioanalysis, the Food and Drug Administration (FDA) considers a technique as suitable if analyte response in at least 5 determinations does not exceed 15% of the relative standard deviation (RSD%) except for those near the limit of detection where ≤ 20% can be accepted (FDA, 2013). While an upper limit of 30% is considered adequate for biomarker discovery (Gika et al., 2007, Want et al., 2010). The reproducibility of the FIE-MS and chip-based infusion analyses was relatively low compared to LC-MS analysis, but it is still within the acceptable limit for biomarker discovery (Gika et al., 2007, Want et al., 2010). Whereas LESA-MS is still counted as the least reproducible method compared to the rest. In LESA-MS, the sample was extracted every time with a solvent using a robotic arm which might introduce another level of variability to the analysis (i.e. technical variability). This variability is mainly due to the possible loss of sample during aspiration in the extraction process and subsequently concentrations of analytes in the sample may vary.

Adequate chromatographic separation (graphs not shown) was obtained with LC-MS, which is comparable to urinary profiles of previously published studies using HILIC column (Gika et al., 2008b, Cubbon et al., 2007, Chen et al., 2015b). The

spread out of analytes in the chromatography reduces ion suppression effect of urine salts and subsequently, increases the ionisation efficiency over FIE-MS (Buszewski and Noga, 2012). Hence, better analyte signals and coverage were attained with artificial urine analysis using LC-MS. In addition, the adequate precision obtained in chromatography had produced a stable flow to ESI source, therefore, low variability in the analytes signal was observed with RSDs less than 8%. Comparable coverage of the artificial urine analytes were obtained with FIE-MS and chip-based infusion to LC-MS, which may be a true reflection of the extensive optimisation of the method for urine metabolomics.

This is the first time that FEI-MS, chip-based infusion and LESA-MS were evaluated and compared with LC-MS for untargeted metabolomics using urine samples spiked with a mixture of standards. However, a published targeted analysis using chip-based infusion for urine analysis had reported a 40% coverage of the compounds under investigation compared to LC-MS method (Froesch et al., 2004). Yang et al., highlighted the possible sources of variability in chip-based infusion for the quantification of methylphenidate in urine (Yang et al., 2004). They reported that the variability in the analysis was directly proportional to the concentration of the analyte in urine with RSD% of 19.1% at 20 ng/mL and improved to 3.5% at 200 ng/mL. Considering these reported findings with the wide concentration range of artificial urine compounds, the obtained results with the developed direct ESI-MS methods demonstrates excellent coverage of urine metabolites with adequate accuracy and consistency in comparison to LC-MS.

2.3.7 Application of the optimised LC-MS and direct ESI-MS methods for urine metabolomics analysis: a human intervention study using green tea

Analytical performance of LC-MS and direct ESI-MS for urine analysis

Different approaches to ensure validity of metabolomics studies have been reported including the use of internal standards, QC samples and test mix (Gika et al., 2007, Gika et al., 2008a, Sangster et al., 2006, Viswanathan et al., 2007). Hence, it was decided to reference the direct ESI-MS results against a valid published protocol for urine profiling using UHPLC-MS analysis (Want et al., 2010). Urine samples from subjects after the ingestion of green tea (n = 6) and controls (n = 6) were analysed in a single analytical run using the developed LC-

MS, FIE-MS, chip-based infusion and LESA-MS. Pooled QC sample was inserted between samples during the analysis. Typical LC-MS base peak chromatograms of urine samples before and after ingestion of green tea are shown in Figure 2-12. Some of the base peaks are marked to give an intuitive display. Similar urinary profiles were obtained from subjects after ingestion of green tea and controls in both modes. Creatinine, *m/z* 114.0663 $[M+H]^+$, was the base peak in the positive mode, while it is *m/z* 187.0073 (p-cresol sulphate) in the negative mode. The separation profile in LC-MS between control urine samples and after consumption of green tea were quite similar in both modes. However the TIC for the control urine samples was slightly higher than the green tea samples. This indicates that the collected urine samples after the ingestion of green tea were more diluted than the control which could be as a result of the mild diuretic effect of caffeine in the green tea (Maughan and Griffin, 2003). To give an overview of the differences in the findings between LC-MS and the developed methods; summed mass spectra of urine samples after the consumption of green tea analysed in positive ion modes are presented in Figure 2-13. Similar spectral profiles were observed with the LC-MS and FIE-MS, while higher peak counts but different profiles were obtained with chip-based infusion and LESA-MS. The high peak counts in chip-based infusion and LESA-MS analyses, suggesting an improved sensitivity of nanoESI over standard ESI which was consistent with previously reported studies (Dethy et al., 2003, Wickremsinhe et al., 2005, Flangea et al., 2011). While detection of different MS profiles by chip-based infusion and LESA-MS compared to LC-MS and FIE-MS, may be a result of using different ESI source configurations (Gika et al., 2010, Trunzer et al., 2007), ESI in LC-MS and FIE-MS, and nanoESI in chip-based infusion and LESA-MS.

Validation of LC-MS and direct ESI-MS for urine analysis

Complete LC-MS, FIE-MS, chip-based infusion and LESA-MS urine metabolomics datasets were acquired for urine samples after the ingestion of green tea (n=6), the healthy control samples (n=6) and the pooled QC sample in in a single analytical run. A representative set of 57 urine metabolites in the QC sample was used to check the stability of the analysis and compare the performance of the developed methods against LC-MS (Appendix D). 44, 43, 45 and 41 of these metabolites were detected with LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively. This indicates adequate coverage of urine metabolites by the developed method compared to LC-MS. LESA-MS and chip-based infusion detected some metabolites which were not found by LC-MS,

indicating that enhanced sensitivity of some metabolites were obtained by nanoESI compared to ESI, consistent with Froesch et el. (Froesch et al., 2004). This finding suggested that the use of nanoESI in urine metabolomics studies may provide complementary information that could not be attained with LC-MS alone.

The maximum RSD% values of peaks counts/areas of the selected urine metabolites in the datasets of the pooled QC sample were 15.0%, 24%, 32% and 34% for LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively. LC-MS retention time shifts were less than 0.03 min with maximum RSDs of 0.65% and mass accuracy was within 5 ppm in both positive and negative ion modes for all analyses. The RSDs values obtained from LC-MS and FIE-MS were within the acceptable limits, less than 30%, stated for the bioanalysis and biomarker discovery (FDA, 2013, Gika et al., 2007, Want et al., 2010). Chip-based infusion and LESA-MS gave a slightly higher level of variability of metabolites in the QC sample compared to FIE-MS. But, 93% and 79% of the selected peaks in the QC samples gave RSD% less than 30% for chip-based infusion and LESA-MS, respectively, which indicates adequate repeatability of the methods for urine metabolomics (Want et al., 2010).

In addition, the quality of the data obtained from LC-MS analysis was also assessed using all peaks present in at least 80% of the pooled QC samples. The RSD% across the mean was less than 30% for 89% of these peaks, which was lower than the recommended threshold for metabolomics analysis (Want et al., 2010). For direct ESI-MS, principal component analysis (PCA) was used to check the quality of the obtained datasets as suggested by Beckmann et al. (Beckmann et al., 2008). PCA score plots of the sample sets (Figure 2-14) showed comparable clustering of the pooled QC sample towards the centre of the plots with all direct ESI-MS methods and LC-MS. Chip-based infusion and LESA-MS showed tight clustering of the pooled QC samples but with a shift towards the urine samples after the ingestion of green tea, yet it is considered adequate in term of analysis stability (Beckmann et al., 2008, Gika et al., 2007). In summary, these results validate and compare the analytical performance of the developed direct ESI-MS methods against LC-MS.

Evaluation of direct ESI-MS throughput for urine analysis

Advanced features and capabilities of the developed methods compared to LC-MS were summarised in Table 2-5. Direct ESI-MS analysis time per sample was less than 1.0 min compared to 15 min in LC-MS. This means that the methods are quite capable of performing large batch analysis of urine samples. For FIE-MS, Beckmann et al., estimated a possible throughput of 240 samples/20 h (576/48 h) and a total of at least two weeks for a full study to be accomplished (Beckmann et al., 2008). While for urine analysis, Gika et al., investigated the stability of urine samples in the autosampler at 4 °C for LC-MS analysis by continuous reanalysis of a pooled urine QC sample over 6 days. They reported that QC sample was stable for 48 h, after which significant changes in the sample were observed, suggesting that this was the maximum time urine samples should be kept for the analysis under such conditions (Gika et al., 2008c). Considering this time limitation for urine analysis, it is worth highlighting that in 48 h the developed direct ESI-MS methods can analyse up to 2,880 samples compared to 192 samples with LC-MS. Thus, a higher throughput in direct ESI-MS analysis with a quick and simple data analysis protocol was obtained. Furthermore, the improved detection sensitivity and adequate reproducibility with chip-based infusion and LESA-MS for urine analysis make them suitable and versatile choices for urine metabolomics. The ease of the analysis with no carry-over effect (Yang et al., 2004) and no sample treatment further the advantages of LESA-MS over conventional FIE-MS and chip-based infusion.

Table 2-5 Overall direct ESI-MS methods comparison with LC-MS

Description	LC-MS	FIE-MS	Chip-based infusion	LESA-MS
Sample preparation (time/sample)	Dilution-method (10-20min)	Dilution-method (15-20min)	Dilution-method (15-20min)	**Spotting (<1min)**
Analysis time/sample	15min	**0.5min**	**0.5min**	**0.5min**
Carryover effect	Possible requires proper wash cycle	Possible requires proper wash cycle or blank injection	**No**	**No**
Wash cycle/sample	2 min	0.5 min (blank injection)	**0 min**	**0min**
Dataset dimensionality	3D (RT, *m/z*, intensity)	**2D (*m/z*, intensity)**	**2D (*m/z*, intensity)**	**2D (*m/z*, intensity)**
Data pre-processing time/sample	~10 min	**< 0.5 min**	**< 0.5 min**	**< 0.5 min**
Overall time/sample	37-47 min	16.5-21.5 min	16-21 min	**<2 min**
Instrumentation	- LC system - HRMS	- LC system - HRMS	- Chip-based infusion device - HRMS	- Chip-based infusion device - HRMS
Number of peaks detected (urine)	4,128	748	**1,064**	**1,611**
Capital cost compared to LC-MS	Standard	Standard	**Lower (no need for LC System)**	**Lower (no need for LC System)**
Solvents and reagents	- Mobile Phases for chromatography - Sample solvent	- Mobile Phases for flow injection - Sample solvent	**- Mobile Phases not required** **- Sample solvent**	**- Mobile Phases not required** **- Extraction solvent**
other cost	-Chromatographic guard/ analytcial column -HPLC vials or 96-well plates	**-NO column required** **-HPLC vials or 96-well plates**	**-NanoESI chip** **-96-well plates**	**-NanoESI chip** **-Slides**
Analysis cost/sample compared to LC-MS	Standard	**Lower: no column, quick run (minimal solvent usage)**	higher (disposable nanoESI chip)	higher (disposable nanoESI chip)

Advanced desirable capabilities of each method compared to others were highlighted in bold.

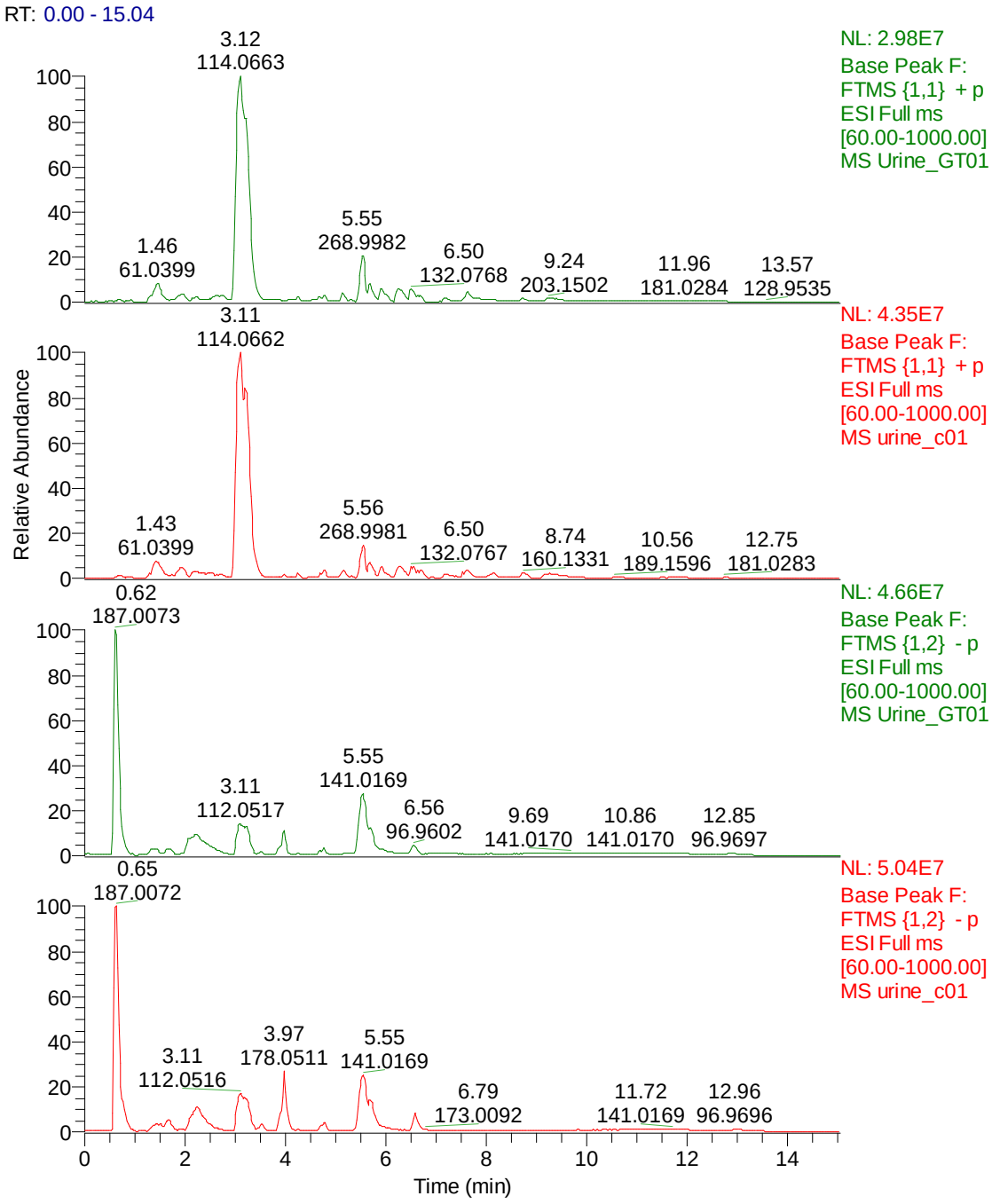


Figure 2-12 LC-MS base peak chromatograms (BPC) of urine samples before and after ingestion of green tea. From top to bottom, BPC of urine samples after green tea consumption (ESI+), control urine (ESI+), after green tea consumption (ESI-) and control urine (ESI-) analysed by LC-MS using BEH HILIC column.

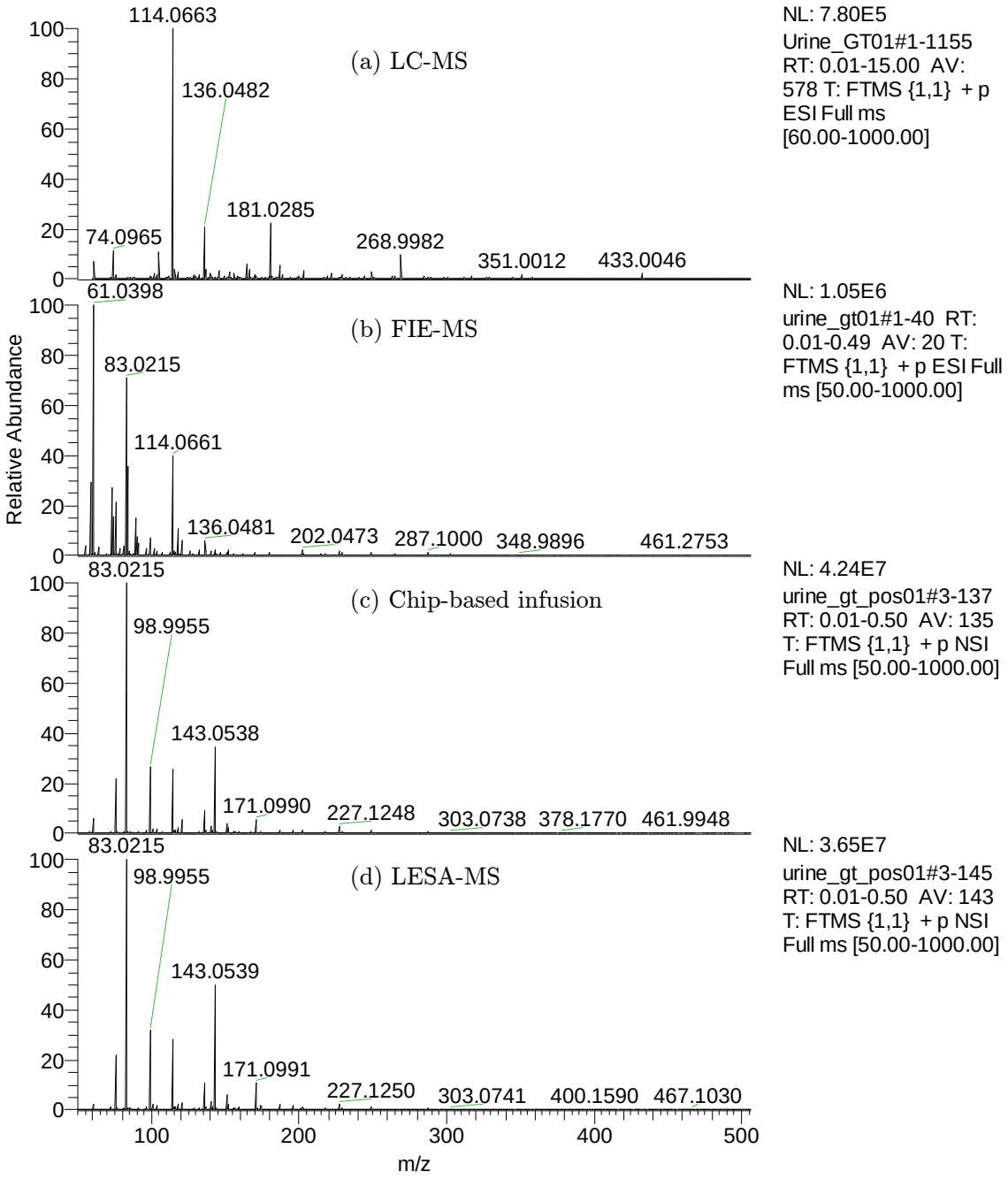


Figure 2-13 Typical metabolic fingerprints obtained from urine samples after the consumption of green tea analysed with LC-MS and the developed direct ESI-MS methods. Figure presents the combined mass spectra (50-500 Da) of urine samples analysed in the positive mode by (a) LC-MS, (b) FIE-MS, (c) chip-based infusion and (d) LESA-MS using Orbital trap MS.

2.3.8 Direct ESI-MS data analysis protocol for high-throughput urine metabolomics

Software Evaluation for high-throughput urine metabolomics

For a comprehensive modelling of the differences between groups in the study, datasets generated by LC-MS and direct ESI-MS methods were submitted for multivariate analysis. Three metabolomics software, MetaboAnalyst (Xia et al., 2012), SpecAlign (Wong et al., 2005) and Progenesis QI (Nonlinear-Dynamics, 2014), were evaluated to establish a high-throughput data analysis protocol (data not shown). In metaboAnalyst, the raw datasets was exported into common data format files (CDF), first before pre-processing. The data analysis was carried out on a web server with an upload capacity limit of 50MB, indicating that it may not be a suitable choice for large batch analysis. In SpecAlign raw data was extracted manually and imported into the software in comma-separated values (.csv) for peak alignment. The processing and conversion of the raw datasets may require time-consuming manual intervention steps. Therefore, the use of Metaboanalyst and SpecAlign might not be suitable for high-throughput data analysis. In Progenesis QI, raw data was imported directly without conversion, indicating that the throughput of the proposed direct ESI-MS methods can be increased with minimum intervention. It also provided a way of comparing LC-MS with the developed direct ESI-MS. Therefore, it was decided to use Progenesis QI for data pre-processing of LC-MS and direct ESI-MS.

Multivariate analysis

The raw datasets obtained from the analysis of urine samples after ingestion of green tea (n = 6) and controls (n = 6) were submitted for peak picking, peak alignment and normalisation using Progenesis QI. The datasets of LC-MS, FIE-MS, chip-based infusion and LESA-MS generated 4,128, 748, 1,064 and 1,611 variables, respectively (Table 2-6). The normalised datasets were exported for multivariate analysis using Simca P+14. Unsupervised principal component analysis (PCA) was used to give an unbiased overview for any possible trends and groupings within the sample datasets. PCA score plots showed complete separation between urine control samples and urine samples after ingestion of green tea with LC-MS (Figure 2-14). Comparable results were obtained with FIE-MS and chip-based infusion, whereas LESA-MS datasets produced a lesser degree of separation between groups. Subsequent supervised multivariate analysis

method using OPLS-DA showed clear separation and adequate clustering of green tea group from the control in all MS methods (Figure 2-15). The clustering trend in FIE-MS is quite similar to that in LC-MS. While in chip-based infusion and LESA-MS, the urine samples after the consumption of the green tea showed lesser clustering trends.

Validation of OPLS-DA models and metabolite selection

These OPLS-DA models were evaluated using cross-validation (leave-one out method) by monitoring the fitness of model (R^2Y) and predictive ability (Q^2) values (Cubbon et al., 2007). Further validation using prediction models based on randomly selected training/test sets was also performed. Comparison of the results of multivariate data analysis and the prediction of the models obtained from the developed methods against the LC-MS are summarised in Table 2-6. Although, a few number of peaks were extracted from the datasets of the developed direct ESI-MS methods but the generated OPLS-DA models gave comparable R^2Y and Q^2 values to LC-MS. Variable importance in the projection (VIP) of the OPLS-DA models is used in conjunction with loadings plots for the extraction of urine metabolites with significant differences between the two group of samples. A VIP score of a variable above 1.0 is considered important for the model (Yin et al., 2009). Student *t*-test (Miller and Miller, 2010) corrected with false discovery rate for multiple testing problem was performed with Progenesis QI across peaks intensities of urine samples and controls. Ions of significant differences (q-value < 0.05) have been considered as potential metabolites related to green tea.

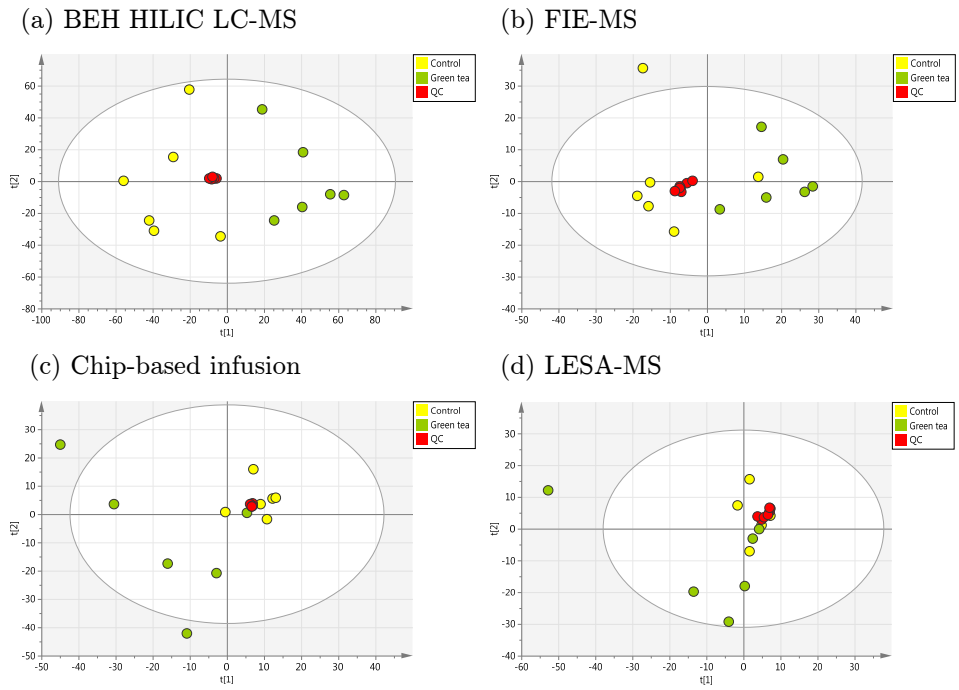


Figure 2-14 PCA score plots overview obtained from urine samples using LC-MS and the developed direct ESI-MS methods. a, b, c and d are PCA score plots of urine samples after consumption of green tea (green circles, n = 6), control samples (yellow circles, n = 6) and pooled QC samples (red circles, n = 6) analysed by LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively, in ESI+ and ESI- modes (combined).

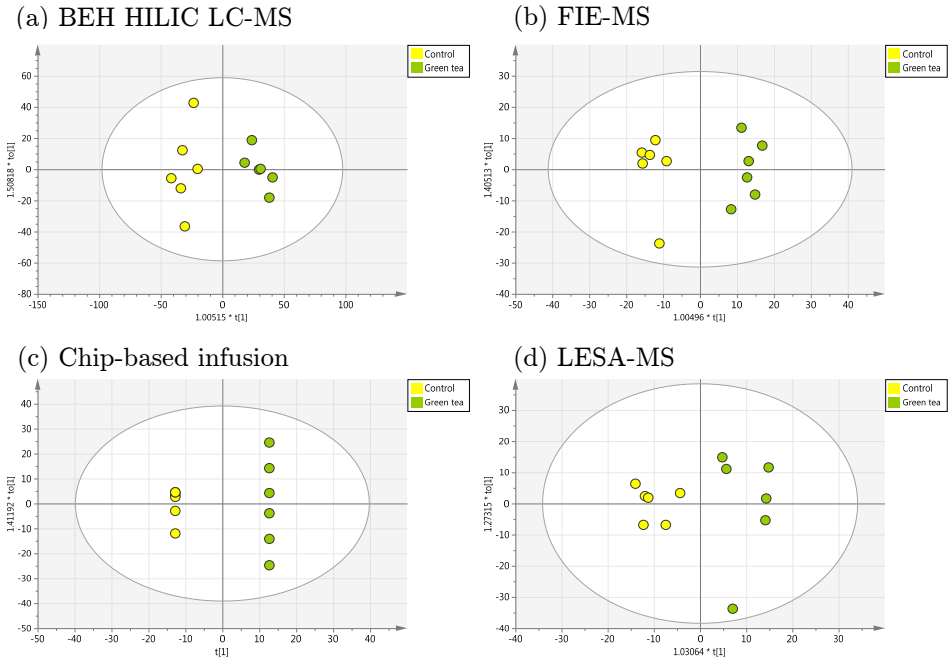


Figure 2-15 OPLS-DA score plots overview obtained from urine samples using LC-MS and the developed direct ESI-MS methods. a, b, c and d are OPLS-DA score plots of urine samples after consumption of green tea (green circles, n = 6) and control samples (yellow circles, n = 6) analysed by LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively, in ESI+ and ESI- modes (combined).

Table 2-6 Validation of OPLS-DA models generated from urine samples after consumption of green tea and control samples using LC-MS and the developed direct ESI-MS methods.

Description	LC-MS	FIE-MS	Chip-based infusion	LESA-MS
1. Peak detected				
ESI+	2,461	325	472	771
ESI-	1667	423	592	840
Total (ESI+ and ESI-)	4,128	748	1,064	1,611
2. Cross-validation				
R^2Y	0.945	0.963	0.995	0.875
Q^2	0.576	0.635	0.598	0.547
3. External validation: classification (training/test models)				
True positive (TP)	3	3	3	3
False positive (FP)	0	0	0	1
True negative (TN)	3	3	3	2
False negative (FN)	0	0	0	1
Sensitivity (%)[1]	100%	100%	100%	75%
Specificity (%)[2]	100%	100%	100%	67%
Accuracy (%)[3]	100%	100%	100%	71%

[1]Sensitivity is the true positive rate (TPR) calculated from the formula, TPR = TP/(TP+FP).
[2]Specificity is the true negative rate (TNR) calculated from the formula, TNR = TN/(TN+FN).
[3]Acuuracy is calculated from the formula, Accuracy (%) = (TP+TN)/(TP+FP+TN+FN).

2.3.9 Direct ESI-MS performance for biomarker discovery

Significantly altered metabolites in urine samples after the consumption of green tea compared to controls were used to evaluate the methods capability for biomarker discovery. A set of metabolites reported in urine after consumption of green tea as caffeine, paraxanthine, L-theanine and flavonoids has been used to generate a local database of green tea metabolites (Ridder et al., 2014, Spencer, 2003, van der Hooft et al., 2012). The exact mass was used to provide possible identification of the altered metabolites in urine samples after the consumption of green tea as detailed in the method section. The identified metabolites by each method were summarised in Table 2-7 together with their identity code from the Human Metabolome database (HMDB) (http://www.hmdb.ca/).

15 metabolites were found with LC-MS compared to 8, 5 and 6 with FIE-MS, chip-based infusion and LESA-MS, respectively. Most of the metabolites detected by FIE-MS were found with either chip-based infusion or LESA-MS, such as caffeine, L-theanine and hippuric acid. The majority of metabolites obtained by direct ESI-MS methods were also detected by LC-MS. Also, some metabolites detected by direct ESI-MS methods such as caffeine were not found in the standard LC-MS method. Surprisingly, LESA-MS detected one of the metabolites as M+K-2H ion instead of the prominent ion M+Na-2H. The detection of different ions and metabolites by the developed direct ESI-MS compared to LC-MS could be due to the very different electrospray source configurations and/or the improved sensitivity observed with nanoESI chip, a point already has been discussed in section 2.3.7. This may also be related to the effect of ion suppression of urine salts on ESI as higher load of urine salts was present in samples for LC-MS (3-fold dilution) compared to urine samples for direct ESI-MS (10-fold dilution). This finding, suggested that direct ESI-MS methods can be used as complementary methods with LC-MS to improve metabolite coverage for certain classes of metabolites. However, the infusion of all metabolites in the samples for ESI in direct ESI-MS methods decreases the detection ability of the methods compared to LC-MS. Thus, the LC-MS approach is quite adequate for comprehensive metabolomics study. While, direct ESI-MS methods are quite suitable for high-throughput applications. In direct ESI-MS, a high-throughput was achieved (i.e. they require ~6.7% of the LC-MS analysis time) and yet about 30-50% information of LC-MS was obtained. These results suggested that the developed direct ESI-MS methods can be used in pilot studies or as rapid diagnostic tools in clinical practice.

Table 2-7 Green tea metabolites detected in urine by LC-MS and direct ESI-MS methods

				LC-MS			FIE-MS		Chip-based infusion		LESA-MS	
Green tea metabolites	Metabolite MW (Da)	Formula	HMDB ID	RT (min)	Ions	Mass difference (mDa)	Ions	Mass difference (mDa)	Ions	Mass difference (mDa)	Ions	Mass difference (mDa)
(epi-)catechin-O-sulphate-O-methyl I	384.0515	$C_{16}H_{16}O_9S$	Not available	0.66	M-H	0.18						
3-Hydroxybenzoic acid	138.0317	$C_7H_6O_3$	HMDB02466	0.89	M-H	0.01						
3-Hydroxyphenylacetic acid	152.0473	$C_8H_8O_3$	HMDB00440	0.68	M+H	0.08						
4-hydroxy-5-(3',5'-dihydroxyphenyl)-valeric acid-O-methyl-O-glucuronide I	416.1319	$C_{18}H_{24}O_{11}$	Not available	2.04	M+H-H_2O	1.81						
4-Hydroxy-5-(phenyl)-valeric acid-O-sulphate	274.0511	$C_{11}H_{14}O_6S$	HMDB59981				M+H+Na	0.67				
5-(3',4',5'-trihydroxyphenyl)-γ-valerolactone-3'-O-glucuronide	400.1006	$C_{17}H_{20}O_{11}$	HMDB59984								M+K-2H	1.67
5-(3',4',5'-trihydroxyphenyl)-γ-valerolactone-4'-O-sulphate	304.0253	$C_{11}H_{12}O_8S$	HMDB59987	1.71	M-H	0.13						
5-(3',4',5'-trihydroxyphenyl)-γ-valerolactone-O-methyl-O-sulphate I	318.0409	$C_{12}H_{14}O_8S$	Not available								M+Na	0.33
5-(3',4'-dihydroxyphenyl)-γ-valerolactone-O-sulphate	288.0304	$C_{11}H_{12}O_7S$	HMDB29191	0.59	M+H	0.60	M+H+Na	0.75				
5-(dihydroxyphenyl)-γ-valerolactone-O-sulphate-O-methyl	302.0460	$C_{12}H_{14}O_7S$	HMDB60031	7.71	M-H_2O-H	0.34						
5-(hydroxyphenyl)-γ-valerolactone-O-sulphate	272.0355	$C_{11}H_{12}O_6S$	HMDB59993	1.14	M+ACN+Na	1.23						
Pyrocatechol-O-glucuronide	286.0689	$C_{12}H_{14}O_8$	Not available	2.46	M+ACN+Na	0.86			M+ACN+Na	1.17		
Pyrocatechol-O-sulphate	189.9936	$C_6H_6O_5S$	HMDB59724	0.56	M+2Na-H	0.59						
Caffeine	194.0804	$C_8H_{10}N_4O_2$	HMDB01847				M+H	0.17			M+2Na-H	1.20
Dihydroferulic acid 4-O-glucuronide	372.1056	$C_{16}H_{20}O_{10}$	HMDB41723	6.28	M+NH_4	0.22	M-H	0.07				
Epicatechin 3-O-(4-methylgallate)	456.1056	$C_{23}H_{20}O_{10}$	HMDB39328			0.00	M+3Na	1.00				
ferulic acid-O-sulphate	274.0147	$C_{10}H_{10}O_7S$	HMDB29200	1.29	M+2Na-H	0.86						
hippuric acid	179.0582	$C_9H_9NO_3$	HMDB00714	1.27	M+H	0.63	M+2Na-H	0.25	2M+Na	0.81	M+CH_3OH-H	1.34
Kaempferol 3-glucuronide	462.0798	$C_{21}H_{18}O_{12}$	HMDB29500						M-3H	0.07		
L-Theanine	174.1004	$C_7H_{14}N_2O_3$	HMDB34365	5.38	M+H	0.03	M+Cl	1.99			3M+H	1.66
Paraxanthine	180.0647	$C_7H_8N_4O_2$	HMDB01860	1.17	M+H	0.03						
Pyrogallol-2-O-glucuronide	302.0638	$C_{12}H_{14}O_9$	HMDB60017				M+Na	2.70				
vanillic acid	168.0423	$C_8H_8O_4$	HMDB00484	3.24	2M+ACN+Na	0.36			M-H	1.26	M+Na	0.21
Vanillin 4-sulphate	232.0042	$C_8H_8O_6S$	HMDB41789						M-H	0.29		

2.4 Conclusion

High-throughput direct ESI-MS approaches have been developed and validated for untargeted metabolomics of human urine. The results were compared with a conventional LC-MS method. The time consumed during sample preparation, especially with large sets of samples, adds another limit to the throughput of the LC-MS method. LESA-MS with no sample preparation step which involves only placing samples directly onto a slide surface prior analysis gives the method the advantage of being high-throughput by reducing pre-analysis time significantly. On the other hand, preparing samples using urine dilution methods for FIE-MS and chip-based infusion have the same time frame as LC-MS. Considering the run time for each sample in direct ESI-MS methods is only 1.0 min compared with several 10s of mins for LC-MS, direct ESI-MS undoubtedly provides a much higher analysis throughput. Chip-based infusion and LESA-MS use a disposable chip with nanoESI nozzles provide analysis without carryover effects. Although, this adds an extra cost to the analysis, but improves the throughput as well. There is no need for a wash cycle between samples as in conventional LC-MS and FIE-MS. In term of overall performance of the developed methods, FIE-MS provided the best throughput solution compared with chip-based infusion and LESA-MS. It had the lowest analysis cost/sample, same analysis time, lower variability and detects more metabolites from urine samples. However, LESA-MS requires no sample preparation step and only involves placing urine samples directly onto a slide surface, gives the method further credit of being very simple and easy to use in clinical practice. Although all direct ESI-MS methods provide high-throughput analysis compared to LC-MS, they lack the ability to resolve isobaric compounds and a fewer number of peaks were detected. In conclusion, direct ESI-MS methods employed a robust model for differentiation between inherent types of samples and it is recommended to be used as a fast analytical tool for clinical urine samples, while LC-MS is necessary when comprehensive biomarker screening is required.

CHAPTER THREE

Metabolic Signatures of Osteoarthritis (OA) in Human Urine

3. Metabolic Signatures of Osteoarthritis (OA) in Human Urine

3.1 Introduction

3.1.1 Osteoarthritis (OA)

Osteoarthritis (OA) is a chronic degenerative joint disorder which is considered as the most common type of arthritis and it is directly linked to ageing and obesity (Shane Anderson and Loeser, 2010). It is developed as a result of progressive erosion of articular cartilage of the joints and the continuous attempts of the body's repair processes. During this slow process, the cartilage roughens and becomes flimsy; the subchondral bone (the underlying bone) thickens and starts forming osteophytes (bone spurs) and the synovial cavity (the gap between bones in the joint) narrows. The synovium (the inner membrane of the joint capsule) may become thick and produce excess fluids, causing the affected joint to swell, while the ligaments and the joint capsule thicken and contract. In severe cases (Figure 3-1), bone friction and wearing away may happen due to complete loss of cartilage which is often associated with chronic pain and physical disability (Loeuille, 2012a). OA affects all type of joints but it is more prominent in the weight bearing joints such as the knees, hips and lower spine than in the shoulders, elbows, wrists and fingers joints (Hinman and Crossley, 2007). World Health Organisation (WHO) estimates that 9.6% of men and 18% of women above 60 years of age have symptomatic OA worldwide (WHO, 2015a). Due to an ongoing ageing population, OA is expected to be the fourth leading cause of disability among elderly by the year 2020. Radiographic surveys among European and United States populations aged above 45 years had showed that knee OA is the most common types of OA: 14.1% (men) and 22.8% (women) (Woolf and Pfleger, 2003). In the UK, according to Arthritis Research UK, 8.75

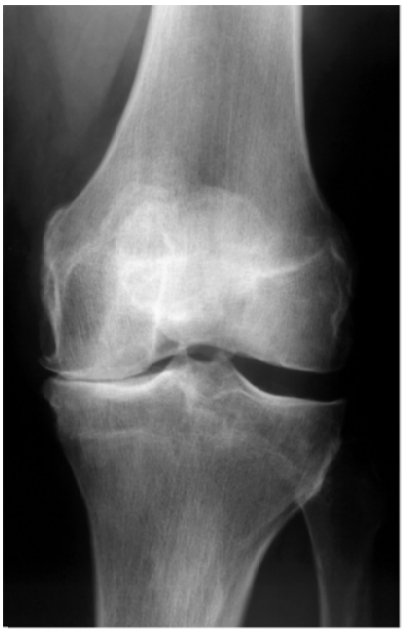


Figure 3-1 Right knee X-ray radiogram of a patient with severe OA (Bliddal and Christensen, 2006).

million people over age of 45 year have sought treatment for OA from which 4.71 million have suffered from OA of the knee and more than 77,000 surgeries were performed for knee replacements (Arthritis Research UK, 2013). Figure 3-2 presents an estimated number and percentage of people in the UK who have sought treatment of OA in the last 7 years according to age group and gender. It shows that around 33% of the total population aged 45 years and over have suffered from OA, from which 49% of women and 42% of men of those aged 75 years and over have been affected the most. Knee OA is the most common type of OA in England followed by the hip; 4.11 million people aged 45 and over have suffered from OA of the knee (18% of total population) whereas 2.46 million people have OA of the hip (11% of total population aged ≥ 45 years) (Arthritis Research UK, 2013).

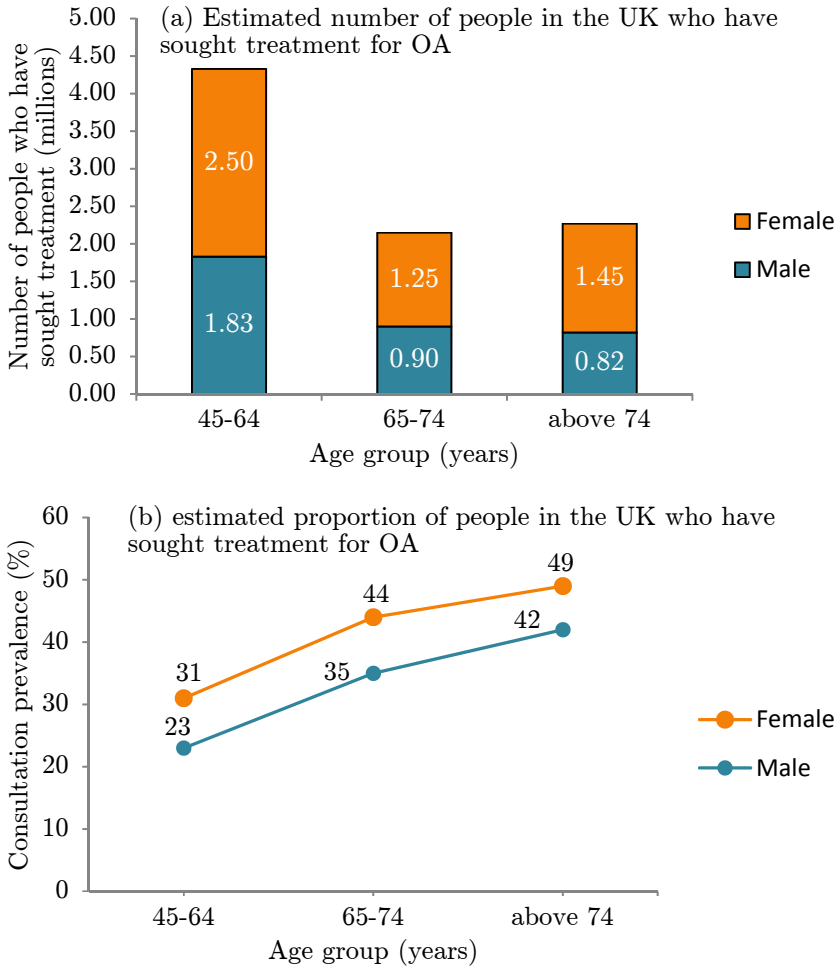


Figure 3-2 Osteoarthritis in the UK. The estimated number and percentage of people in the UK who have asked for OA treatment. The statistics based on 7 years consultation prevalence in UK general practice (Arthritis Research UK, 2013).

3.1.2 Osteoarthritis: aetiology and symptoms

The exact aetiology of OA is not yet fully understood. However, a range of factors such as age, obesity, gender, ethnicity, sport, physical activity, inherited susceptibility such as hereditary metabolic disorders, muscle weakness and rheumatoid arthritis have been shown to predispose to the development of OA, and it is often a combination of these factors that increases the risk of having OA (De Ceuninck et al., 2011, Li et al., 2010). In overweight populations, OA may develop as a result of cartilage damage due to repeated excess weight load on the joints or due to physical injury. In addition, the metabolic alterations associated with obesity along with the continuous production of inflammatory mediators by the adipose tissues are believed to be the major contributing factors in the progression of the OA (Yusuf et al., 2010). Recently, there are some new findings that link OA pathogenesis to synovitis; the local release of a wide variety of inflammatory mediators such as cytokines and chemokines associated with synovial inflammation has shown to have a direct metabolic effect on chondrocytes, the healthy cells that produce and maintain the cartilaginous matrix (Scanzello and Goldring, 2012).

OA is clinically characterised not only by the classical features of cartilage damage, subchondral sclerosis, bone cysts and formation of osteophytes, but also associated with low to moderate synovitis, ligamentous laxity, malalignment and meniscal degeneration (Cibere, 2006). Therefore, the clinical manifestations of OA are quite varied in terms of disease progression rate, joint involved, severity and onset of symptoms. In general, the major OA clinical symptoms are pain, stiffness and loss of function (De Ceuninck and Berenbaum, 2009). Pain is the primary symptom of OA and it is the main reason for patients seeking medical help. The mechanism of pain in OA is still not clear as cartilage is aneural (i.e. not innervated), however, bone marrow lesions observed with magnetic resonance imaging (MRI) has been associated with pain in knee OA (Felson et al., 2001). Other presumed causes of OA pain include synovial inflammation, elevated interosseous pressure due to vascular congestion, osteophytic periosteal elevation, muscle fatigue and joint contracture (NFMCPA, 2015).

3.1.3 Diagnosis and management of osteoarthritis

OA is currently diagnosed with clinical symptoms and radiography, in which X-ray is considered the “gold standard” radiographic method. The main X-ray

features for the OA diagnosis are: presence of osteophytes, space narrowing in the joint, formation of subchondral cysts and development of subchondral sclerosis. Although X-ray imaging is quite capable of detecting those features, recent MRI studies have demonstrated that X-ray as a plain radiography technique only provides a two dimensional image for the three dimensional structural features associated with cartilage and bone damage and therefore fails to detect hidden osteophytes due to overlapped bony features (Cibere, 2006). Magnetic resonance imaging (MRI) has the capability to provide a multiplanar tomographic imaging of cartilage, bone synovial effusion and ligaments, hence avoids the problem of undetected structural features on X-ray radiography. However, it is not commonly used due to its high capital cost and limited availability (Menashe et al., 2012, Loeuille, 2012b). Alternatively, arthroscopy provides a direct view of the affected joint features but its use is quite limited for OA diagnosis as it involves an invasive surgical procedure and therefore, cannot be routinely applied for OA clinical screening (Felson, 2010). Although all these techniques are quite capable of giving a direct insight on the impaired structural features associated with OA, they lack the ability to provide pathophysiological information at early stages of OA development. Therefore, non-invasive OA diagnostic tests are urgently needed for early prognosis and they would further the development of alternative treatment strategies and clinical evaluation of the disease.

To date, no disease-modifying agent has been approved yet to effectively treat OA and therefore, treatment strategies concentrate on improving the quality of life of OA patients through management of associated symptoms including pain control and improvement of lost function (De Ceuninck et al., 2011). Current guidelines for OA therapy as recommended by the European League against Rheumatism (EULAR) are defined in the following order: (1) non-pharmacological treatment, i.e. behavioral interventions such as exercise, life style changes and education, (2) pharmacological treatment with simple analgesic e.g. paracetamol, non-steroidal anti-inflammatory drugs (NSAIDs) or topical analgesics, (3) intra-articular corticosteroids or hyaluronic acid injections and (4) surgical interventions, i.e. arthroplasty such as joint replacement (Jordan et al., 2003). Most of these treatment strategies involve only long term symptomatic relief of OA and different side effects have been reported for the majority of patients (Goldring and Berenbaum, 2015). Current OA research aims are directed towards understanding the aetiologies associated with OA development and progression in terms of the common and distinct mechanisms at the molecular level that would be amendable for the development of novel disease-modifying

OA drugs (DMOAD) for OA treatment (Le Graverand-Gastineau, 2010). Current DMOAD research is sub-divided into inflammatory pathway targets, remodelling of subchondral bone and cartilage catabolism/anabolism (Yu and Hunter, 2015). Nonetheless, in order to find suitable drug targets, assess the therapeutic efficacy of new chemical entities and reliably monitor OA from early stages of progression, revealing the identity of biomarkers associated with OA pathophysiology are desperately needed. Chemical biomarkers such as proteins, peptides and metabolites that can be detected directly from different body fluids, e.g. blood, plasma, serum, synovial fluid and urine are excellent candidates to fulfill the above purposes.

3.1.4 The current value of biomarkers in OA research

The suitable candidates for biomarkers in OA are most likely to be biomolecules or fragments related to pathophysiological or structural changes of cartilage, bone or synovium during OA development and progression. They may be specific to certain type of joint tissue or common to all of them and may be involved in tissue degradation or synthesis pathways and can be measured in biological fluids such as synovial fluid, plasma or urine (Lotz et al., 2013). OA Biomarker Network established by National Institutes of Health (NIH) proposed a classification scheme to facilitate biomarker research in OA; it includes six categories which are highlighted in the acronym (BIPEDS): **B**urden of disease, **I**nvestigative, **P**rognostic, **E**fficacy of intervention, **D**iagnostic, and **S**afety (Bauer et al., 2006). The most important category for DMOAD development, as mentioned in the previous section, is prognostic biomarkers as it helps to predict progression and future incidence of OA. Current investigated biomarkers for the evaluation of OA are mainly linked to cartilage/bone collagen metabolism, cartilage aggrecan metabolism, non-collagenous proteins, inflammation and fibrosis. Comprehensive reviews of these biomarkers have been reported in the literature (Rousseau and Delmas, 2007, Mobasheri, 2012, Loeuille, 2012a, Lotz et al., 2013, Goldring and Berenbaum, 2015). Table 3-1 lists the most interesting biomolecules under investigation for the evaluation of OA; some of them are quite promising OA biomarkers such as serum cartilage oligomeric protein and urinary C-terminal telopeptide of collagen type II. However, none of them sufficiently discriminate between OA patients under different stages of progression, i.e. OA active (symptomatic) and OA inactive (asymptomatic) and healthy controls, predict prognosis of OA or perform consistently with OA patients so they can be

used as a surrogate biomarkers for further clinical trial investigations (Lotz et al., 2013).

Table 3-1 Selected biomarkers currently investigated for the evaluation of OA

Biomarker	Description/Type
1. Biomarkers related to collagen metabolism	
CTX-II	C-terminal telopeptide of collagen type II
α-CTX-II	Type II collagen α chains collagenase neoepitope
PIINP, PIIANP, PIIBNP, PIICP & CPII	Type II collagen propeptide
Pyridinoline & Glc-Gal-PYD	Cross-linking compound of collagen fibres
C2C	Type II collagen cleavage product
C2M	Collagen type II-specific neoepitope
CTX-I & α-CTX-I	C-terminal telopeptide of collagen type I
NTX-I	N-terminal telopeptide of collagen type I
PINP	Aminoterminal propeptide of collagen type I
C1 and C2	Types I & II collagen cleavage neoepitope
2. Biomarkers related to aggrecan metabolism	
Aggrecan neoepitopes, ARGS & FFGV	Core protein fragments
Chondroitin sulfate epitope 846	Linear heteropolysaccharide
3B3(–)	Monoclonal antibody 3B3(–)
Keratan sulphate	sulphated mucopolysaccharide
3. Biomarkers related to other non-collagenous proteins	
COMP & D-COMP	Cartilage oligomeric matrix proteins
Fib3, Fib3-1& Fib3-2	Peptides of fibulin
FSTL-1	Follistatin-like protein 1
Hyaluronic acid	Hyaluronan
MMP-1, MMP-3, MMP-9, MMP-13	Matrix metalloproteinases
TIMPs	Matrix metalloproteinases
YKL-40 (HC gp-39)	Human cartilage glycoprotein 39
sRAGE	Soluble receptor for advanced glycation end products
4. Biomarkers related to other processes	
hs-CRP, IL-1β, IL-6 & COX-5	Inflammatory biomarkers
Adiponectin, leptin & visfatin	Adipokines
sOB-Rb	Soluble leptin receptor (long form)
periostin	Cellular interactions in bone
DKKs & SOST	Wnt inhibitor
Uric acid	Heterocyclic purine derivative

COMP: cartilage oligomeric protein, D-COMP: deaminated COMP, COX-2: cyclo-oxygenase-2, hs-CRP: high sensitivity C reactive protein, DKK1: dickkopf-related protein 1, Fib: fibulin, Glc-Gal-PYD: glucosyl–galactosyl–pyridinoline, IL: interleukin, MMP: matrix metalloproteinase, sOB-Rb: soluble leptin receptor, PIIANP: N-propeptide IIA of type II collagen, PIIBNP: N-propeptide IIB of type II collagen, PIICP: C-propeptide of collagen type II, PIINP: N-propeptide II of type II collagen, SOST: sclerostin, TIMP: tissue inhibitor of matrix metalloproteinase. Table was reproduced from (Lotz et al., 2013).

3.1.5 The current role of urinary biomarkers in OA research

Urine is simple, readily available and non-invasive samples for metabolomics studies and it also provides an end metabolite pool of the body. Hence, it has the potential to increase our understanding of metabolic variation associated with OA development and progression. However, there are few studies in the literature related to the use of urine metabolomics in OA research (Li et al., 2010, Nepple et al., 2015, Lamers et al., 2003a, Lamers et al., 2005). Table 3-2 summaries a selected list of reported metabolites that have shown significant differences in urinary profiles of OA patients and animal models compared to healthy controls or OA patients at a different stage of disease progression. Lamers et al, performed urinary metabolomics experiment to study OA progression in guinea pigs. They reported disturbances in the level of malic acid, lactic acid, alanine and hypoxanthine which were found significantly contributed to the OA metabolic profiles (Lamers et al., 2003a). The same group further studied the urine of human from OA patients (n = 45) and healthy volunteers (n = 47) and they were able to discriminate between the two groups using NMR and multivariate analysis (Lamers et al., 2005). Li and co-workers performed GC-MS metabolomics approach on urine from subjects with and without knee OA. They reported a list of metabolites that showed significant differences between OA patients and healthy controls (Table 3-2). Elevated levels of aconitic acid, citric acid and isocitric acid were measured in the urine of OA patients compared to healthy controls, which might be linked to the enhanced activity level of tricarboxylic acid cycle (TCA) associated with the perturbed metabolism of chondrocytes and cartilage (Li et al., 2010). Although these urinary metabolites are quite promising for the prognosis of OA and could be logically linked to the disturbed biochemical pathways in OA, but still they lack the proper validation to be considered for clinical use. Therefore, more research is still needed to further identify, confirm, validate and characterise OA urinary biomarkers.

State-of-the-art high-resolution MS (HRMS) techniques have the potential to help not only in better understanding of OA progression but also in revealing a new sets of biomarkers and would further the development of novel OA therapeutic strategies. The application of direct ESI-MS metabolomics techniques such as FIE-MS is still missing in OA research. The versatility of using new direct ESI mass spectrometric techniques such as chip-based infusion and LESA-MS with their potential as a quick diagnostic tool for clinical use has not been

explored yet. Therefore, we intend to apply the developed LC-MS, FIE-MS, chip-based infusion and LESA-MS for urinary metabolomics study of OA.

Table 3-2 Selected list of urinary metabolites reported as potential biomarkers for OA

Metabolites	Type of the study	Analysis	Reference
4-hydroxy hippurate	knee OA vs. control	GC-MS	(Li et al., 2010)
4-methyl phenol	knee OA vs. control	GC-MS	(Li et al., 2010)
4-Methyleneproline	knee OA vs. control	GC-MS	(Li et al., 2010)
Acetoacetic acid	knee OA vs. control	GC-MS	(Li et al., 2010)
Aconitic acid	knee OA vs. control	GC-MS	(Li et al., 2010)
Alanine	OA in guinea pigs	NMR	(Lamers et al., 2003a)
Citric acid	knee OA vs. control	GC-MS	(Li et al., 2010)
Creatine	knee OA vs. control	NMR	(Lamers et al., 2005)
Creatinine	knee OA vs. control	NMR	(Lamers et al., 2005)
Glutamine	knee OA vs. control	GC-MS	(Li et al., 2010)
Glycerol	knee OA vs. control	NMR	(Lamers et al., 2005)
Glycine	knee OA vs. control	GC-MS	(Li et al., 2010)
Hippuric acid	knee OA vs. control	GC-MS	(Li et al., 2010)
Histamine	knee OA vs. control	GC-MS	(Li et al., 2010)
Histidine	knee OA vs. control	GC-MS	(Li et al., 2010)
Histidine	knee OA vs. control	NMR	(Lamers et al., 2005)
Homovanillic acid	knee OA vs. control	GC-MS	(Li et al., 2010)
Hydroxybutyrate	knee OA vs. control	NMR	(Lamers et al., 2005)
Hypoxanthine	OA in guinea pigs	NMR	(Lamers et al., 2003a)
Isocitric acid	knee OA vs. control	GC-MS	(Li et al., 2010)
Lactic acid	OA in guinea pigs	NMR	(Lamers et al., 2003a)
Malic acid	OA in guinea pigs	NMR	(Lamers et al., 2003a)
Methylhistidine	knee OA vs. control	NMR	(Lamers et al., 2005)
N-phenylacetyl glutamine	knee OA severity comparison	GC-MS	(Li et al., 2010)
Pyruvate	knee OA vs. control	NMR	(Lamers et al., 2005)
Tryptophan	knee OA severity comparison	GC-MS	(Li et al., 2010)

3.1.6 Aims and objectives of the chapter

- To investigate urinary profiles of OA patients in active and inactive stages of the disease using LC-MS and direct ESI-MS methods.
- To identify urinary biomarkers, which can differentiate between OA active patients and healthy controls.

3.2 Materials and methods

3.2.1 Reagents and chemicals

Mass spectrometry (MS) grade water was obtained from Fluka, Sigma-Aldrich, Switzerland. Formic acid, 0.1% formic acid in water and leucine enkephalin are MS grade and were purchased from Sigma-Aldrich, Germany. Methanol (LC-MS grade) was obtained from Fisher Scientific, UK. Acetonitrile (LC-MS grade) was supplied from VWR international, EU. Ammonium acetate (MS grade) was obtained from Fluka, Sigma-Aldrich, Netherland. Chemicals used for the preparation of artificial urine (a standard urine metabolite mixture) were either HPLC or MS grade; their description and sources were previously summarised in chapter 2, section 2.2.1, Table 2-2.

3.2.2 LC-MS and direct ESI-MS analyses

Orbital trap mass spectrometer (Exactive-Orbitrap, Thermo Fisher Scientific, USA) equipped with Accela UHPLC system (Thermo Fisher, USA) was used for LC-MS and FIE-MS analyses as detailed in section 2.2.3 and 2.2.4, respectively. LC-MS analysis was carried out on BEH HILIC UHPLC column (2.1 x 100 mm, 1.7 µm particle size, Waters, USA) with the optimised multistep gradient conditions detailed previously in chapter 2, section 2.2.3. Triversa NanoMate (Advion, USA) coupled to orbital trap MS was used for chip-based infusion and LESA-MS analyses as detailed in section 2.2.4. NanoESI spray conditions comprised a nitrogen gas head pressure of 0.7 psi and 1.5 kV (ESI+), 1.8 kV (ESI-) spray voltage controlled by Chipsoft Manager software (version 8.3.3, Advion BioSciences, USA). Urine dried spots positions on the cell culture slide were mapped for LESA-MS analysis using LESA-Points software (version 1.1, Advion BioSciences, USA). All analyses were carried out with the optimised orbital trap MS source parameters as described in section 2.2.2.

3.2.3 Preparation of artificial urine standard mixture

A mixture of 35 compounds was prepared in MS grade water in the concentration range of 2.7-67.8 µg/mL as "artificial urine", as described in chapter 2, section 2.2.6. Artificial urine was used to optimise flow injection parameters, MS source parameters, urine dilution, study the effect of ion suppression of urine salts on

ESI and to check the analytical stability of the MS system during metabolomics analysis.

3.2.4 Samples collection and storage

Seventy four OA patients and sixty eight healthy volunteers from Nottingham City Hospital, UK, were enrolled with consent in this study. The study was undertaken by Professor M Doherty (School of Medicine, University of Nottingham) and was passed by the appropriate University and Hospital Ethics Committees. Some patients were in the active stage of the disease, while the rest were inactive. OA patients were diagnosed and classified according to the American Rheumatism Association 1987 revised criteria (Arnett et al., 1988). The criteria used for OA diagnosis were based on expression of pain, limited motion in one or more of the joint and presence of pathological lesions of OA in at least one joint using conventional radiography. Then, patients were classified into active and inactive OA based on the clinical and laboratory data for the presence of active inflammatory process. OA active patients expressed high erythrocyte sedimentation rate (ESR), swelling and local hyperthermia in at least one joint, whereas OA inactive patients lacked such symptoms. All urine samples were collected in urinary collection vessels without the use of preservatives. Samples were then, aliquoted in triplicates of 2.0 mL into pre-labelled cryotubes and stored immediately in -80 ˚C freezer.

3.2.5 Preparation of urine samples for metabolomics analysis

Thawed urine samples of OA patient and healthy controls were prepared with the optimised urine dilution protocol for LC-MS and direct ESI-MS analyses as described in section 2.2.7. For metabolomics analysis, a pooled QC sample was prepared by mixing 20 μL aliquots taken from each urine sample and was treated the same as described for the samples in section 2.2.7.

3.2.6 Urine metabolomics analysis of OA patients and controls

Participants' urine samples were randomised and all urine samples including OA patients and healthy controls were analysed in a single run in both positive and negative ion modes with LC-MS, FIE-MS, chip-based infusion and LESA-MS. For LC-MS analysis, six injections of pooled QC sample were analysed at the beginning of the run to condition (i.e. stabilise) the column prior the analysis of

participants' clinical samples. Artificial urine triplicates were injected in the beginning and at the end of the run as a reference test mix to check the stability of instrument accurate mass performance. Pooled QC sample was injected once after each 10 samples for the purpose of monitoring the stability, robustness, repeatability and performance of the analytical platform. To monitor mass accuracy within each run a reference standard of leucine enkephalin, *m/z* 556.2771 (ESI+), 554.2615 (ESI-) was spiked in every sample to give final concentration of 0.2 µg/mL. Blank samples were injected after each sample in FIE-MS analysis to minimise the carryover effect, if any.

3.2.7 Multivariate analysis and metabolite identification

The raw datasets for LC-MS, FIE-MS, chip-based infusion and LESA-MS analyses were acquired using Xcalibur v2.1 software (Thermo Scientific, USA). The performance of the analytical methods were validated by monitoring a representative set of urine metabolites in pooled quality control (QC) samples for retention time (RT) shifts (LC-MS), mass accuracy, relative standard deviations (RSD%) of peak areas (LC-MS) and peak counts (direct ESI-MS). Xcalibur Quan Browser was used to integrate and extract peaks areas of (RT, *m/z*) metabolites pairs from the LC-MS raw data. While, Thermo ToxID v2.12.57 software (Thermo Scientific, USA) was used to extract intensities of *m/z* values of metabolites of interest from the direct ES-MS raw datasets. In addition, the quality of the datasets obtained was assessed using methodology proposed by Want et al., (Want et al., 2010) and Beckmann et al., (Beckmann et al., 2008) for LC-MS and direct ESI-MS analysis, respectively, as detailed in chapter 2.

For metabolomics data analysis, raw datasets from OA active patients, OA inactive patients and healthy controls groups were imported into Progenesis QI software (Waters, USA) for peak picking, peak alignments, adducts and isotopes deconvolution, normalisation and transformation as previously detailed in chapter 2, section 2.2.11. The normalised abundance of extracted metabolites peaks (RT, *m/z*) pairs (LC-MS) and *m/z* (direct ESI-MS) were imported into Simca P+14 (Umetrics AB, Sweden) for multivariate analysis (MVA). The data were mean-centred and scaled to unit variance (UV). The mean-centering technique subtracts the mean of the variables intensity and hence, shifts the data towards the mean. UV scaling procedure gives the weight of each variable by its standard deviation, which shrinks the weight of intense features and stretches the weight of smaller ones so that all features rest with equal weights (Eriksson et al.,

2006b). UV scaling was found to generate multivariate models with higher fitness of model (R^2Y) and predictive ability (Q^2) values (Eriksson et al., 2006b) compared to pareto scaling, therefore, it was decided to use such scaling for further analysis. Any variables associated with metabolites due to analgesic medication received by OA patients were excluded to eliminate the potential confounding effect of medication in the classification of groups under investigation. Principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) were used to investigate metabolic changes between different OA patients and healthy controls. To maximise the classification between OA groups and healthy controls, successive OPLS-DA models were generated comprising only two classes; datasets from either OA active patients (n = 22) or OA inactive patients (n = 52) and the healthy controls (n = 68) were used to build the OPLS-DA. Cross-validation using leave-one out method (1 out of 7) was used to evaluate the robustness of the models by monitoring the fitness of model (R^2Y) and predictive ability (Q^2) values. Permutation test was used to evaluate the validity of the predictive ability (Q^2 values) of the models.

3.2.7.1 Balancing sample size for OA classification

In order to match OA active patients (n = 22) and healthy controls (68) sample size difference, a multivariate approach was implemented to select a representative set (n = 22) from the healthy control samples using a combination of multivariate design, OPLS-DA models and shared and unique structures (SUS) plots. In multivariate design, a single class analysis of the whole set of samples is generated using PCA, then any strong outliers should be removed to achieve symmetrical distribution of the samples within the confidence intervals of the PCA score plot. Finally, a representative set of samples is assigned from the PCA by judicious selection of a well distributed group of samples throughout the score plot area (Wold et al., 2004) and then, SUS plot can be used to check the validity of the selection. The SUS-plot is used in multivariate analysis for multiple comparisons of classification models (e.g. OPLS-DA) based on a common reference class, e.g. control group (Wiklund et al., 2008). The dataset selected in accordance with a multivariate design will have, in general, a good spread of the latent variables as the original dataset, and this, in turn, will be beneficial for the subsequent multivariate analysis (Eriksson et al., 2006b).

Initially, healthy controls were sub-divided into 3 subsets using multivariate design based on PCA single class analysis (n = 22 or 23) as illustrated in Figure 3-3, then OPLS-DA models were subsequently generated from each subset against OA active patients sample set. SUS plot was used to monitor the similarity of the generated OPLS-DA models. This procedure was repeated in an iterative manner (10-11 times) until all healthy controls groups (i.e. the 3 subsets) generated adequately similar OPLS-DA models with OA active patients. One of the balanced OPLS-DA models was then selected for further analysis. Successive permutation test was used to evaluate the predictive ability (Q^2 values) of the model. External validation (prediction method) based on randomly selected training (50%) and test sets (50%) of samples with correct group classification in the OPLS-DA model were reported. The specificity and selectivity of the prediction models were tested using area under the ROC (receiver operating characteristic) curve (AUC).

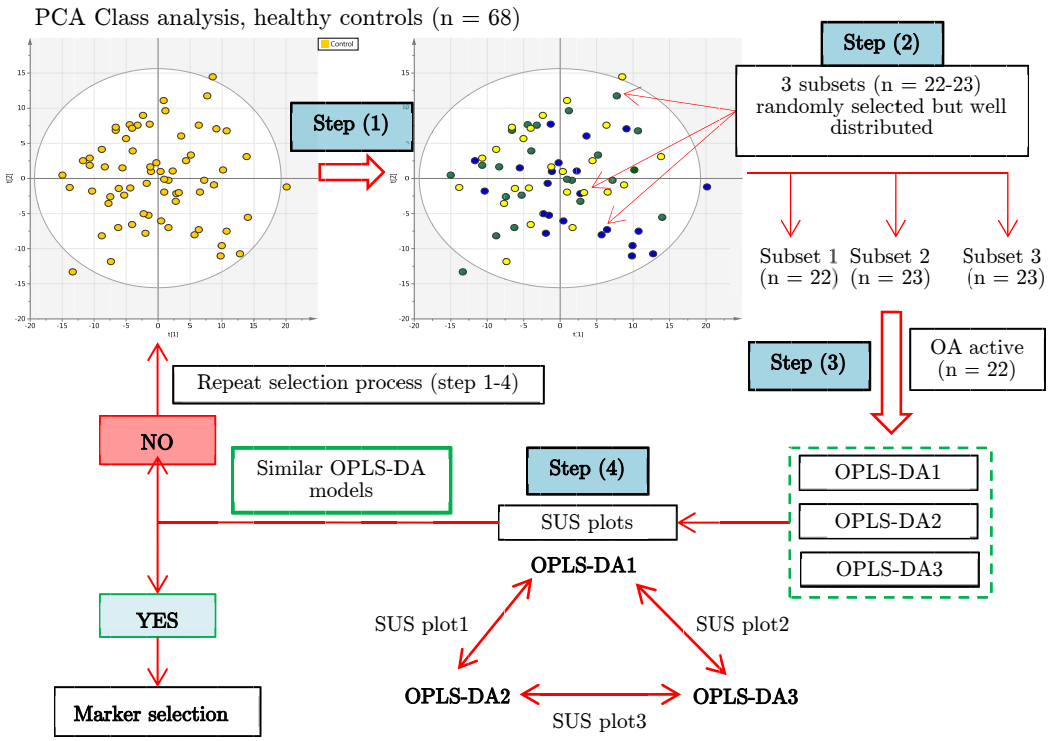


Figure 3-3 Workflow of balancing OA study groups for multivariate analysis. Healthy controls (n = 68) were sub-divided into 3 subsets: subset 1, 2 and 3, using multivariate design based on PCA single class analysis. 3 OPLS-DA models were generated from each dataset (i.e. subset 1, 2 or 3) against OA active patients dataset. SUS plot was used to monitor the similarity of the generated OPLS-DA models. SUS plots were generated for 2 models at a time (i.e. OPLS-DA 1 vs OPLS-DA 2, OPLS-DA 1 vs OPLS-DA 3 and OPLS-DA 2 vs OPLS-DA 3). This procedure was repeated until the selected subsets of the healthy controls generated adequately similar OPLS-DA models with OA active patients.

3.2.7.2 OA urinary metabolite identification

The ions responsible for the class separation in OPLS-DA models between OA active patients and healthy controls were selected by means of Variable Importance for the Projection (VIP) and variables loadings plots. To increase the confidence in the selection of those ions, univariate analysis using Student's *t*-test was computed in parallel to test the significant difference of the selected ions between the two groups. Prior to univariate analysis, ArcSinh transformation (Jones, 2008) was performed for the normalised data to restore normality using Progenesis QI. Accounting for multiple testing problems, p-values were adjusted using false discovery rate (FDR) technique, which was computed using Progenesis QI. Tentative identification of urine biomarkers was achieved using Metascope search engine of Progenesis QI to search for possible identification based on accurate mass measurements within 3 mDa error range from the Human Metabolome Database (HMDB) (Wishart et al., 2013).

3.3 Results

3.3.1 Demographic data assessment for metabolomics study

Participants' clinical and demographic data are listed in Table 3-3. The median age of cases was 68 years old and within the range of 50-91 and 52-88 for OA patients and healthy controls respectively, indicating adequate age matching for metabolomics study. Statistical Student *t*-test was computed to check whether the distribution of body mass index (BMI) is significantly different between the participants groups. There was no significant difference in the BMI between OA inactive patients and healthy controls, indicating that they were well matched for the study. However, there were significant differences between all OA patients, OA active patients and healthy controls were observed.

Table 3-3 Demographic data comparing OA patients to healthy participants

Description	OA patients	Healthy controls
1. Number of subjects		
a. All	74	68
Male	26	30
Female	48	38
b. OA active patients	22	-
Male	5	-
Female	17	-
c. OA inactive patients	52	-
Male	21	-
Female	31	-
2. Age:		
a. All		
Median	68	68
Range	50-91	52-88
b. OA active patients		
Median	69	-
Range	54-86	-
c. OA inactive patients		
Median	68	-
Range	50-91	-
3. BMI:		
a. All		
Mean (p-value = 0.05)	30.23	28.34
Median	29.27	27.62
Range	20.40-51.04	20.28-45.52
b. OA active patients vs healthy controls		
Mean (p-value = 0.02)	31.55	28.34
Median	29.41	27.62
Range	24.77-46.85	20.28-45.52
c. OA inactive patients vs healthy controls		
Mean (p-value = 0.19)	29.67	28.34
Median	28.99	27.62
Range	20.40-51.04	20.28-45.52

p-values were computed using a two tail Student *t*-test at 95% confidence limits assuming equal variance between groups.

3.3.2 Validation of LC-MS and direct ESI-MS analytical performance for osteoarthritis urine metabolomics

The quality of datasets generated from metabolomics analysis is of paramount importance as it has a direct reflection on the final biological outcome obtained from the study. Therefore, the use of pooled QC approach was used (Want et al., 2010, Gika et al., 2007) to assess the quality of the acquired datasets from OA patients and healthy controls. All samples were analysed in a single LC-MS, FIE-MS, chip-based infusion and LESA-MS with pooled QC samples being inserted after each 10 samples in the analysis. Six injections of pooled QC sample were analysed at the beginning of the run to stabilise the system prior analysis of clinical urine samples. A representative subset of 54 metabolites ions in the pooled QC urine sample including those detected in artificial urine (30 in ESI+ and 24 in ESI-), covering a range of retention times (RT) (LC-MS), *m/z* values and peak areas were selected to monitor the stability of the LC-MS and direct ESI-MS analyses. The analytical variability of these metabolites was measured across the datasets of the pooled QCs injected throughout the runs (n = 15), discarding the first 6 pooled QC injections used for equilibration. Appendix E reports the percentage of relative standard deviation (%RSD) of the RT, peak areas (LC-MS) and peak intensity (direct ESI-MS methods) for the selected metabolites; a summary of the results are presented in Table 3-4. The mean RSD% values of the selected metabolites peak areas/counts were 15% (range 7 - 18%), 15% (range 6 - 20%), 16% (range 8 - 25%) and 21% (range 10 - 34%) for LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively, and mean %RSD across LC-MS retention times was 0.52% (range 0.11 – 1.31%). Adequate mass accuracy for these metabolites was also observed with all methods with mass error within 5 ppm. Food and Drug Administration (FDA) has suggested an upper limit of 20% of the RSD% for the analyte response in at least 5 analytical replicates to consider the analytical technique as suitable for bioanalysis (FDA, 2013). Whereas a 30% of the RSD% is considered adequate for biomarker discovery (Gika et al., 2007). For LC-MS, retention times variability within 1% is considered acceptable for metabolomics analysis (Want et al., 2010). The observed results from LC-MS, FIE-MS and chip-based infusion were within these acceptable limits. While in LESA-MS a slightly higher level of variability in the metabolites response was observed. This might be due to the fact that in LESA-MS analysis another level of variability has been introduced as a result of sampling different urine spots (i.e. technical variability), which increases the

overall variability (Kertesz and Van Berkel, 2010), a point already has been discussed in chapter 2, section 2.3.6. In addition, the quality of the datasets obtained with LC-MS was assessed using the pooled QC sample. The RSD% across the mean peak areas of peaks present in at least 80% of the pooled QC samples was less than 30% for 71% of these peaks, which was lower than the recommended threshold for metabolomics analysis (Want et al., 2010). These results indicate satisfactory stability and validate the LC-MS and direct ESI-MS analytical performance for urine metabolomics.

Table 3-4 Validation summary of LC-MS and direct ESI-MS urine analyses for metabolomics profiling of OA patients and healthy controls using selected urine metabolites peaks from pooled QC sample.

Description	LC-MS	FIE-MS	Chip-based infusion	LESA-MS
1. Peak detection				
Total number of peaks	54	54	54	54
Number of peaks detected	54	42	39	40
% of peaks detected	100%	78%	72%	74%
Number of artificial urine metabolites detected	17	13	13	13
2. Retention time (RT) variability				
RT RSD% (mean)	0.52%	-	-	-
RT RSD% (range)	0.11 - 1.31%	-	-	-
3. Peak intensity variability				
Peak area/count RSD% (mean)	15%	15%	16%	21%
Peak area/count RSD% (range)	7 - 18%	6 - 20%	8 - 25%	10 - 34%

Retention time and peak area/count statistics of the selected peaks were calculated from the analysis of the pooled QC urine sample (n = 15).

3.3.3 OA metabolomics analysis of urine samples using LC-MS and developed direct ESI-MS methods

Visual examination of LC-MS base peak chromatograms (BPC) of urine samples showed differences between OA active patients, OA inactive patients and healthy controls in both positive (Figure 3-4, left side) and negative (Figure 3-4, right side) ion modes. For example, in Figure 3-4, the marked regions (A) and (B) show an increased concentration level of the metabolite, *m/z* 232.0274 (unknown) and *m/z* 152.0706 (unknown) in OA active patients compared to OA inactive patients and healthy controls. While the concentration of creatinine, *m/z* 112.0515 (ESI-, marked region C) was slightly higher in healthy controls compared to OA patients. Combined mass spectra generated from the analysis of OA active patients, OA inactive patients and healthy controls urine samples

analysed by LC-MS, FIE-MS, chip-based infusion and LESA-MS are presented in Figure 3-5, Figure 3-6, Figure 3-7 and Figure 3-8, respectively. Visual examination of MS spectra showed metabolic profile differences between the OA active patients, OA inactive patients and healthy controls urine samples analysed by all methods. For instance in LC-MS, the metabolite ion, *m/z* 181.0286 (ESI+) and, *m/z* 286.2642 (ESI-) showed higher peak counts in the OA active patients compared to OA inactive patients and healthy controls. Taking all molecular features into account, most of the detected ions were concentrated in the lower mass range (*m/z* 60-300) with all methods in the positive mode, while relatively higher mass ions were detected in the negative modes. Creatinine, *m/z* 114.0667 (ESI+) was found to be the most abundant ion in the healthy control urine sample analysed by all methods which indicates comparable results of the developed methods with LC-MS.

For a comprehensive modelling of the differences between groups in the study, datasets generated by LC-MS and direct ESI-MS methods were submitted for multivariate analysis. The raw datasets obtained from the analysis of OA active patients (n = 22), OA inactive patients (n = 52) and healthy controls (n = 68) in a single analytical run with pooled QCs (n = 15) interspaced in the analysis were submitted for peak picking, peak alignment and normalisation using Progenesis QI (Nonlinear-Dynamics, 2014). The datasets of LC-MS, FIE-MS, chip-based infusion and LESA-MS generated 7405, 484, 576 and 743 variables, respectively (Table 3-5). The normalised datasets were exported to Simca P+14 for multivariate analysis. Unsupervised principal component analysis (PCA) was used to give an unbiased overview for any possible trends and groupings within the samples datasets. PCA score plots showed no separation or clustering trends between the OA study groups with all MS methods (Figure 3-9). Nonetheless, the analytical runs demonstrated adequate stability as indicated by sufficient clustering of the pooled QC samples towards the centre of the PCA score plots with all methods (Beckmann et al., 2008, Gika et al., 2007). Subsequent supervised orthogonal partial least square-discriminant analysis (OPLS-DA) models were constructed to find differences and discrimination between OA patients and healthy controls. Visual evidence of separation and clustering between the OA active patients, OA inactive patients and the healthy controls were observed with all MS methods as shown in the score plots presented in Figure 3-10.

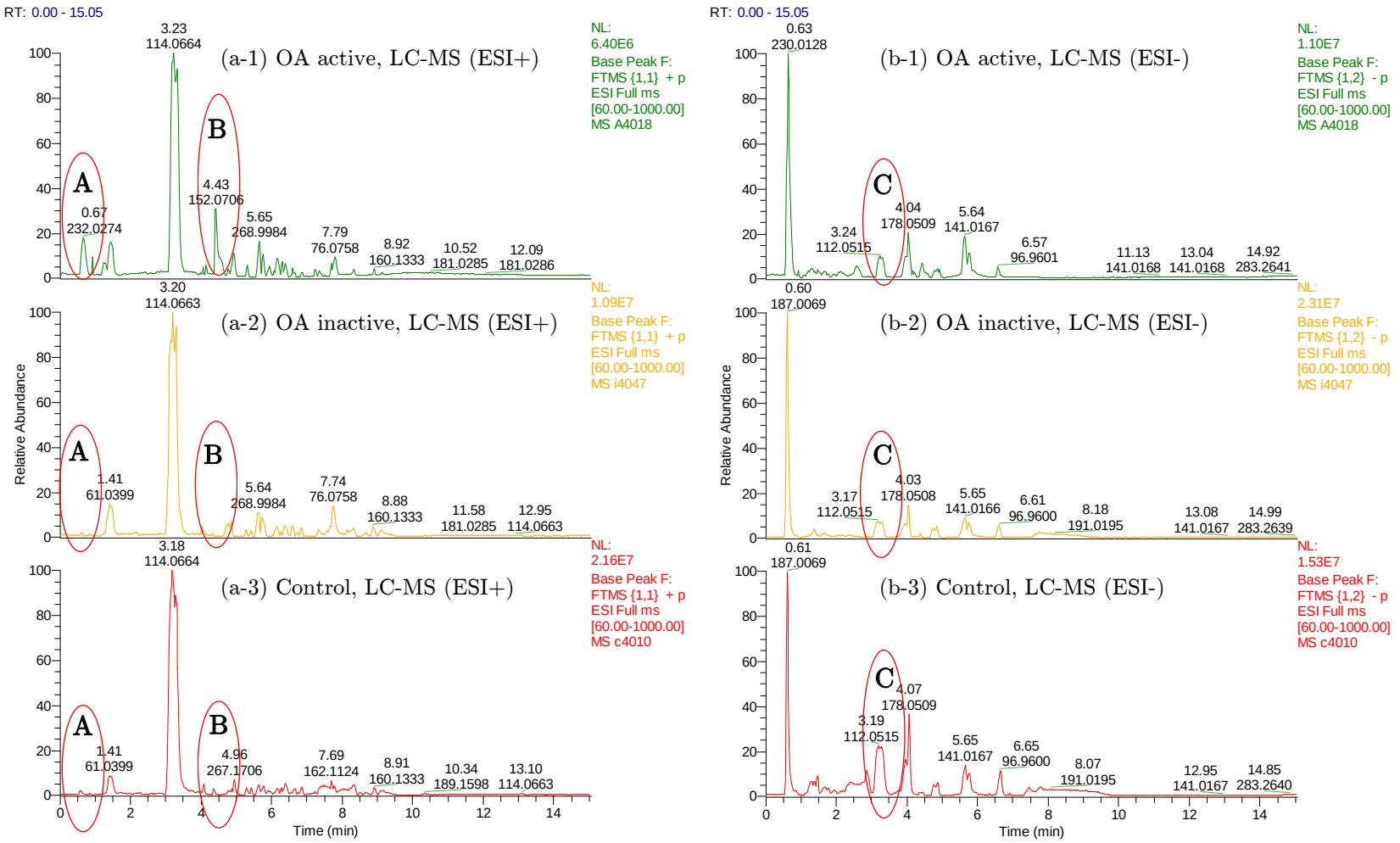


Figure 3-4 LC-MS base peak chromatograms (BPC) of urine samples of OA patients and healthy controls. BPCs showed difference between the urine profiles of (a-1 and b-1): OA active patients, (a-2 and b-2): OA inactive patients and (a-3 and b-3) healthy controls of urine samples in positive (a) and negative (b) ion modes, respectively, analysed by LC-MS using BEH HILIC column. A, B and C are selected regions with obvious differences.

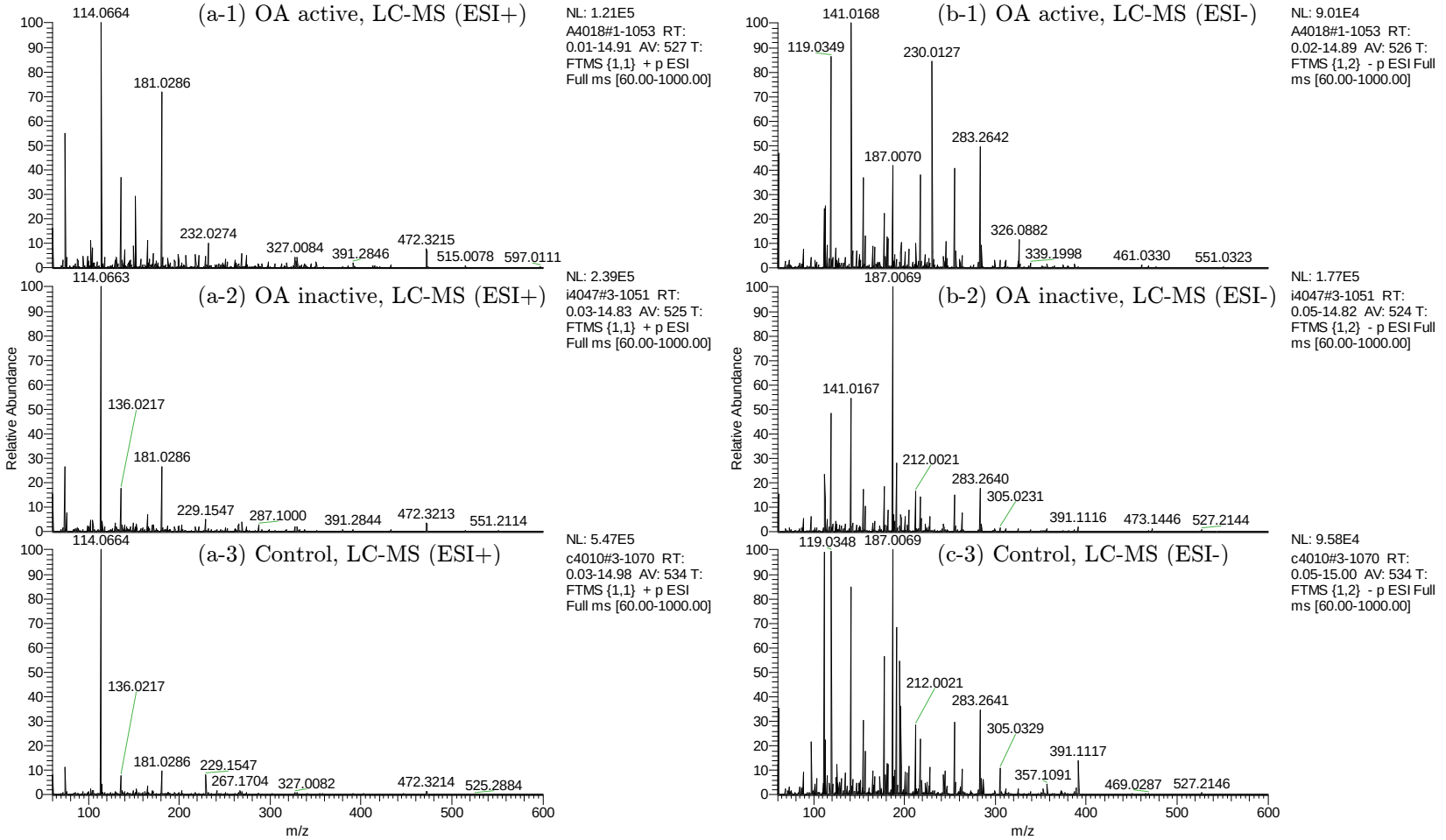


Figure 3-5 Typical metabolic fingerprints obtained from urine samples of OA patients and healthy controls analysed with LC-MS method. Combined mass spectra (*m/z* 60-600) showed difference between the urine profiles of (a-1 and b-1): OA active patients, (a-2 and b-2): OA inactive patients and (a-3 and b-3) healthy controls of urine samples in positive (a) and negative (b) ion modes, respectively, analysed using BEH HILIC column.

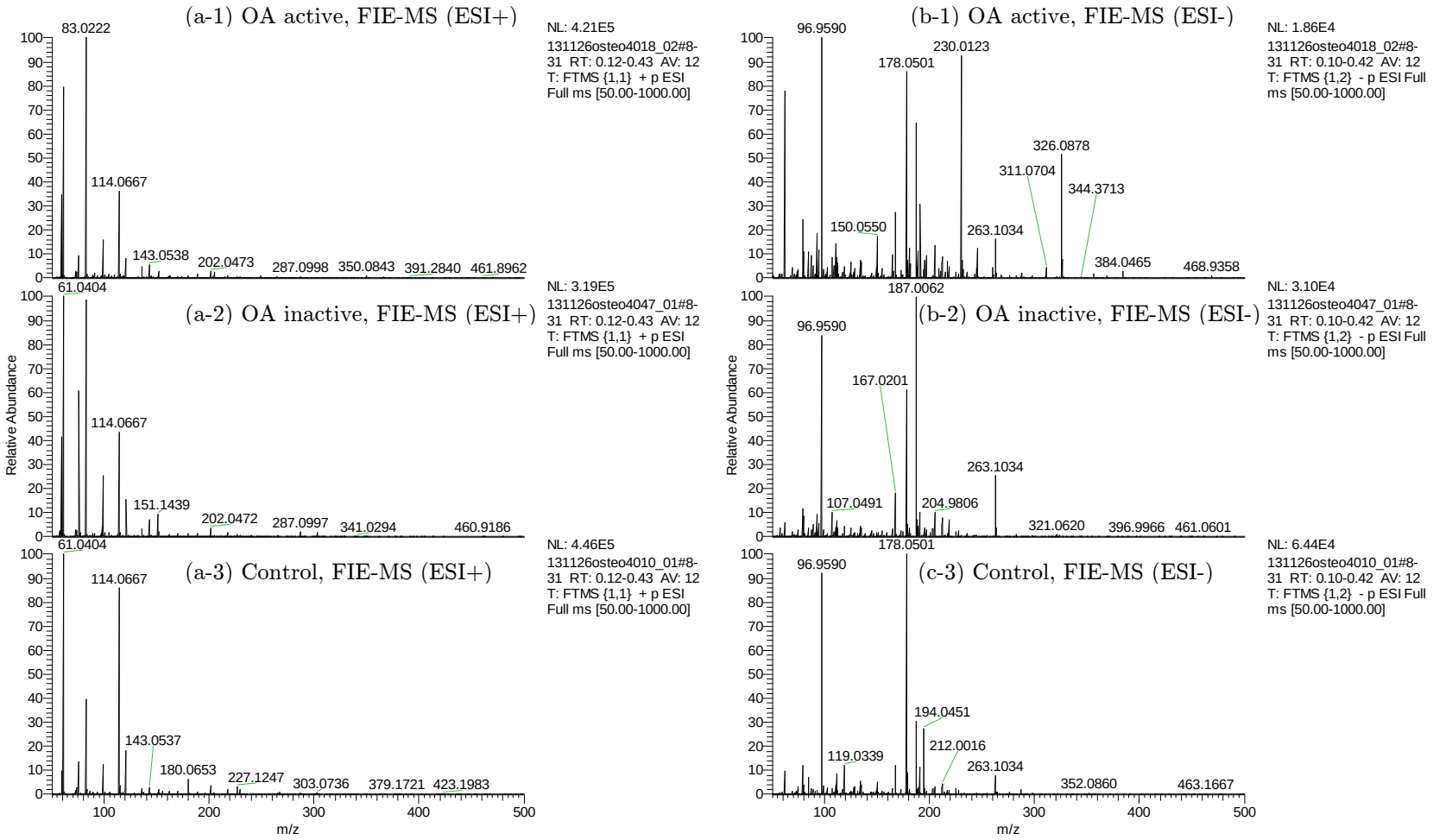


Figure 3-6 Typical metabolic fingerprints obtained from urine samples of OA patients and healthy controls analysed with FIE-MS method. Combined mass spectra (*m/z* 50-500) showed difference between the urine profiles of (a-1 and b-1): OA active patients, (a-2 and b-2): OA inactive patients and (a-3 and b-3) healthy controls of urine samples in positive (a) and negative (b) ion modes, respectively.

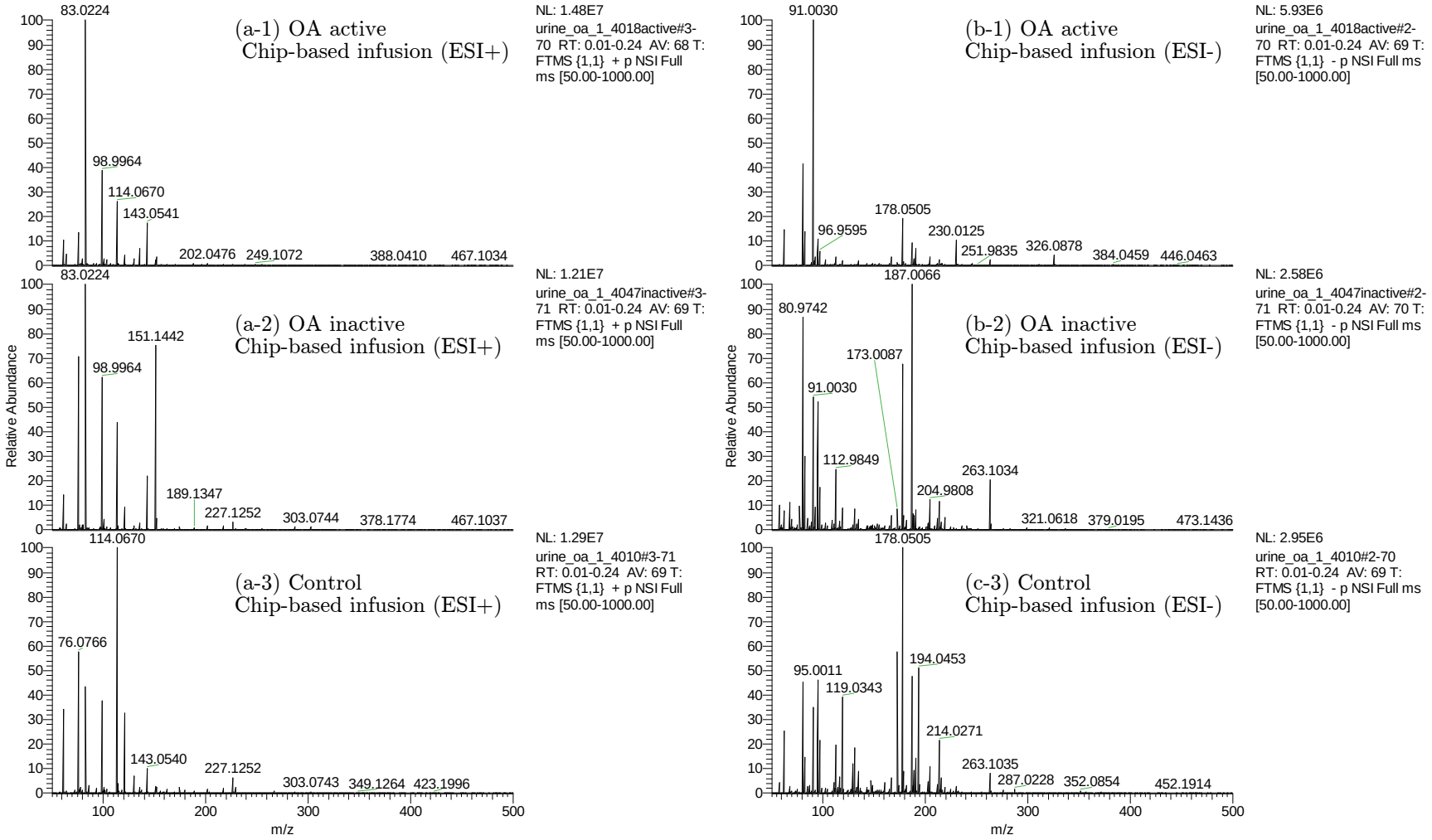


Figure 3-7 Typical metabolic fingerprints obtained from urine samples of OA patients and healthy controls analysed with chip-based infusion MS method. Combined mass spectra (*m/z* 50-500) showed difference between the urine profiles of (a-1 and b-1): OA active patients, (a-2 and b-2): OA inactive patients and (a-3 and b-3) healthy controls of urine samples in positive (a) and negative (b) ion modes, respectively.

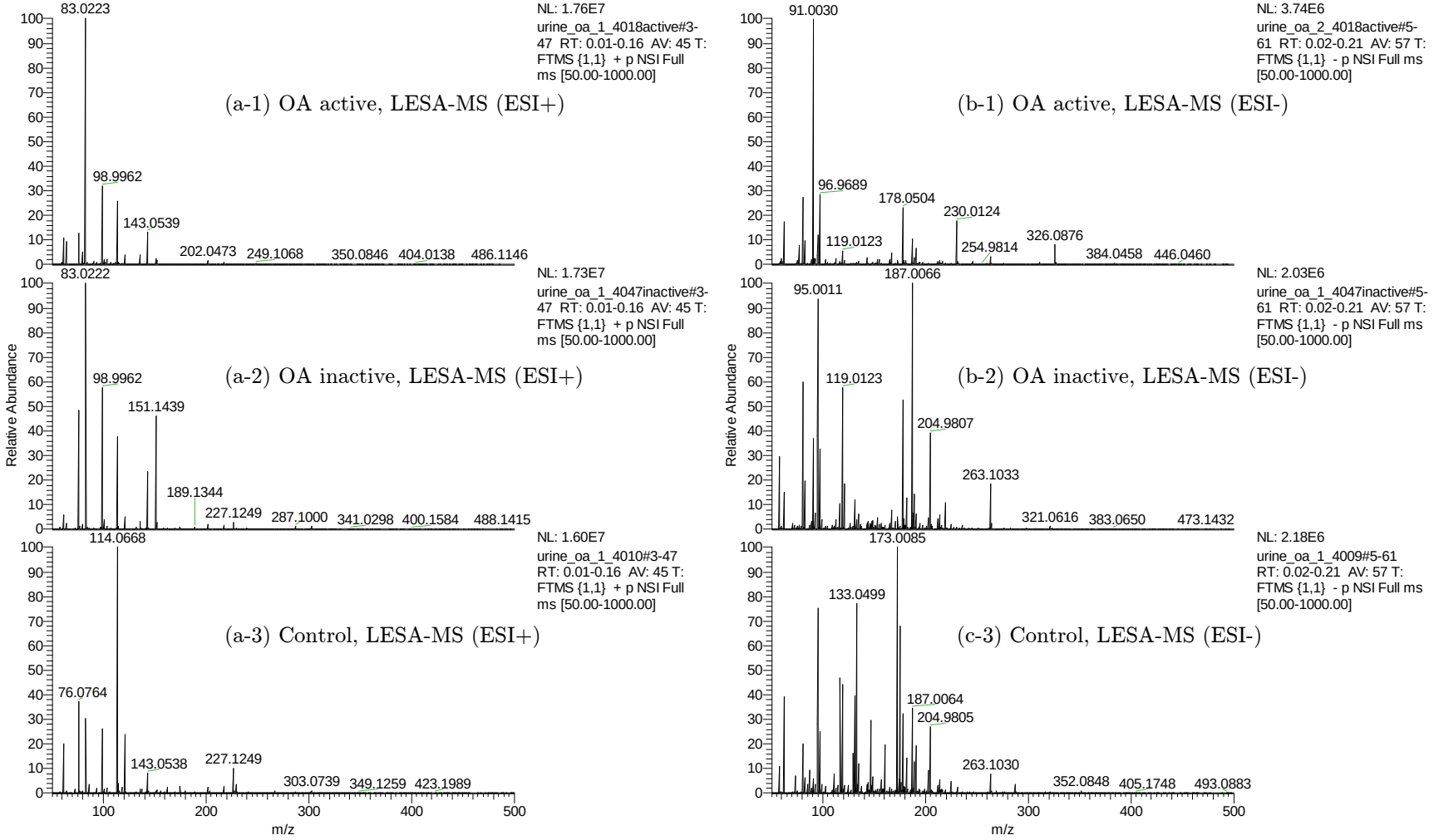


Figure 3-8 Typical metabolic fingerprints obtained from urine samples of OA patients and healthy controls analysed with LESA-MS method. Combined mass spectra (*m/z* 50-500) showed difference between the urine profiles of (a-1 and b-1): OA active patients, (a-2 and b-2): OA inactive patients and (a-3 and b-3) healthy controls of urine samples in positive (a) and negative (b) ion modes, respectively.

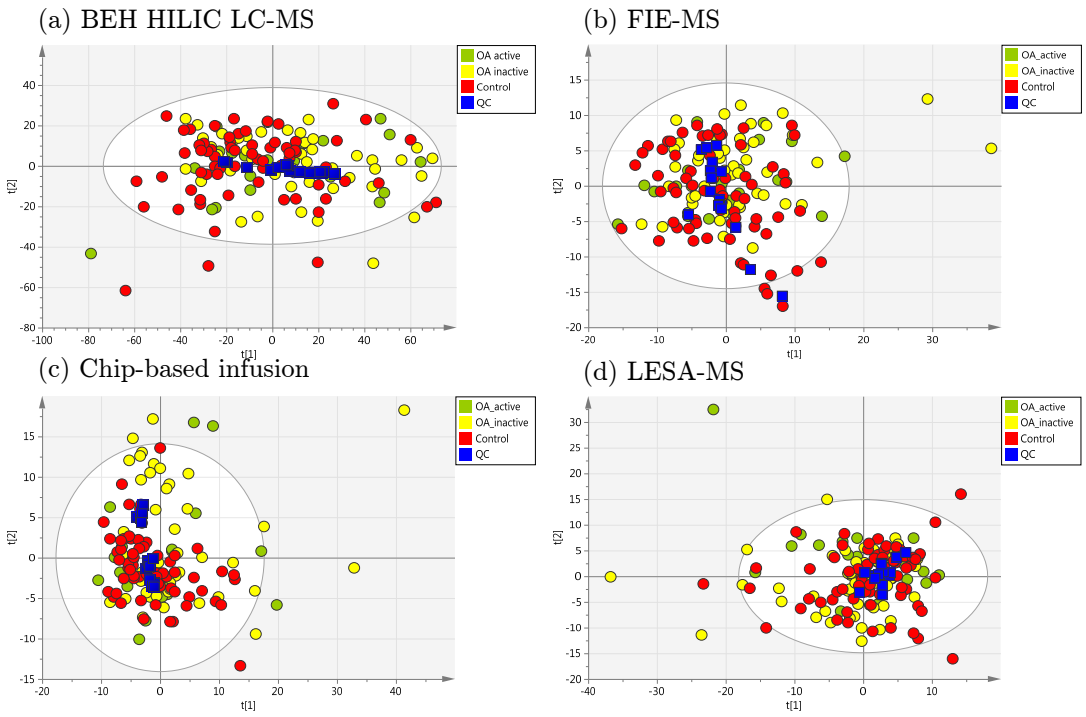


Figure 3-9 PCA score plots overview obtained from all OA patients and healthy controls. Healthy controls (red circles, n = 68), OA active patients (green circles, n = 22), OA inactive patients (yellow circles, n = 52) and pooled QC (dark blue squares, n = 15) analysed by (a) LC-MS (b) FIE-MS, (c) chip-based infusion and (d) LESA-MS.

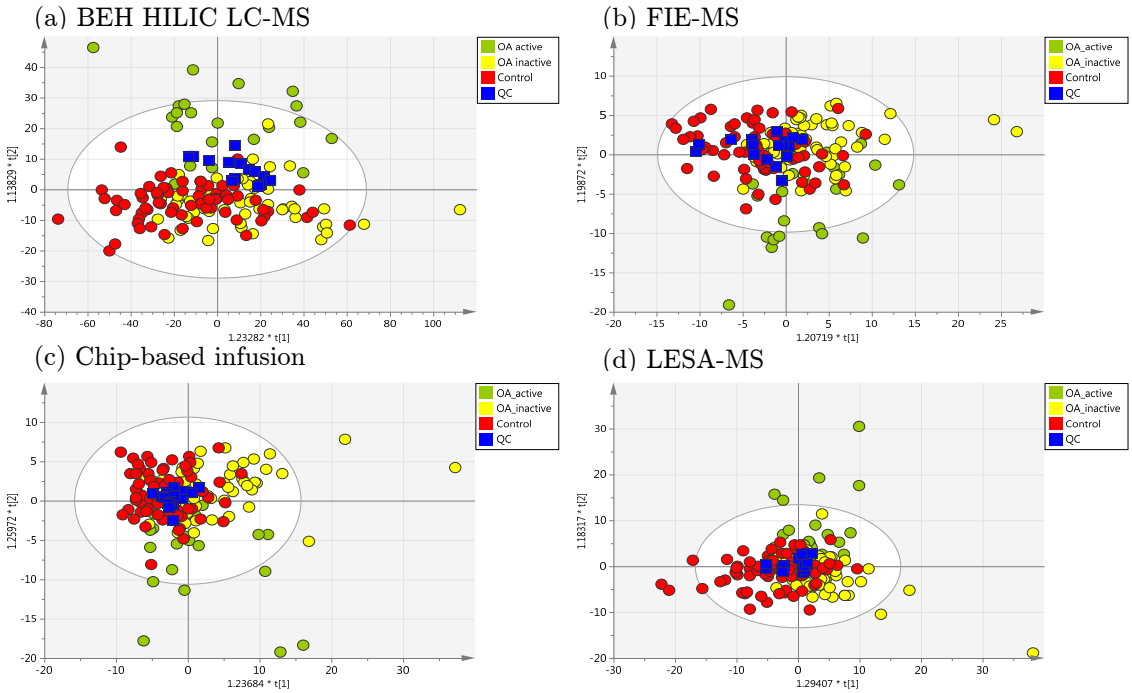


Figure 3-10 OPLS-DA score plots of OA patients and healthy controls. Healthy controls (red circles, n = 68), OA active patients (green circles, n = 22), OA inactive patients (yellow circles, n = 52) and pooled QC (dark blue squares, n = 15) analysed by (a) LC-MS (b) FIE-MS, (c) chip-based infusion and (d) LESA-MS.

3.3.4 Urine metabolomics study of active and inactive OA

To maximise the difference in classification between OA patients and healthy controls, successive OPLS-DA models were generated from only two classes from the study's samples. Figure 3-11 presents the OPLS-DA score plots of OA inactive patients (n = 52) and healthy controls (n = 68); whereas, Figure 3-12 presents the OPLS-DA score plots of OA active patients (n = 22) and healthy controls (n = 68) analysed with LC-MS, FIE-MS, chip-based infusion and LESA-MS.

OA inactive patients versus healthy controls

Insufficient separation between OA inactive patients and healthy controls was observed with all methods. These models were evaluated using cross-validation and permutation test. The OPLS-DA classification models of OA inactive patients and healthy controls gave satisfactory fitness of the model (R^2Y) values in the range of 0.482 - 0.638, but they showed very poor predictive ability with Q^2 values in the range of -0.221 to 0.026, and inadequate permutation test results (Table 3-5), indicating no significance difference between OA inactive patients and healthy controls classes.

OA active patients versus healthy controls

On the other hand, complete separation between OA active patients and healthy controls was observed with datasets from LC-MS analysis, and an improved, but not sufficient, separation between OA active patients and healthy controls was observed with all direct ESI-MS methods. These models gave good R^2Y and Q^2 values with LC-MS (0.913 and 0.347 respectively) than FIE-MS (R^2Y= 0.590, Q^2 = 0.167), chip-based infusion (R^2Y = 0.620, Q^2 = -0.067) and LESA-MS (R^2Y = 0.642, Q^2 = 0.073) (Table 3-5). All OA active patients and healthy controls OPLS-DA models gave good fitness of the model but they showed low predictive ability values which increases the possibility of separation between classes to model overfitting. The low prediction of these models may be attributed to the fact that the OPLS-DA models were built with two imbalanced groups in terms of sample size, i.e. OA active patients (n = 22) and healthy controls (n = 68), hence, balancing groups size may improve the classification and prediction of the OPLS-DA models.

Balancing sample size of OA active patients and healthy controls classes

OPLS-DA model performs best when the datasets are fairly symmetrically distributed and have a fairly constant "error variance". Therefore, in case of having different number of samples in each class, the reference point in OPLS-DA will not be in the middle of the model and the obtained results might be misleading for predictions and interpretation (Eriksson et al., 2006b, Bylesjo et al., 2006). A simple remedy of this problem is to balance the number of samples in each group, however, it is important to select a smaller set of samples that is representative of the whole set. The selection of smaller representative set from a bigger group of samples can be carried out using multivariate design based on PCA (also known as statistical molecular design) (Linusson et al., 2001, Gabrielsson et al., 2002). Using multivariate design, a representative set (n = 22) from the healthy control samples (n = 68) was selected as detailed in the method section 3.2.7. A subsequent balanced OPLS-DA models were obtained from the metabolic profiles of the OA active patients (n = 22) and healthy controls (n = 22) using the datasets generated by LC-MS and direct ESI-MS methods as shown in Figure 3-13. The obtained OPLS-DA models showed a clear separation and clustering of the OA active patients from the healthy controls with improved cross-validation R^2Y and Q^2 values for all methods compared to the previous models with the imbalanced group size (Table 3-5).

Validation of OPLS-DA of OA active patients and healthy controls

Cross-validation

The cross-validation results of the OPLS-DA models of OA active patients and healthy controls after balancing group size (Table 3-5) indicated adequate fitness and predictive ability of the models. However, these results give no statement about the statistical significance of the estimated predictive power of the OPLS-DA models (Eriksson et al., 2006b). Therefore, further permutation test was carried out for each OPLS-DA model to test if the generated models were spurious, i.e. the good predictive ability of the model is not due to data overfitting. The produced Q^2-intercept values of the regression lines of the Y-permuted Q^2 values of all OPLS-DA models were less than zero and less than the Q^2 value of the tested model (Table 3-5), indicating reliable predictive power of the generated models and it is not due to overfitting of the datasets.

External validation

In addition, a typical validation using prediction models based on randomly selected 50% training and 50% test sets of OA active patients and healthy controls datasets was also performed as shown in Table 3-5. The obtained OPLS-DA prediction models showed satisfactory results with all methods. The sensitivity and specificity of these generated prediction models were assessed by calculating area under receiver operating characteristic (ROC) curve (AUC) using Simca P +14. The value of the AUC of the OPLS-DA models was 0.76, 0.88, 0.79 and 0.86 for LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively. A rough guide based on AUC for assessing the clinical utility of a biomarker is as follows: 0.5–0.6 (fail), 0.6–0.7 (poor), 0.7–0.8 (fair); 0.8–0.9 (good) and 0.9–1.0 (excellent) (Xia et al., 2013). These results validate the OPLS-DA models of OA active patients and healthy controls for biomarker discovery. Although few variables were detected by direct ESI-MS compared to LC-MS, the direct ESI-MS datasets still generated as robust a model as LC-MS.

Table 3-5 Validation of OPLS-DA models of OA active patients and healthy controls

Description	LC-MS	FIE-MS	Chip-based infusion	LESA-MS
1. Peak detected				
ESI+	4,405	273	238	291
ESI-	3,000	211	338	452
Total (ESI+ and ESI-)	7,405	484	576	743
2. OPLS-DA of OA inactive patients (n = 52) and healthy controls (n = 68)				
R^2Y	0.638	0.482	0.548	0.554
Q^2	-0.221	-0.031	-0.071	0.026
Permutation (intercept)[1]	-0.141	-0.255	-0.194	-0.192
3. OPLS-DA of OA active patients and healthy controls				
a. Different group size: OA active patients (n = 22), healthy controls (n = 68)				
R^2Y	0.913	0.590	0.620	0.642
Q^2	0.347	0.167	-0.067	0.073
Permutation (intercept)	-0.204	-0.215	-0.233	-0.219
b. Balanced group size: OA active patients (n = 22), healthy controls (n = 22)				
R^2Y	0.874	0.975	0.972	0.945
Q^2	0.465	0.562	0.472	0.493
Permutation (intercept)	-0.130	-0.401	-0.240	-0.262
External validation: classification (training/test models)				
True positive (TP)	7	7	5	9
False positive (FP)	1	2	1	2
True negative (TN)	10	9	10	9
False negative (FN)	4	4	6	2
Sensitivity (%)[2]	88%	78%	83%	82%
Specificity (%)[3]	71%	69%	63%	82%
Accuracy (%)[4]	77%	73%	68%	82%
Area under receiver operative characteristic (ROC) curve (AUC)				
AUC (TPR vs FPR)[5]	0.76	0.88	0.79	0.86

[1]The OPLS-DA model is considered robust when the regression line of the permuted Q^2 values intercept at, or below zero and less than the Q^2 value of the model. [2]Sensitivity is the true positive rate (TPR) calculated from the formula, TPR = TP/(TP+FP). [3]Specificity is the true negative rate (TNR) calculated from the formula, TNR = TN/(TN+FN). [4]Acuuracy is calculated from the formula, Accuracy (%) = (TP+TN)/(TP+FP+TN+FN). [5]AUC: Area under receiver operating characteristic curve, ideal model gives AUC = 1 (Eriksson et al., 2006b, Xia et al., 2013).

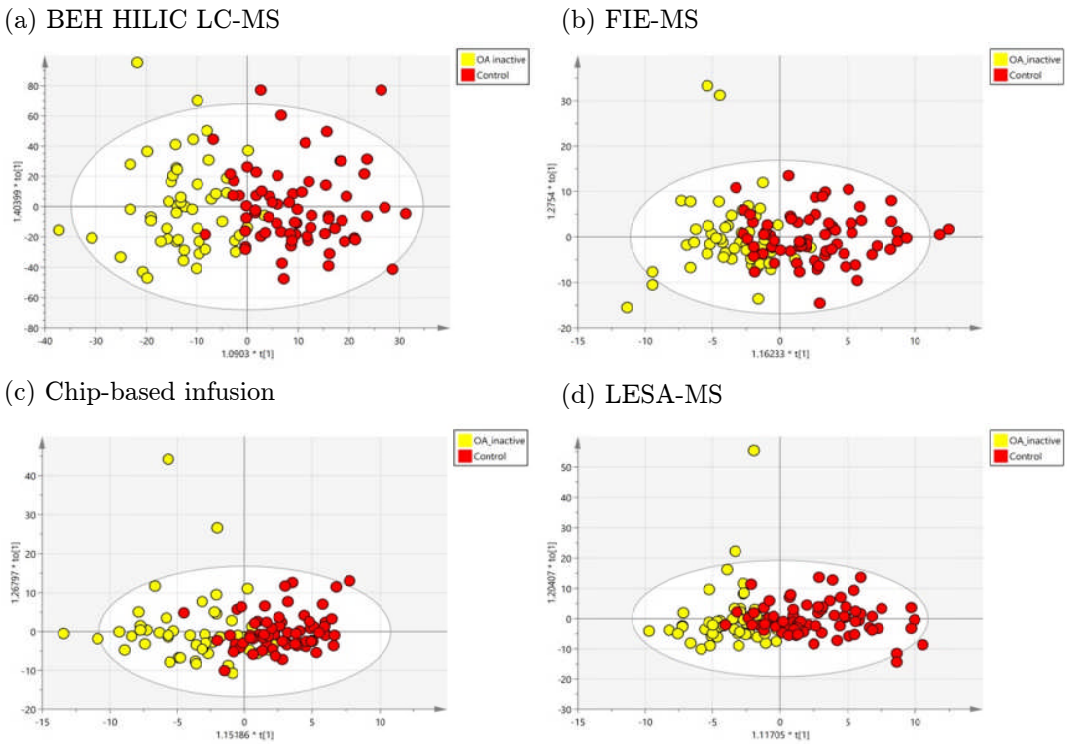


Figure 3-11 OPLS-DA score plots obtained from OA inactive patients and healthy controls. Healthy controls (red circles, n = 68) and OA inactive patients (yellow circles, n = 52) analysed by (a) LC-MS (b) FIE-MS, (c) chip-based infusion and (d) LESA-MS.

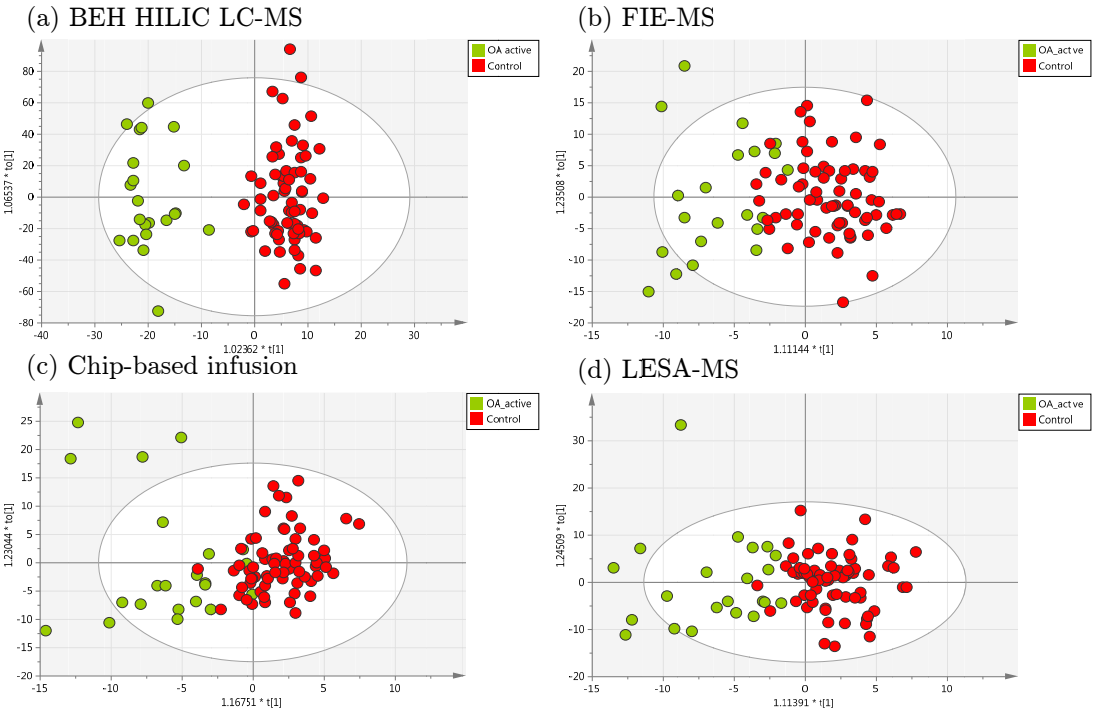


Figure 3-12 OPLS-DA score plots obtained from all OA active patients and healthy controls. Healthy controls (red circles, n = 68) and OA active patients (green circles, n = 22) analysed by (a) LC-MS (b) FIE-MS, (c) chip-based infusion and (d) LESA-MS.

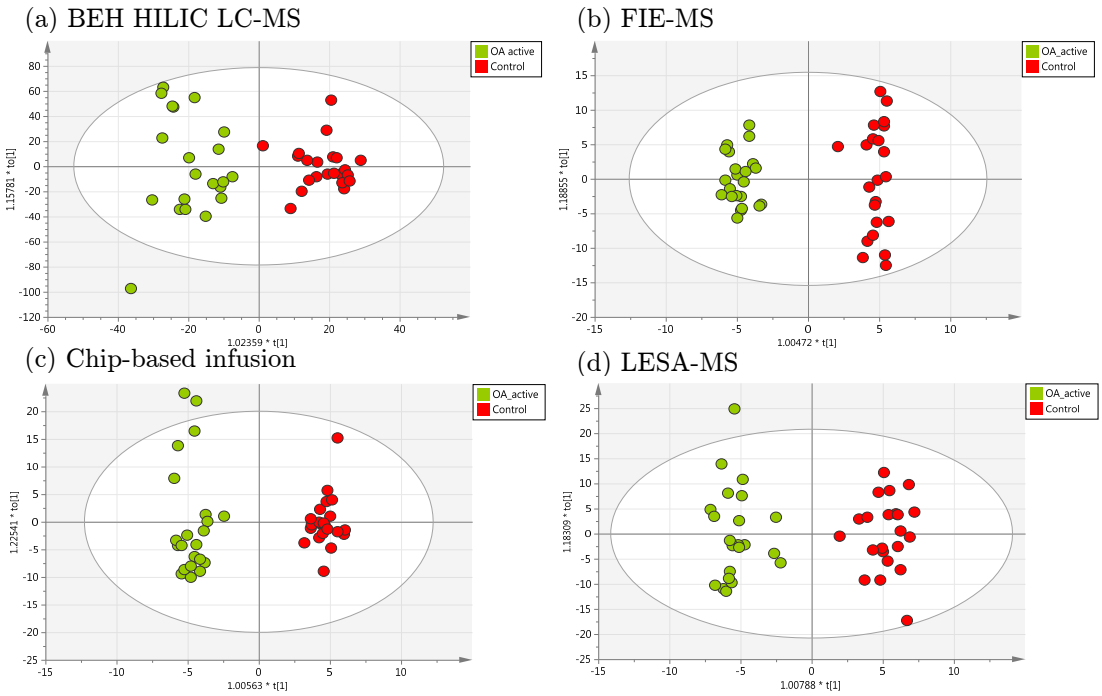


Figure 3-13 OPLS-DA score plots obtained from OA active patients and balanced healthy controls urine samples. Healthy Controls (red circles, n = 22) and OA active patients (green circles, n = 22) analysed by (a) LC-MS (b) FIE-MS, (c) chip-based infusion and (d) LESA-MS.

3.3.5 Tentative identification of potential biomarkers of OA in urine

Selection of potential OA biomarkers

Metabolites responsible for the separation between OA active patients and healthy controls were selected from the balanced group size OPLS-DA models (Figure 3-13). Variable importance for the projection (VIP) and loadings plots were used for the extraction of tentative OA biomarkers with significant differences between OA active patients and healthy controls urine samples. A VIP score of a variable above 1.0 is considered important for the model (Yin et al., 2009). As a result, 2366, 153, 207 and 237 variables were extracted from the models with LC-MS, FIE-MS, chip-based infusion and LESA-MS, respectively. Subsequent Student *t*-test with false discovery rate (FDR) for multiple testing problem was performed across peaks intensities of OA active patients and healthy controls datasets using Progenesis QI. Top ions of significant differences (q-value ≤ 0.05) have been considered as potential biomarkers. As a result, 27, 9, 12 and 12 variables were considered as potential OA biomarkers from the LC-MS, FIE-MS, chip-based infusion and LESA-MS datasets, respectively.

Identification of OA biomarkers

The exact mass of the biomarkers was then used to interrogate the Human Metabolome Database (HMDB) to provide tentative identification of the selected urine metabolites. Table 3-6 lists the tentative OA biomarkers identified with LC-MS and direct ESI-MS methods. Considering the limitation of database's repositories some metabolites remained unknown, however, they were listed in the table due to their importance for the study. Most of the metabolites obtained by FIE-MS were also detected with LC-MS such as L-homoserine, pyruvic acid and 2-hydroxybutyric acid. However, some urinary metabolites were exclusively detected with chip-based infusion and/or LESA-MS such as 2-keto-glutaramic acid, nervonyl carnitine and carbonic acid. Structurally, these metabolites are small molecular weight polar organic acids or amines (62-145 Da), suggesting that chip-based infusion and LESA-MS have high sensitivity for such metabolites. The detection of different ions and metabolites by chip-based infusion and LESA-MS compared to LC-MS and FIE-MS could be due to the very different electrospray source configurations and improved sensitivity observed with nanoESI chip, a point already has been discussed in chapter 2, section 2.3.7. This may also be due to the higher ion suppression effect of urine salts in LC-MS as samples were prepared with 3-fold dilution compared to 10-fold dilution in direct ESI-MS methods. This demonstrated that, chip-based infusion and LESA-MS could provide complementary information to LC-MS for metabolomics studies.

Overall direct ESI-MS methods comparison have shown that FIE-MS, chip-based infusion and LESA-MS detected about 33%, 44% and 44% respectively, of the LC-MS OA biomarkers. Whereas, the majority of metabolites detected with direct ESI-MS methods were also found by LC-MS. Thus, LC-MS analysis is the best option for a comprehensive metabolomics study, while direct ESI-MS methods with their high-throughput power might provide a quick screening approach for OA and a potential future diagnostic tool.

Table 3-6 Tentative identification of urinary biomarkers showing differences between OA active patients and healthy controls

Metabolite MW (Da)	q-value[a]	Fold change[b]	Tenative identification	HMDB ID	Formula	Mass Error (mDa)	MS method			
							LC-MS	FIE-MS	Chip-based infusion	LESA-MS
60.9926	1.33E-02	2.61	Hydrogen carbonate	HMDB00595	CHO_3	0.79			✓✓	✓
62.0004	4.59E-03	2.43	Carbonic acid	HMDB03538	CH_2O_3	0.72			✓✓	✓
70.0055	4.47E-02	1.92	Propynoic acid	HMDB06804	$C3H_2O_2$	0.48		✓✓		
74.0735	6.25E-04	9.21	Unknown				✓		✓✓	
83.0483	4.18E-05	174.65	5-Aminoimidazole	HMDB03929	$C_3H_5N_3$	0.13	✓		✓✓	
88.0160	1.63E-02	-1.56	Pyruvate	HMDB00243	$C_3H_4O_3$	0.12	✓	✓✓		
91.0029	3.84E-02	-14.78	Unknown							✓✓
93.0299	9.76E-03	1.30	Unknown						✓✓	✓
97.9769	5.78E-03	1.78	Phosphoric acid	HMDB02142	H_3O_4P	1.26	✓✓	✓		
98.9667	4.59E-03	14.83	Unknown				✓		✓✓	
98.9892	3.75E-02	-16.15	Unknown							✓✓
102.1283	4.53E-02	-4.08	Nervonyl carnitine	HMDB06509	$C_6H_{16}N$	1.61			✓✓	
103.0269	2.71E-02	-1.53	3-Oxoalanine	HMDB11602	$C_3H_5NO_3$	0.12	✓✓			
104.0473	5.05E-03	-1.39	2-Hydroxybutyric acid	HMDB00008	$C_4H_8O_3$	1.17	✓	✓✓		
104.0586	3.59E-03	1.91	2,3-Diaminopropionic acid	HMDB02006	$C_3H_8N_2O_2$	0.07	✓✓			
111.0433	1.14E-02	-1.41	Cytosine	HMDB00630	$C_4H_5N_3O$	0.05	✓✓			
113.8855	2.08E-02	-2.91	Unknown				✓			✓✓
119.0582	5.63E-03	-1.29	L-Homoserine	HMDB00719	$C_4H_9NO_3$	0.09	✓✓		✓	✓
129.0790	3.76E-02	-8.43	Pipecolic acid	HMDB00070	$C_6H_{11}NO_2$	0.96	✓			✓✓
136.0524	1.12E-03	1.68	Phenylacetic acid	HMDB01326	$C_8H_8O_2$	0.11	✓✓			
139.9875	1.03E-02	1.32	Acetylphosphate	HMDB01494	$C_2H_5O_5P$	1.76	✓	✓✓		
140.9505	2.19E-02	1.24	Unknown					✓✓		
145.0375	4.59E-03	8.24	2-Keto-glutaramic acid	HMDB01552	$C_5H_7NO_4$	0.08			✓✓	
157.0739	3.81E-03	-2.05	3-Methylcrotonylglycine	HMDB00459	$C_7H_{11}NO_3$	0.34	✓✓			

Metabolite MW (Da)	q-value[a]	Fold change[b]	Tenative identification	HMDB ID	Formula	Mass Error (mDa)	MS method			
							LC-MS	FIE-MS	Chip-based infusion	LESA-MS
161.0688	5.46E-03	-1.32	Aminoadipic acid	HMDB00510	$C_6H_{11}NO_4$	0.05	✓✓	✓		✓
166.0630	3.29E-02	-3.44	4-Methoxyphenylacetic acid	HMDB02072	$C_9H_{10}O_3$	0.04	✓✓			
167.0252	5.78E-03	-3.51	Homocysteinesulfinic acid	HMDB06462	$C_4H_9NO_4S$	0.01	✓✓			
173.0688	6.51E-04	-1.59	N-Acetyl-L-glutamate 5-semialdehyde	HMDB06488	$C_7H_{11}NO_4$	0.01	✓✓	✓		
174.0892	4.59E-03	1.50	Suberic acid	HMDB00893	$C_8H_{14}O_4$	0.19	✓		✓✓	✓
197.9929	4.20E-02	5.61	1-Acylglycerone 3-phosphate	HMDB11750	$C_4H_7O_7P$	0.85	✓		✓✓	✓
204.0899	5.78E-03	-1.44	Tryptophan	HMDB30396	$C_{11}H_{12}N_2O_2$	0.44	✓✓			
224.0797	6.85E-03	-2.56	Hydroxykynurenine	HMDB01152	$C_{10}H_{12}N_2O_4$	0.51	✓✓		✓	✓
226.0590	1.01E-02	1.40	3-Nitrotyrosine	HMDB01904	$C_9H_{10}N_2O_5$	0.17	✓✓	✓		
243.0981	8.19E-03	2.29	Prolyl-Glutamate	HMDB29016	$C_{10}H_{15}N_2O_5$	1.12	✓✓			
264.1110	8.57E-03	-1.38	Alpha-N-Phenylacetyl-L-glutamine	HMDB06344	$C_{13}H_{16}N_2O_4$	0.68	✓✓			
379.1049	2.43E-02	1.52	S-Lactoylglutathione	HMDB01066	$C_{13}H_{21}N_3O_8S$	0.54	✓✓			

[(a)] q-value: is the adjusted student's *t*-test p-value using false discovery rate (FDR), [(b)] the positive value of fold change means a higher level of metabolite in OA active patients compared to healthy controls, whereas the negative value represents a lower level of metabolite. The q-value, fold change and mass error were calculated for each metabolite from its MS dataset generated by the MS method with the highest VIP value (primary MS method). Primary method used for the extraction of each biomarker is highlighted as (✓✓) in the table, whereas (✓) symbol in the MS method columns indicates that the biomarker was also detected by that method.

3.4 Discussion

GC-MS and NMR-based metabolomics have been reported for the measurement of urinary metabolites to reflect the OA process (Li et al., 2010, Lamers et al., 2005). However, other metabolomics techniques such HILIC LC-MS, which are more suited for the analysis of urinary polar and semi-polar metabolites, have not been applied in OA urinary biomarker research. In addition, the versatility of using direct ESI mass spectrometric techniques such as chip-based infusion and LESA-MS with their high-throughput potential, minimum or no sample preparation step and enhanced ionisation sensitivity (Flangea et al., 2011) have not been explored yet. Comprehensive understanding of metabolic variations associated with OA progression would be helpful in finding relevant OA biomarkers which will assist in the diagnosis and allow proper management of the disease. Therefore, the main objective of this study was to conclude whether the urinary metabolomics analysis using HILIC LC-MS or direct ESI-MS methods could establish an alternative direction for understanding the biological mechanism behind OA progression and assist in the diagnosis and proper management of the disease. The observed metabolic changes associated with OA patients were relatively small; therefore, detailed multivariate analysis was used to separate the OA patients from the healthy controls.

There were no significant variations observed between the metabolic profiles of urine samples from OA inactive patients and healthy controls with all MS methods. This may be due to the fact that the metabolic changes associated with OA in the asymptomatic stage (inactive stage) are quite small due to slow progression of the disease over years compared to the acute symptomatic stage of the disease (active) (Loeuille, 2012a). On the other hand, significant differences in the urinary metabolomics profile were found between OA active patients and healthy controls. Urine samples of OA active patients showed significantly increased levels of metabolites such as propynoic acid, phenylacetate, carbonic acid, 2-keto-glutaramic acid, 1-acylglycerone 3-phosphate, s-lactoylglutathione, 3-nitrotyrosine and acetylphosphate when compared with healthy controls. On the other hand, some metabolites showed significantly decreased levels in OA patients compared to healthy controls such as pyruvate, 2-hydroxybutyric acid, L-homoserine, pipecolic acid, 3-methylcrotonylglycine, aminoadipic acid, tryptophan and alpha-N-phenylacetyl-L-glutamine. The decreased level of pyruvate, tryptophan and alpha-N-phenylacetyl-l-glutamine in OA was consistent with previously reported studies (Li et al., 2010, Mickiewicz et al., 2015).

The identification of metabolites from urinary metabolomics studies of OA patients enhances our understanding about the possible metabolic changes associated with OA. For example, increased levels of s-lactoylglutathione, carbonic acid and acetylphosphate were found in the urine of OA active patients, which indicating enhanced activity of the pyruvate metabolic pathway and the tricarboxylic acid (TCA) cycle as a result of the perturbed metabolism in the cartilage cells. As most of the enzymes involves in the pyruvate metabolism and TCA, including lactoylglutathione lyase and phosphate acetyltransferase, are located inside the mitochondrial matrix of the cartilage cells, abnormal urinary excretion of intermediates of pyruvate metabolism and TCA cycle may provide metabolic evidence of cartilage cells mitochondrial dysfunction in OA. The increased activity of the TCA cycle in the mitochondria of the cartilage cells of the OA patients was previously reported in the literature (Gavriilidis et al., 2013, Blanco et al., 2004). Li and co-workers, linked the detection of abnormal levels of TCA substances such as aconitic acid and citric acid in the urine of knee OA patients to the enhanced activity of the TCA cycle (Li et al., 2010). However, these metabolites were not detected with LC-MS or direct ESI-MS methods, which may be attributed to the fact that they used GC-MS for the analysis, which have different source and MS configurations and hence, different sensitivity. The finding of increased activity in the pyruvate metabolic pathway was consistent with the low level of pyruvate found in the urine of OA patients compared to healthy controls, which indicates that pyruvate was highly consumed and therefore, lower levels were detected in the urine of OA patients.

Propynioc acid is an intermediate in propanoate metabolism and mitochondrial fatty acid β-oxidation (the major mitochondrial pathway for the breakdown of fatty acids (lipolysis) for energy production) (Harvey et al., 2011). The increased level of urinary propynoic acid in OA patients, as part of a regulatory mechanism affecting lipids metabolism to maintain certain intermediates for the purpose of biosynthesis and energy production (Ardawi and Newsholme, 1984), suggested that lipolysis may play an important role as an alternative source of energy in cartilage cells of OA patients. This was consistent with previously reported studies (Li et al., 2010, Damyanovich et al., 1999).

Glycerophospholipids are one of the essential building blocks of cell membrane lipid bilayers and are intimately involved in regulation of membrane trafficking, cell signalling and many other membrane-related activities (Berridge and Irvine, 1989, Farooqui et al., 2000). 1-acylglycerone-3-phosphate is a glycerophospholipid

intermediate in the glycerophospholipids metabolism and ether lipid metabolism pathways. In the glycerophospholipids metabolic pathway, 1-acylglycerone-3-phosphate is produced from glycerone phosphate by the action of glyceronephosphate O-acyltransferase enzyme. It is then converted to 1-acyl-sn-3-glycercol-phosphate catalysed by 1-acylglycerone phosphate reductase or enters ether lipid metabolism (Purich and Allison, 1999). The increased levels of 1-acylglycerone 3-phosphate in OA patients, suggested that glycerophospholipids metabolism might be altered in OA. Studies reported by Hills, indicated that alterations in the concentrations and compositions of phospholipids are associated with the development of OA (Hills, 2002). Kosinska et al., reported a 3.6 fold increase in the levels of some glycerophospholipids (two lysophosphatidylglycerol and five phosphatidylglycerol species) in the synovial fluids of late OA patients compared to healthy controls (Kosinska et al., 2014). Zhang et. al., also reported a distinct separation and clustering between two OA phenotypes due to significantly increased levels of glycerophospholipids in synovial fluids between OA patients subgroups (Zhang et al., 2014b). Our results are quite consistent with the above findings.

Oxidative stress and inflammation processes are believed to be one of the primary degenerative mechanisms in the development and progression of OA (De Ceuninck et al., 2011). Amino acids, the structural building blocks of proteins, plays an important role in the regulation of these processes. For instance, L-cysteine is essential for the production of the antioxidant glutathione, which is thought to help in scavenge the destructive oxygen-free radicals produced during normal cell metabolism. It also plays an important role in the inflammatory response by influencing the production of phagocytes in OA (Surapaneni and Venkataramana, 2007). Under metabolic stress, L-cysteine is produced from homocysteine and 2-hydroxybutyrate is released as a byproduct. No significant difference was observed of L-cysteine between OA active patients and healthy controls, but significantly low level of 2-hydroxybutyrate was found in the urine of OA patients, which may indicate impaired production of the necessary L-cysteine in OA. Similarly, abnormal concentrations of urinary amino acids or their metabolites may provide an evidence of oxidative stress and/or inflammation in OA patients. Low levels of tryptophan, pipecolic acid and aminoadipic acid (lysine metabolites), L-homoserine (serine metabolite) and methylcrotonylglycine (glycine metabolite) were found in the urine of OA patients compared to healthy controls, signalling the possibility of altered metabolic pathways and the biological functions of these amino acids in OA.

Glutamine, an essential amino acid for energy production in connective tissues, is required by articular cartilage cells as an amino group donor for the biosynthesis of cartilage collagen (Handley et al., 1980). The increased urinary excretion of 2-keto-glutaramic acid, a deaminated metabolite of glutamine, in OA patients may be an indication of the disturbed glutamine metabolism in the chondrocytes and cartilage cells. Normally, excess glutamine conjugates with active phenyl acetate (i.e. phenyl acetyl-CoA) to form alpha-phenylacetyl-glutamine and Coenzyme A. The end product, alpha-phenylacetyl-glutamine is then excreted in urine as a normal metabolite of glutamine and phenyl acetate (Shockcor et al., 1996), alternatively under abnormal conditions, glutamine deaminated to form 2-keto-glutaramic acid. The increased levels of phenyl acetate and 2-keto-glutaramic acid in addition to the decreased level of alpha-phenylacetyl-glutamine found in the urine of OA active patients, may give further evidence of the altered glutamine metabolic pathways in the cartilage cells as illustrated in Figure 3-14. Altered glutamine level in the urine of knee OA patients was previously reported in the literature (Li et al., 2010)

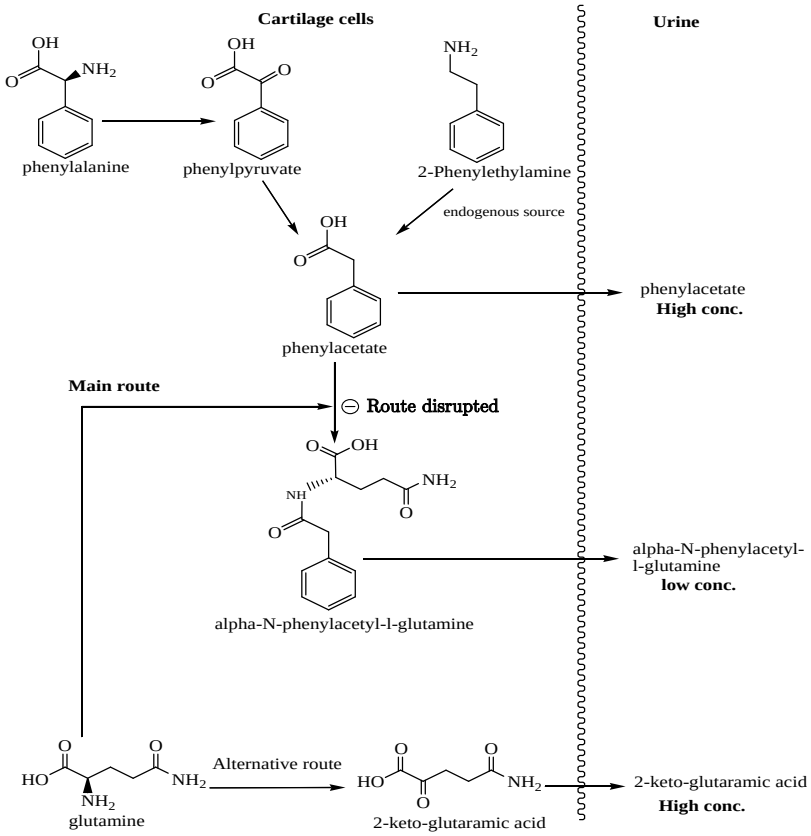


Figure 3-14 Proposed mechanism of altered glutamine fate in the cartilage cells of OA active patients compared to healthy controls and the interlink between glutamine and phenyl acetate during urinary elimination process.

Peroxynitrite is a potent short-lived oxidant species produced by the cell during its normal metabolism. The peroxynitrite anion ($ONOO^-$) is produced in the reaction between nitric oxide ($^{\bullet}NO$) and superoxide anion ($O2^{\bullet-}$) radicals. It is a strong oxidising and nitrating agent, which can cause damage to proteins, DNA and other cellular structures. Peroxynitrite is believed to be involved in many pathological conditions such as inflammation, pain, arteriosclerosis, cardiovascular and neurodegenerative disorders (Pacher et al., 2007). Szabo et al. summarised the biological and pathological effects of peroxynitrite exposure such as calcium dysregulation, mitochondrial dysfunction, inhibition of prostaglandin formation, imbalance of anti-inflammatory mediator pathways and amino acids nitration (Szabo et al., 2007). 3-Nitrotyrosine is detected in human urine and plasma by GC-MS and LC-MS and it is strongly believed to be one of the biomarkers of oxidative damage of peroxynitrite (Tsikas et al., 2012, Gaut et al., 2002). There was significantly increased level of 3-nitrotyrosine in the urine of OA patients compared to healthy controls detected by LC-MS and FIE-MS, which may be an indication of an increased oxidative damage of cartilage cells due to peroxynitrite.

3.5 Conclusion

In this chapter, a non-invasive metabolomics study using the developed LC-MS and direct ESI-MS methods combined with a multivariate analysis was able to find significant differences between the metabolic profiles of OA patients in the active stage and healthy controls. This is the first study that used urinary metabolomics approach to classify OA patients from healthy controls using HILIC LC-MS, FIE-MS, chip-based infusion and LESA-MS. Moreover, a novel approach using PCA multivariate design and SUS plot is used to balance the number of samples in each group in the study for multivariate analysis. The OPLS-DA models generated from the datasets of OA active patients and healthy controls, demonstrated good classification of the disease from the healthy controls with all MS methods. Considering the high-throughput of the direct ESI-MS; FIE-MS, chip-based infusion and LESA-MS were able to provide 33%, 44% and 44% respectively, of the LC-MS information, indicating their great potential of diagnostic application in OA, while LC-MS is necessary when comprehensive biomarker screening is required. The biological interpretation of the metabolites associated with the separation in the OPLS-DA models, revealed an increased activity in TCA cycle and pyruvate metabolism, which supports the previously reported literature of mitochondrial dysfunction in the cartilage cells of OA patients. In addition, an increased activity in the β-oxidation pathway may provide an evidence of the use of lipolysis as an alternative source of energy in the chondrocytes and cartilage cells in OA. Also, altered levels of glycerophospholipids and amino acids metabolites were found in the urine of OA patients compared to healthy controls, indicating perturbed metabolism of these biomolecules which can be linked to inflammation, oxidative stress and collagen destruction.

CHAPTER FOUR

A Case Study Based in Ethiopia to Investigate the Effect of Malaria on Urinary Metabolic Profiles and for Biomarker Discovery

4. A case Study Based in Ethiopia to Investigate the Effect of Malaria on Urinary Metabolic Profiles and for Biomarker Discovery

4.1 Introduction

4.1.1 Malaria: a global health and economic burden

Malaria is an infectious disease caused by parasite species belonging to the genus *Plasmodium*. Four of these species: *P. falciparum*, *P. vivax, P. malariae and P. ovale* are known to cause malaria in human, which is transmitted from one person to another by the bite of female *Anopheles* mosquitoes. There are more than 400 *Anopheles* species worldwide, but only 30 of these are vectors of major importance (Jain et al., 2014). Malaria is endemic in 104 tropical and subtropical countries, in which the causative agent, *P. falciparum* is the most virulent among *Plasmodium* species, and the most common cause of death from malaria in Africa (Gething et al., 2011). *P. vivax* has a wider geographic distribution worldwide, but its risk of infection is lower than *P. falciparum* due to lack of a specific gene known as the "Duffy" gene in many African populations, which helps in the production of an essential protein for the parasite to invade human red blood cells (RBCs) (Barnwell et al., 1989). The proportion of malaria infection due to *P. malariae and P. ovale* is very low compared to *P. falciparum* and *P. vivax* worldwide, but populations at risk of being infected with *P. malariae* are distributed over a wide geographic area including sub-Saharan Africa, South-East Asia, Amazonian Basin and Western Pacific islands (Collins and Jeffery, 2007). *P. ovale* is distributed mainly in Africa and Asia-Pacific regions; however, its prevalence is deemed to be lower than *P. malariae* (Sutherland et al., 2010). In 2008, some cases of malaria have been reported in certain forested areas of southeast Asia due *P. knowlesi,* a causative species of malaria in monkeys (Cox-Singh et al., 2008).

The World Malaria Report released by World Health Organisation (WHO) estimated that 3.3 billion people (half of the world's population) are at risk of being infected with malaria, and 1.2 billion at high risk in 2013 (WHO, 2014). Figure 4-1 presents the WHO recent estimate of the number and percentage of people who live at risk of being infected with malaria worldwide. The highest

disease burden is in the African continent, where an estimated 85% of the population are at risk of being infected with malaria, with 69% of them considered to be at high risk. Globally, 198 million cases with mortality of about 584,000 people attributed to malaria have been reported in the year 2013. In Africa only, an estimate of 90% of the mortality occurred, and 78% of all deaths were children under the age of 5 years. The estimated global annual cost to attain malaria control and elimination is $5.1 billion, however a maximum funding of 53% ($2.7 billion) was achieved in 2013, which represents three times the amount spent in the year 2005. Although there are continuous international efforts to increase the available fund for malaria control and elimination, the total expected annual amount by 2020 would still fall short of the estimated value to achieve international targets (WHO, 2014). This indicates that the fight against malaria is still one of the unresolved economic burdens on the globe.

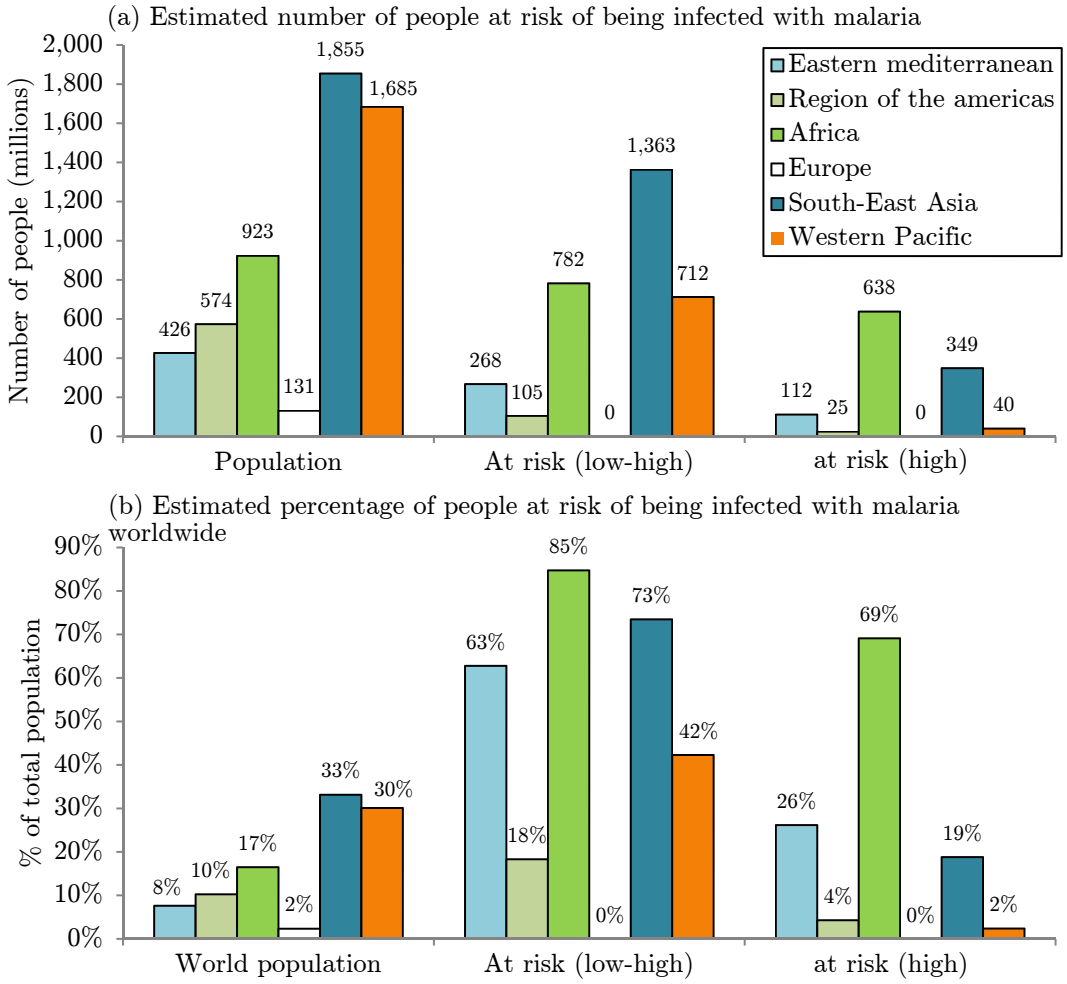


Figure 4-1 Estimated global number and percentage of people at risk of being infected with malaria in the year 2013. This data was reproduced from World Malaria report 2014 (WHO, 2014).

4.1.2 The life cycle of *Plasmodium* and symptoms of malaria

The life cycle of *Plasmodium* parasite is quite complex (Figure 4-2), which involves two hosts: a mosquito vector (sexual cycle) and a vertebrate host (asexual cycle). Sexual development of the parasite (the sporogonic cycle) takes place in the midgut wall of a female *Anopheles* mosquito, where the sexual gametocytes fuse and develop to from the human infective stage "the sporozoite". The infected-malaria mosquito inoculates sporozoites into the vertebrate host (e.g. human) during a blood meal, where the asexual cycle of the parasite starts. In the human host, sporozoites migrate and infect hepatocytes and mature into hypnozoites (exo-erythrocytic cycle) and then undergo asexual multiplication in the red blood cells (erythrocytic cycle). At a certain point, some parasites during the erythrocytic stage start to differentiate into sexual gametocytes. During the bite of a female *Anopheles* mosquito, gametocytes transfer into the stomach of the mosquito where the sexual life cycle of the parasite begins again (Pain and Hertz-Fowler, 2009).

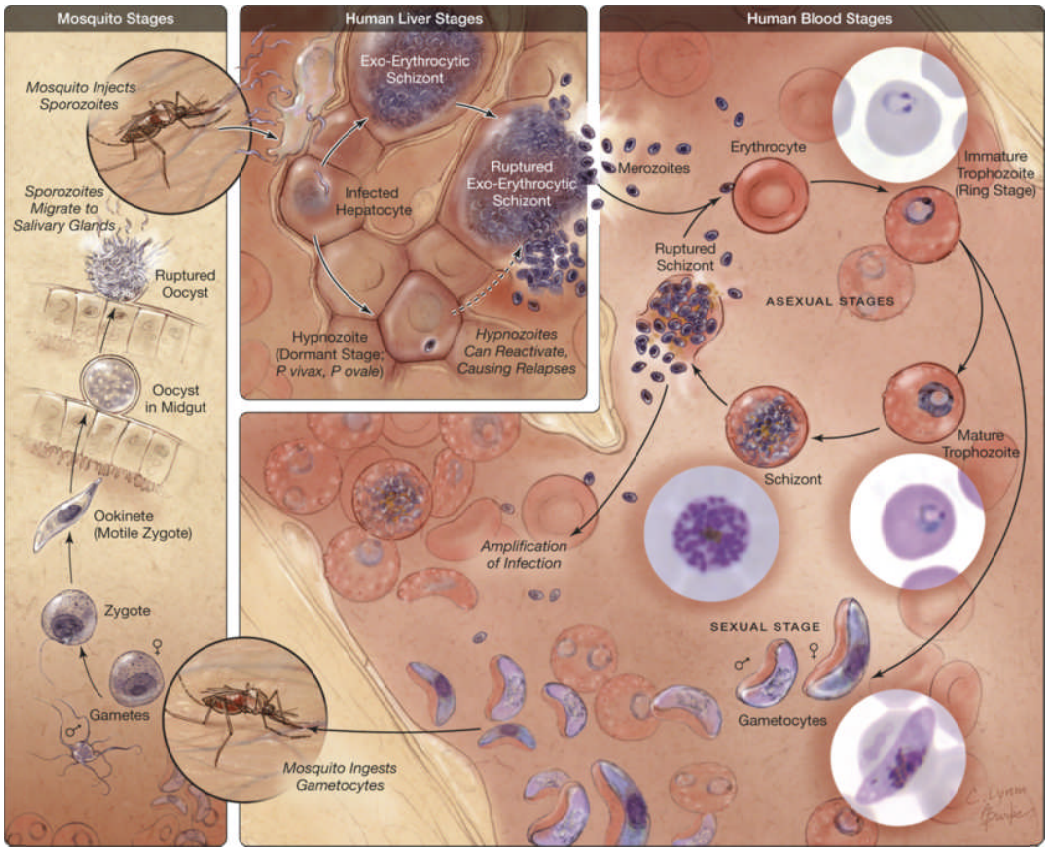


Figure 4-2 Diagram of the plasmodium's life cycle showing the sexual and the asexual development stages of the parasite (Griffith et al., 2007).

The blood stage of the parasite is responsible for the clinical manifestations of malaria. Human infection with *Plasmodium* parasites affects all age groups, and patients present with a wide variety of clinical symptoms, ranging from asymptomatic or very mild to severe symptoms and even death. Typical uncomplicated malaria symptoms include recurrent episodes of high fever, chills and rigors. The frequency of the recurrence of the fever is 48 h in *P. falciparum*, *P. ovale*, or *P. vivax*, 72 h in *P. malariae* and 24 h in *P. knowlesi* infections. The liver stage of the *P. vivax* parasite (the hypnozoites) can stay dormant in the hepatocytes and may cause disease recurrence (CDC, 2015). The underlying mechanisms of infection severity outcomes remain poorly understood, however, in severe infection with *P. falciparum*, the infected erythrocytes were reported to exhibit sequestration in the blood vessels (CDC, 2015, Dondorp et al., 2004). This sequestration of the infected erythrocytes in turns affects the blood flow in the microvasculature of different tissues and organs such as liver, kidneys, brain, intestines, heart, adipose tissue and eyes (Dondorp et al., 2000, MacPherson et al., 1985), and increases the possibility of a localised metabolic stress. Patients with severe malaria may exhibit a variety of severe clinical manifestations possibly through inflammatory immune responses, including vital organs failure (e.g. liver, kidney), abnormalities in blood coagulation, severe anaemia, hypoglycaemia, acidosis, cerebral malaria, respiratory distress and coma, which are the common causes of death from malaria (Lakshmanan et al., 2011, Haldar et al., 2007, Planche and Krishna, 2006, Trampuz et al., 2003). The process of transition from one clinical disease stage into another is still not fully understood and till present there are no reliable predictors available. However, drastic metabolic changes in the host are associated with the different manifestations of the disease (Basant et al., 2010). The study of metabolites from different tissues and body fluids during malaria progression may provide invaluable information that can be used to investigate the perturbed metabolic pathways in the host as a response to *Plasmodium* infection.

4.1.3 Diagnosis and treatment of malaria

Microscopy has been used for the diagnosis of malaria since 1904 when Gustav Giemsa detected the erythrocyte stage of the *Plasmodium* parasite from a thin film of blood from a malaria patient stained with a mixture of methylene blue, azure B (trimethylthionine) and eosin (the stain mixture also known as Giemsa's solution) (Giemsa, 1904). Although microscopy is considered as the "gold standard" for the diagnosis of malaria, the method is time-consuming, requires

intensive labour, invasive procedures and expert skills. Recent scientific reports revealed that there are more sensitive and specific techniques than microscopy for the diagnosis of malaria (Goncalves et al., 2012, Jain et al., 2014). These techniques include polymerase chain reaction (PCR) (Morassin et al., 2002), nested PCR (Li et al., 2014), loop-mediated isothermal amplification (Lau et al., 2011), flow cytometry (Malleret et al., 2011), laser desorption mass spectrometry (Demirev et al., 2002) and enzyme-linked immunosorbent assay (Martin et al., 2009). However, most of these techniques are either more expensive than microscopy, require high labour skills, not field deployable, or prone to false results and cross-contamination. Therefore, they are not widespread and their use is limited to research settings and blood bank screening.

Apart from microscopy and the above-mentioned techniques, rapid diagnostic tests (RDTs) are also available as dipsticks for quick diagnosis of malaria (Bjorkman and Martensson, 2010, Jain et al., 2014). RDTs are immunochromatographic diagnostic tests (ICTs), and they are used for diagnosis of malaria based on parasite specific antigen-antibody interactions and enzymatic assays from a peripheral blood sample from malaria patients. They mainly target plasmodial lactate dehydrogenase (pLDH), histidine-rich protein II (HRP II) and aldolase for the diagnosis of *P. falciparum* and *P. vivax* infections (Tjitra et al., 1999, Eisen and Saul, 2000). Nonetheless, they demonstrated very poor sensitivity for *P. ovale* and *P. malariae* infections (Maltha et al., 2010). Despite the fact that ICTs are widely used as quick, portable and cheap diagnostic tools for malaria and can be used by semi-skilled workers in rural settings; they suffer from certain limitations such as high variability, false positive/negative results, and loss of activity notably in tropical areas (Jorgensen et al., 2006). The loss of activity is mainly attributed to the denaturation of antibodies or nitrocellulose membrane damage as a result of high temperature and humidity (Chiodini et al., 2007).

Advanced biosensor techniques including electrochemical immunosensor (Sharma et al., 2008), piezoelectric biosensors (Sharma et al., 2011) and optical biosensors (Piper et al., 1999) are also reported for the diagnosis of malaria. A biosensor is a self-integrated device that is capable of measuring a specific biological entity using a bio-recognition element such as an enzyme, an aptamer (a single-stranded DNA or RNA) or an antibody, which is maintained in spatial contact with a transduction element (Eggins, 2008). In general, biosensors provide a quick, sensitive and selective determination of bio-molecules of clinical importance, however, the available biosensors for the diagnosis of malaria are mainly limited

for the detection of the parasite's pLDH (Lee et al., 2012) and HRP II (Piper et al., 1999). These biosensors are still in nascent stages and they mainly use antibodies and aptamers as bio-recognition elements, hence, the loss of activity due to denaturation of antibodies and nucleic acids is still a limiting factor in hot climates.

Malaria is a curable disease if diagnosed early and there are different chemotherapeutic agents available for the treatment of malaria. However, drug resistance has drastically increased in recent years especially for *P. falciparum* infections (Katzung et al., 2012). Drug resistance has developed against chloroquine and sulphadoxine-pyrimethamine combinations, the most widely used therapies to treat malaria (Ridley, 2003). Other drugs such as mefloquine and halofantrine are quite effective against resistant *P. falciparum* infections. Nevertheless, mefloquine may cause serious neuropsychiatric adverse effects such as unusual behaviour, hallucinations, depression, anxiety and suicidal ideations, whereas halofantrine can be associated with cardiac arrhythmias at therapeutic dosages. Therefore, certain measures should be considered to determine the risk/benefit ratio of administering those drugs (Katzung et al., 2012). For such reasons, the World Health Organisation (WHO) recommends artemisinin-based combination therapies (ACTs) as the new international standard for treatment of malaria (WHO, 2015b). ACTs should be followed by primaquine to eradicate the liver form in *P. vivax* and *P. ovale* infections, while in case of severe *P. falciparum* infection, intravenous quinine, quinidine or artesunate is recommended as an alternative therapy for comatose patients (Katzung et al., 2012).

The misdiagnosis or over-diagnosis of malaria remains another caveat; it is often in endemic areas that malaria is misdiagnosed or over-diagnosed leading to mismanagement of malarial or non-malarial fevers and over-prescription of malaria treatment (Nankabirwa et al., 2009). The most important factors that have a direct contribution to this are the unfamiliarity of the microscopic diagnostic procedure in rural settings and the inherent lower sensitivity of the technique (Hanscheid, 2003). Therefore, there is an urgent need to develop new diagnostic techniques which are accurate, non-invasive, simple to use and comparatively rapid. RDTs and biosensors research holds the promise of providing a robust, stable and field-deployable device for the diagnosis of malaria due to their quick detection capability and portable nature. However, in order to develop an efficient and stable RDTs or biosensors to attain the desired target in

the diagnosis and treatment of malaria, surrogate biomarkers of malaria need to be exploited further. This requires a profound understanding of the physiological and metabolic processes of the parasite and the host during the course of the infection, and, therefore, a biomarker discovery approach is important to reach such goals.

4.1.4 Metabolomics and malaria: the current role in understanding the host response and parasite biology

Different metabolomics studies have been reported in the literature to identify biomarkers of *Plasmodium* infections using NMR, GC-MS or LC-MS. Table 4-1 summarises a selected set of reported metabolites that have been shown a significant difference between *Plasmodium*-infected RBCs, animal models or patients and healthy controls.

4.1.4.1 In vitro studies

In vitro, parasite infected RBCs have been extensively used to study *Plasmodium spp.* metabolome against non-infected RBCs (Sana et al., 2013, Mehta et al., 2005, Olszewski et al., 2009). Teng et al., identified more than 40 parasite-specific intracellular metabolites from *P. falciparum*-infected RBCs using ^{1}H-NMR based untargeted approach (Teng et al., 2009). Similarly, Lian et al., studied the *P. falciparum* growth in an oxygen-limited environment using ^{13}C-labelled precursor which showed that infected RBCs produced glycerol and glycerol-3-phosphate as major glucose metabolites (Lian et al., 2009). This unexpected result uncovered an important aspect of parasite biology as glycolysis was believed to be the only energy source for the parasite in which glucose is incompletely oxidized to lactic acid and excreted (Mehta et al., 2006, Olszewski and Llinas, 2011).

LC-MS/MS targeted analysis was used to investigate the host-parasite interactions using *P. falciparum*-infected RBCs (Olszewski et al., 2009). In this study, arginine was found to be converted into ornithine by parasite arginase enzyme, indicating a presence of a potential link between parasite-induced arginine depletion and cerebral malaria pathogenesis. Using targeted metabolomics profiling of infected RBCs, central carbon metabolism (Olszewski and Llinas, 2011), mitochondrial metabolism (MacRae et al., 2013), mitochondrial redox balance (Storm et al., 2014), tricarboxylic acid (TCA) cycle, sphingolipid and fatty acid metabolism (Sana et al., 2013), isoprenoid pathway

(Couto et al., 1999) and carotenoid biosynthesis (Tonhosolo et al., 2009) of *P. falciparum* have also been explored.

4.1.4.2 In vivo animal model studies

The study of the host-parasite responses can provide invaluable information which may help in understanding the parasite biology, disease prognosis and possible alternative drug targets as there is a constant complex interplay of metabolites between the parasite and the host. The dynamic nature of the host-parasite interactions during the course of infection may perturb the biochemical profiles of both, the parasite and the host. For such a dual system, it is quite challenging to profile individual metabolomes; however, this co-metabolome can be quite informative if a healthy host is available as a baseline for metabolomics comparisons. For instance, urine and plasma samples of *P. berghei*-infected mice were compared with healthy controls using ^{1}H-NMR; as a result increased levels of urinary pipecolic acid, phenyl acetylglycine and dimethylamine were observed, indicating a gut microbial community disturbance caused by the parasite (Li et al., 2008). Similarly, Basant et al., studied the metabolic alterations during disease progression in urine, serum and brain samples from *P. berghei*-infected mice using ^{1}H-NMR (Basant et al., 2010). The altered levels of these metabolites (Table 4-1) were correlated with sexual dimorphism and were suggested to affect the prognosis and the treatment of malaria.

4.1.4.3 In vivo human studies

Few studies have been reported for clinical metabolomics of malaria (Lakshmanan et al., 2012, Sengupta et al., 2011b, Sengupta et al., 2015). Sengupta and co-workers, delineated the urinary metabolic profile of *P. vivax* infection using ^{1}H-NMR; altered levels of different metabolites were observed including pipecolic acid, ornithine, acetate and phenylpyruvate, indicating a perturbed urea cycle and impaired liver functions (Sengupta et al., 2011b). Few years later, the same group compared the urinary profile of *P. vivax*-infected patients with viral fever patients to healthy controls (Sengupta et al., 2015). Altered urinary levels of taurine, glycine, hippurate, citrate, alanine and 3-methylhistidine were observed in malaria patients compared to healthy controls. While, taurine and hippuric acid were significantly elevated in malaria compared to viral fever patients, and they were correlated to parasitemia levels in malaria, hence, they have been suggested to gauge anaemia status in malaria.

These metabolites are quite promising for prognosis of malaria and they could be further validated for clinical diagnosis. However, they were obtained from in vitro infected RBCs, rodent models or clinical samples of the less virulent parasite, *P. vivax* (i.e. it is not common in Africa where the disease global burden is the heaviest). In addition, the majority of these studies used NMR for metabolomics analysis which is known to be less sensitive than high resolution mass spectrometry (HRMS) (Dettmer et al., 2007). Therefore, more clinical metabolomics studies are still needed to find surrogate biomarkers for the most known virulent malaria parasite, *P. falciparum*. To date, the use of HILIC LC-MS and FIE-MS for clinical *P. falciparum* infections has not been explored yet. The application of these methodologies could enhance our understanding of host-parasite metabolic responses during the course of infection and foster efforts towards finding surrogate biomarkers or drug targets for the rapid clinical diagnosis or treatment of malaria.

Table 4-1 Selected list of reported metabolites as potential biomarkers for *Plasmodium* infections

Reference	Species/method	Sample type	Metabolites
(Basant et al., 2010)	*P. berghei,* NMR	Brain extract, mice	2-Hydroxy-2-methylbutyrate Glycerol
		Serum, mice	Alanine Kynurenic acid Lactate Lysine Quinolinic acid
		Urine, mice	Ureidopropionate Asparagine Carnitine Dimethyl-glycine
(Lakshmanan et al., 2012)	*P. falciparum,* LC-MS	Plasma, Infected RBCs	3-Hexenal Hypoxanthine Methyl jasmonate Oxononadienal Taurine Traumatic acid Traumatin
(Li et al., 2008)	*P. berghei,* NMR	Plasma, mice	Creatine Glucose Gluycerophoshoryl choline Lactate Pyruvate
		Urine, mice	Dimethyl amine Pipecolic acid
(Lian et al., 2009)	*P. falciparum,* NMR	Infected RBCs	Alanine Glycerol Glycerol-3-phosphate
(Lian et al., 2009)	*P. falciparum,* NMR	Infected RBCs	Lactate Pyruvate
(Mehta et al., 2006)	*P. falciparum,* NMR	Infected RBCs	Lactate
(Olszewski et al., 2009)	*P. falciparum,*	Infected RBCs	Arginine

Reference	Species/method	Sample type	Metabolites
	NMR, LC-MS and GC-MS		Ornithine
(Sana et al., 2013)	*P. falciparum,* LC-MS and GC-MS	Infected RBCs	2-Oxoglutarate
			Adenine
			Adenosine mono- & di-phosphate
			Arginine
			Citrulline
			D-Myo-inositol 1,4-bisphosphate
			Fumarate
			Guanosine diphosphofucose
			Inositol phosphate
			Ornithine
			S-Succinyl dihydrolipoamide
			Succinate
(Sengupta et al., 2015)	*P. vivax*, NMR	Urine	3-Methyl-histidine
			Alanine
			Citrate
			Glycine
			Hippurate
			Taurine
(Sengupta et al., 2011b)	*P. vivax,* NMR	Urine	3-Hydroxybutyrate
			Acetate
			Alanine
			Butyrate
			Creatine
			Hippurate
			N-Acetylglutamine
			Ornithine
			Phenylalanine
			Phenylpyruvate
			Phosphocreatine
			Pipecolic acid
			Salicyluric acid
			Tyrosine
(Teng et al., 2009)	*P. falciparum,* NMR, LC-MS and GC-MS	Infected RBCs	Arginine
			Aspartate
			Glutamate
			Lysine
			Phosphocholine
			Phosphoethanolamine
(Tritten et al., 2013)	*P. berghei* , NMR	Urine/mice	Pipecolic acid
			Unknown 1
			Unknown 2

Unknown 1 and Unknown 2 were structurally elucidated using UPLC-TOF-MS/MS and LC-NMR/TOF-MS as 4-amino-1-[3-hydroxy-5-(hydroxymethyl)-2,3-dihydrofuran-2-yl]pyrimidin-2(1H)-one, and 2-amino-4-(–[5-(4-amino-2-oxopyrimidin-1(2H)-yl)-4-hydroxy-4,5-dihydrofuran-2-ylmethyl sulfanyl)butanoic acid, respectively (Tritten et al., 2013).

4.1.5 Aims and objectives

The aim of this chapter is to use the high resolution mass spectrometry methods developed in the previous chapters to identify potential urinary biomarkers of malaria (*Plasmodium falciparum*). The specific objectives required to fulfil this aim are as follows:

- To plan and establish the collections of urine samples from healthy volunteers and malaria patients in a clinical setting where malaria is endemic (Ethiopia)
- To investigate urinary profiles of malaria patients and healthy controls using LC-MS and FIE-MS methods.
- To identify urinary biomarkers, which can differentiate between malaria patients and healthy controls.
- To confirm the identity of some tentatively identified urinary biomarkers of malaria using LC-MS/MS.

4.2 Materials and experimental

4.2.1 Materials and reagents

Ultra-pure LC-MS grade ammonium acetate was supplied from Fluka, Sigma-Aldrich, Netherland. Water (VWR international, EC), formic acid, 0.1% formic acid in water (Sigma-Aldrich, Germany), leucine enkephalin, acetonitrile and methanol (Fisher Scientific, UK) are MS grade and minimum handlings were performed to minimise any possible contamination. Artificial urine was prepared using chemicals supplied from different sources as detailed elsewhere (chapter 2, table 2-2). L-Threonine (Acros Organics, USA), urea (Aldrich-Chemie, Germany) and creatinine (Alfa Aesar, UK) are either HPLC or MS grade. Taurine, succinic acid, Uridine and inosine were purchased from Sigma-Aldrich, Germany.

4.2.2 LC-MS analysis

Chromatography was performed using Accela UHPLC system (Thermo Fisher, USA) on BEH HILIC column (2.1 x 100 mm, 1.7 µm particle size, Waters, USA) as detailed previously in chapter 2, section 2.2.3. Orbital trap mass spectrometer (Exactive-Orbitrap, Thermo Fisher Scientific, USA) was used in simultaneous ESI+ and ESI- modes for LC-MS. The operational parameters were: spray voltage 3.2 kV (ESI+), 2.4 kV (ESI-), capillary voltage 25 V (ESI+), -27 V (ESI), sheath, auxiliary and sweep gas flow rate were: 20, 5 and 5 (arbitrary unit), respectively, for both modes. Capillary and heater temperature were maintained at 350 and 120 ° C, respectively. Data were acquired in full scan mode with resolution 25,000 from *m/z* 60-1000 with 4 Hz scan rate.

4.2.3 FIE-MS analysis

Flow injection ESI-MS (FIE-MS) analysis used the above LC system to inject 10 µL of the sample as detailed in chapter 2, section 2.2.4. The same mass spectrometer was used for FIE-MS analysis with the following parameters: spray voltage 1.5 kV (ESI+), 1.8 kV (ESI-), capillary voltage 25 V (ESI+), -40 V (ESI-), sheath, auxiliary and sweep gas flow rates were: 50, 5 and 5 (arbitrary unit) (ESI+/-), respectively. Capillary and heater temperature were maintained at 250 and 120 °C, respectively, in both modes. Data were acquired in full scan mode with resolution 25,000 from *m/z* 50-1000 with 4 Hz scan rate.

4.2.4 LC-MS/MS analysis

The identity of some urinary metabolites was confirmed using liquid chromatography-tandem mass spectrometry (LC-MS/MS). LTQ Velos mass spectrometer equipped with Accela UHPLC system (Thermo Fischer Scientific, USA) was used for LC-MS/MS analysis. The LC and MS conditions were the same as those described in section 4.2.2. Tentative biomarkers of malaria from orbital trap LC-MS and FIE-MS analyses and their expected retention times were used to generate parent ions list of the analytes for LC-MS/MS. Product ions were detected over a scanning range of *m/z* 50–1000. The MS/MS (MS2) analysis carried out using data global mass list scanning mode with 35 eV collision energy and helium was used as the collision gas. Metabolite identification was confirmed by comparison with pure authentic standards.

4.2.5 Urine samples collection, transport and storage

The study was undertaken by Dr Wakagri Deressa (School of Public Health, Adis Ababa University, Ethiopia), Dr Andrew Fogarty (School of Medicine, University of Nottingham, UK), Prof Gail Davey (Brighton and Sussex Medical School, UK) and Prof David Barrett (School of Pharmacy, University of Nottingham, UK). Ethical approval has been received to conduct the research from the Ethiopian Ministry of Science and Technology, the Institutional Review Board of the College of Health Sciences, Addis Ababa University and University of Nottingham Ethics Committees. 21 Malaria patients (17 male and 4 female) and 25 healthy volunteers (20 male and 5 female) from Ethiopia were enrolled with consent in this study. Four sets of mid-stream urine samples were collected in Ethiopia, Addis Ababa during September-October 2013: control samples from healthy volunteers (C1), a 2nd round collection after 4 weeks (C2), samples from malaria patients infected with *P. falciparum* (PF1) and a 2nd collection after 4 weeks (PF2). The subjects were classified as healthy or infected with *P. falciparum* using microscopy. All urine samples were collected in urinary collection vessels without the use of preservatives and kept at -20 ° C. After transport to UK, samples were aliquoted into cryotubes (6x1.0 mL) and stored in -80 ° C freezer. During the MS analysis the researcher involved was blinded to the study.

4.2.6 Dipstick urinalysis

Simple urinalysis was performed on urine samples in the study using reagent strips for urinalysis (SureScreen Diagnostics, UK) to detect presence of possible RBCs and measure urinary pH, glucose, nitrite, urobilinogen, proteins, specific gravity, ketones and bilirubin. Reagent strip was dipped in each thawed urine sample for few seconds and then reading was taken after 1 min.

4.2.7 Preparation of urine samples and authentic standards

Thawed urine samples of malaria patients and healthy controls were prepared with the optimised urine dilution protocol for LC-MS and direct ESI-MS analyses as described in chapter 2, section 2.2.7. For metabolomics analysis, a pooled QC sample was prepared by mixing 20 µL aliquots taken from each urine sample in the study and was treated the same as described for the samples. Artificial urine was prepared as described in chapter 2, section 2.2.6. Authentic standards for LC-MS/MS were prepared in the concentration range of 23 - 68 µg/mL in MS grade water.

4.2.8 LC-MS and FIE-MS metabolomics analysis

Urine samples from malaria patients and healthy controls were randomised and analysed in a single analytical run using simultaneous positive and negative ion modes with LC-MS and FIE-MS. Six injections of pooled QC sample were analysed at the beginning of the run to equilibrate the column (LC-MS) prior the analysis. Pooled QC and artificial urine sample were interspaced throughout the run to check the stability, robustness, repeatability and performance of the analytical system. Leucine enkephalin, *m/z* 556.2771 (ESI+), 554.2615 (ESI-) was spiked in every sample (0.2 µg/mL) to monitor mass accuracy within each run. Blanks were injected after each sample in FIE-MS analysis to minimise the carryover effect, if any.

4.2.9 Data analysis and metabolite identification

The raw data for LC-MS and FIE-MS were acquired and visualised with Xcalibur v2.1 software (Thermo Scientific, USA). The performance of the analytical methods was validated by monitoring a representative set of urine metabolites in the pooled quality QC as detailed previously in chapter 3, section 3.2.7. In addition, the quality of the datasets obtained was assessed using methodology

proposed by Want et al., (Want et al., 2010) for LC-MS, and Beckmann et al., (Beckmann et al., 2008) for FIE-MS analysis, as detailed in chapter 2, section 2.2.11.

For metabolomics analysis, full datasets from malaria patients and healthy controls from both methods were imported and pre-processed by Progenesis QI software. Pre-processing steps include: peak picking, peak alignments, normalisation and transformation as previously detailed in chapter 2, section 2.2.11. Automatic deconvolution of the extracted *m/z* variables was carried out using Progenesis QI to remove isotopes, adducts, and other confounding peaks resulting from MS detection. Multivariate data analysis using Principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) were used to investigate metabolic changes between all groups datasets using Simca P +14 (Umetrics AB, Sweden). OPLS-DA models based on the datasets from malaria patients and healthy controls were validated using cross-validation and permutation test. Successive external validation using a prediction method based on randomly selected training (50%) and test sets (50%) of samples were performed. The specificity and selectivity of the prediction models were tested using area under the ROC (receiver operating characteristic) curve (AUC). The ions responsible for the class separation in the OPLS-DA models of malaria patients (PF1) and healthy controls (C1) were selected using Variable Importance for the Projection (VIP) and variables loadings plots. Student's *t*-test was performed to test the significant difference of the selected ions between the two groups. Prior to Student's *t*-test, ArcSinh transformation (Jones, 2008) was applied to restore normality. The p-values of the Student's *t*-test were adjusted using false discovery rate (FDR) for multiple testing problem (Figure 4-3). Tentative identification of malaria biomarkers was achieved by interrogating the Human Metabolome Database (HMDB), for possible identification based on accurate mass measurements within 5 ppm mass error (Wishart et al., 2013). In-house urine metabolite database (built using the same LC-MS described in this study) was also used for identification based on accurate mass determination and retention times. Confirmation of some biomarker identities was performed by means of tandem mass spectrometry as detailed in section 4.2.4, using pooled QC sample. The MS/MS spectra of the possible metabolites were compared with MS/MS spectra obtained from authentic standards.

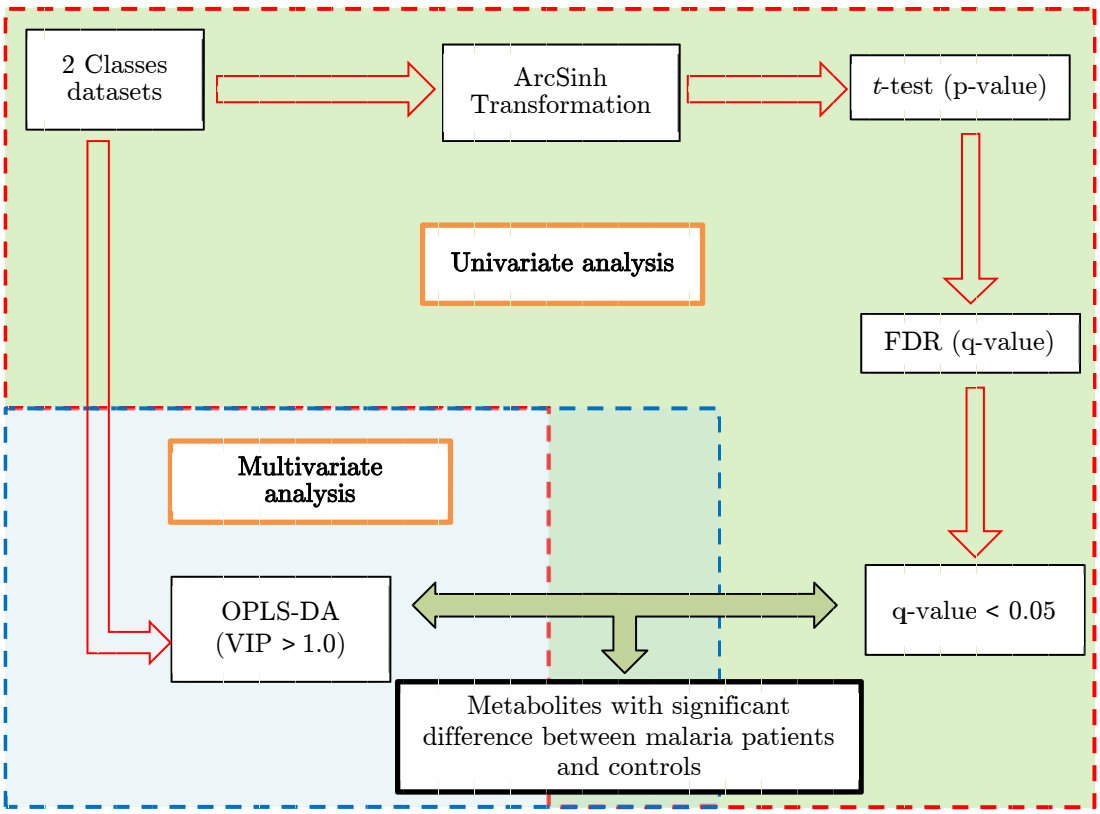


Figure 4-3 Workflow of the selection of potential urinary biomarkers of malaria. Parallel multivariate and univariate analyses were performed to extract the metabolites with significant difference between malaria patients and healthy controls. FDR: false discovery rate, q-value: adjusted p-value using FDR.

4.3 Results

4.3.1 Assessment of urine quality for metabolomics study

Visual examination of urine samples of malaria patients and healthy participants showed that some of the urine samples had turned slightly red in colour. Therefore, in order to check the quality of the urine after transportation, and to determine possible confounding factors that may contribute in the classification of the study groups, urinalysis using simple dipstick was performed. The analysis was carried out to check for possible RBCs that may contribute to the observed red colour in urine and to estimate urinary pH, glucose, nitrite, urobilinogen, proteins, specific gravity, ketones and bilirubin. Table 4-2 presents the results of urinalysis performed on all urine samples in the study.

Table 4-2 Urinalysis dipstick test of malaria patients and healthy controls

	Number of urine samples			
Urinalysis strip test	PF1 (n=21)	PF2 (n=20)	C1 (n=25)	C2 (n=22)
1. Blood				
Normal (< 3 RBCs/μL)	**17**	**19**	**23**	**22**
Trace Haemolysis (> 3 RBCs/μL)	4	1	2	0
2. pH				
6.0 (slightly acidic)	18	12	24	19
7.0 (Neutral)	3	2	0	2
above 7.0 (basic)	0	6	1	1
Normal (5-9)	**21**	**20**	**25**	**22**
3. Glucose				
Normal (< 100 mg/dL)	**20**	**20**	**23**	**20**
Abnormal (> 100 mg/dL)	1	0	2	2
4. Nitrite				
Normal	**21**	**20**	**25**	**22**
Abnormal	0	0	0	0
5. Proteins				
Negative (< 0.2g/L)	**21**	**20**	**25**	**22**
Positive (> 0.2g/L)	0	0	0	0
6. specific gravity				
Normal (1.001-1.035)	**21**	**20**	**25**	**22**
Abnormal (<1.001 and >1.035)	0	0	0	0
7. Ketones				
Negative (normal)	**19**	**20**	**25**	**22**
Positive	2	0	0	0
8. Bilirubin				
Negative (normal)	**21**	**20**	**24**	**21**
Positive	0	0	1	1
9. Urobilinogen				
Normal (0.1-2.0 mg/dL)	**20**	**20**	**24**	**22**
Abnormal (> 2.0mg/dL)	1	0	1	0

PF1: urine samples collected from malaria patients in the beginning, PF2: urine samples collected from malaria patients after 4 weeks, C1: urine samples collected from healthy controls at the beginning, C2: urine samples of healthy controls collected after 4 weeks.

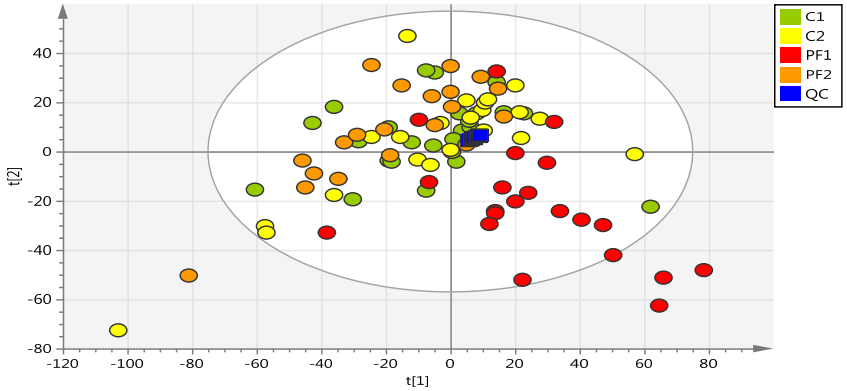
All urine samples showed either zero or only trace amount of RBCs, nitrite, proteins, ketones bilirubin and urobilinogen; the pH and specific gravity of urine samples were within the normal range (Mundt and Shanahan, 2010). This indicates that the red colour observed in some samples was not due to haematuria (i.e. presence of RBCs in urine) and the transportation of the samples did not affect the quality of the urine samples in the study for metabolomics analysis. Cross group examination showed similar ranges of all tests, indicating adequate matching of study groups for metabolomics analysis and therefore, the possible classification would not be as a result of any confounding effects associated with the presence of these metabolites or cells in the samples.

4.3.2 Validation of LC-MS and FIE-MS performance

The analytical performance of LC-MS and FIE-MS for metabolomics analysis was evaluated using pooled quality control (QC) samples. All samples were analysed in a single LC-MS and FIE-MS with pooled QC samples being inserted after each 10 samples in the analysis. In the pooled QC samples datasets the mean RSD% values of selected urine metabolites peak areas (LC-MS) and peak counts (FIE-MS) were 4.6 % (range: 0.1 - 12.1%) and 10.9% (range: 3.8 - 23.3%), respectively. LC-MS retention time shifts were less than 0.06 min (≤ 1% RSDs) and mass accuracy shift was less than 5 ppm in both positive and negative ion modes (for detailed list of the selected metabolites statistics refer to Appendix F). There is no general consensus on accepting analytical technique as suitable for metabolomics analysis. However, for conventional bioanalysis, the Food and Drug Administration (FDA) considers a technique suitable if analyte response ($n \geq 5$) not exceed 15% of the relative standard deviation (RSD %) except for those near the limit of detection where ≤ 20% can be accepted (FDA, 2013). For biomarker discovery an upper limit up to 30% is considered adequate (Gika et al., 2007, Want et al., 2010). The obtained results with the developed LC-MS and FIE-MS were within those limits. In addition, the quality of the datasets obtained with LC-MS was assessed using all peaks present in at least 80% of the pooled QC samples. The RSD% across the mean peak areas was less than 30% for 73% of those peaks, which was lower than the recommended threshold for metabolomics analysis (Want et al., 2010). Unsupervised PCA score plots of the sample sets showed adequate clustering of the QC samples towards the centre of the plots in both LC-MS and FIE-MS, as shown in Figure 4-4. These results validate the LC-MS and FIE-MS analytical performance for both methods. Also, for the purpose

of evaluating the FIE-MS method against the LC-MS method, a direct comparison of the LC-MS pooled QC samples selected peaks was performed. FIE-MS was able to detect 67% of the peaks found by the LC-MS. The above results demonstrated that the FIE-MS consistency, metabolic coverage and reproducibility for the metabolomic analysis of the malaria urine samples were acceptable and as concluded in chapter 2.

1. HILIC LC-MS



2. FIE-MS

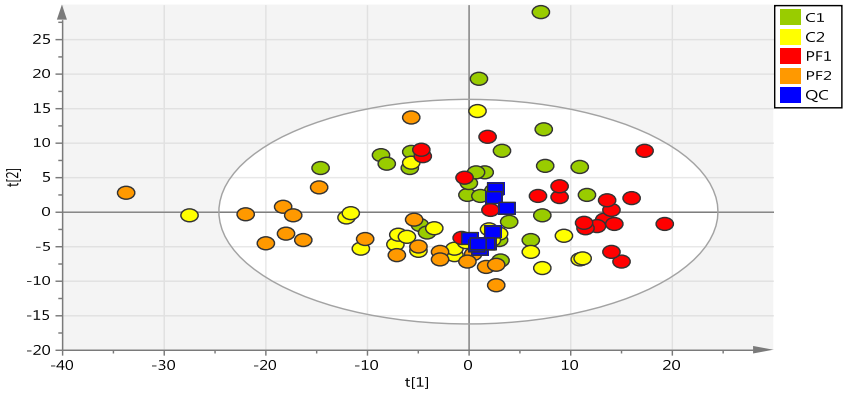


Figure 4-4 PCA score plots overview obtained from all malaria and control urine samples. Control samples: C1 (green circles, n=25), C2 (yellow circles, n=22), malaria samples: PF1 (red circles, n=21) and PF2 (light brown circles, n=20) and pooled QC (dark blue squares) analysed by LC-MS (upper) and FIE-MS (lower).

4.3.3 Urine metabolomics analysis of malaria and healthy controls

LC-MS and FIE-MS analyses

Complete LC-MS and FIE-MS urine metabolomics datasets were acquired for malaria patients (PF1, n = 21), malaria patients after 4 weeks (PF2, n = 20), healthy controls (C1, n = 25) and healthy controls after 4 weeks (C2, n = 22). Typical LC-MS base peak chromatograms obtained from malaria patients and healthy controls urine samples are shown in Figure 4-5. Some of the base peaks are marked to give an intuitive display. Adequate chromatographic separation was attained with most of metabolite peaks eluted within 9 min (Appendix F). The metabolites observed in the chromatograms mainly comprised a range of organic acids, amino acids and pyrimidine nucleosides. Amino acids such as L-alanine, L-tryptophan, tyrosine and phenylalanine were eluted within the retention time range of 5-6.5 min, whereas, organic acids such as 4-aminohippuric acid, homovanillic acid, lactic acid, uric acid and 2-hydroxyisobutyric acid were detected within a wider retention time window (0.5-5 min). Some urinary pyrimidine nucleosides such as cytidine and uridine were eluted within 2.5 min. To give an overview of the differences in the findings between LC-MS and the FIE-MS; summed mass spectra obtained from malaria urine samples analysed by LC-MS and FIE-MS in both positive and negative ion modes are presented in Figure 4-6 and Figure 4-7, respectively. Similar MS spectra were obtained with LC-MS and FIE-MS from the urine samples of malaria patients and healthy controls. The observed metabolic changes associated with malaria patients compared to healthy controls were relatively small; therefore, detailed multivariate analysis was used to separate and classify the clinical groups in the study.

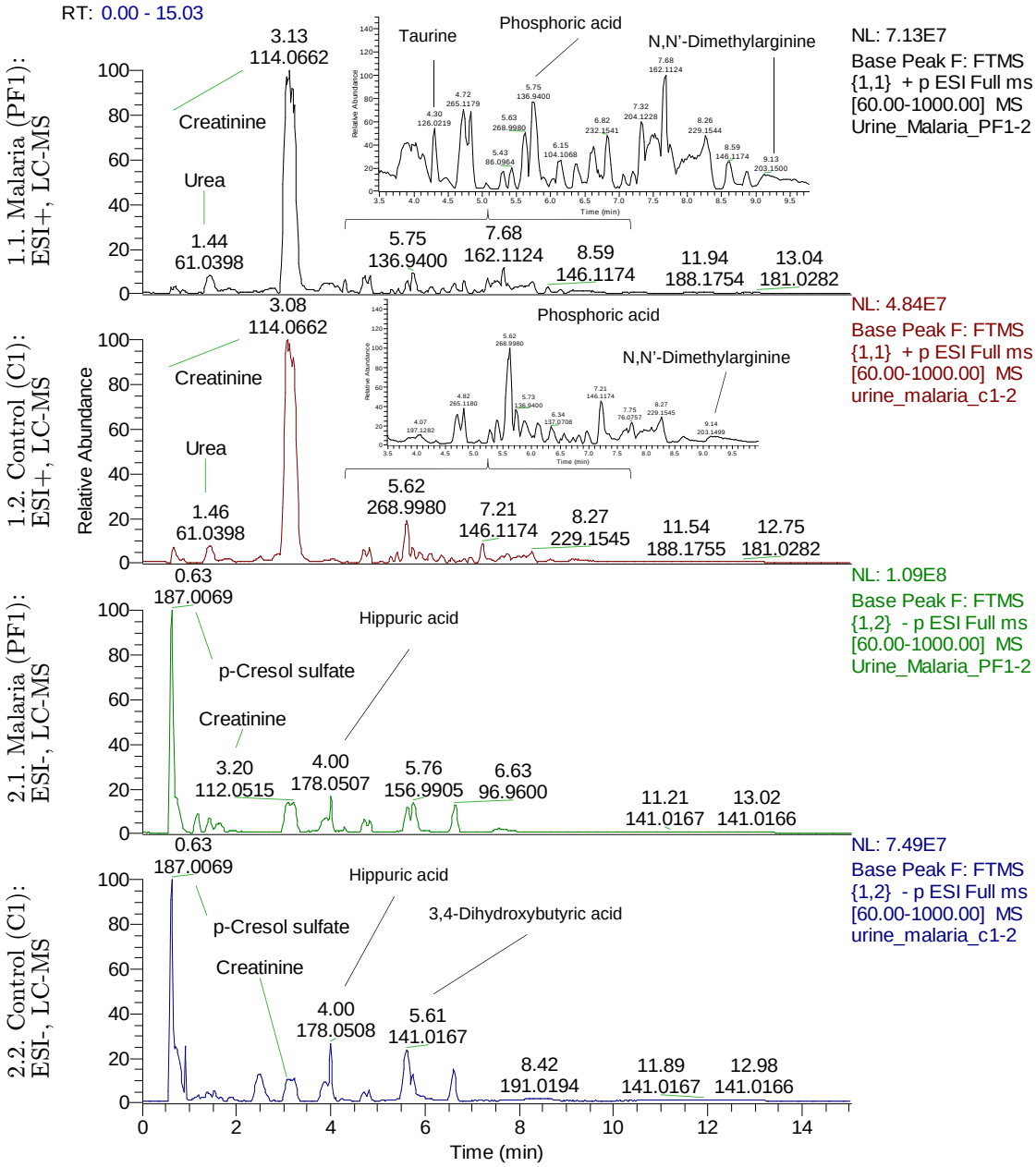


Figure 4-5 LC-MS base peak chromatograms (BPC) obtained from malaria and control urine samples. BPC of (1.1.) malaria patients (ESI+), (1.2) healthy controls (ESI+), (2.1.) malaria patients (ESI-) and (2.2.) healthy controls (ESI-) analysed with LC-MS using HILIC column.

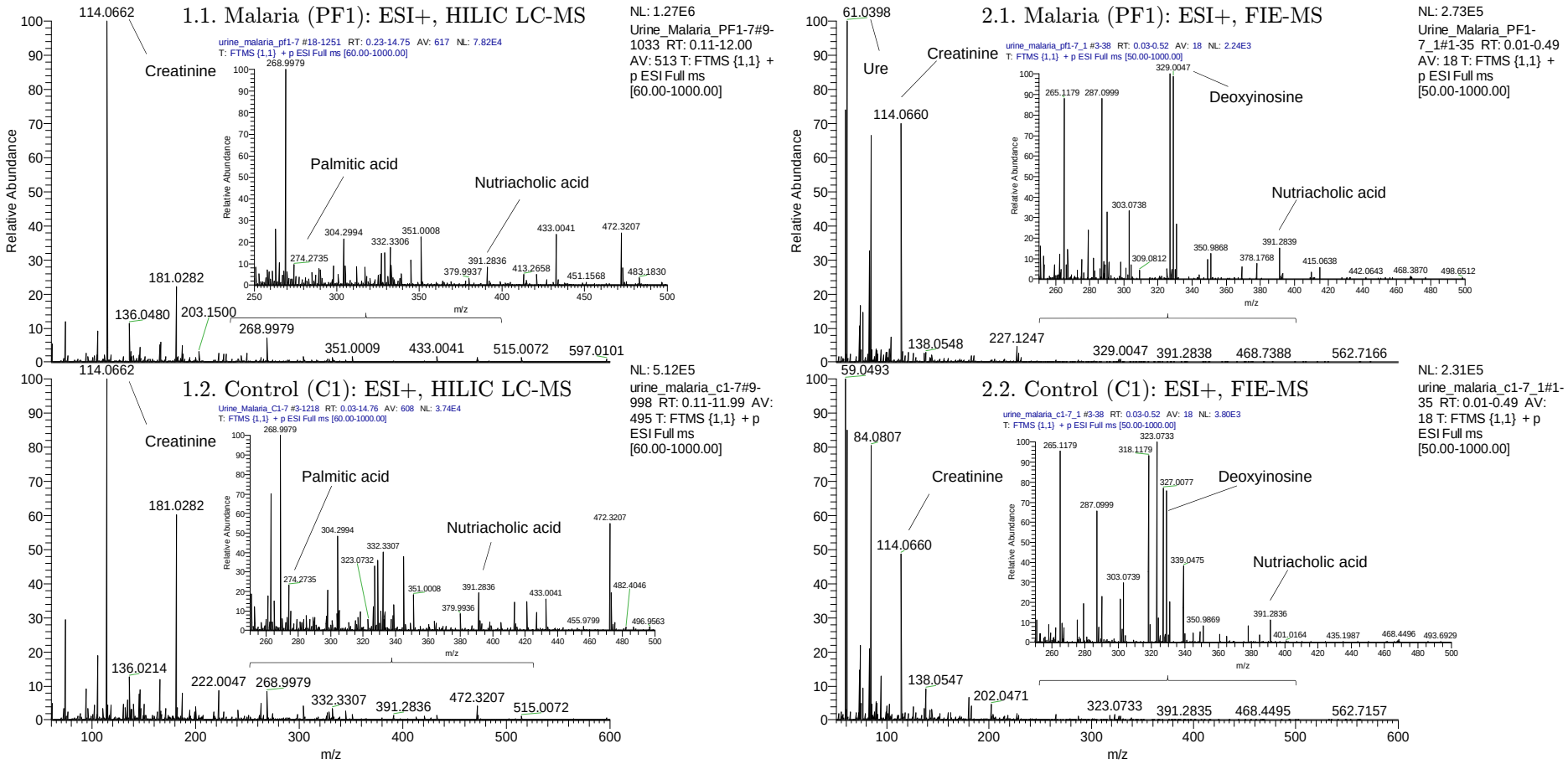


Figure 4-6 Typical urinary MS metabolic fingerprints of malaria patients and healthy controls in positive ion mode. Combined mass spectra of urine sample from: (1.1.) Malaria (PF1): ESI+, HILIC LC-MS, (1.2.) Control (C1): ESI+, HILIC LC-MS, (2.1.) Malaria (PF1): ESI+, FIE-MS and (2.2.) Control (C1): ESI+, FIE-MS.

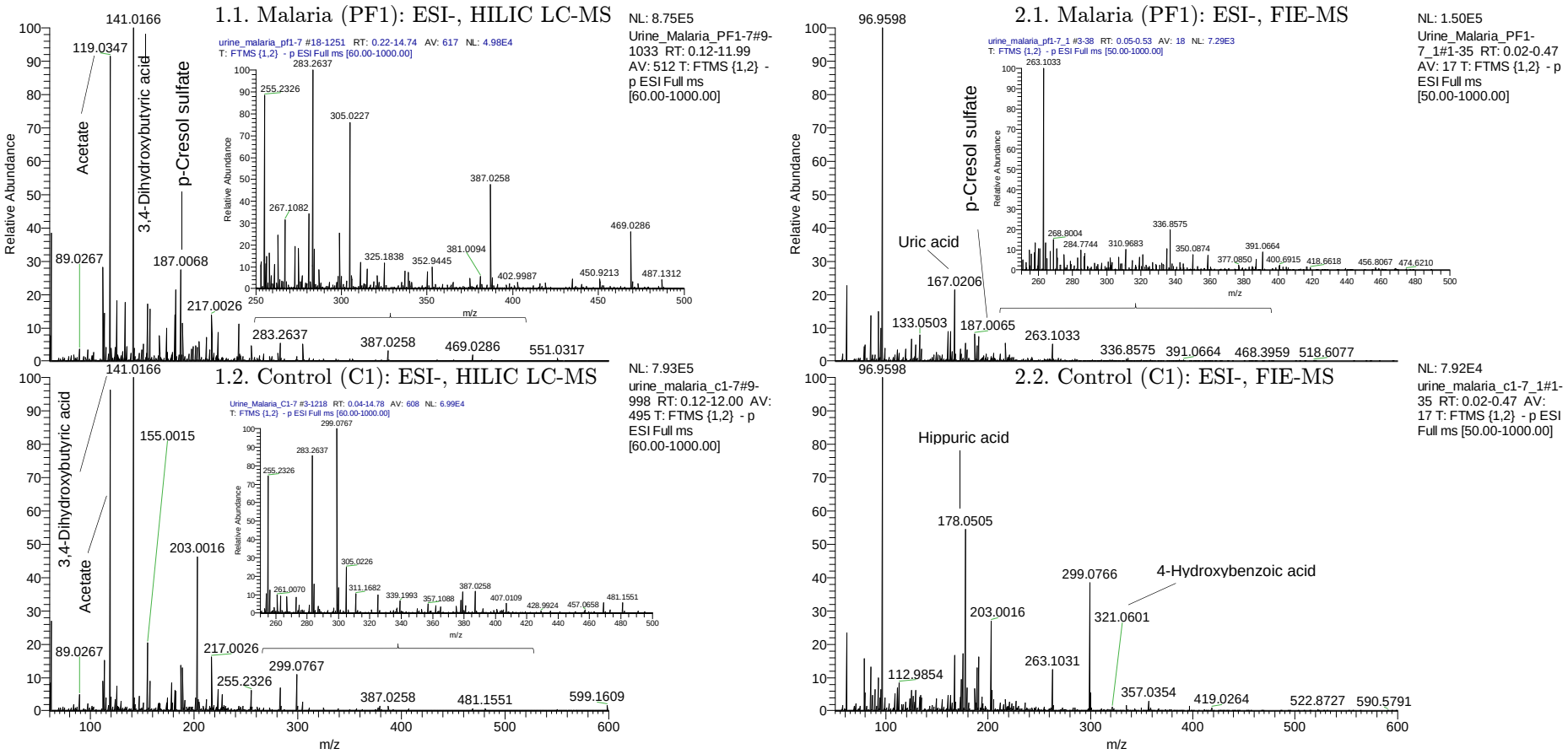


Figure 4-7 Typical urinary MS metabolic fingerprints of malaria patients and healthy controls in negative ion mode. Combined mass spectra of urine sample from: (1.1.) Malaria (PF1): ESI-, HILIC LC-MS, (1.2.) Control (C1): ESI-, HILIC LC-MS, (2.1.) Malaria (PF1): ESI-, FIE-MS and (2.2.) Control (C1): ESI-, FIE-MS.

Multivariate analysis

Complete LC-MS and FIE-MS datasets were acquired from a single analytical run for PF1 (n = 21), PF2 (n = 20), C1 (n = 25) and C2 (n = 22) urine samples. These raw datasets were submitted for peak picking, peak alignment and normalisation using Progenesis QI (Nonlinear-Dynamics, 2014). The datasets of LC-MS and FIE-MS analyses generated 9,744 and 576 variables, respectively (Table 4-3). The normalised datasets were exported to Simca P+14 for multivariate analysis. No differences in urine metabolic profiles could be distinguished between the two control groups (C1 and C2) and malaria samples collected after 4 weeks (PF2) using principal component analysis (PCA) as shown previously in (Figure 4-4). However, some degree of separation and clustering of malaria samples (PF1) from the rest of the groups was observed in both methods. Supervised multivariate analysis using OPLS-DA showed clear separation and clustering of PF1 group from the rest in both LC-MS and FIE-MS methods (Figure 4-8). As the analysis was conducted blindly, the clustering of malaria samples after 4 weeks (PF2) within the control groups region in the models indicates no significant differences between those samples. Later, this observation was confirmed as those patients had received treatment during this period and no longer had malaria. This gives an added confidence of the model for malaria metabolic profiling. Subsequent supervised OPLS-DA models were obtained from malaria patients (PF1) and healthy controls (C1) datasets only, total separation between the two groups were observed with both methods (Figure 4-9).

Validation of OPLS-DA models of malaria patients and healthy controls

Cross-validation

OPLS-DA models of malaria patients (PF1) and healthy controls (C1) were evaluated using cross-validation by monitoring the fitness of model (R^2Y) and predictive ability (Q^2) values. The R^2Y and Q^2 values from FIE-MS were 0.810 and 0.538, respectively, which are comparable to the values of $R^2Y = 0.993$ and $Q^2 = 0.583$ achieved by LC-MS (Table 4-3). Successive permutation test was carried out for each OPLS-DA model to test if the good predictive ability of the model is not due to data overfitting (Eriksson et al., 2006b). The produced Q^2-intercept values of the regression lines of the Y-permuted Q^2 values (Figure 4-10) were less than Q^2 values of the tested OPLS-DA models (Figure 4-9) and they

were intercepted at -0.268 and -0.396 for LC-MS and FIE-MS, respectively, indicating a reliable predictive power of the generated models and it is not because of data overfitting.

External validation

A rigorous testing of the classification performance of the OPLS-DA models of malaria patients and healthy controls was also performed using prediction models based on 50% randomly selected training/test sets (i.e. external validation) as shown in (Figure 4-9, 1.2 and 2.2). Overall correct classification (predictive accuracy) of the OPLS-DA models was 78% for both LC-MS and FIE-MS. The sensitivity (true positive rate) and specificity (true negative rate) of these models were 80% and 77% for LC-MS and FIE-MS, respectively (Table 4-3), indicating a reliable and comparable predictive power of both models. Sensitivity reflects the probability of a positive result from an actual positive clinical outcome, whereas selectivity estimates the probability of a negative result from an actual negative outcome (Xia et al., 2013). The use of sensitivity and selectivity helps to assess the predictability of the OPLS-DA models for clinical use. For ease of interpretation, in clinical practice using a given biomarker for the diagnosis of malaria with a specified decision boundary (e.g. model score or biomarker concentration), a sensitivity of 80% and a specificity of 77% indicate that: (1) there is an 80% chance that the new subject with a test score higher than the given boundary is correctly diagnosed with malaria, whereas, (2) there is a 77% chance that the subject with a test score less than the boundary is correctly classified as healthy.

The predictive boundary of these OPLS-DA models to classify a subject as healthy (negative) or infected with malaria (positive) was based on judicious selection; hence, there will always be some degree of uncertainty attached to the predictability of any reported outcome. Therefore, the robustness of the generated OPLS-DA models was further validated using an unbiased approach, the area under receiver operating characteristic (ROC) curve (AUC). ROC curve is a non-parametric test which is generated by computing the area under the curve of true positive rate (TPR) (sensitivity) against false positive rate (FPR) (1-specificity) of the prediction models. Ideal model gives AUC equals 1.0 (perfect classifier), whereas AUC equals 0.5 is equivalent to randomly classifying subjects as either positive or negative (i.e. the classifier is of no clinical utility). A rough guide based on AUC for assessing the clinical utility of a biomarker is as follows:

0.5–0.6 (fail), 0.6–0.7 (poor), 0.7–0.8 (fair); 0.8–0.9 (good) and 0.9–1.0 (excellent) (Xia et al., 2013). The OPLS-DA models gave comparable results to those obtained using training/test sets with AUC of 0.83 for both LC-MS and FIE-MS (Figure 4-11), indicating that the predictability of the models were robust and valid. These results validate the OPLS-DA models of malaria (PF1) and healthy controls (C1) for diagnosis and biomarker discovery of malaria. Although few variables were detected by FIE-MS compared to LC-MS, the FIE-MS datasets still generated a robust model as LC-MS.

Table 4-3 Multivariate analysis and validation of OPLS-DA models of malaria patients and healthy controls

Description	LC-MS	FIE-MS
1. Peak detected		
ESI+	6,278	248
ESI-	3,466	328
ESI+ and ESI-	9,744	576
2. Cross-validation		
R^2Y	0.993	0.810
Q^2	0.583	0.538
3. Permutation test		
Intercept[1]	-0.268	-0.396
4. External validation: classification (training/test models)		
Malaria patients		
True positive (TP)	8	8
False positive (FP)	2	2
True negative (TN)	10	10
False negative (FN)	3	3
Sensitivity (%) (true positive rate (TPR) =TP/(TP+FP)	80%	80%
Specificity (%) (true negative rate (TNR) =TN/(TN+FN)	77%	77%
Accuracy (%) = (TP+TN)/(TP+FP+TN+FN)	78%	78%
5. Area under receiver operative characteristic (ROC) curve (AUC)		
AUC[2] (TPR vs FPR)	0.83	0.83

[1]The model is considered valid when the regression line of the permuted Q^2 values intercept at, or below zero, [2]AUC: Area under receiver operating characteristic curve, which is the total area under the curve of sensitivity "true positive rate (TPR)" vs 1-specificity "false positive rate (FPR)", ideal model gives AUC = 1 (Eriksson et al., 2006b, Xia et al., 2013).

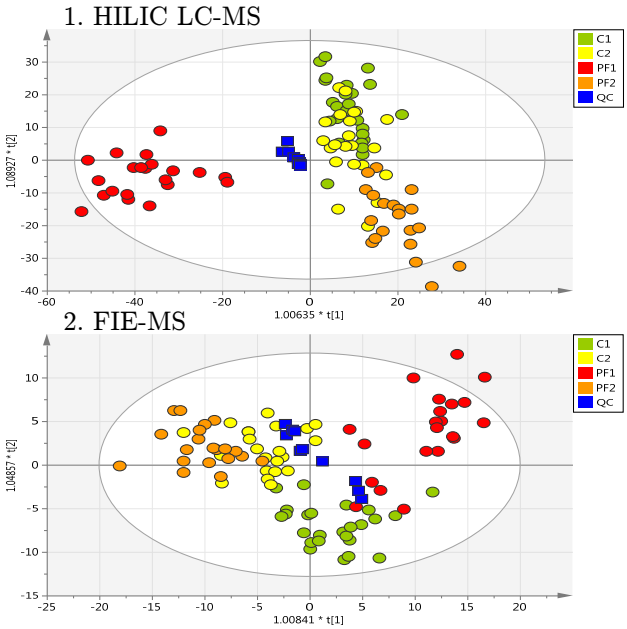


Figure 4-8 OPLS-DA score plots overview obtained from all malaria and control samples. Control samples: C1 (green circles, n = 25), C2 (yellow circles, n = 22), malaria samples: PF1 (red circles, n = 21) and PF2 (light brown circles, n = 20) and pooled QC (blue squares) analysed by LC-MS (upper) and FIE-MS (lower) (combined ESI+/-).

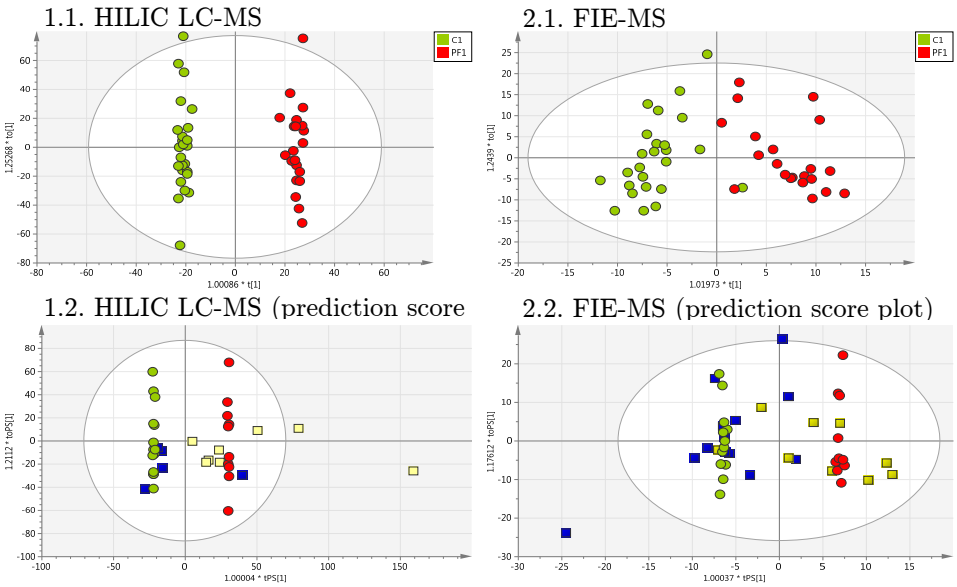
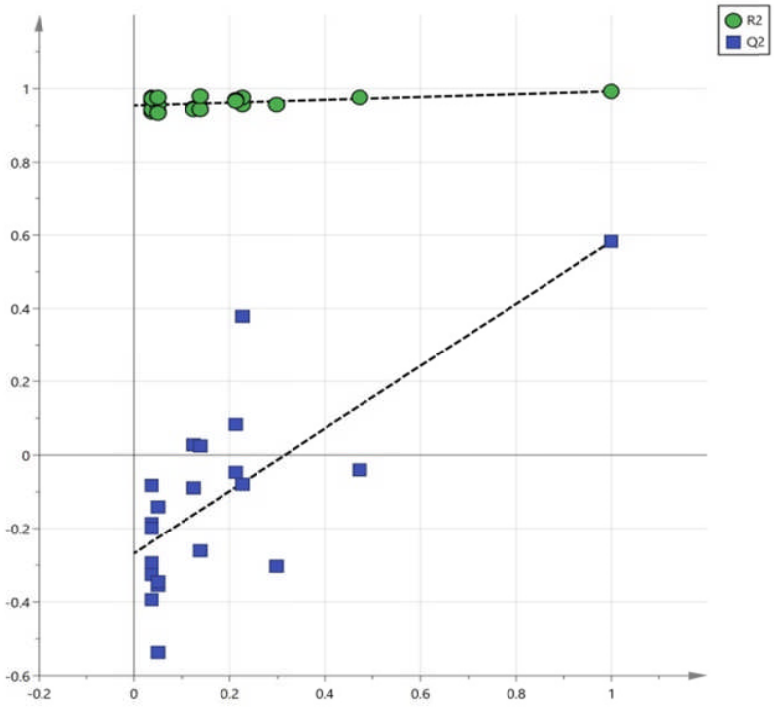


Figure 4-9 OPLS-DA score plots obtained from malaria and healthy controls. OPLS-DA of malaria (PF1, red circles, n = 21) and healthy controls (C1, green circles, n = 25) analysed by (1.1.) LC-MS (upper left) and (2.1.) FIE-MS (upper right). Lower OPLS-DA models represent prediction score plots (test set) obtained from 50% randomly selected samples of PF1 (yellow squares) and C1 (blue squares) analysed by 1.2. LC-MS (lower left) and 2.2. FIE-MS (lower right). Models were built using the complementary malaria (PF1) and control (C1) datasets (training sets) (combined ESI+/-).

1. HILIC LC-MS



2. FIE-MS

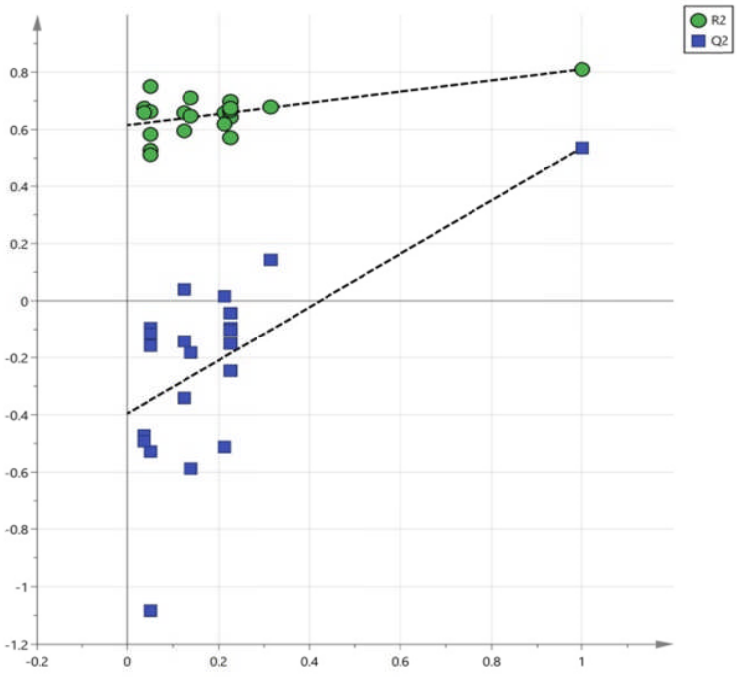
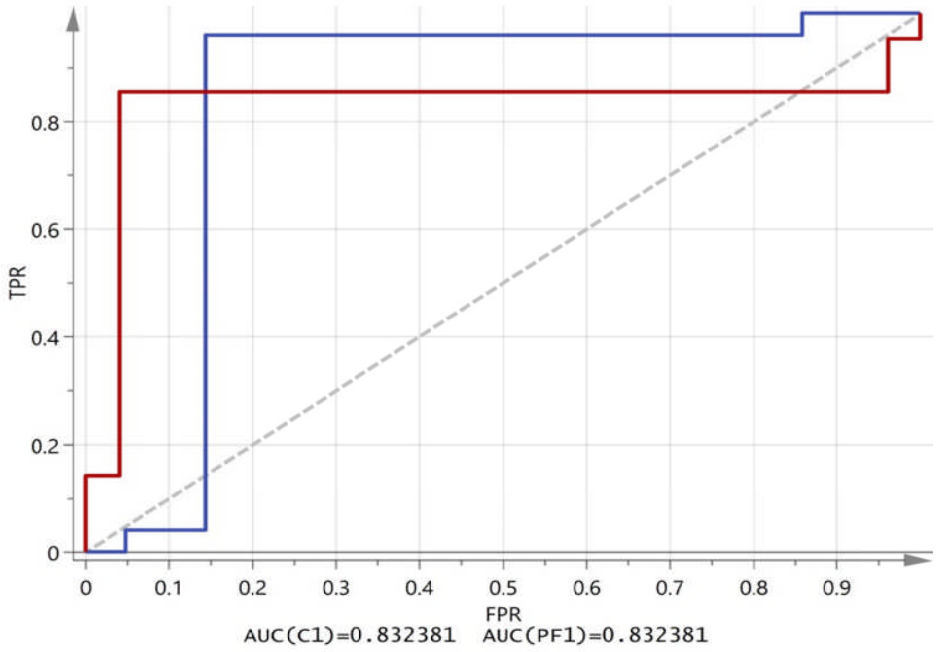


Figure 4-10 Validation of OPLS-DA models of malaria patients and healthy controls using permutation test (n = 46). R^2 is the model fit, and Q^2 is the predictive ability of the model. The Q^2 intercept values of the regression line of randomly permuted Q^2 values were - 0.268 and - 0.396 for HILIC LC-MS and FIE-MS, respectively, indicating that the model is statistically sound and the high prediction ability of the model is not due to data overfitting.

1. HILIC LC-MS



2. FIE-MS

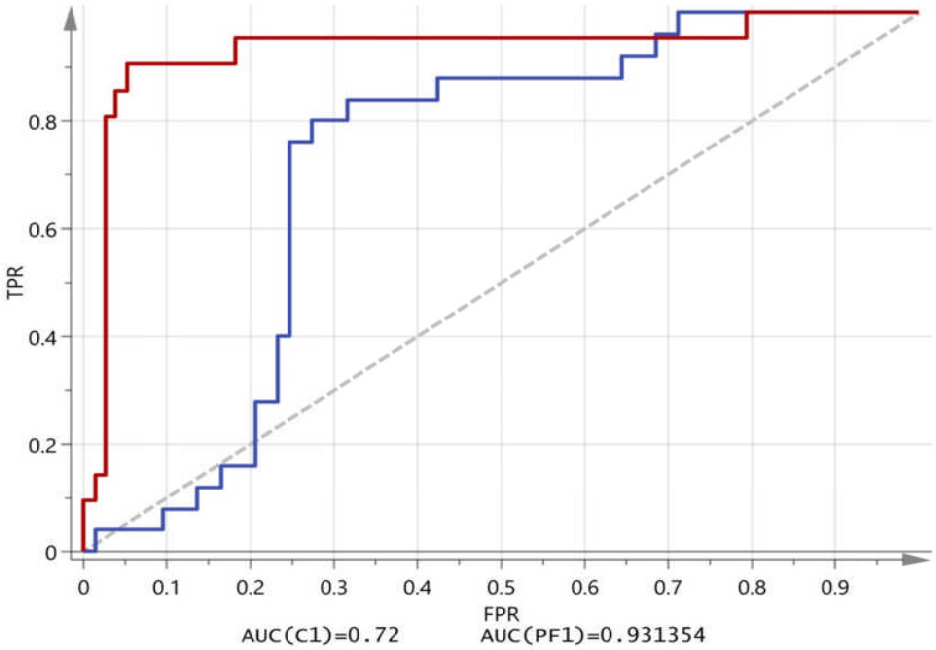


Figure 4-11 ROC curve shows how the sensitivity (true positive rate) and 1-specificity (false positive rate) of the OPLS-DA models change as the classification decision boundary is varied across the range of available biomarker scores. The above graph shows the ROC curve of the OPLS-DA models of LC-MS and FIE-MS of malaria (PF1, brown line) and healthy controls (C1, blue line) datasets. The areas under the ROC curve (AUC) of the OPLS-DA models were 0.83 and 0.83 for LC-MS and FIE-MS respectively, indicating high sensitivity and specificity of the generated models.

4.3.4 Tentative identification of malaria urinary biomarkers

Metabolites responsible for the classification between malaria patients (PF1) and healthy controls (C1) were selected according to VIP values from the OPLS-DA models. Variables (metabolites ions) with VIP score more than 1.0 were chosen. The selected ion intensities across malaria patients and healthy controls samples were subjected to Student's *t*-test and the generated p-values were adjusted using false discovery rate for multiple testing problem. Top ions that differed significantly (q-value ≤ 0.05) have been considered as potential biomarkers, and as a result, 30 and 17 variables were considered as potential urinary biomarkers of malaria from the LC-MS and FIE-MS datasets, respectively. These biomarkers were tentatively identified; the exact mass of the ions (i.e. urinary biomarkers) was used to interrogate the Human Metabolome Database (HMDB) with 5 ppm mass window for possible identification. The list of tentatively identified malaria biomarkers with their molecular weight, fold change, mass error (ppm) and HMDB ID are summarised in Table 4-4. Some of the ions remain unknown due to the limitation of the databases, but they were also listed because of their importance as potential biomarkers of malaria. To provide a more intuitive display of the metabolic difference between malaria patients and healthy controls, box and whisker plots of some representative biomarkers are shown in Figure 4-12 and Figure 4-13. Most of the metabolites obtained by FIE-MS were also detected with LC-MS such as taurine, glycolic acid, pipecolic acid and creatinine. However, some metabolites such as butyric acid and 3-hydroxypicolinic acid were not detected with LC-MS. This may be due to the higher ion suppression effect of urine salts on the ionisation of these metabolites in LC-MS compared to FIE-MS or due to instability of these metabolites during the long run of LC-MS analysis. Urine samples in LC-MS were prepared with 3-fold dilution compared to 10-fold dilution in FIE-MS, therefore, the salt content in LC-MS samples were higher than FIE-MS and subsequently, the ion suppression effect on these metabolites might be higher than FIE-MS. This suggested that, FIE-MS method could be used with LC-MS to improve metabolite coverage, or the optimised sample preparation protocol could be used with the LC-MS to decrease the ion suppression effect of urine salts on the ESI. In addition, FIE-MS detected about 50% of LC-MS urinary biomarkers, thus, the LC-MS is a suitable choice for comprehensive urinary metabolomics study, while FIE-MS as a high-throughput method might be used as a quick approach for clinical screening.

Table 4-4 Tentative identification of urinary biomarkers showing differences between malaria patients and healthy controls

				HILIC LC-MS				FIE-MS			
Metabolite MW (Da)	Tentative identification	HMDB ID	Formula	RT (min)	q-value	Fold Change	Mass error (ppm)	q-value	Fold Change	Mass error (ppm)	Pathway/process
60.0325	Urea	HMDB00294	CH_4N_2O	1.45	2.51E-05	1.70	3.06	2.68E-03	1.39	1.64	Arginine and proline metabolism, urea cycle [1,2]
76.0160	Glycolic acid	HMDB03035	$C_2H_4O_2$	3.88	1.50E-02	1.34	2.72	1.27E-03	1.75	4.39	Glyoxylate and dicarboxylate metabolism [1]
88.0524	Butyric acid	HMDB00039	$C_4H_8O_2$					1.77E-03	2.14	3.00	Butanoate metabolism [1]
101.5395	Unknown			1.51	1.83E-04	8.03					-
113.0589	Creatinine	HMDB00562	$C_4H_7N_3O$	3.11	1.43E-02	1.13	0.50	2.39E-03	1.42	1.45	Arginine and proline metabolism [1]
118.0266	Succinic acid	HMDB00254	$C_4H_6O_4$	3.08	4.99E-02	1.08	1.63	1.27E-03	1.86	2.43	Tricarboxylic acid cycle, Propanoate metabolism [1,2]
119.0582	L-Threonine	HMDB00167	$C_4H_9NO_3$	6.53	4.08E-04	3.02	1.33				Aminoacyl-tRNA biosynthesis [1,2]
125.0147	Taurine	HMDB00251	$C_2H_7NO_3S$	4.29	2.96E-05	5.39	2.20	1.45E-05	14.50	0.80	Taurine metabolism, Nitrogen metabolism [1]
128.0837	1,3-Diacetylpropane	HMDB29165	$C_7H_{12}O_2$	6.46	7.62E-04	35.89	1.81	1.77E-03	3.68	2.54	Polyamine metabolism [2]
129.0790	Pipecolic acid	HMDB00070	$C_6H_{11}NO_2$	2.18	3.28E-02	1.34	0.11	1.59E-02	2.78	0.89	Lysine degradation [1]
130.1106	N-Acetylputrescine	HMDB02064	$C_6H_{14}N_2O$	6.62	3.23E-04	1.69	1.45	1.59E-02	3.31	1.81	Arginine and proline metabolism [1,2]
139.0269	3-Hydroxypicolinic acid	HMDB13188	$C_6H_5NO_3$					1.01E-03	-3.42	2.37	Tryptophan catabolism [1]
139.0633	3,4-Dihydroxybenzylamine	HMDB12153	$C_7H_9NO_2$	3.25	7.57E-04	2.00	2.54	5.17E-03	1.99	2.55	Not available
162.0528	3-Hydroxyadipic acid	HMDB00345	$C_6H_{10}O_5$	1.50	2.66E-05	3.40	0.46	4.99E-02	1.24	4.16	Fatty acid metabolism [1]
172.0736	2-Octenedioic acid	HMDB00341	$C_8H_{12}O_4$	6.46	1.52E-04	7.20	1.84	4.79E-03	2.74	2.12	Fatty acid metabolism [1]
174.1546	N-Acetylasparagine	HMDB06028	$C_6H_{10}N_2O_4$	3.61	7.33E-06	2.27	0.86				Asparagine catabolism [1]
187.1685	N-Acetylspermidine	HMDB01276	$C_9H_{21}N_3O$	11.90	9.18E-05	1.72	1.09	2.39E-03	4.17	1.48	Polyamine metabolism [1,2]
188.0797	N-Acetylglutamine	HMDB06029	$C_7H_{12}N_2O_4$	3.59	2.30E-04	3.33	2.20				Not available
188.1525	Trimethyl-L-lysine	HMDB01325	$C_9H_{20}N_2O_2$	7.41	9.44E-07	2.60	1.89	2.68E-03	5.84	-1.89	Carnitine biosynthesis [1]
193.0582	Unknown			1.51	8.03E-07	2.52					-
195.0532	3-Hydroxyhippuric acid	HMDB06116	$C_9H_9NO_4$	4.05	2.05E-03	-3.00	2.98	9.06E-04	-2.32	3.07	Fatty acid metabolism [1]
208.0955	Unknown			6.80	1.46E-03	4.44					-
209.0434	Unknown			1.41	3.80E-07	2.18					-
210.0528	Vanilpyruvic acid	HMDB11714	$C_{10}H_{10}O_5$	1.46	1.56E-03	4.36	0.90				Vanilactic acid biosynthesis [1]
212.0794	Unknown			3.88	1.43E-05	3.36					-
217.1063	Alanyl-Glutamine	HMDB28685	$C_8H_{15}N_3O_4$	6.87	2.45E-05	5.45	2.19				Protein catabolism [1]
244.0694	Uridine	HMDB00296	$C_9H_{12}N_2O_6$	1.64	6.83E-07	1.63	0.57				Pyrimidine metabolism [1,2]
252.1222	Prolyl-Histidine	HMDB29019	$C_{11}H_{16}N_4O_3$	2.93	8.29E-05	5.86	1.23				Protein catabolism [1]
268.0808	Inosine	HMDB00195	$C_{10}H_{12}N_4O_5$	1.42	2.17E-06	2.62	0.82				Purine metabolism [1,2]
281.1124	1-Methyladenosine	HMDB03331	$C_{11}H_{15}N_5O_4$	5.68	8.07E-05	1.51	1.34	9.60E-03	3.16	1.76	Not available
282.0961	1-Methylinosine	HMDB02721	$C_{11}H_{14}N_4O_5$	1.78	1.26E-03	2.15	1.26				Not available
285.0961	N4-Acetylcytidine	HMDB05923	$C_{11}H_{15}N_3O_6$	1.56	5.76E-05	3.37	0.01				Degradation of transfer ribonucleic acid (tRNA) [1]

RT: retention time, MW: molecular weight, q-value: is the adjusted Student's *t*-test p-value using false discovery rate (FDR), the positive value of fold change means a higher level of metabolite in malaria patients compared to healthy controls, whereas the negative value represents a lower level of metabolite. Pathways existence: [1] human and [2] *P. falciparum*.

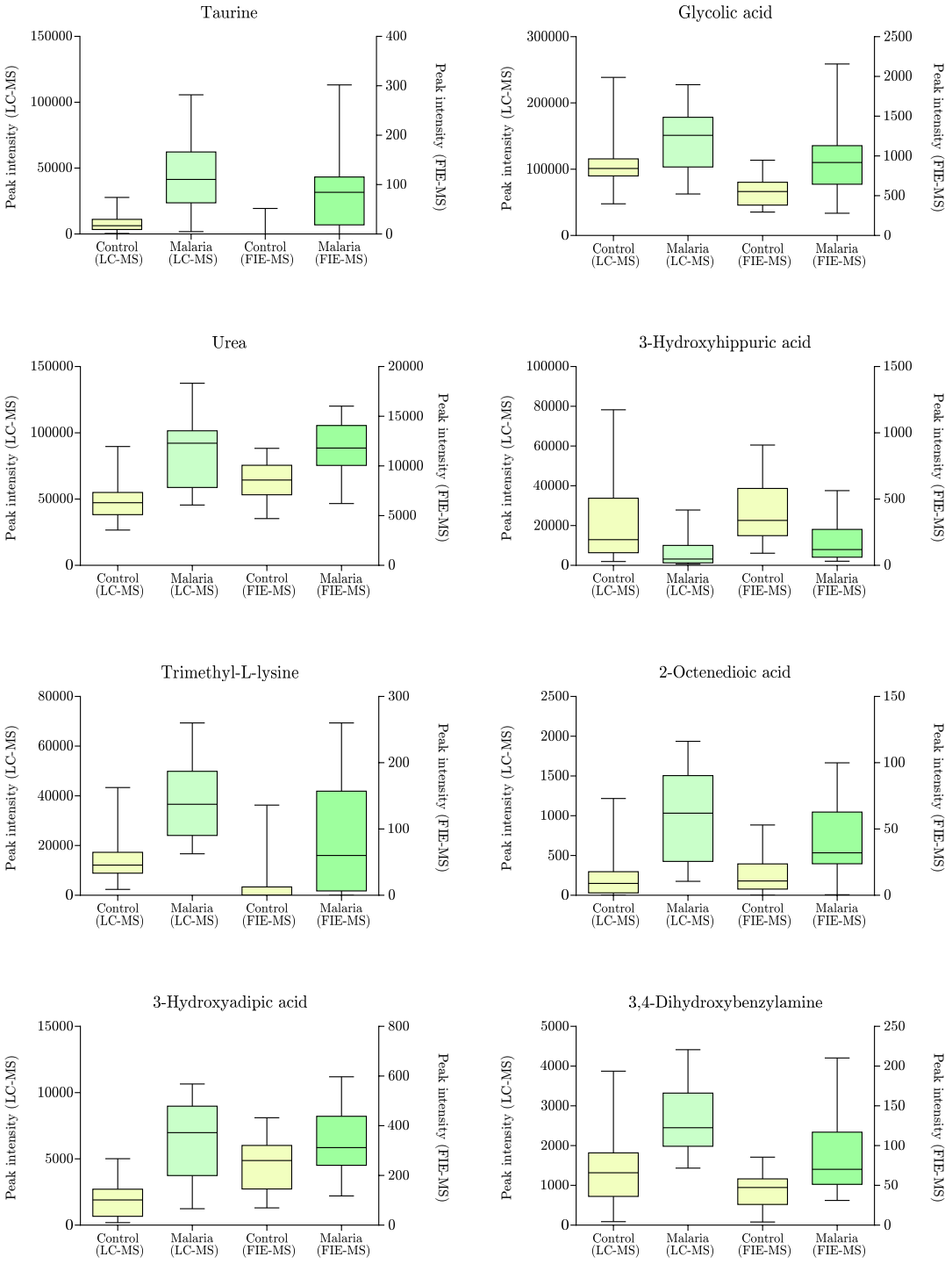


Figure 4-12 Box and whisker plots demonstrated significantly altered levels of representative urinary biomarkers ($q < 0.05$) between malaria patients and healthy controls using LC-MS and FIE-MS. The boxes extend from the 25th to 75th percentiles in the biomarker intensity distribution, while the 50th percentile (the median) is the black horizontal line in the box. The whiskers extend from the upper and lower fence values.

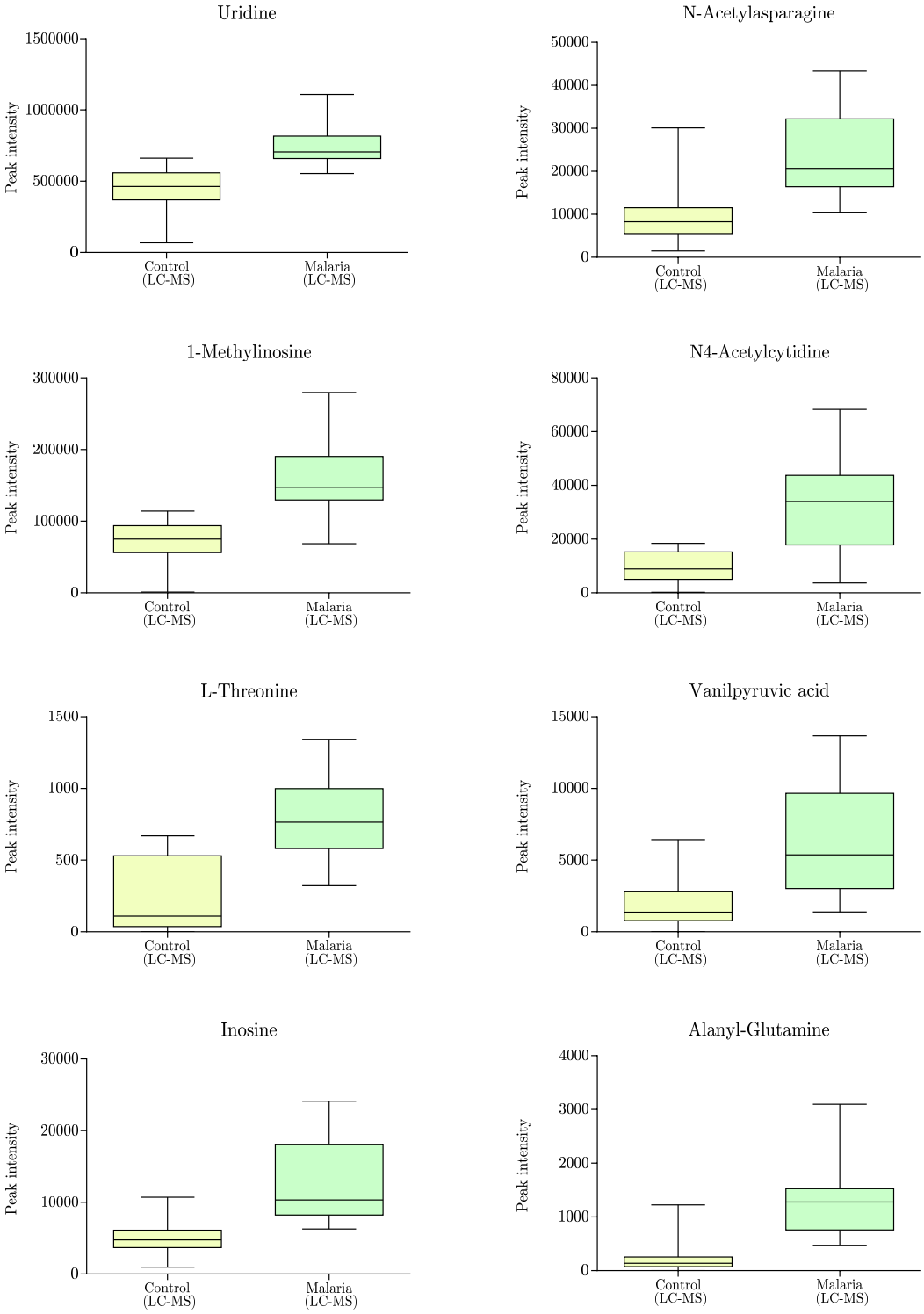


Figure 4-13 Box and whisker plots demonstrated significantly altered levels of representative urinary biomarkers ($q < 0.05$) between malaria patients and healthy controls using LC-MS. The boxes extend from the 25^{th} to 75^{th} percentiles in the biomarker intensity distribution, while the 50^{th} percentile (the median) is the black horizontal line in the box. The whiskers extend from the upper and lower fence values.

4.3.5 Identity confirmation of urinary biomarkers of malaria

Initial confirmation of identity started with comparing the retention times and accurate masses of metabolites (i.e. malaria biomarkers) in the urine sample with those of the standards (data not shown). Then, the identity of some of urinary malaria biomarkers were confirmed by means of tandem mass spectrometry, in which pure authentic standards were injected along with the sample and their MS/MS spectra were compared. Absolute identity confirmation using LC-MS/MS was carried out for taurine, L-threonine, inosine, uridine, creatinine, urea and succinic acid against their authentic standards. The MS/MS spectra of the selected metabolites matched their standards, thus, confirming their identity as illustrated in Figure 4-14 and Figure 4-15. However, urea (*m/z* 61.0398) was confirmed by comparing only its retention time and exact mass with the standard; this is because the expected MS/MS fragments of urea are lower in mass than the minimum mass range of the MS/MS instrument used in this study.

Taurine:

In LC-MS/MS analysis, taurine standard ($C_2H_7NO_3S$) fragmented to produce ions at *m/z* 107.00 and 94.98, which is attributed to the fragments $[C_2H_5NO_2S]^-$ and $[CH_3O_3S]^-$, respectively. Extra loss of the side chain methyl group $[CH_3]$ or $[CH_2]$ from the $[CH_3O_3S]^-$ fragment produced ions at *m/z* 79.91 and 80.97, respectively (Chaimbault et al., 2004, Tang et al., 2014). The tentatively identified taurine ion in the sample fragmented to produce those peaks, thus, confirming the identity of the compound as illustrated in Figure 4-14.

Inosine:

Inosine is relatively large molecule (MW 268.0808) which is readily fragmented in LC-MS/MS to produce many peaks. However, characteristic fragment peaks were observed at *m/z* 250.01 (postulated as $[M\text{-}H_2O\text{-}H]^-$), 230.82, 204.96 (postulated as $[C_9H_9N_4O_2]^-$, 183.96 and 135.00 (postulated as $[C_5H_3N_4O]^-$. The presence of these fragment ions in the MS/MS spectrum of the sample (Figure 4-14) strongly suggests that the putatively identified metabolite is inosine.

Lactic acid and glycolic acid

A metabolite at *m/z* 135.0295 (ESI-) found with a significant difference between malaria patients and healthy controls was firstly identified as lactic acid. The retention time and MS/MS spectrum of lactic acid standard did not match that of the sample (data not shown), therefore the tentative identification of the metabolite as lactic acid was ruled out. Nonetheless, the matching of the exact mass of the unknown metabolite still suggests that possibly glycolic acid might be the one present in the sample according to the human metabolome database

(HMDB) (Wishart et al., 2013). However, the confirmation of the identity of this metabolite requires further LC-MS/MS verification with glycolic acid authentic standard which was beyond the scope of this thesis.

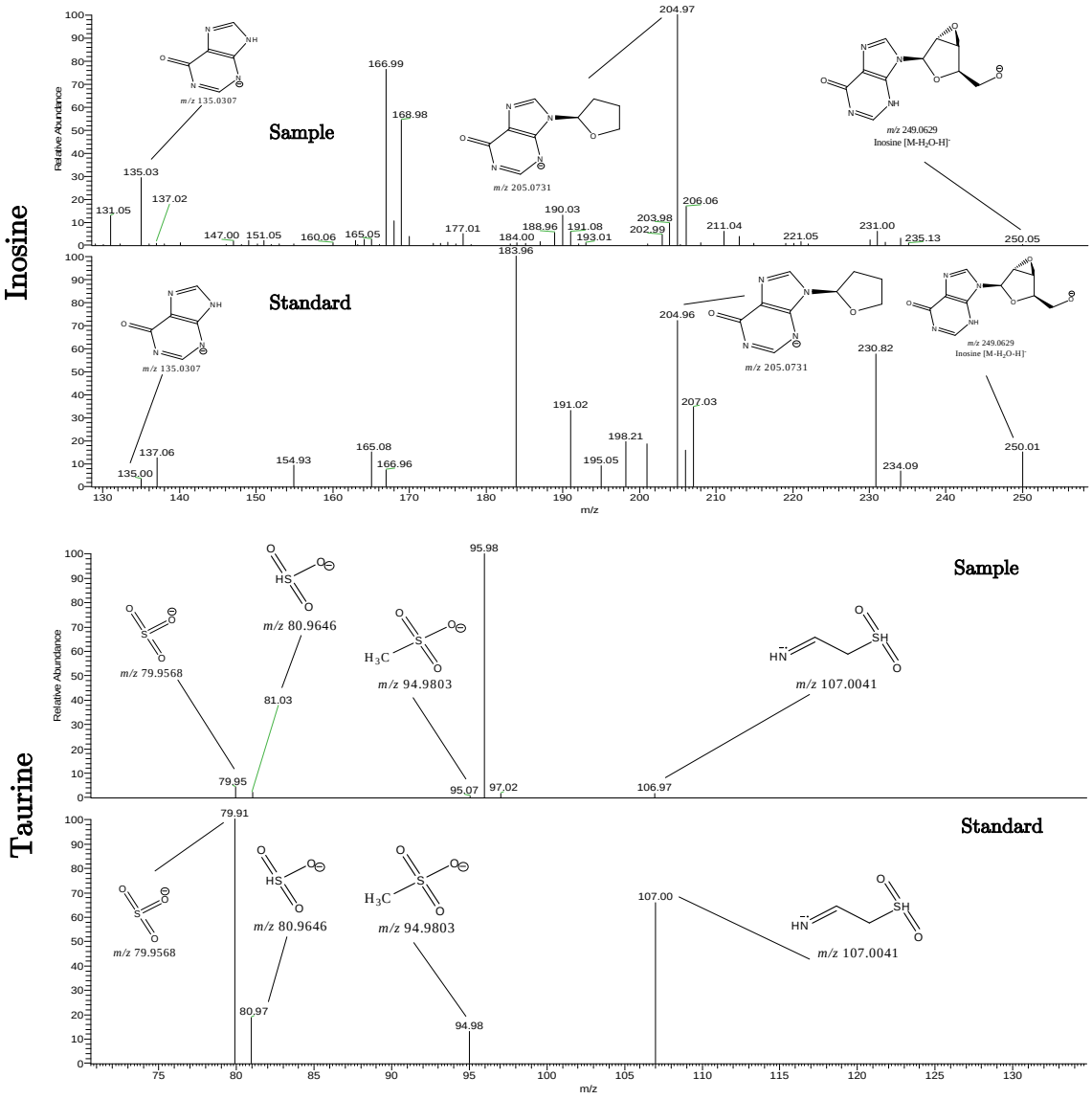


Figure 4-14 MS/MS spectra comparison of tentitively identified inosine and taurine in the urine sample of malaria patients against their authentic standards. The upper two spectra present the matching of ionsine fragnment peaks (standard) against the fragment peaks obtained from metabolite in the sample (sample), while the lower two spectra present those of taurine.

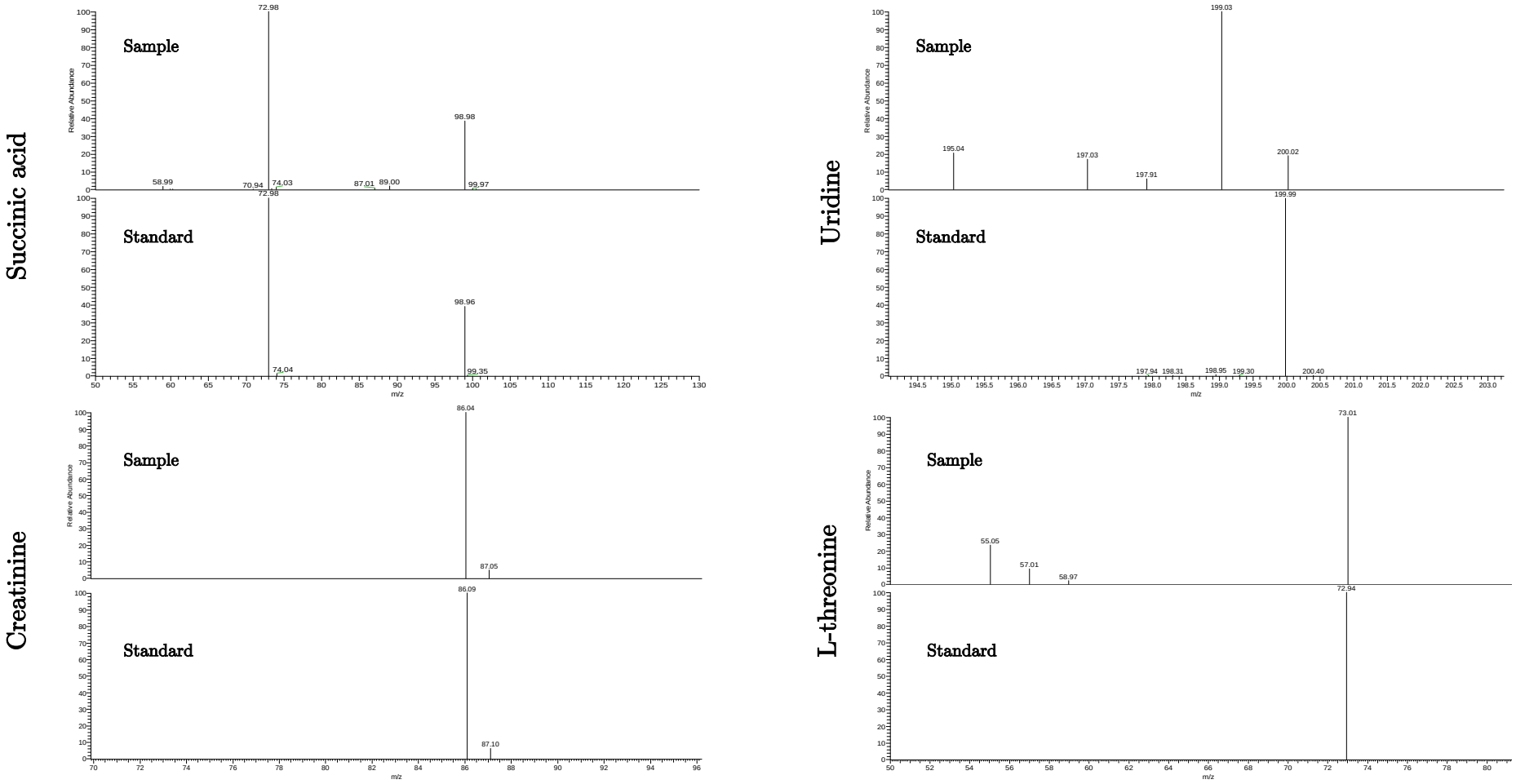


Figure 4-15 MS/MS spectra confirmation of tentitively identified succinic acid, creatinine, uridine and L-threonine in the urine sample of malaria patients against their authentic standards. The spectra present the matching of standards fragnment peaks (standard) against the fragment peaks obtained from the metabolites in the sample (sample).

4.4 Discussion

OPLS-DA models generated from urinary datasets from LC-MS and FIE-MS analyses showed clear segregation between malaria patients and healthy controls. Urine samples of malaria patients showed significantly increased levels of metabolites such as amino acids (taurine, L-threonine, alanine), dipeptides (alanyl-glutamine, prolyl-histidine), amino acids metabolites (pipecolic acid, N-acetylasparagine, N-acetylglutamine, trimethyl-L-lysine, 3-hydroxypicolinic acid, 3,4-dihydroxybenzylamine), a tricarboxylic acid (TCA) cycle intermediate (succinic acid), acetylated polyamines metabolites (N-acetylputrescine, N-acetylspermidine and 1,3-diacetylpropane), fatty acids urinary metabolites (2-octenedioic acid, 3-hydroxyadipic acid), purine and pyrimidine metabolism intermediates (inosine, 1-methyladenosine, 1-methylinosine, N-acetylcytidine, uridine), urea and creatinine, when compared with healthy controls, signalling the possibility of altered metabolic pathways and biological functions of these metabolites in malaria. The identity of taurine, uridine, succinic acid, L-threonine, creatinine and inosine were confirmed by comparing their MS/MS spectra with those obtained from pure authentic standards using LC-MS/MS. The increased level of succinic acid, taurine, alanine and pipecolic acid in malaria patients was consistent with previously reported studies (Sengupta et al., 2015, Sengupta et al., 2011b, Gardner et al., 2002, Li et al., 2008, Tritten et al., 2013). However, the altered level of metabolites such as 1,3-diacetylpropane (acetylated polyamine), N-acetylspermidine and N-acetylputrescine in urine of malaria patients compared to healthy controls were found for the first time, suggesting that they may provide surrogate urinary biomarkers of malaria.

There is a constant dynamic metabolic interplay between the host and the parasite during the course of infection which may perturb the biochemical profiles of both, the parasite and the host. The parasite invasion induces a constellation of responses by the host which are collectively known as “acute-phase responses” (Dinarello, 1984). This phase is characterised by metabolic, immunological, neuro-endocrine and behavioural alterations to the host (Kushner, 1988). Hence, the altered level of metabolites observed in malaria patients compared to healthy controls might be a direct signal of parasite activity (parasite-specific metabolites) or from the host as a response to the effect of the parasite on different organs during the acute-phase response. Moreover, the parasite during the course of infection releases certain metabolites which induce

the host metabolic response and subsequently metabolites of parasite-specific molecules may accumulate in different body fluids. The metabolites directly related to the parasite are good biomarker candidates of the infection; however, their altered levels in different body fluids depend on the level of parasitemia and the severity of the disease and they might not be detected in early stage of the disease (Surowiec et al., 2015). Therefore, biomarkers related to organ failure rather than specifically for the parasite have the potential to be more suited for the early diagnosis of malaria from different body fluids such as urine.

Changes in the glycolysis pathway

No significant difference in the level of urinary lactic acid, a marker of enhanced glycolysis pathway of the parasite (Delic et al., 2010, Ghosh et al., 2012), was observed between malaria patients and healthy controls. This finding was consistent with previous clinical studies, in which no significant difference in the level of lactic acid was found in urine and plasma of patients infected with *P. vivax* (Sengupta et al., 2015, Sengupta et al., 2011b) and *P. falciparum* (Gardner et al., 2002), respectively, compared to healthy controls. However, from in vitro study an elevated level of lactic acid in *Plasmodium*-infected RBCs compared to non-infected ones has been previously reported in the literature as a result of enhanced glycolysis by the parasite during the infection (Mehta et al., 2006). Also, in rodent models, blood lactic acidosis was reported as a result of enhanced parasitic glycolysis in severe malaria (Basant et al., 2010). *Plasmodium spp.* as an intracellular parasite has a reduced metabolic capacity compared to higher organisms and depends mainly on the continuous supply of nutrition from the host. Therefore, *Plasmodium* parasite depends mainly on glucose supplied by the host as a major energy source to meet the essential requirements of proliferation and growth (Lakshmanan et al., 2011, Olszewski and Llinas, 2011, Olszewski et al., 2009). On the contrary, human erythrocytes as non-proliferative cells have modest energetic needs which are mainly provided through glucose fermentation. Most of glucose absorbed by the parasite enters *Plasmodium* anaerobic Embden–Meyerhof–Parnas (EMP) pathway of glycolysis and is incompletely oxidised to lactic acid (Delic et al., 2010, Ghosh et al., 2012). Lactic acid is then converted to pyruvate and may be deaminated to alanine (Sengupta et al., 2011a).

An increased level of alanine was observed in the urine of malaria patients compared to healthy controls, suggesting that lactic acid was converted to alanine and may provide evidence of enhanced glycolysis pathway of the parasite

during the course of infection. The increased level of alanine in malaria was previously reported in the literature. Surowiec et al. reported significantly altered level of alanine but not lactic acid in the plasma of *P. falciparum* patients compared to healthy controls, which was linked to hypoglycaemia observed in malaria (Surowiec et al., 2015). This result is quite consistent with the above finding. However, alanine, an amino acid, is also an essential precursor for gluconeogenesis (production of glucose from non-carbohydrate source) in the liver and its elevated level may also be an indication of an impaired hepatic gluconeogenesis or a perturbed amino acid metabolism as a result hepatic dysfunction in malaria. The increased level of alanine in malaria patients might also be a result of different diet, life-style and genetics of malaria patients compared to healthy controls, however, this seems unlikely as alanine level was found elevated in a well-controlled rodent model infected with *Plasmodium* parasite (Sengupta et al., 2011a).

Enhanced activity of tricarboxylic acid (TCA) cycle in malaria

The level of succinate, a human and parasite TCA cycle intermediate, was significantly elevated in the urine of malaria patients compared to healthy controls, indicating an enhanced metabolic TCA cycle activity by the parasite during the course of infection. The increased level of succinate in malaria patients may also indicate an increased TCA cycle activity by the host to meet the increased energy demand caused by the infection, indicating a perturbed energy metabolism in malaria. The increased level of succinate in *Plasmodium* infection was consistent with previous in vitro studies (Sana et al., 2013, Teng et al., 2009). Recently, Sengupta et al., reported an altered level of succinate in the urine of *P. vivax* infected patients (Sengupta et al., 2015, Sengupta et al., 2011b). This result was consistent with the above finding, suggesting succinate as a potential urinary biomarker of malaria.

As malaria parasite depends mainly on anaerobic glycolysis, it has dampened tricarboxylic acid (TCA) cycle that is largely dissociated from carbohydrate metabolism (Olszewski and Llinas, 2011). This is mainly correlated to the absence of some fundamental enzymes such as malate dehydrogenase from the mitochondrion, which are required to interlink glycolysis and TCA (Danks et al., 1975). Alternatively, glutamine and glutamate were found to be the major metabolites that entered the TCA cycle of the parasite as a carbon source to produce acetyl-CoA moiety (Plecko et al., 2000). No significant difference in the urinary levels of glutamine and glutamate were observed between the malaria

patients and healthy controls. This might be attributed to the fact that these metabolites were supplied by the host and, therefore, their expected depletion in the biofluids of malaria patients may be masked due to their abundant presence in the host serum. Sana and co-workers reported increased level of glutamine and glutamate in erythrocytes infected with *P. falciparum* compared to non-infected ones. On the contrary, clinical studies reported no significant difference in the urinary level of glutamine and glutamate (Sengupta et al., 2015, Sengupta et al., 2011b). A recent study reported an elevated level of glutamine in the plasma of *P. falciparum* infected patients, which was observed only at high level of parasitaemia in malaria patients (Surowiec et al., 2015). This results support the finding in this study and demonstrating that metabolites due to parasite signals are difficult to detect in body fluids and depend mainly on the degree of parasitemia and severity of the disease.

Perturbations of amino acids metabolism during malaria

Abnormal levels of urinary amino acids, amino acid metabolites and dipeptides during the course of *P. falciparum* infection may also indicate a perturbed amino acids metabolism and protein catabolism as a result of the enhanced parasite activity or due to host response to the infection. High levels of alanyl-glutamine, prolyl-histidine, pipecolic acid (metabolite of lysine), trimethyl-L-lysine (methylated derivative of lysine), alanine, L-threonine, N-acetylglutamine (metabolite of glutamine), N-acetylasparagine (metabolite of asparagine) and 3-hydroxypicolinic acid (catabolite of tryptophan) were observed in urine of malaria patients compared to healthy controls. This finding was consistent with previously reported studies, in which abnormal levels of amino acids were found in urine and plasma of patients infected with *P. vivax* (Sengupta et al., 2015, Sengupta et al., 2011b) and *P. falciparum* (Surowiec et al., 2015), respectively.

Malaria parasites are incapable of *de novo* amino acid synthesis and rely instead on scavenging them from haemoglobin catabolism and uptake from host serum, which generates an excess of amino acids in the bloodstream of the host (Gardner et al., 2002). Significantly altered levels of urinary alanyl-glutamine and prolyl-histidine may indicate excessive incomplete breakdown of proteins, which may provide evidence of the increased haemoglobin catabolism by the parasite to meet its amino acids demand for proliferation and growth. The altered level of these dipeptides may also signal the possibility of increased RBCs breakdown during the course of infection. Although the altered level of these peptides was found for the first time in urine of malaria patients, but their presence as an indication of

increased protein catabolism is further supported by the excessive urinary execration of urea in malaria patients. Urea, a principal urea cycle product, is formed in the liver from ammonia produced by the deamination of amino acids; it is the main urinary waste product of protein and amino acid catabolism in the body. The increased execration of urea in malaria patients may indicate an enhanced urea cycle activity as a response to the increased body demand of eliminating excess amino acids generated from RBCs breakdown. Indeed acute kidney injury during malaria infection resulting in partial-renal impairment may also lead to the increased level of urea, resulting in ammonia toxicity. Renal dysfunction was previously reported in severe malaria (Day et al., 1999). This finding was quite consistent with previously reported literature, in which a high level of urea was observed in the plasma of malaria patients infected with *P. falciparum* (Surowiec et al., 2015).

Abnormal level of taurine (a sulphur amino acid) was observed in the urine of malaria patients compared to healthy controls. Taurine is known to play an important role in the liver for detoxification of ammonia in malaria (Delic et al., 2010), suggesting that it was up-regulated in the liver as a response to the increased body demand for ammonia elimination. Taurine is mainly metabolised in the liver (de la Puerta et al., 2010), therefore, it is elevated level may also indicate liver dysfunction in malaria. Diet or enhanced parasite activity may also increase taurine level in malaria patients, however, the enhanced level of taurine was previously found in the sera of *Plasmodium*-infected mice (Ghosh et al., 2012) and in the urine of *P. vivax* infected patients (Sengupta et al., 2015) but not *Plasmodium*-infected RBCs (Teng et al., 2009). These results support the above finding, suggesting taurine as a potential biomarker of malaria.

Pipecolic acid, a metabolite of lysine, showed an increased level in the urine of malaria patients compared to healthy controls, which may also signalling liver dysfunction in malaria. This was consistent with the high level of pipecolic acid previously reported in the urine of *P. berghei* infected mice (Li et al., 2008, Tritten et al., 2013) and *P. vivax* infected patients (Sengupta et al., 2011b). Pipecolic acid is derived mainly from the catabolism of lysine by the intestinal microbiota or diet (e.g. fermented beverages and dairy products). Fujita et al., found that urinary pipecolic acid is mainly produced by gut microbiota rather than diet (Fujita et al., 1999b), hence, food intake seems unlikely to contribute to the increased level of pipecolic acid in malaria patients. However, the accumulation of pipecolic acid was found to be associated with a range of

different diseases such as Zellweger syndrome (Danks et al., 1975), chronic liver disease (Fujita et al., 1999a), neurological damage (Nomura et al., 1981) and pyridoxine epilepsy (Plecko et al., 2000). The biological relevance of pipecolic acid in malaria is still unclear, since liver dysfunction, neurological damage and disturbance of microbiota ecology are all associated with malaria (Li et al., 2008). However, its elevated level in urine may provide a potential biomarker of organ failure in malaria.

Changes in polyamines metabolism due to enhanced parasitic activity

Significantly high levels of acetylated polyamines such as 1,3-diacetylpropane, N-acetylspermidine and N-acetylputrescine were found in urine of malaria patients compared to healthy controls by both LC-MS and FIE-MS methods. This is the first time that altered levels of acetylated polyamines were detected in urine of malaria patients which may provide potential surrogate biomarkers of malaria. Polyamines are low molecular weight biogenic molecules that exist in different types of organisms including parasites (e.g. *plasmodium*), bacteria, plants and animals including human. They mediate many essential biological processes such as cell development, membrane stabilisation, proliferation and stress responses (Kusano et al., 2008). In *Plasmodium spp.*, polyamines constitute up to 14% of the total parasite's metabolome and serve as important and critical biomolecules for the rapid proliferation and growth of the parasite (Muller et al., 2001). Despite the reduced metabolic capacity, malaria parasite is still capable of *de novo* synthesis of polyamines from arginine using unique metabolic machinery (Muller et al., 2000). On the contrary, human erythrocytes, as non-proliferative cells, lack the ability to synthesise polyamines and very small amount of them exist in the cytoplasm (Lakshmanan et al., 2011).

Assaraf et al., performed in vitro experiment using infected human erythrocytes to study the production of polyamines by *P. falciparum* (Assaraf et al., 1984). They reported that synthesis of polyamines by the parasite reached the peak maximum in the early trophozoite stage, concluding that the parasite is critically dependant on high concentrations of polyamines for proliferation. In contrast, polyamines such as putrescine and spermidine are known to be toxic at high levels in human and, therefore, excess amount are detoxified mainly by oxidative deamination or acetylation by diamine acetyltransferase to N-acetylputrescine and N-acetylspermidine (Wunderlichova et al., 2014). No significant difference in the levels of putrescine and spermidine were observed between malaria patients

and healthy controls in this study, consistent with the previous clinical studies (Sengupta et al., 2015, Sengupta et al., 2011b, Surowiec et al., 2015). However, the altered levels of 1,3-diacetylpropane, N-acetylspermidine and N-acetylputrescine in urine of malaria patients, suggesting that putrescine and spermidine have been continuously detoxified by the body before execration as a response to their excessive production by the parasite (Figure 4-16). Putrescine and spermidine are derived mainly from diet (e.g. meat) or catabolism of amino acids (Patocka and Kuehn, 2000). Teng et al., reported significantly elevated levels of putrescine and spermidine in *Plasmodium*-infected erythrocytes compared to non-infected ones (Teng et al., 2009). Hence, dietary intake in malaria seems unlikely to influence the execration of acetylated polyamines in urine, whereas perturbed amino acids metabolism still a valid possibility.

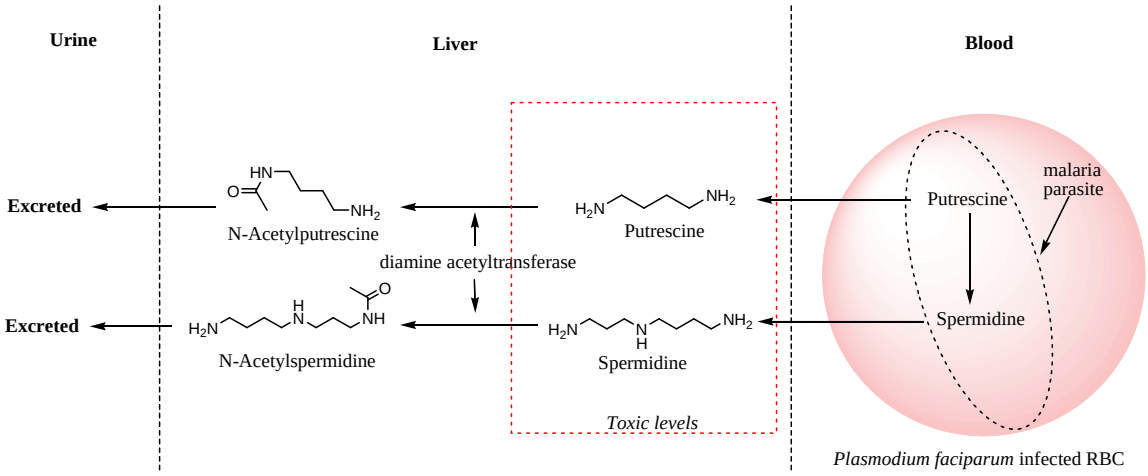


Figure 4-16 Suggested possible fate of the toxic levels of polyamines produced by *P. falciparum* during malaria infection. Putrescine and spermidine are excessively produced by *Plasmodium spp.* during erythrocytic cycle to meet its demand for proliferation and growth (Teng et al., 2009). In human, when the extracellular levels of these polyamines become abnormal, excess amount are detoxified by diamine acetyltransferase to N-acetylputrescine and N-acetylspermidine and excreted in urine (Wunderlichova et al., 2014).

Different NMR, GC-MS and LC-MS based metabolomics approaches have been used to study the *Plasmodium* metabolome (Sana et al., 2013, Mehta et al., 2005, Olszewski et al., 2009) and the host-parasite responses (Lakshmanan et al., 2012, Sengupta et al., 2011b, Sengupta et al., 2015) during malaria infections. The majority of these studies were concentrated in understanding the parasite biology or the host response using in vitro or rodent models. Metabolomics analysis of clinical samples such as urine, serum and plasma give more insight than in vitro or rodent model studies in understanding the host-parasite biology and provide a powerful tool for biomarker discovery and phenotypic biology. However, few

studies used clinical samples for untargeted metabolic profiling of *Plasmodium* infections (Sengupta et al., 2015, Sengupta et al., 2011b, Lakshmanan et al., 2012) and there is no published work using urine to explore the perturbed human-parasite metabolome during the course of *P. falciparum* infection.

Urine, as a readily available sample, has its potential as a convenient clinical sample to diagnose and study *P. falciparum* infection. The infection mainly occurred in rural remote areas in developing countries (WHO, 2014), where the essential facility, equipment and personnel expertise for invasive sample collection, (e.g. blood) and microscopic examination is quite difficult to establish and maintain for the diagnosis of malaria. In contrast, urine collection is fairly simple, non-invasive and no sophisticated facility or high skills are required if a suitable urinary diagnostic test is available for malaria. Therefore, in this study, an untargeted approach was implemented to analyse urine samples from *P. falciparum*-infected patients and healthy controls using HRMS and multivariate analysis. Two HRMS methods, HILIC LC-MS and FIE-MS were employed to identify possible surrogate biomarkers of malaria, those will complement the current strategies used to diagnose and treat malaria. This is the first time that an untargeted approach is employed to develop a high-throughput diagnostic method and find diagnostic urinary biomarkers of *P falciparum* infections.

4.5 Conclusion

LC-MS and direct ESI-MS metabolomics analysis of urine samples from malaria patients and healthy controls successfully demonstrated metabolic changes as a result of malaria and identified potential biomarkers. Confirmation of identity of some biomarkers was performed by LC-MS/MS with reference to pure authentic standards. Multivariate analysis using OPLS-DA, demonstrated good classification of the disease from healthy controls with both LC-MS and FIE-MS methods. Considering the throughput of the direct ESI-MS, FIE-MS was able to cover 50% of the LC-MS information, indicating its great potential as a diagnostic tool for malaria. Interestingly, FIE-MS detected all the potential candidates detected with LC-MS, such as pipecolic acid, taurine, N-acetylspermidine, N-acetylputrescine and 1,3-diacetylpropane, indicating that FIE-MS performed as well as LC-MS for urinary metabolomics of malaria. The biological interpretation of the metabolites significantly altered in malaria, revealed an increased activity in TCA cycle, which was consistent with previous reports. Furthermore, taurine and pipecolic acid which are previously suggested in the literature as potential biomarkers for malaria were also found associated with malaria. For the first time, altered levels of acetylated polyamines were detected in urine of malaria patients with LC-MS and FIE-MS which may provide new potential biomarkers and suggest FIE-MS as a quick non-invasive diagnostic tool of malaria. However, some urine samples of patients turned red during transportation from Ethiopia to UK, which flagged a caveat that the differences found between malaria patients and healthy controls may be influenced by presence of blood in urine or changes of urine chemistry during transportation. Nevertheless, simple dipstick urinalysis showed no blood or any significant differences in urine chemistry between the study groups. Also, many parameters such as diet, life-style, genetic disposition and sample collection may influence the metabolic profiles of malaria patients and healthy controls and subsequently the outcome. In order to circumvent these possible confounding factors, metabolic profiling of urine samples (24 h pooled urine sample is preferential, if possible) collected and stored properly from cohorts of malaria patients and healthy subjects with a longitudinal follow-up and a controlled dietary intake may provide additional specific information directly related to disease and metabolic changes associated with *P. falciparum* infection in human.

CHAPTER FIVE

The Influence of Dietary Oligosaccharides on Gut Bacterial Metabolites Detected in Urine

5. The Influence of Dietary Oligosaccharides on Gut Bacterial Metabolites Detected in Urine

5.1 Introduction

5.1.1 Human-gut microbiota

The human gastrointestinal (GI) microbiota (also known as gut flora or microflora) can be described as a complex collection of microorganisms, mainly bacteria but also viruses, archaea and protozoans, that mainly colonise the large intestine of the digestive tract (Scholtens et al., 2012). It is estimated that the human intestinal microbiota ecosystem contains about 10^{14} bacterial cells comprising 500-1000 different species, outnumbering the human cells by ten times (Sommer and Backhed, 2013). This vast and complex population of the microbiota that inhabits the human GI tract suggests that the microbiome (the collective set of microbiota genome) encodes 100-fold more unique genes than that of the human genome (Qin et al., 2010). Moreover, the gut microbiome is constantly changing within and between individuals as a result of different human processes (e.g. GI secretions, gut physiology, innate and adaptive immunity), diet, life events, use of drugs, disease state and environmental factors which affect the composition and function of the gut microbiota (Macfarlane and Macfarlane, 2012).

The main bacterial phyla of microbiota colonise the human GI tract are: *Firmicutes* (e.g. *Enterococcus* and *Staphylococcus*), *Bacteroidetes* (e.g. *Bacteroides*), *Proteobacteria* (e.g. *Escherichia*), *Verrumicrobia*, *Fusobacteria*, *Cyanobacteria*, and *Actinobacteria* (e.g. *Bifidobacterium*) (Johnson and Versalovic, 2012, Guarner and Malagelada, 2003). In general, human intestinal bacteria exert beneficial effects (also known as commensal bacteria, e.g. *Lactobacillus spp.*) or harmful effects (pathogenic bacteria, e.g. *Staphylococcus spp.*) (Figure 5-1). Some of the commensal bacteria such as *Escherichia coli* (*E. coli*) have the potential to be pathogenic under compromised health conditions (Gibson and Roberfroid, 1995). Although the commensal bacteria are isolated from the human circulating system by the GI mucosal barrier, there are constant metabolic interactions between the two systems allowing mutual benefits. The health-promoting effects of the gut bacteria include direct contributions to different metabolic, protective and trophic functions of the human body. The metabolic functions of microbiota mainly promote digestion and absorption of

essential nutrients, synthesis of vitamins and production of short-chain fatty acids from non-digestible carbohydrates (Guarner and Malagelada, 2003). Microbiota also inhibit the growth of harmful pathogens (protective functions) and contribute to the maturation of the immune system (trophic function) (Meyer and Stasse-Wolthuis, 2009).

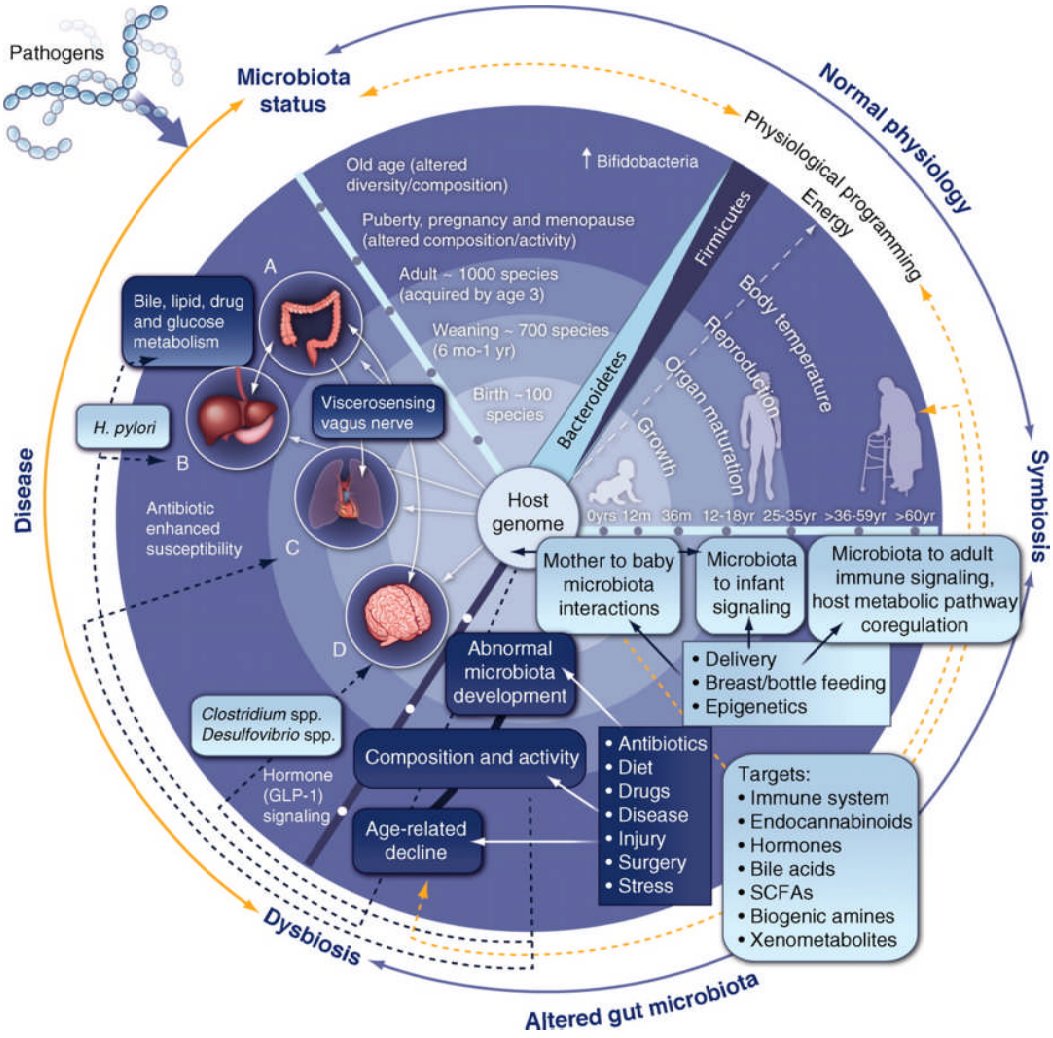


Figure 5-1 The effect and influence of gut microbiota on human health and disease from birth to old age. At birth, the method of delivery, nutrition, epigenetic factors play an important role in the development of a healthy gut microbiota (symbionts) of the fetus. The composition of microbiota continuously changes from childhood to adulthood in which *Bacteroidetes* predominate in childhood while *Firmicutes* predominate in adulthood. The intestinal microbiota involves in many essential physiological and energy production processes such as growth, organ maturation, reproduction and temperature regulation in the human body. Dysbiosis (alteration to the human gut microbiota) may lead to different variety of organ related diseases such as: (A) colon: e.g. irritable bowel syndrome (IBS), (B) liver: e.g. fatty liver disease, (C) lungs: e.g. asthma, and (D) brain: mood and behavioural abnormalities through, e.g. GLP-1 hormone signalling (Nicholson et al., 2012).

Recent studies have suggested that alteration to the human gut microbiota, known as dysbiosis or dysbacteriosis, may play an important role in the pathogenesis of many diseases. There is strong evidence that dysbiosis has been associated with irritable bowel syndrome (IBS) (Le Gall et al., 2011), ulcerative colitis (Williams et al., 2009, Le Gall et al., 2011), Crohn's disease (Williams et al., 2009), multiple sclerosis (Cantarel et al., 2015), diabetes (Musso et al., 2010) and obesity (Delzenne and Cani, 2011). In many circumstances, the host-microbe relationships and processes are not fully understood and further studies are needed for better understanding of the role of dysbiosis in human health and disease. Initial studies suggested that modulating the composition of the gut microbiota using certain nutrients may promote or maintain health over diseases and give rise to the concept of prebiotics (Gibson and Roberfroid, 1995, Roberfroid et al., 2010).

5.1.2 Introduction to the concept of prebiotics

The term "prebiotic" was first introduced by Gibson and Roberfroid in 1995 to describe "nondigestible food ingredient that beneficially affects the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon, and thus improves host health" (Gibson and Roberfroid, 1995). The concept of prebiotics is based on the use nutrients such as non-digestible fermentable carbohydrates to favourably fortify and stabilise the composition of health-promoting bacteria in the colon through selective fermentation. The effectiveness of prebiotics depends mainly on their ability to overcome digestion in the upper GI tract and reach the colon where microbiota utilises them for fermentation and energy production. Therefore, certain criteria should be met to consider a dietary nutrient as a prebiotic. These criteria mainly include the following: (1) resistance to the gastric digestive environment such as acidic pH, enzymatic hydrolysis and absorption, (2) readily fermentable by intestinal microbiota, and (3) selective stimulation of growth and activity of health-promoting intestinal bacteria (Roberfroid, 2007). Among the different types of dietary ingredients, some lipids, proteins, peptides and carbohydrates are good candidates to be prebiotics (Gibson and Roberfroid, 1995). However, most of the interest in food industry is focused on a certain class of carbohydrates, mainly nondigestible oligosaccharides such as inulin and oligofructose which are considered as the most potential prebiotics substrates (Rycroft et al., 1999). Table 5-1 lists some of oligosaccharides and their prebiotics properties.

Maltodextrin is digestible and easily absorbable carbohydrate and lacks the criteria to be considered as a prebiotic. However, it has been widely used in food industry as a food additive to substitute fats, modify texture and as a sweetener. Also, maltodextrin, being a digestible and non-fermentable carbohydrate, has been used as a placebo (baseline control) for oligosaccharide prebiotics (e.g. oligofructose) dietary intervention studies (Dewulf et al., 2013, Cani et al., 2005, Verhoef et al., 2011, Cani et al., 2009). In contrast, inulin and oligofructose (heterogeneous blends of fructose polymers) are non-digestible oligosaccharides that pass through the upper GI tract and reach the colon where they can be fermented by the intestinal microbiota. They have the basic characteristic of being non-digestible and fermentable by the intestinal flora. Recent studies reported that they modulated the microbiota ecology and host physiology in animals and humans (Delzenne et al., 2011a, Delzenne et al., 2011b), hence they are considered as the most likely prebiotic candidates amongst the dietary oligosaccharides.

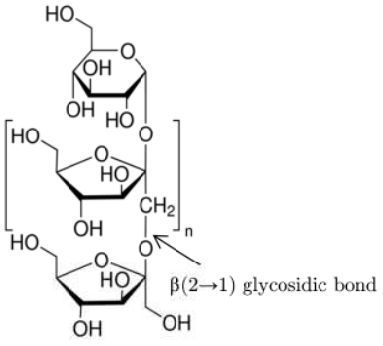
Table 5-1 Summary of the prebiotic effects of some oligosaccharides

	Prebiotics criteria			
Carbohydrate	Nondigestible	Fermentable	Selective	Prebiotics status
Galactooligosaccharides	Probably	Preliminary data	Yes	Yes
Glucooligosaccharides	No available data	No available data	No available data	No
Inulin and oligofructose	Yes	Yes	Yes	Yes
Isomaltooligosaccharides	Partially	Yes	Promising	No
Lactosucrose	No available data	No available data	Promising	No
Lactulose	Probably	Preliminary data	Yes	Yes
Maltodextrin	No	No	No	No
Soybean oligosaccharides	No available data	No available data	No available data	No
Xylooligosaccharides	No available data	No available data	Promising	No

Table reproduced from Roberfroid, 2007 (Roberfroid, 2007).

5.1.3 Oligofructose: chemistry and sources

Oligofructose (also known as fructooligosaccharides) is a subgroup of inulin, and defined chemically as a short chain of 2 - 10 fructose monosaccharide residues connected by a β(2→1) glycosidic bond (Figure 5-2) (IUB-IUPAC, 1982). Inulin and oligofructose are naturally occurring food ingredients which are present in more than 36,000 plant species including chicory, wheat, onion, bananas and garlic (Carpita et al., 1989). They are commonly used as dietary fibre in processed food to



Oligofructose n = 2-10, inulin n ≈ 35

Figure 5-2 Chemical structure of inulin and oligofructose

improve the texture of fat-free products or to reduce calorie intake. Most of the available oligofructose in food industry market is synthesised from sucrose (a disaccharide of glucose and fructose), extracted or enzymatically derived from the longer chain inulin polymers in chicory roots. Chicory root of *Cichorium intybus* contains 5-10% oligofructose and 15-20% inulin (Niness, 1999).

5.1.4 The current research on the effect of oligofructose on intestinal microbiota in health and disease state

The β(2 - 1) glycosidic bonds (Figure 5-2) prevent oligofructose from being digested like a typical carbohydrate in the upper GI tract and, therefore, it passes through to the colon where it is fermented by the microbiota. The fermentation process of oligofructose in the colon produces gases (e.g. hydrogen and methane) and short chain fatty acids (SCFAs). The fermentation products of oligofructose in the colon are believed to induce a variety of health-promoting prebiotic effects, for instance, SCFAs such as butyrate, lactate, propionate and acetate provide essential cellular nutrients that enhance the proliferation of epithelial cells and improve the absorption capacity of colonocytes for minerals and hence, improve bone health (Coxam, 2007, Scholz-Ahrens et al., 2001). Also, these SCFAs have been reported to improve intestinal motility (Sabater-Molina et al., 2009) and modulate the immune system (Arslanoglu et al., 2007).

The presence of oligofructose in the colon was reported to alter the composition of the gut microbiota (Blaut, 2002). Different dietary interventions have been reported to investigate the effect of oligofructose and its fermentation products on intestinal microbiota in health and disease state. Table 5-2 summaries a selected list of reported metabolites that have shown significant differences in the metabolic profile of different body fluids due to oligofructose dietary intervention compared to controls under different health and disease conditions. Dewulf et al., reported a ^{1}H-NMR-based metabolomics study to investigate the modulation of inulin/oligofructose 50:50 supplement on gut microbiota and host metabolism in obese women using urine and plasma samples (Dewulf et al., 2013). Simultaneous analysis of faecal samples using qPCR and human intestinal tract chip (a phylogenetic microarray) were also performed to assess the composition of the gut microbiota. They reported increased concentrations of *Bifidobacterium* and *Faecalibacterium prausnitzii* bacteria; which were negatively correlated to the altered level of serum lipopolysaccharide. Inulin/oligofructose supplement also decreased *Bacteroides intestinalis, Bacteroides vulgatus* and *Propionibacterium;*

an effect associated with the altered levels of bacteria-related metabolites such as lactate, hippurate and phosphatidylcholine in obese women compared to controls. Similarly, Vitali et al., used GC-MS to investigate the effect of oligofructose on the metabolic profile of gut microbiota in healthy subjects (Vitali et al., 2010). Altered levels of SCFAs, ketones, carbon disulphide and methyl acetate were measured in the faecal extract of oligofructose treated subjects compared to controls, indicating a perturbed metabolism of gut microbiota due to oligofructose.

De Preter and co-workers performed a GC-MS metabolomics approach on faecal samples from Crohn's disease patients and healthy controls before and after oligofructose/inulin dietary intervention (De Preter et al., 2013). They reported a list of metabolites that showed significant differences between groups in the study (Table 5-2). Significantly decreased levels of butyrate, hexanoate, pentanoate, heptanoate and p-cresol were found in Crohn's disease patients compared to healthy controls. Whereas, an elevated level of butyrate was found in Crohn's disease patients after receiving oligofructose/inulin diet compared to baseline, indicating that oligofructose fermentation in the colon upregulated butyrate to the level in healthy subjects. Presence of butyrate and other SCFAs in the colon is believed to improve health and welfare; for instance, they decrease the pH in the colon and, therefore, they selectively stimulate the proliferation of health-promoting *bifidobacteria* and *lactobacilli* which are perfectly adaptable to such an acidic environment. Moreover, the decrease in the pH increases the resistance to the growth of enteric pathogens (Gibson et al., 1995).

Table 5-2 A selected list of metabolites reported as potential biomarkers for oligofructose dietary intervention under different health and disease state

Author	Sample type	Method	Type of intervention	Altered metabolites
(Dewulf et al., 2013)	Urine	NMR	Metabolic changes in obesity	Creatinine Hippurate
	Plasma			Lactate Lipopolysaccharide Phosphatidylcholine
(Keller et al., 2011)	Plasma/horses	NMR	Laminitis in horses	Acetate Alanine Glutamine Glycine Isobutyrate Lactate Oligofructose Phosphatidylcholine
(Lamers et al., 2003b)	Cell culture	NMR	Caco-2 cells profiling	Alanine Glucose Glutamate
(Lamers et al., 2003b)	Cell culture	NMR	Caco-2 cells profiling	Isoleucine

Author	Sample type	Method	Type of intervention	Altered metabolites
(continued)				Lactate
				Leucine
				Phenylalanine
				Tyrosine
				Valine
(Meijers et al., 2010)	Plasma	HPLC	Targeted analysis in haemodialysis patients	p-Cresol sulphate
(De Preter et al., 2010)	Faecal sample	GC-MS	Effect on appetite in healthy subjects	Acetate
				Butyrate
				Dimethyl sulphide
				p-Cresol
				Propionate
(De Preter et al., 2011)	Faecal extract	GC-MS	Metabolic changes in healthy subjects	Acetate
				Butyrate
				Dimethyl bisulphide
				Dimethyl sulphide
				Dimethyl trisulphide
				Ethyl benzene
				Pentanoate
(De Preter et al., 2013)	Faecal extract	GC-MS	Profiling in Crohn's disease	1-Hexanol
				2,2,4-Trimethyl-pentane
				2,5-Dimethyl furan
				2-Butanone
				2-Methyl-2-propenyl-benzene
				2-Methylbutyrate
				2-Methylpropanal
				2-Pentyl furan
				3-Methyl-1H-indole
				3-Methyl-2-pentene
				Acetone
				Acetophenone
				Benzeneacetaldehyde
				Butyrate
				Carbon disulphide
				Dimethyl sulphide
				Furan
				Heptanoate
				Hexane
				Hexanoate
				Methanethiol
				Methyl 2-propenyl disulfide
				Methyl alcohol
				Methyl propyl disulphide
				Nonyl-cyclopropane
				Octanoate
				Pentanoate
				Phenol
				Phenol, 4-methyl-alpha-terpinene
(Verbrugghe et al., 2010)	Urine	GC-MS	Metabolic changes in domestic cats	Propionylcarnitine
(Vitali et al., 2010)	Faecal extract	GC-MS	Metabolic changes in healthy subjects	Acetate
				Valerate
				2-Propanone
				Carbon disulphide
				Methyl acetate
				Indole
				Thiophene
(Windey et al., 2014)	Faecal water	GC-MS	high dose in healthy subjects	Isovalerate
				p-Cresol

Most of the reported metabolomics studies investigating the effect of oligofructose on the metabolic signature of gut microbiota (Table 5-2) relied on the analysis of faecal samples due to practical inaccessibility to sample different parts of the colon (Figure 5-3). However, several metabolites produced by the intestinal microbiota are rapidly absorbed, metabolised and eliminated by the host depending on GI transit time (Lewis and Cochrane, 2007). For example, more than 95% of SCFAs produced from prebiotics fermentation in the gut are absorbed by the host, making their determination a difficult task to attain from faecal samples. Therefore, the use of other body fluids such as urine or plasma may provide a better alternative and improve metabolic profiling of gut microbiota. Urine, compared to plasma, is simple, readily available and non-invasive samples for metabolomics studies and provides an end metabolite pool of the body. Moreover, the kind of small, polar metabolites which are produced by the gut microflora are generally excreted efficiently in the urine, and hence this is an added advantage. Therefore, urine profiling has the potential to increase our understanding of metabolites variation associated with the perturbed metabolism and composition of gut microbiota due to oligofructose intervention. However, simultaneous faecal sample analysis is desirable to estimate and correlate the composition of the gut microbiota to the altered metabolites that might be found in urine.

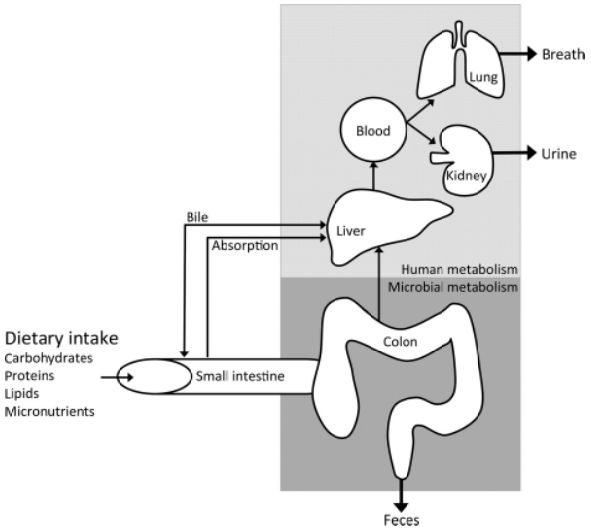


Figure 5-3 A schematic representation of site and fate of the expected metabolites produced by the interplay between human and intestinal microbiota due to oligofructose. Oligofructose passes-through the small intestine and reaches the colon, where it is fermented by the microbiota. This process produces specific microbial metabolites which are either execrated in faeces or absorbed. The absorbed metabolites pass to the liver where they are subject to human metabolism and can be execrated in urine and breath (Hamer et al., 2011).

5.1.5 The current role of FODMAP diet on gastrointestinal physiology and microbiota

Oligofructose and inulin enriched diet has shown to promote health and welfare by selectively stimulating the activity and growth of intestinal microbiota in health and disease (Verhoef et al., 2011, De Preter et al., 2010, De Preter et al., 2011, Meijers et al., 2010). However, some studies proposed that such poorly absorbed short-chain carbohydrates and polyols (designated as FODMAP: Fermentable Oligo-, Di-, Mono-saccharides And Polyols) exacerbated the symptoms of inflammatory bowel disease such as Crohn's disease and irritable bowel syndrome (IBS) including abdominal discomfort and bloating (Gibson and Shepherd, 2005, Gibson and Shepherd, 2010). An enriched FODMAP diet is thought to draw water into the small intestine by osmosis and increase the fermentable carbohydrates in the proximal colon, which is likely to induce luminal distension and metabolic effects on motility (Gibson and Shepherd, 2005). Luminal distension induces pain, bloating and clinical abdominal destination along with possible secondary motility changes (Gibson and Shepherd, 2010). Recent studies demonstrated that dietary supplements with no or low FODMAP contents alleviated the GI symptoms in IBS patients, and hence low FODMAP diet has been proposed as a treatment of IBS (Chumpitazi et al., 2015, Halmos et al., 2014, Magge and Lembo, 2012).

Recently, Hoad et al., investigated the actual effects of FODMAP on gastrointestinal physiology by measuring free mobile water in the small bowel and colonic gas using a valid MRI protocol (Hoad et al., 2007) developed by the same group. They reported that a single, high dose of inulin (40 g) distended the colon with gases compared to fructose, glucose or a mixture of both, while fructose but not inulin increased water in small intestine (Murray et al., 2014). Such a high dose of inulin is beyond the normal range of dietary variation, but these results supported the hypothesis that a FODMAP diet increases the production of colonic gases and the causes of luminal distension. However, no correlation was found between the extent of luminal distension and the reported symptoms of bloating and pain in participants, suggesting that the colon naturally accommodates excess gases without causing clinical GI symptoms. This interesting finding indicates that the ingestion of FODMAP enriched diet is well tolerated by healthy subjects and, therefore, their use still might be advantageous for exerting the anticipated prebiotic effect on gut microbiota. However, to draw

such conclusion, further assessment of FODMAP such as oligofructose on gastrointestinal form, function and microbiota are essentially needed.

5.1.6 Aims and objectives

The aim of this chapter is to use the high resolution mass spectrometry methods developed in the previous chapters to explore the changes in the urinary metabolic profiles of healthy subjects treated with a low FODMAP and oligofructose diet. The specific objectives required to fulfil this aim are as follows:

- To investigate the urinary metabolic profile changes in healthy subjects due to the effect of low FODMAP and oligofructose diet on gut microbiota using LC-MS and FIE-MS.
- To identify urinary biomarkers, which can differentiate between low FODMAP and oligofructose diet study group and controls.

5.2 Materials and experimental

5.2.1 Materials and reagents

MS grade acetonitrile and methanol were supplied from Fisher Scientific, UK. Water (VWR international, EC), formic acid, 0.1% formic acid in water (Sigma-Aldrich, Germany), leucine enkephalin, are MS grade and minimum handlings were performed to minimise any possible contamination. Ultra-pure LC-MS grade ammonium acetate was supplied from Fluka, Sigma-Aldrich, Netherland. Artificial urine was prepared using a mixture of 35 compounds and they were supplied from different sources as detailed in chapter 2, Table 2-2.

5.2.2 LC-MS and FIE-MS analyses

Chromatography was performed using Accela UHPLC system (Thermo Fisher, USA) on a BEH HILIC column (2.1 x 100 mm, 1.7 μm particle size, Waters, USA) as detailed previously in chapter 2, section 2.2.3. Orbital trap mass spectrometer (Exactive-Orbitrap, Thermo Fisher Scientific, USA) was used in simultaneous ESI+ and ESI- modes for LC-MS. The operational parameters were: spray voltage 3.2 kV (ESI+), 2.4 kV (ESI-), capillary voltage 25 V (ESI+), -27 V (ESI-), sheath, auxiliary and sweep gas flow rate were: 20, 5 and 5 (arbitrary unit), respectively, for both modes. Capillary and heater temperature were maintained at 350 and 120 ° C, respectively. Data were acquired in full scan mode with resolution 25,000 from *m/z* 60-1000 with 4 Hz scan rate.

Flow injection ESI-MS (FIE-MS) analysis used the same Accela UHPLC system to inject 10 μL of the sample without a column as detailed in chapter 2, section 2.2.4. The above orbital trap mass spectrometer was also used in simultaneous ESI+ and ESI- modes for FIE-MS. The MS parameters were: spray voltage 1.5 kV (ESI+), 1.8 kV (ESI-), capillary voltage 25 V (ESI+), -40 V (ESI-), sheath, auxiliary and sweep gas flow rates were: 50, 5 and 5 (arbitrary unit) (ESI+/-), respectively. Capillary and heater temperature were maintained at 250 and 120 °C, respectively, in both modes. Data were acquired in full scan mode with resolution 25,000 from *m/z* 50-1000 with 4 Hz scan rate.

5.2.3 Subject recruitment

This study, was undertaken as a part of a randomised double-blind controlled trial to study the effect of low FODMAP diet and dietary oligofructose on gastrointestinal form, function and microbiota in healthy volunteers (abbreviated as FOG study) using a diagnostic MRI protocol (Hoad et al., 2007). FOG study was designed by Prof Robin Spiller (study chair), Dr Giles Major, Dr Luca Marcian, Ms Shanthi Krishnasamy (Nottingham Digestive Diseases Centre (NDDC), Queen's Medical Centre (QMC), Nottingham, UK), Prof Penny Gowland (Sir Peter Mansfield Magnetic Resonance Centre (SPMMRC), University of Nottingham, UK), Dr Mathew Diggle, Dr Tim Sloan (Department of Microbiology, University of Nottingham, UK), Prof David Barrett (Centre for Analytical Biosciences, University of Nottingham, UK) and Dr Miranda Lomer, (Division of Nutrition and Dietetics, King's College London, UK). The role in this study was to investigate the influence of oligofructose and low FODMAP diet on gut bacterial metabolites detected in urine.

Subject recruitment, dietary intervention and urine collection were undertaken by NDDC and SPMMRC, Nottingham, UK. 48 healthy subjects were voluntarily recruited by local advertisements for a 2-week double-blind dietary intervention study. Participants were 18 years old or above and gave informed consent. Exclusion criteria were: self-declared vegetarian, vegan or kosher/halal diet as the study involves consumption of carmine red dye of animal origin, pre-existing bowel-complaints or diseases such as IBS, intestinal stoma and resection of oesophagus, stomach or intestine (excluding the appendix), health conditions that contraindicate the use of MRI scanning such as pacemakers, metallic implants or above MRI scanner weight limit (i.e. more than 120 kg), any medical conditions will compromise the participation in the study (e.g. diabetes mellitus) and use of medicines that affect bowel functions including antibiotics. Also, participants were asked to limit their alcohol intake to less than or equal to 35 units/week and less than or equal to 8 units/day. A diet evaluation using food-frequency questionnaire was also used to evaluate and exclude subjects with unusual fibre intake. 11 subjects were excluded after the consent or withdrew from the study, whereas the remaining 37 subjects were included in the analysis. Ethical approval was provided by University of Nottingham and QMC Ethics Committees and appropriate written informed consent was obtained from each participant.

5.2.4 FOG dietary intervention

The study's participants lived at home and they attended 5 visits (visit 1-5) throughout the 2-week intervention period as illustrated in Figure 5-4. Morning spot urine (visit 2) was collected from subjects after 1 week of normal diet plan (spot urine controls), and then a 24 h pooled urine sample was collected in the next day (visit 3, 24 h urine controls). Previous in vivo and in vitro intervention studies demonstrated that oligofructose and inulin exert their prebiotic effect and change the colonic microbiota community within 12 h to 7 days (Euler et al., 2005, Bouhnik et al., 1999, Kolida et al., 2002, Gibson and Roberfroid, 1995). Therefore it was decided to undertake oligofructose intervention for a week. All participants went on a low FODMAP diet by reducing their intake of poorly digested and fermentable carbohydrates for 7 days. As part of the diet, subjects were advised to avoid food containing fructans, fructose, galacto-oligosaccharides and polyols containing food and limit their dietary lactose to less than 4 g/meal (Table 5-3). For ease of regimen education, food types were colour-coded for participants using the traffic light system; red (food to avoid), yellow (food in moderation) and green (allowed food). The type of dietary intake to avoid was listed to common food groups, i.e. fruits, vegetables, bread/rice/cereal, meat/poultry/eggs and milk/milk products. Considering that most of the foods contain more than one poorly absorbable and fermentable carbohydrate, listing them as food groups made understanding easier. As the diet was restrictive and could be expensive, subjects were given tips on food preparation and shopping, and they were also given a loaf of gluten free bread and a packet of rice to take away.

Table 5-3 Some food examples that should be avoided during the 2nd week of the dietary intervention

Fructans	Galacto-oligosaccharides	Fructose	Polyols	Lactose
Fructo-oligosaccharide Garlic Inulin Leek Mushroom Oligofructose Onion Wheat/Rye	Cashews Chickpeas (baked beans, kidney beans etc.) Legumes Lentils Pistachios	Apple Concentrated fruit juices Dried fruits Honey	Cauliflower Mannitol Mushroom Sorbitol Sweet potato Watermelon	Ice-cream Milk Soft cheese Yogurt

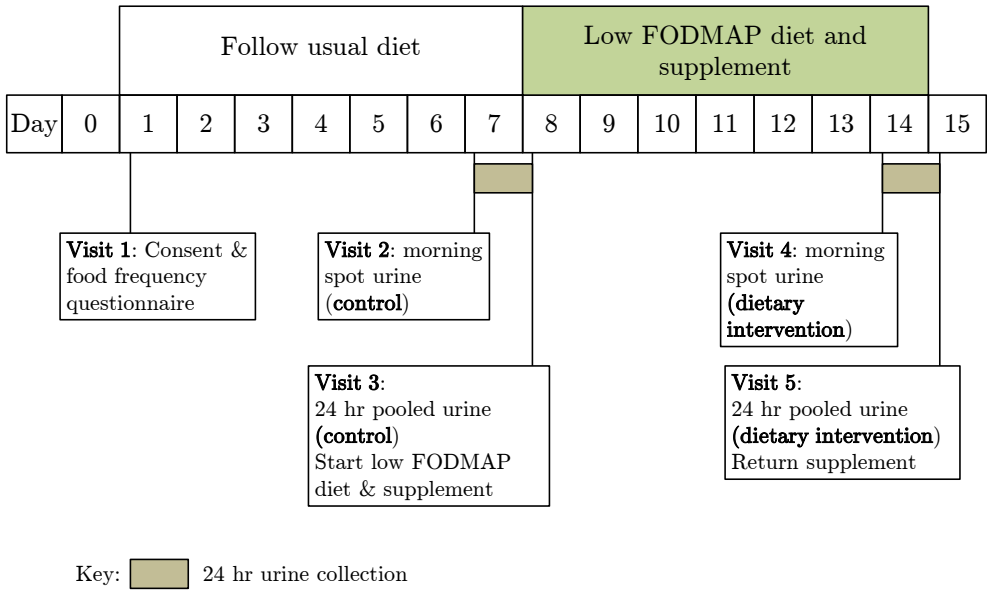


Figure 5-4 Low FOODMAP diet and supplement intervention study diagram. Visit 1, 2 and 4 took place in the Nottingham Digestive Diseases Centre (NDDC), Queen's Medical Centre, Nottingham, UK. Visit 3 and 5 took place at the 1.5T scanner of the Sir Peter Mansfield Magnetic Resonance Centre (SPMMRC), University of Nottingham, Nottingham, UK. Urine samples for visit 2-5 were collected at 8:30am after morning fasting.

During the low FODMAP dietary period, subjects (n = 37) were assigned to receive 7 grams twice a day of either a placebo (maltodextrin, an easily digestible long-chain glucose polysaccharide) (group A) or oligofructose (group B) supplement using a systematic method, in which half of the study's participants received the oligofructose supplement (n = 19) and the other half remained in the placebo group (n = 18). Maltodextrin powder has a physical appearance and a taste similar to oligofructose and has been used in previously reported dietary interventions as a placebo for inulin-type fructans such as oligofructose (Dewulf et al., 2013, Cani et al., 2005, Verhoef et al., 2011, Cani et al., 2009). Carmine red dye, a food dye of animal origin, was mixed with the placebo and oligofructose supplement for colour matching, and they were both provided in identical opaque packages. The supplement powder was served as a warm drink. During visit 4 (i.e. after 6 days of low FOODMAP and supplement), morning urine was collected from the subjects and then pooled 24 h urine was collected in the next morning (visit 5). To assess the compliance to the supplement, subjects provided in visit 2 (i.e. the start of dietary intervention) with a pre-weighed supplement bag, which returned back at the end for monitoring. The subjects, caregivers, investigators and outcomes assessors involved were blinded to the

intervention; however for ease of classification, eight sets of urine samples were collected:

1. **A2:** A (maltodextrin, control) is spot urine samples (n = 18) collected from subjects in maltodextrin supplement group (A) during visit 2.
2. **A3:** A (maltodextrin, control) is 24 h urine samples (n = 18) collected from subjects in maltodextrin supplement group (A) during visit 3.
3. **A4:** A (maltodextrin) is spot urine samples (n = 18) collected from subjects in maltodextrin supplement group (A) during visit 4 (i.e. after 6 days of low FODMAP and maltodextrin supplement).
4. **A5:** A (maltodextrin) is 24 h urine samples (n = 18) collected from subjects in maltodextrin supplement group (A) during visit 5 (i.e. after 7 days of low FODMAP and maltodextrin supplement).
5. **B2:** B (oligofructose, control) is spot urine samples (n = 19) collected from subjects in oligofructose supplement group (B) during visit 2.
6. **B3:** B (oligofructose, control) is 24 h urine samples (n = 19) collected from subjects in oligofructose supplement group (B) during visit 3.
7. **B4:** B (oligofructose) is spot urine samples (n = 19) collected from subjects in oligofructose supplement group (B) during visit 4 (i.e. after 6 days of low FODMAP and oligofructose supplement).
8. **B5:** B (oligofructose) is 24 h urine samples (n = 19) collected from subjects in oligofructose supplement group (B) during visit 5 (i.e. after 7 days of low FODMAP and oligofructose supplement).

All urine samples were collected in urinary collection vessels without the use of preservatives. Samples were then, aliquoted in triplicates of 2.0 mL into pre-labelled cryotubes and stored immediately in -80 ˚C freezer.

5.2.5 Preparation of urine samples and artificial urine for metabolomics analysis

Thawed spot and 24 h pooled urine samples of subjects were prepared with the optimised urine dilution protocol for LC-MS and direct ESI-MS analyses as described in chapter 2, section 2.2.7. For metabolomics analysis, pooled QC samples for spot urine and 24 h urine were prepared separately by mixing 20 µL aliquots taken from each urine sample (i.e. spot or 24 h urine) and was treated the same as described for the samples. Artificial urine was prepared as described in chapter 2 section 2.2.6.

5.2.6 LC-MS and FIE-MS metabolomics analysis

Spot urine and 24 h urine samples were analysed in two separate analytical runs, however, they both treated the same during the preparation and analysis. Spot/24 h urine samples collected after the dietary intervention and controls were randomised and analysed in a single analytical run using simultaneous positive and negative ion modes with LC-MS and FIE-MS. Six injections of pooled QC sample were analysed at the beginning of the run to equilibrate the column (LC-MS) prior the analysis. Pooled QC urine sample were interspaced throughout the run to check the stability, robustness, repeatability and performance of the analytical system. Leucine enkephalin, *m/z* 556.2771 (ESI+), 554.2615 (ESI-) was spiked in every sample (0.2 µg/mL) to monitor mass accuracy within each run. Blanks were injected after each sample in FIE-MS analysis to minimise the carryover effect, if any.

5.2.7 Data analysis and metabolite identification

The raw data for LC-MS and FIE-MS were acquired and visualised using Xcalibur v2.1 software (Thermo Scientific, USA). The performance of the analytical methods was validated by monitoring a representative set of urine metabolites in the pooled QC sample as described previously in chapter 3, section 3.2.7. In addition, the quality of the datasets obtained was assessed using methodology proposed by Want et al., (Want et al., 2010) for LC-MS, and Beckmann et al., (Beckmann et al., 2008) for FIE-MS, as detailed in chapter 2, section 2.2.11.

For metabolomics analysis, full datasets acquired by LC-MS and FIE-MS from urine samples before (controls) and after dietary intervention were imported and pre-processed by Progenesis QI software using within-subject design. Within-subject design is a time series experiment which is suitable for experiment where each subject has been sampled at different time points. Such design eliminates or reduces individual differences as a source of between condition differences, which helps to create a more powerful statistical test (Nonlinear-Dynamics, 2014). Pre-processing steps include: peak picking, peak alignments, normalisation and transformation as previously detailed in section 2.2.11. Automatic deconvolution of the extracted *m/z* variables was carried out using Progenesis QI to remove isotopes, adducts, and other confounding peaks resulting from MS detection.

Normalised abundance of extracted metabolites peaks (RT, *m/z*) pairs (LC-MS) and *m/z* (FIE-MS) were exported for multivariate analysis (MVA). Multivariate data analysis using principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) were used to investigate metabolic changes between A (maltodextrin), A (maltodextrin, control), B (oligofructose) and B (oligofructose, control) datasets using Simca P +14 (Umetrics AB, Sweden). To maximise the differences between the study groups, successive OPLS-DA models based on two groups: A (maltodextrin) versus A (maltodextrin, control) and B (oligofructose) versus B (oligofructose, control), were generated. The obtained OPLS-DA models were validated using cross-validation and permutation test. The validation results were used to assess the suitability of the models for further metabolomics data analysis based on the type of urine samples used (i.e. spot urine or 24 h urine). The more suitable OPLS-DA models were then validated using a prediction method based on randomly selected training (50%) and test sets (50%) of samples. The specificity and selectivity of the prediction models were further tested using area under the ROC (receiver operating characteristic) curve (AUC).

The ions responsible for the class separation in the OPLS-DA models of dietary intervened urine samples and controls were selected using Variable Importance for the Projection (VIP) and variables loadings plots. Student's *t*-test was performed to test the significant difference of the selected ions between the two groups. Prior to student's *t*-test, ArcSinh transformation (Jones, 2008) was applied to restore normality. The p-values of the student's *t*-test were adjusted using false discovery rate (FDR) for multiple testing problems. Tentative identification of significantly altered urinary metabolites due to dietary intervention was achieved by interrogating the Human Metabolome Database (HMDB), for possible identification based on accurate mass measurements within 5 ppm mass error (Wishart et al., 2013). In-house urine metabolite database (built using the same LC-MS described in this study) was also used for identification based on accurate mass determination and retention times. For ease of biological interpretation and to concentrate on the metabolites related to the effect of low FOODMAP and oligofructose supplement, the common metabolites related to maltodextrin (baseline) and diet were excluded.

5.3 Results

5.3.1 Subjects' compliance to dietary supplement

Qualitative assessment by the Nottingham Digestive Diseases Centre (NDDC) researchers during the study reported a compliance with the supplement ranged from 82% - 147%. These results indicate good compliance for both maltodextrin and oligofructose during the dietary intervention.

5.3.2 Validation of LC-MS and FIE-MS performance

The quality of datasets obtained from mass spectrometry analysis is of paramount importance as it has a direct impact on the final biological outcome obtained from a metabolomics study. Therefore, a pooled QC approach (Want et al., 2010, Gika et al., 2007) was used to assess the quality of the acquired MS datasets from FOG dietary intervention urine samples. All spot/24 h urine samples of FOG dietary intervention were analysed in a single LC-MS and FIE-MS analytical runs with pooled QC samples being interspaced throughout the analysis. Appendix G reports the percentage of relative standard deviation (%RSD) of the retention times (RT), peak areas (LC-MS) and peak intensities (FIE-MS) of selected set of 58 metabolites found in the pooled QC sample. A summary of the mean and the range of these RSD% values are listed in Table 5-4. In the pooled QC sample datasets of spot urine, the mean RSD% values of the selected urine metabolites peak areas (LC-MS) and peak counts (FIE-MS) were 7 % (range: 3 - 20%) and 10% (range: 2 - 22%), respectively. Whereas they were 9 % (range: 3 - 15%) and 10% (range: 3 - 22%), for peak areas (LC-MS) and peak counts (FIE-MS), respectively, for 24 h urine. LC-MS retention time shifts for both spot and 24 h urine were less than 0.07 min (≤ 1% RSDs) and mass accuracy shift was less than 5 ppm in both positive and negative ion modes for all analyses. These values were within the recommended acceptable limits of the Food and Drug Administration (FDA) for bioanalysis (FDA, 2013) and biomarker discovery (Gika et al., 2007, Want et al., 2010). In addition, the quality of the datasets obtained with LC-MS was assessed using the peaks present in at least 80% of QC samples. The RSD% of mean peak areas was less than 30% for 73% (spot urine) and 75% (24 h urine) of these peaks, which was lower than the suggested threshold for metabolomics analysis (Want et al., 2010). The analytical runs demonstrated adequate stability as indicated by sufficient clustering of the pooled QC samples towards the centre of the PCA score plots

with both methods (Figure 5-11). These results validate the analytical performance of LC-MS and FIE-MS for metabolomic analysis of FOG dietary intervention study spot and 24 h urine samples. Also, for the purpose of evaluating the FIE-MS against the LC-MS, a direct comparison of the LC-MS pooled QC samples selected peaks was performed. FIE-MS was able to detect 78% (spot urine) and 74% (24 h urine) of the peaks found by the LC-MS. The above results demonstrated that the FIE-MS consistency, metabolic coverage and reproducibility for the metabolomic analysis of the dietary intervention urine samples were acceptable and as concluded in chapter (chapter 2).

Table 5-4 A validation summary of LC-MS and FIE-MS analyses for metabolomics profiling of FOG dietary intervention spot/24 h urine samples using selected urine metabolites peaks from pooled QC samples.

	LC-MS (BEH-HILIC)		FIE-MS		Acceptable limits[1]
Description	Spot urine	24 h urine	Spot urine	24 h urine	
1. Peak detection					
Total number of peaks	58	58	58	58	-
Number of peaks detected	57	57	45	43	-
% of peaks detected	98%	98%	78%	74%	-
Artificial urine metabolites detected	18	18	15	14	-
2. Retention time (RT) variability					
RT RSD% (mean)	0.35%	0.46%	-	-	1%
RT RSD% (range)	0.06 - 1.36%	0.11 - 0.97%	-	-	0.00 - 1.00%
RT Shift (mean)	0.03	0.04	-	-	-
RT Shift (range)	0.01 - 0.07	0.00 - 0.07	-	-	-
3. Peak intensity variability					
Peak area/count RSD% (mean)	7%	9%	10%	10%	30%
Peak area/count RSD% (range)	3 - 20%	3 - 15%	2 - 22%	3 - 22%	0 - 30%

Selected peaks retention times and peak areas/counts statistics were calculated from the pooled QC urine sample injections throughout the analytical run (n = 9). [1]Acceptable limits based on FDA regulation for bioanalysis (FDA, 2013) and the recommendations for a reliable metabolomics analysis (Gika et al., 2007, Want et al., 2010) .

5.3.3 Metabolomics analysis of FOG urine samples: a human dietary intervention using low FODMAP and oligofructose diet

LC-MS and FIE-MS analyses

Spot urine and 24 h urine samples of FOG dietary intervention classes were analysed in two separate analytical runs. In each run, complete LC-MS and FIE-MS datasets were acquired for urine samples of A (maltodextrin, n = 19), A (maltodextrin control, n = 18), B (oligofructose, n = 19) and B (maltodextrin control, n = 18). Standard ESI+ and ESI- LC-MS base peak chromatograms (BPCs) obtained from these classes are shown in Figure 5-5 and Figure 5-6, respectively. Visual examination of chromatography showed an adequate

separation with most of urinary metabolites eluted within 9 min as shown in appendix G. The metabolites observed in chromatography were mainly comprised a range of amino acids, acylcarnitines, nitrogen containing compounds, organic acids, purines and pyrimidines (Table 5-5). Amino acids such as phenylalanine, L-alanine and L-tryptophan were eluted within a narrow retention time window of 5.0-6.7 min, whereas acylcarnitines including acetylcarnitine and butyrylcarnitine were eluted within a slightly higher retention time window (6.7-7.5 min). Organic acids such as hippuric acid, homovanillic acid, 5-hydroxyindole-3-acetic acid, 4-Aminohippuric acid, lactic acid and uric acid were detected within a wider retention time window (0.5-5 min). To give an overview of the differences in the findings between LC-MS and the FIE-MS; summed mass spectra obtained from A (maltodextrin), A (maltodextrin control), B (oligofructose) and B (oligofructose control) urine samples are presented in Figure 5-7 (LC-MS, ESI+), Figure 5-8 (LC-MS, ESI-), Figure 5-9 (FIE-MS, ESI+) and Figure 5-10 (FIE-MS, ESI-). For ease of comparison, some of the known urinary metabolites are annotated in the spectra. Organic acids such as 3,4-dihydroxybutyric acid, lactic acid and butyric acid were mainly detected in the negative ESI mode in both LC-MS and FIE-MS; however, some of them such as hippuric acid were observed in both ESI modes of FIE-MS. Taking all molecular features into account, most of the detected ions were concentrated in the lower mass range (*m/z* 60-300) in FIE-MS, while relatively higher mass ions (*m/z* 60-400) were detected with LC-MS. Similar MS spectra were obtained with LC-MS and FIE-MS from the urine samples of A (maltodextrin), B (oligofructose) and controls. The observed metabolic changes associated with maltodextrin) and oligofructose compared to controls were relatively small; therefore, detailed multivariate analysis using PCA and OPLS-DA was used to separate and classify urine samples in the study.

Table 5-5 Some of urinary metabolites detected with LC-MS from the analysis of urine samples of FOG dietary intervention study

Metabolites	RT(min)	*m/z*	Polarity	Metabolites	RT(min)	*m/z*	Polarity
1. Amino acids				Lactic acid	2.33	89.0268	ESI-
L-alanine	6.65	88.0399	ESI-	MHPG sulphate	0.75	263.0225	ESI-
Phenylalanine	5.47	166.0868	ESI+	Uric acid	1.49	167.0205	ESI-
L-Tryptophan	5.39	203.0821	ESI-	**4. Nitrogen containing compounds**			
2. acylcarnitines				Urea	1.46	61.0399	ESI+
Acetylcarnitine	7.29	204.1230	ESI+	Creatine	6.65	132.0769	ESI+
Butyrylcarnitine	6.74	232.1546	ESI+	Creatinine	3.15	114.0664	ESI+
3. Organic acids				Trimethylamine N-oxide	7.71	76.0758	ESI+
4-Aminohippuric acid	3.96	193.0613	ESI-	**5. Purines and pyrimidines**			
5-Hydroxyindole-3-acetic acid	3.02	190.0504	ESI-	Hypoxanthine	1.98	137.0459	ESI+
Hippuric acid	4.06	178.0509	ESI-	Cytidine	2.40	226.0828	ESI+
2-Hydroxyisobutyric acid	3.26	103.0395	ESI-	Uridine	1.71	243.0621	ESI-
Homovanillic acid	3.15	181.0501	ESI-				
4-Hydroxy-3-methoxy-mandelic acid	2.80	197.0450	ESI-				

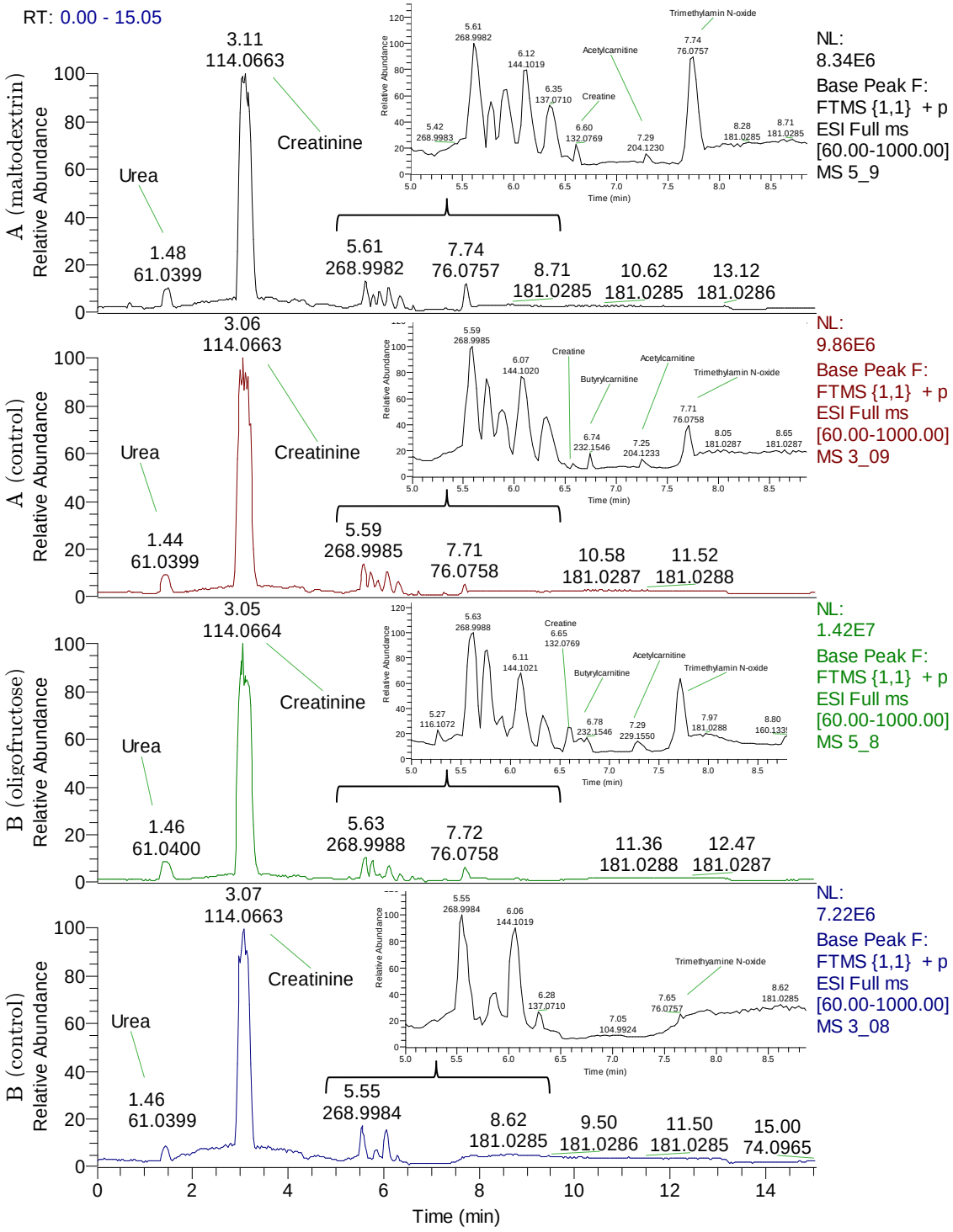


Figure 5-5 LC-MS base peak chromatograms (BPC) of 24 h urine samples before and after FOG dietary intervention. From top to bottom, BPC of urine samples of A (maltodextrin), A (control), B (oligofructose) and B (control) groups analysed by LC-MS using BEH HILIC column in ESI positive mode.

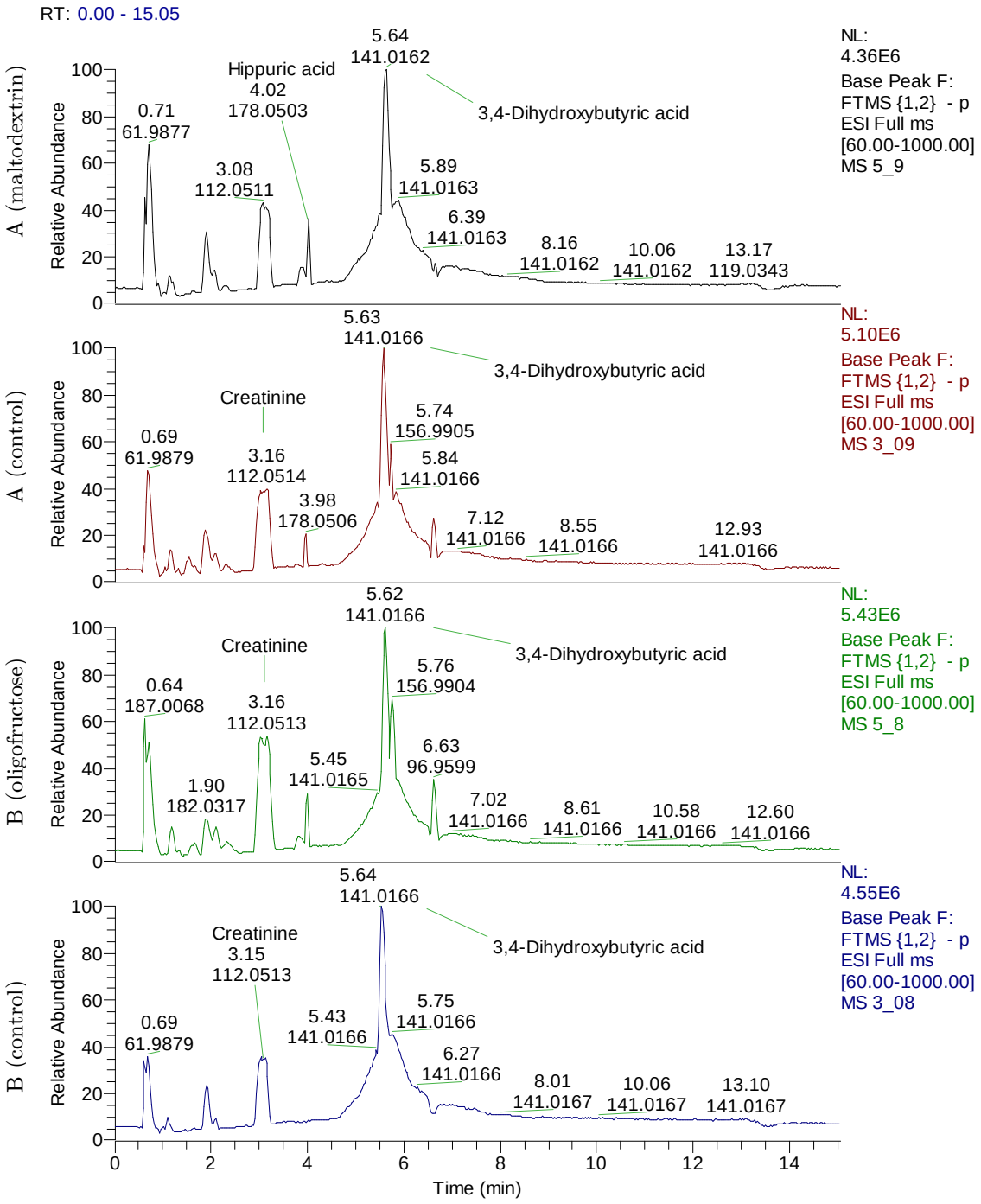


Figure 5-6 LC-MS base peak chromatograms (BPC) of 24 h urine samples before and after FOG dietary intervention. From top to bottom, BPC of urine samples of A (maltodextrin), A (control), B (oligofructose) and B (control) groups analysed by LC-MS using BEH HILIC column in ESI negative mode.

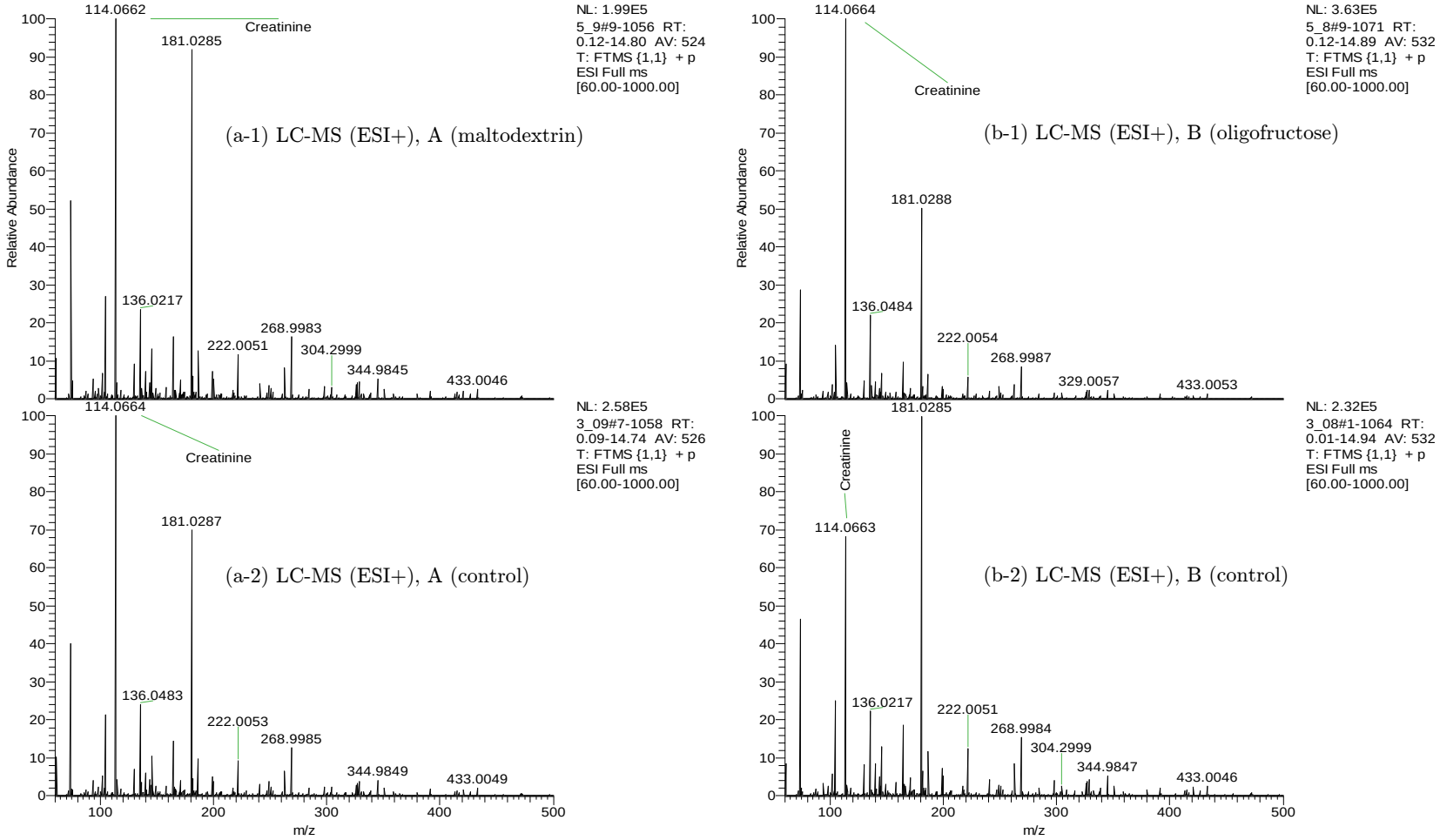


Figure 5-7 Typical LC-MS metabolic fingerprints of 24 h urine samples before and after FOG dietary intervention in positive ion mode. HILIC LC-MS combined mass spectra (*m/z* 60-500) present the urine profile of: (a-1) A (maltodextrin), (a-2) A (control), (b-1) B (oligofructose) and (b-2) B (control).

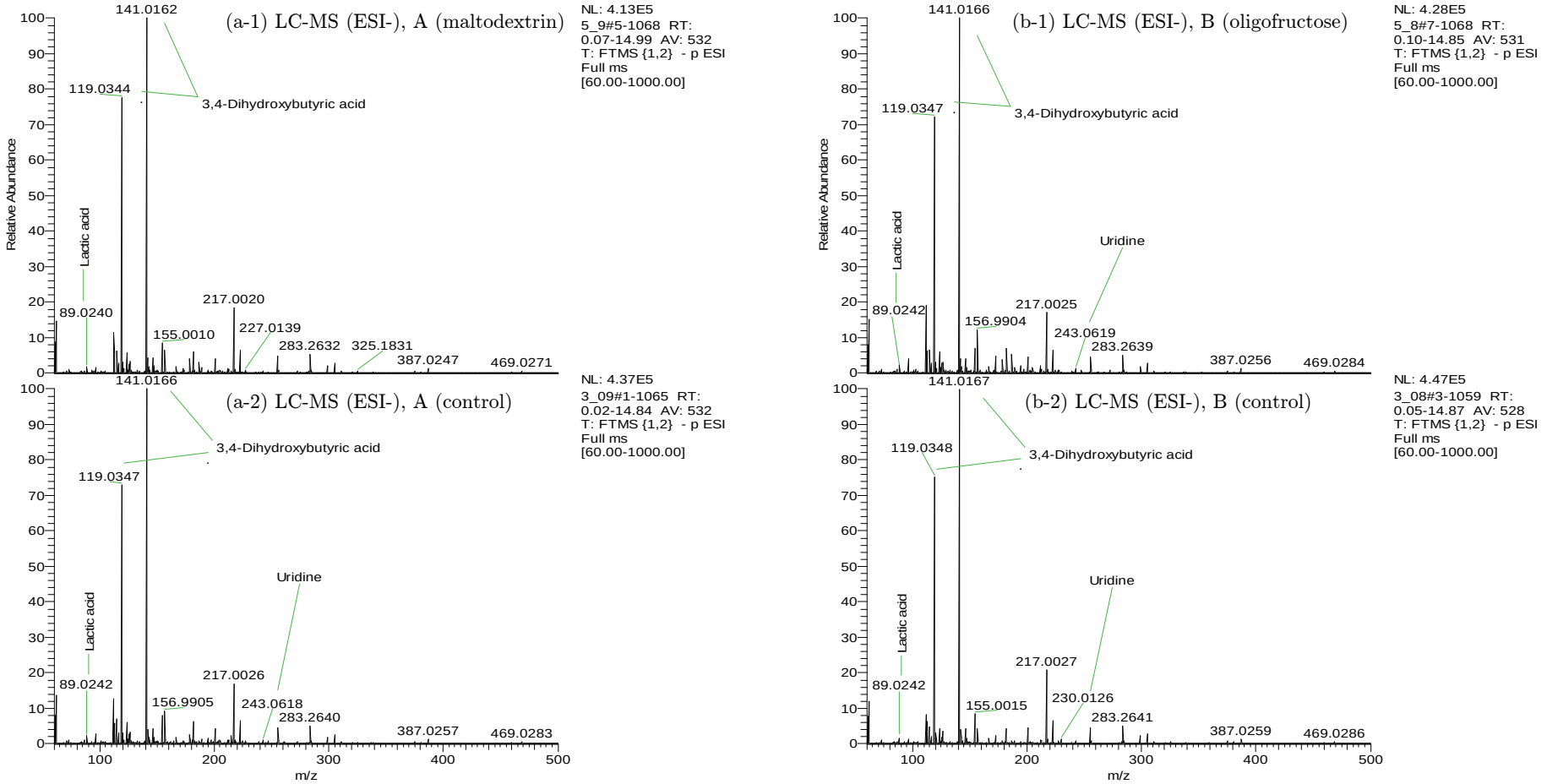


Figure 5-8 Typical LC-MS metabolic fingerprints of 24 h urine samples before and after FOG dietary intervention in negative ion mode. HILIC LC-MS combined mass spectra (m/z 60-500) present the urine profile of: (a-1) A (maltodextrin), (a-2) A (control), (b-1) B (oligofructose), (b-2) B (control).

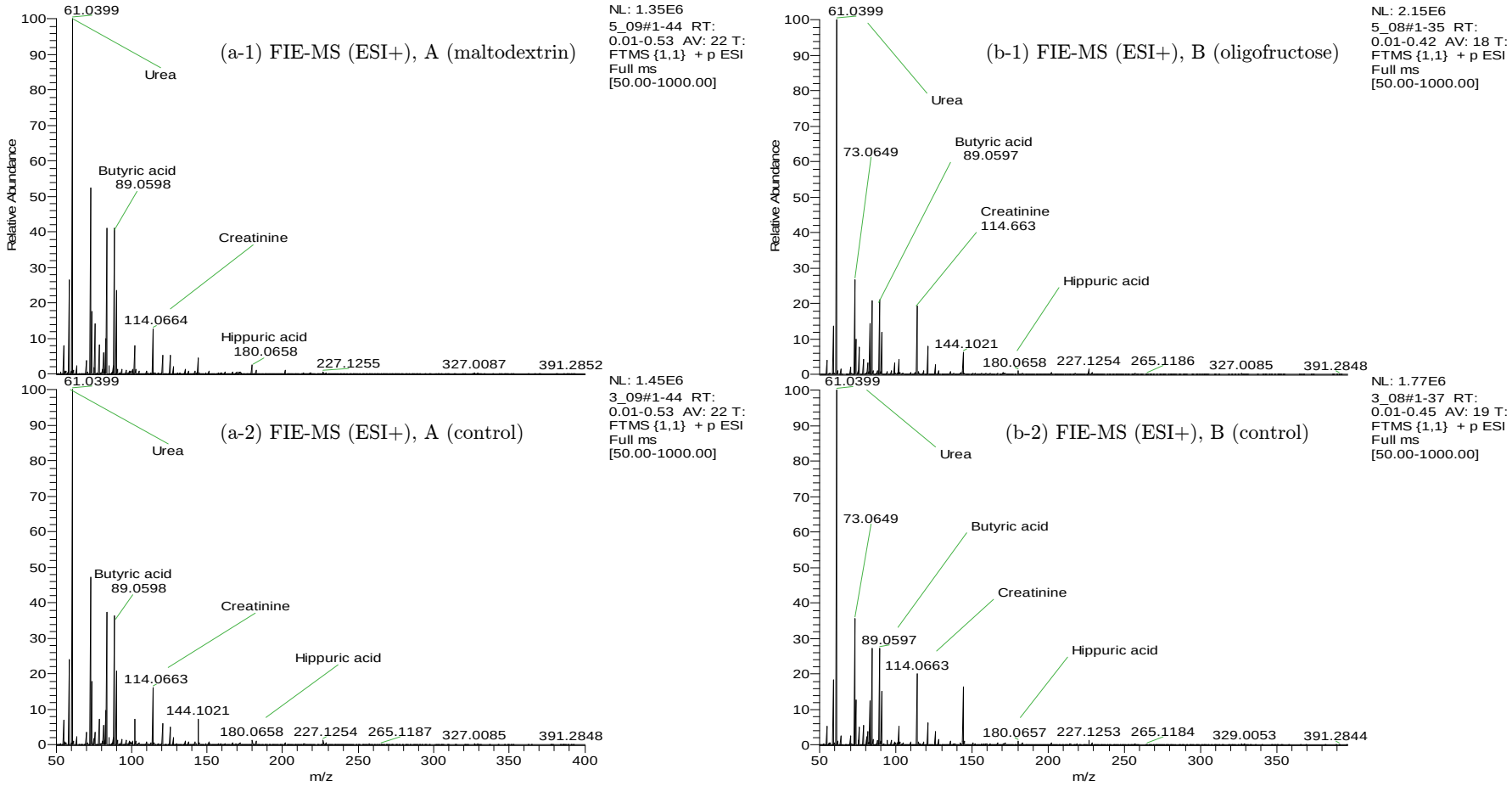


Figure 5-9 Typical FIE-MS metabolic fingerprints of 24 h urine samples before and after FOG dietary intervention in positive ion mode. Combined mass spectra (*m/z* 50-400) present the urine profile of: (a-1) A (maltodextrin), (a-2) A (control), (b-1) B (oligofructose) and (b-2) B (control).

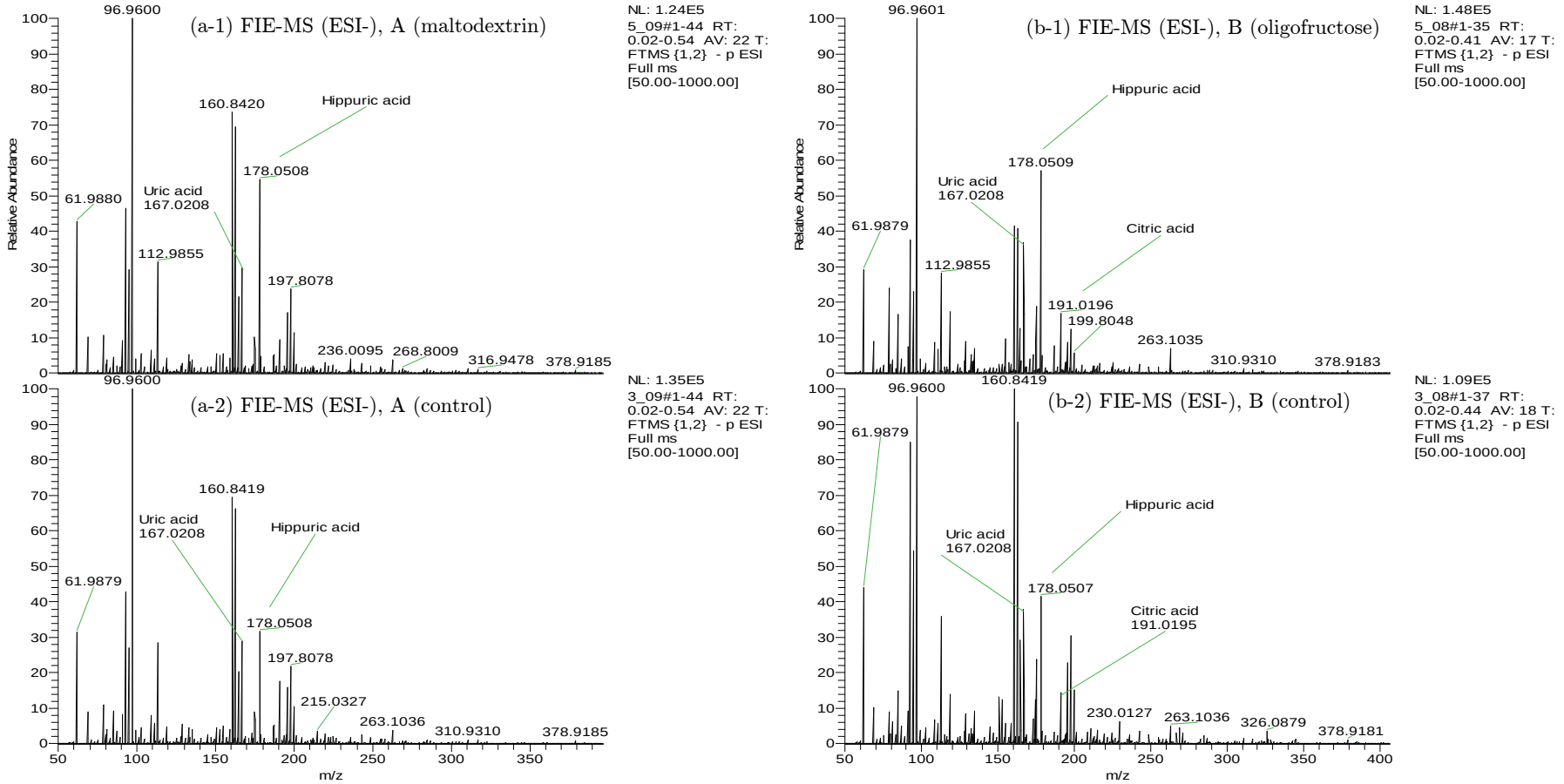


Figure 5-10 Typical FIE-MS metabolic fingerprints of 24 h urine samples before and after FOG dietary intervention in negative ion mode. Combined mass spectra (*m/z* 50-400) present the urine profile of: (a-1) A (maltodextrin), (a-2) A (control), (b-1) B (oligofructose), (b-2) B (control).

Evaluation of spot and 24 h urine samples of FOG dietary intervention for urinary metabolomics using multivariate analysis

For a comprehensive modelling of the differences between the groups in the study, datasets generated by LC-MS and FIE-MS were submitted for multivariate analysis. Spot urine and 24 h urine samples were acquired in two separate analytical runs. The raw datasets obtained from the analysis of spot urine or 24 h urine samples of A (maltodextrin, n = 18), A (maltodextrin control, n = 18), B (oligofructose, n = 19) and B (oligofructose control, n = 19) in a single analytical run with pooled QCs (n = 9) interspaced in the analysis were submitted for peak picking, peak alignment and normalisation using Progenesis QI (Nonlinear-Dynamics, 2014). The datasets of spot urine analysed by LC-MS and FIE-MS generated 7,690 and 1,053 variables, respectively, whereas 24 h urine samples generated 5378 and 434 from LC-MS and FIE-MS, respectively (Table 5-6). The normalised datasets were exported to Simca P+14 for multivariate analysis. Unsupervised principal component analysis (PCA) was used to give an unbiased overview for any possible trends and clustering within the sample datasets. PCA score plots of spot or 24 h urine samples showed no separation or clustering trends between the FOG study groups with LC-MS and FIE-MS (Figure 5-11).

Supervised orthogonal partial least square-discriminant analysis (OPLS-DA) models were constructed from spot/24 h urine samples to find differences and discrimination between FOG dietary intervention groups. Figure 5-12 presents the OPLS-DA score plots of A (maltodextrin), A (maltodextrin control), B (oligofructose) and B (oligofructose control). Separation and clustering was observed between A (maltodextrin), B (oligofructose) and their controls with spot urine datasets from both LC-MS and FIE-MS. Complete separation and clustering between all study groups was observed with 24 h urine datasets from both methods. However, poor predictive ability (Q^2) values and unexpected separation of control sample groups were observed with these models (Table 5-6), indicating model overfitting for both type of samples (spot and 24 h urine samples). Therefore, subsequent OPLS-DA models were generated with only 2 classes to overcome overfitting problem (Eriksson et al., 2006a) as detailed in the next section.

The high level of variability observed with spot urine ($Q^2 = 0.140$) compared to 24 h urine ($Q^2 = 0.272$) is quite consistent with previous studies (Saude et al.,

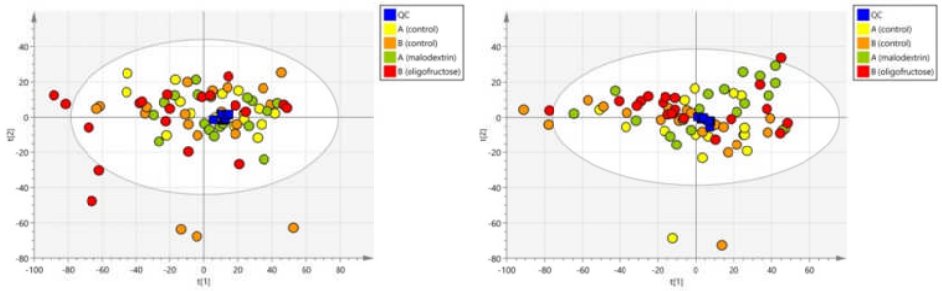
2007, Fave et al., 2011, Araki et al., 1990, Carrieri et al., 2000). Saude et al. investigated the variability associated with using spot urine for metabolomics analysis; they observed high variance between metabolites in the samples (Saude et al., 2007). Similarly, Fave et al., demonstrated that spot urine collection at different times of the day gave distinctly different metabolic profiles (Fave et al., 2011). Therefore, when overall individual metabolic status is investigated (e.g. dietary intervention study), 24 h pooled samples are preferred, if possible, to minimise the variability in the urinary profiles obtained with shorter collection periods (Fernandez-Peralbo and Luque de Castro, 2012). Hence, the 24 h urine samples obtained from the study groups were selected over the spot urine samples for further analysis.

Table 5-6 Multivariate analysis and validation of OPLS-DA models of FOG dietary intervention urine samples

	LC-MS		FIE-MS	
Description	Spot urine	24 h urine	Spot urine	24 h urine
1. Peak detected				
ESI+	4,911	3,266	483	167
ESI-	2,779	2,112	570	267
Total (ESI+ and ESI-)	7,690	5,378	1,053	434
2. Cross-validation				
a. OPLS-DA (all classes)				
R^2Y	0.540	0.598	0.569	0.583
Q^2	0.140	0.272	-0.019	0.141
b. OPLS-DA (maltodextrin)				
R^2Y	0.901	0.813	0.907	0.995
Q^2	-0.046	0.337	-0.053	0.156
c. OPLS-DA (oligofructose)				
R^2Y	0.983	0.822	0.810	0.958
Q^2	0.143	0.411	-0.215	0.012
3. Permutation test[1]				
Intercept (maltodextrin)	-0.1935	-0.2617	-0.1511	-0.1100
Intercept (oligofructose)	-0.2119	-0.2134	-0.1465	-0.2405
4. External validation: classification (training/test models): oligofructose				
R^2Y (training set)	0.892	0.989	0.947	0.999
Q^2 (training set)	-0.103	0.479	-0.119	0.431
True positive (TP)	2	7	5	7
False positive (FP)	2	1	5	4
True negative (TN)	7	8	4	5
False negative (FN)	7	2	4	2
Sensitivity (%)[2]	50%	88%	50%	64%
Specificity (%)[3]	50%	80%	50%	71%
Accuracy (%)[4]	50%	83%	50%	67%
5. Area under receiver operative characteristic (ROC) curve (AUC)[5]				
AUC oligofructose (TPR vs FPR)	0.56	0.85	0.41	0.70

[1]The OPLS-DA model is considered robust when the regression line of the permuted Q^2 values intercept at, or below zero and less than the Q^2 value of the model. [2]Sensitivity is the true positive rate (TPR) calculated from the formula, TPR = TP/(TP+FP). [3]Specificity is the true negative rate (TNR) calculated from the formula, TNR = TN/(TN+FN). [4]Accuracy is calculated from the formula, Accuracy (%) = (TP+TN)/(TP+FP+TN+FN). [5]AUC: Area under receiver operating characteristic curve is the total area under the curve of sensitivity (TPR) vs 1-specificity "false positive rate (FPR)", ideal model gives AUC = 1 (Eriksson et al., 2006b, Xia et al., 2013).

(a) Spot urine (HILIC LC-MS) (b) 24 hr urine (HILIC LC-MS)



(c) Spot urine (FIE-MS) (d) 24 hr urine (FIE-MS)

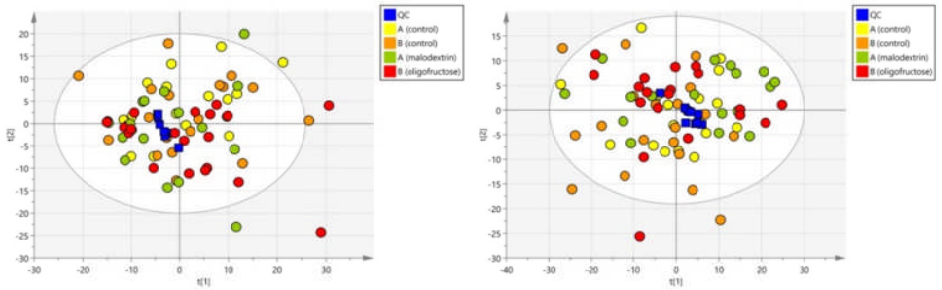
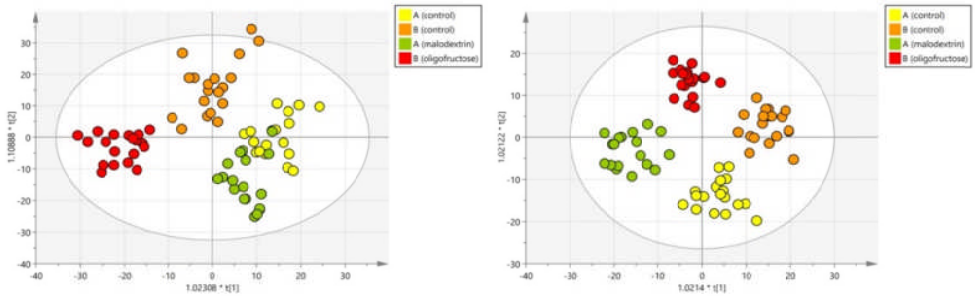


Figure 5-11 PCA score plots obtained from all spot/24 h urine samples of FOG dietary study groups. Sample groups are: A (maltodextrin): green circles, n = 18, A (control): yellow circles, n = 18, B (oligofructose): red circles, n = 19, B (control): light brown circles, n = 19 and pooled QC: dark blue squares, n = 9, were analysed by LC-MS (upper) and FIE-MS (lower) of spot (left) and 24 h (right) urine samples.

(a) Spot urine (HILIC LC-MS) (b) 24 h urine (HILIC LC-MS)



(c) Spot urine (FIE-MS) (d) 24 h urine (FIE-MS)

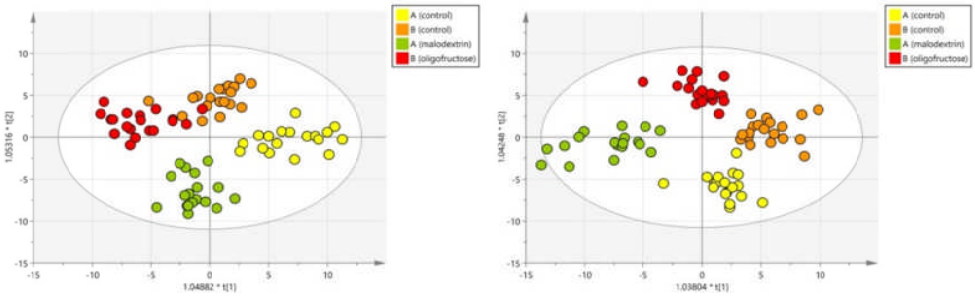


Figure 5-12 OPLS-DA score plots overview obtained from all spot/24 h urine samples of FOG dietary intervention groups. Sample groups are: A (maltodextrin): green circles, n = 18, A (control): yellow circles, n = 18, B (oligofructose): red circles, n = 19 and B (control): light brown circles, n = 19, were analysed by LC-MS (upper) and FIE-MS (lower) of spot (left) and 24 h (right) urine samples.

Validation of oligofructose, maltodextrin and controls OPLS-DA models

To maximise the difference in classification between A (maltodextrin), B (oligofructose) and their controls, successive OPLS-DA models were generated from only two classes in the study using 24 h urine samples. Two OPLS-DA models were constructed from LC-MS and FIE-MS datasets: A (maltodextrin) versus A (maltodextrin control) and B (oligofructose) versus B (oligofructose control) as shown in Figure 5-13. OPLS-DA score plots of these models showed clear separation and clustering of each group from its control in both LC-MS and FIE-MS. To further validate the differences observed between the controls and oligofructose samples, the OPLS-DA models were validated using a randomly selected training set of 10 subjects from B (oligofructose) and 10 subjects from controls and a test set of 9 subjects from B (oligofructose) and 9 subjects from controls for both LC-MS and FIE-MS. The obtained R^2Y/Q^2 cross-validation values from the training set OPLS-DA models were 0.989/0.479 and 0.999/0.431 for LC-MS and FIE-MS, respectively, demonstrating an acceptable model. The OPLS-DA models built using the training set were then used to predict the status of the test set of subjects. In LC-MS prediction models, 7 out of 9 of B (oligofructose) samples were predicted correctly and 8 out of 9 controls samples were predicted correctly, whereas in FIE-MS, 7 out of 9 of B (oligofructose) samples were predicted correctly and 5 out of 9 controls samples were predicted correctly. The sensitivity/specificity of these models were 88%/80% and 64%/71% for LC-MS and FIE-MS, respectively (Table 5-6), indicating a reliable predictive power of both models.

The robustness of the prediction models was further assessed using the area under receiver operating characteristic (ROC) curve (AUC). The value of the AUC of the OPLS-DA models was 0.85 and 0.70 for LC-MS and FIE-MS, respectively. Ideal model gives AUC equals 1.0 (perfect classifier), whereas AUC equals 0.5 is equivalent to randomly classifying subjects as either positive or negative (i.e. the classifier is of no clinical importance). A rough guide based on AUC for assessing the clinical utility of a biomarker is as follows: 0.5–0.6 (fail), 0.6–0.7 (poor), 0.7–0.8 (fair); 0.8–0.9 (good) and 0.9–1.0 (excellent) (Xia et al., 2013). These reference values indicate that the predictability of the LC-MS models is good, whereas it is adequately fair for FIE-MS for FOG study sample classification. Although few numbers of variables were detected by FIE-MS compared to LC-MS, the FIE-MS datasets generated a robust model as LC-MS.

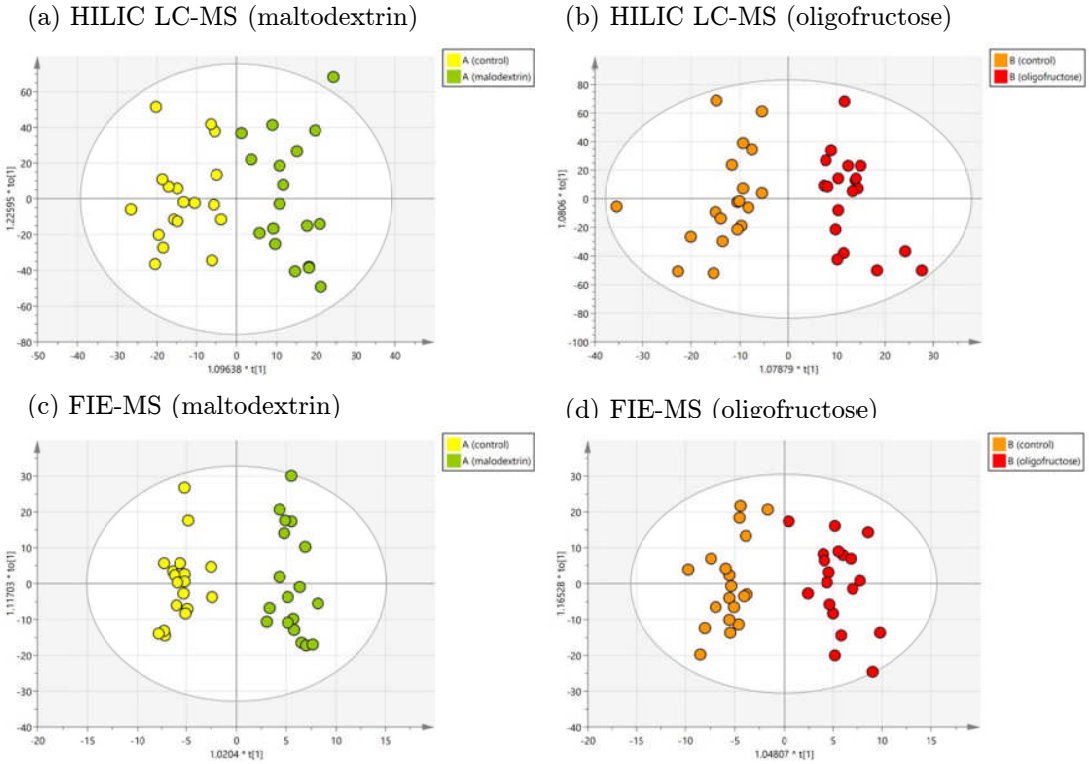


Figure 5-13 Two-classes OPLS-DA score plots obtained from the 24 h urine samples of FOG dietary intervention groups. OPLS-DA models are: A (maltodextrin): green circles, n = 18, versus A (control): yellow circles, n = 18, and B (oligofructose): red circles, n = 19, versus B (control): light brown circles, n = 19, were analysed by LC-MS (upper) and FIE-MS (lower).

5.3.4 Tentative identification of significantly altered urinary metabolites due low FODMAP and oligofructose dietary intervention

Metabolites responsible for the separation and clustering between subjects treated with oligofructose/maltodextrin and controls were selected from the OPLS-DA models generated from the datasets of 24 h urine samples (Figure 5-13) based on loadings plots and variable importance for the projection (VIP) scores. A VIP score of a variable above 1.0 is considered important for the model (Yin et al., 2009). Subsequent Student's *t*-test with false discovery rate (FDR) for multiple testing problem was performed across peak intensities of oligofructose/maltodextrin group and controls datasets using Progenesis QI. The number of metabolites found with significant difference (q-value $\leqslant$ 0.05) between oligofructose/maltodextrin group and controls were listed in Table 5-7. The exact mass of these metabolites was used to interrogate the Human Metabolome database (HMDB) (http://www.hmdb.ca/) to provide tentative identification of

the selected urine metabolites. Then, the altered metabolite lists from oligofructose and maltodextrin (baseline) interventions were compared. As a result, 113 and 51 metabolites were found unique to oligofructose but not maltodextrin with LC-MS and FIE-MS, respectively, demonstrating that FIE-MS provided about 45% of LC-MS information obtained from oligofructose intervention. Thus, LC-MS is a suitable choice for comprehensive urinary metabolomics study while FIE-MS as a high-throughput method might be used as a quick approach for screening.

Table 5-7 Significantly altered metabolites between oligofructose/maltodextrin (baseline) subjects and controls

	Metabolites significantly altered due to dietary intervention against controls		
Method	Oligofructose	Maltodextrin	Unique to oligofructose but not maltodextrin
LC-MS	187	141	113
FIE-MS	83	26	51

In order to investigate the metabolites related only to the effect of low FOODMAP and oligofructose supplement, the common urinary metabolites related to the effect of maltodextrin (the baseline) and diet were excluded from the selected list of metabolites related to oligofructose. Top ions of significant differences in LC-MS and FIE-MS have been considered as potential urinary metabolites related to oligofructose dietary intervention. Table 5-8 lists the tentatively identified metabolites from LC-MS and FIE-MS methods related to the effect of low FOODMAP and oligofructose supplement but not maltodextrin. Most of the metabolites obtained by FIE-MS were also detected with LC-MS such as creatinine, propanoic acid, 2-octenoic acid, glutaric acid, hydantoin-5-propanoic acid and acetylcarnitine. However, some metabolites such as benzoic acid were detected with FIE-MS but not LC-MS. This might be attributed to the ion suppression of urine salt on metabolite ionisation as higher load of urine salt was present in samples prepared for LC-MS (3 fold dilution) compared to samples for FIE-MS (10 fold dilution) as explained elsewhere (chapter 4, section 4.3.4).

Table 5-8 Tentative identification of urinary biomarkers showing significant differences between oligofructose subjects and controls but not maltodextrin

Metabolite MW (Da)	Tentative identification	HMDB ID	Formula	q Value[a]	Fold Change[b]	Mass Error (ppm)	LC-MS	FIE-MS	Pathway/Description	Pathway/process existence
74.0732	1-Butanol	HMDB04327	$C_4H_{10}O$	1.05E-02	1.69	0.50	✓✓		Butanoate metabolism	Human and bacteria (fermentation)
283.2606	7-Methylinosine	HMDB03950	$C_{11}H_{15}N_4O_5$	1.95E-02	2.34	1.64	✓✓	✓	Methylated purine	Human
148.0736	(R) 2,3-Dihydroxy-3-methylvalerate	HMDB12140	$C_6H_{12}O_4$	8.14E-03	-6.83	0.95	✓✓		2-Oxocarboxylic acid metabolism	Human and bacteria
146.0579	2-Methylglutaric acid	HMDB00422	$C_6H_{10}O_4$	2.73E-02	2.26	0.99	✓✓		Metabolite of succinic acid	Human
142.0994	2-Octenoic acid	HMDB00392	$C_8H_{14}O_2$	3.77E-02	2.89	0.28	✓✓	✓	metabolite of aliphatic aldehydes	Human
285.1940	2-Octenoylcarnitine	HMDB13324	$C_{15}H_{27}NO_4$	7.62E-03	1.46	0.07	✓✓		Fatty acid metabolism/acylcarnitine	Human
144.0423	3-Hexenedioic acid	HMDB00393	$C_6H_8O_4$	8.14E-03	2.70	0.05	✓✓		Fatty acid metabolism/dicarboxylic acid metabolite	Human
118.0630	3-Hydroxyvaleric acid	HMDB00531	$C_5H_{10}O_3$	3.35E-03	1.48	2.08	✓✓		Short chain fatty acid metabolite	Human
138.0317	4-Hydroxybenzoic acid	HMDB00500	$C_7H_6O_3$	3.77E-02	1.53	0.63	✓✓		Phenylalanine metabolism	Human
									Ubiquinone and other terpenoid-quinone biosynthesis	Human and bacteria
204.0746	5-Glutamylglycine	HMDB11667	$C_7H_{12}N_2O_5$	1.08E-02	4.46	0.24	✓✓		Protein catabolism/dipeptide	Human
329.2202	6-Keto-decanoylcarnitine	HMDB13202	$C_{17}H_{31}NO_5$	1.49E-02	1.49	0.33	✓✓		Fatty acid metabolism/ acylcarnitine	Human
117.0426	Acetylglycine	HMDB00532	$C_4H_7NO_3$	1.53E-02	5.43	0.40	✓✓		Protein and amino acid biosynthesis/acylglycine	Human
175.0957	Citrulline	HMDB00904	$C_6H_{13}N_3O_3$	2.72E-02	1.78	0.22	✓✓		Arginine and proline metabolism	Human and bacteria
113.0589	Creatinine	HMDB00562	$C_4H_7N_3O$	1.56E-02	1.10	2.90	✓✓	✓	Arginine and proline metabolism	Human
146.0691	Glutamine	HMDB00641	$C_5H_{10}N_2O_3$	1.16E-02	4.97	2.49	✓✓		Alanine, aspartate and glutamate metabolism	Human and bacteria
									Arginine and proline metabolism	Human and bacteria
									D-Glutamine and D-glutamate metabolism	Human and bacteria
									Purine metabolism	Human and bacteria
									Nitrogen metabolism	Human and bacteria
									Pyrimidine metabolism	Human and bacteria

Metabolite MW (Da)	Tentative identification	HMDB ID	Formula	q Value[a]	Fold Change[b]	Mass Error (ppm)	LC-MS	FIE-MS	Pathway/Description	Pathway/process existence
132.0423	Glutaric acid	HMDB00661	$C_5H_8O_4$	4.40E-02	-1.58	2.91	✓✓	✓	Lysine degradation Fatty acid metabolism	Human Human
172.0484	Hydantoin-5- propanoic acid	HMDB01212	$C_6H_8N_2O_4$	4.39E-03	3.19	1.04	✓	✓✓	Histidine metabolism	Human
203.1158	Acetylcarnitine	HMDB00201	$C_9H_{17}NO_4$	8.14E-03	3.28	0.57	✓✓	✓	Fatty acid metabolism/acylcarnitine	Human
133.0375	Aspartic acid	HMDB00191	$C_4H_7NO_4$	2.14E-02	2.92	2.89	✓✓		Alanine, aspartate and glutamate metabolism	Human
259.1784	Hexanoylcarnitine	HMDB00756	$C_{13}H_{25}NO_4$	4.10E-02	9.80	2.54	✓✓		Fatty acid metabolism/ acylcarnitine	Human
215.1158	Propenoylcarnitine	HMDB13124	$C_{10}H_{17}NO_4$	8.97E-03	1.78	0.52	✓✓		Fatty acid metabolism/acylcarnitine	Human
74.0368	Propanoic acid	HMDB00237	$C_3H_6O_2$	8.14E-03	1.73	1.03	✓✓	✓	Propanoate metabolism Nicotinate and nicotinamide metabolism	Human and bacteria Human
123.0320	Picolinic acid	HMDB02243	$C_6H_5NO_2$	2.28E-02	2.41	0.31	✓✓	✓	Tryptophan catabolism	Human
136.0749	Tetrahydropteridine	HMDB01216	$C_6H_8N_4$	8.76E-03	2.40	4.69	✓✓		Cofactor in amino acids hydroxylation	Human
168.0283	Uric acid	HMDB00289	$C_5H_4N_4O_3$	6.92E-03	-6.37	1.97	✓✓	✓	Purine metabolism	Human and bacteria
102.0681	Valeric acid	HMDB00892	$C_5H_{10}O_2$	3.74E-02	1.73	0.71	✓✓	✓	Fatty acid metabolism/short chain fatty acid	Human and bacteria (fermentation)
122.0368	Benzoic acid	HMDB01870		2.90E-02	1.49	0.85		✓✓	Phenylalanine metabolism Benzoate degradation via CoA ligation	Human Bacteria

[a]q-value: is the adjusted Student's *t*-test p-value using false discovery rate (FDR), [b]the positive value of fold change means a higher level of metabolite in oligofructose subjects compared to controls, whereas the negative value represents a lower level of metabolite. The q-value, fold change and mass error were calculated for each metabolite from its MS dataset generated by the MS method with the highest VIP value (primary MS method). Primary method used for the extraction of each biomarker is highlighted as (✓✓) in the table, whereas (✓) symbol in the MS method columns indicates that the biomarker was also detected by that method.

5.4 Discussion

The increased levels of valeric acid, creatinine, glutamine and propionic acid in the urine of oligofructose subjects was consistent with previously reported animal and human studies (Keller et al., 2011, Windey et al., 2014, De Preter et al., 2010, Dewulf et al., 2013).

The effect of oligofructose on the production of short chain fatty acids (SCFAs) by gut microbiota

The identified metabolites from urinary FOG metabolomics study enhance our understanding about the possible metabolic changes associated with the effect of oligofructose on gut microbiota and the subsequent host response to such changes. For instance, significantly increased level of SCFAs such as propanoic acid and valeric acid (pentanoic acid) along with their metabolites: 3-hydroxyvaleric acid and hydantoin-5- propanoic acid, respectively, were found in the urine of oligofructose subjects compared to placebo (maltodextrin) and controls, indicating microbial fermentation of oligofructose and concomitant production of SCFAs in the colon. This finding is quite consistent with previously reported literature as increased levels of propanoic acid, valeric acid and other SCFAs (e.g. acetate and butyrate) were found in faecal samples of healthy subjects treated with oligofructose (Vitali et al., 2010, De Preter et al., 2011, De Preter et al., 2010, De Preter et al., 2013). However, acetate and butyrate were not detected in the urine samples of oligofructose subjects, suggesting that they may be quickly metabolised after being absorbed by the host. Also, acetate and butyrate are volatile, and they may be lost by evaporation during analysis.

Recently, Dewulf et al., performed a urinary metabolomics study to investigate the urinary metabolic signature of obese women after treatment with oligofructose (Dewulf et al., 2013). They reported an increased level of acetate in plasma but not urine, and no altered level of butyrate was found in plasma or urine; this is quite consistent with the above findings. However, another in vitro study reported that inulin was selectively fermented by *bifidobacteria* in the colon and butyrate was not produced from carbohydrates fermentation (Falony et al., 2009). This result suggests that oligofructose, in this study, was mainly fermented by *bifidobacteria*, hence no altered level of butyrate was observed.

Nonetheless, another scenario may also be possible; reduction of butyrate to 1-butanol has been previously reported by gut microbes in order to maintain the bacterial redox balance (Muller, 2001). A significantly increased level of 1-butanol was found in the urine of oligofructose subjects compared to placebo (maltodextrin) and controls, suggesting that butyrate may be produced during oligofructose fermentation in the colon but it was rapidly converted by gut microbes to 1-butanol.

In all these scenarios, regardless of the type of SCFAs produced, the increased synthesis of such biomolecules by microbiota is considered as one of the beneficial effect of prebiotics on gut microbial activity (Cummings, 1981). SCFAs are known to create an acidic environment in the colon (Ndagijimana et al., 2009), which enhance the proliferation of *bifidobacteria* and *lactobacilli* (Gibson et al., 1995), increase the resistance to the growth of enteric pathogens (Topping et al., 2003), decrease the production of secondary bile acids (Zampa et al., 2004), enhance the absorption of minerals and hence, improve bone health (Coxam, 2007, Scholz-Ahrens et al., 2001) and reduce gut protein fermentation (proteolytic fermentation) as a result of impaired activity of protease enzyme at low pH (Macfarlane et al., 1988).

Changes in fatty acid metabolism

Consumption of oligofructose diet resulted in an increased fatty acid metabolism by the host, which was evident from significantly increased levels of different urinary acylcarnitines such as acetylcarnitine, propenoylcarnitine, hexanoylcarnitine, 2-octenoylcarnitine and 6-keto-decanoylcarnitine. Acylcarnitine are formed from conjugation of carnitine and fatty acids, a rate limiting step in the cellular transport of fatty acid from the cytosol into the mitochondrial matrix, where they enter β-oxidation pathway, the major fatty acid metabolic pathway in human (Harvey et al., 2011). Abnormal urinary excretion of acylcarnitine may provide metabolic evidence of enhanced mitochondrial fatty acid metabolism in subjects treated with oligofructose. The increased level of urinary acylcarnitines and mitochondrial β-oxidation activity were previously reported in animal models fed with oligofructose/inulin diet, suggesting colonic fermentation and fatty acid absorption (Verbrugghe et al., 2010, Verbrugghe et al., 2009). The finding of increased activity in the fatty acid

metabolism through β-oxidation due to oligofructose supplement was consistent with the high levels of fatty acids such as propanoic acid, valeric acid, 2-octenoic acid and 3-hexenedioic acid found in the urine of subjects treated with oligofructose compared to controls.

The effect of oligofructose on the activity of gut microbiota

Further analysis of the metabolic profiles showed increased levels of gut bacteria related metabolites such as benzoic acid and 4-hydroxybenzoic acid in the urine of oligofructose treated subjects compared to controls, which may indicate an increased metabolic activity of gut microbiota due to oligofructose. Benzenoids such as benzoic acid and 4-hydroxybenzoic acid have previously been identified in faecal samples and were thought to be related to environmental pollution (Garner et al., 2007). However, some prebiotic intervention studies reported increased levels of benzoic acid and 4-hydroxybenzoic acid as a result of enhanced activity of the gut microbial benzoate degradation pathway (Lacombe et al., 2013) and ubiquinone and other terpenoid-quinone biosynthesis pathway (Monagas et al., 2009). In these studies, the increased level of benzenoids was correlated to the decreased populations of the genera *Lactobacillus* and *Enterococcus* and increased population of *Bifidobacterium*. The increased levels of benzoic acid and 4-hydroxybenzoic acid were quite consistent with the above results, suggesting an increased level of health-promoting *Bifidobacterium* and a decreased level of *Lactobacillus* and *Enterococcus* due to oligofructose. However, simultaneous estimation of microbiota composition in subjects treated with oligofructose compared to controls is an essential prerequisite to correlate and conclude such findings.

Changes in amino acid biosynthesis

Increased levels of amino acids and amino acid metabolites such as glutamine, citrulline, aspartic acid, picolinic acid (catabolite of tryptophan) and acetylglycine (acetylated glycine) were found in the urine of oligofructose subjects compared to controls, signalling the possibility of enhanced microbial amino acid biosynthesis in the colon. From previous studies, none of these amino acids and metabolites were detected in body fluids of healthy subjects treated with oligofructose (Windey et al., 2014, De Preter et al., 2010, De Preter et al., 2011, De Preter et al., 2013, Vitali et al., 2010, Dewulf et al., 2013). However, in vitro

and animal models showed an increased level of glutamine and other amino acids by gut microbes as a result of enhanced microbial amino acid biosynthesis, especially in the presence of oligofructose (Keller et al., 2011, Maczulak et al., 1985). Indeed, increased breakdown of proteins in the colon or dietary amino acids may increase the level of amino acids and their metabolites in urine. However, in this study the dietary intake was adequately controlled before and after the intervention, therefore, it is unlikely that increased level of amino acids and metabolites in oligofructose treated subjects compared to controls were due to diet. Moreover, a decreased level of glutaric acid, a byproduct of protein fermentation in general and of lysine in particular (Muller and Kolker, 2004, Rist et al., 2013), was observed in the urine of subjects treated with oligofructose, indicating that oligofructose may reduce protein fermentation in the gut; hence the increased levels of amino acids and metabolites in the urine of oligofructose subjects were most likely a result of enhanced microbial amino acids biosynthesis. The finding of increased microbial amino acid production was consistent with the high level of tetrahydropteridine, a cofactor in amino acid hydroxylation produced in the liver and GI tract, found in urine of oligofructose subjects compared to controls, which indicates that tetrahydropteridine was up-regulated by the host as a metabolic response to the increased production of amino acids in the colon.

5.5 Conclusion

In this chapter, a non-invasive metabolomics study using the developed LC-MS and FIE-MS methods combined with a multivariate analysis was able to find significant differences between the urinary metabolic profiles of subjects in a low FODMAP and oligofructose diet for a week and controls. This is the first time that HILIC LC-MS and FIE-MS based metabolomics have been used to study the effect of oligofructose on gut microbiota using urine samples from healthy volunteers. The use of spot urine and 24 h urine samples were evaluated for the analysis, in which lower variability was observed in the 24 h urine profiles compared to spot urine. The OPLS-DA models generated from the datasets of oligofructose subjects and controls, demonstrated good classification of the study groups with both MS methods. FIE-MS as a high-throughput method was able to provide about 45% of the information obtained with LC-MS, indicating its potential as a quick screening tool in dietary interventions, while LC-MS is necessary when comprehensive screening is required. The biological interpretation of the metabolites associated with the separation of study groups in the OPLS-DA models, revealed an increased production of SCFAs, fatty acids and amino acids by gut microbiota as a result of oligofructose intake. Also, an increased activity in fatty acid metabolism was evident from the increased levels of different types of acylcarnitines in the subjects treated with oligofructose. Comparison of these finding with previously reported literature, suggesting an increased population of a health-promoting *Bifidobacterium* and decreased *Lactobacillus* and *Enterococcus* genera in the colon due to oligofructose. However, further investigation of the effect of oligofructose on gut microbial composition is necessary to conclude such findings.

CHAPTER SIX

Conclusions and Future Work

6. Conclusions and Future Work

6.1 Methods development and validation for urine metabolomics

In this thesis, a validated HILIC LC-MS and three high-throughput direct ESI-MS approaches: FIE-MS, chip-based infusion and LESA-MS, using high resolution mass spectrometry were developed for urine untargeted metabolomics. A urine dilution protocol was developed for sample preparation, in which the ion suppression effect of urine salts on ESI was studied and optimised to enhance detection of urinary metabolites using artificial urine. Great attention to details were applied in the design, development and execution of the MS methods alongside a rigorous validation of the analytical methods and multivariate models (Want et al., 2010, Gika et al., 2007, Beckmann et al., 2008), which lends credibility to the quality of results obtained. The developed methods were capable of investigating metabolic profiles and studying the altered pathways in various diseases such as osteoarthritis and malaria and the effect of oligofructose diet on the gut microbial community. However, there are many potentially wider applications of the developed methods; they could be applied into any study involving urine analysis, not just those applications in this thesis.

Urine as a readily available biofluid has been an attractive choice for more comprehensive biomarker investigations using metabolomics approaches. However, the high level of day-to-day variability and the salt content of urine, makes its analysis with direct ESI-MS method a complex and a difficult area of research, indicated by the relatively small number of untargeted metabolomics studies reported (Hasegawa et al., 2007, Hasegawa et al., 2010, Lloyd et al., 2011a, Lloyd et al., 2011b, Beckmann et al., 2010, Gonzalez-Dominguez et al., 2014). Considering the above challenges, the developed methods have made many improvements over the existing methodologies; for instance, ion suppression effect of urine salts on ESI was minimised by optimising the urine dilution protocol, and the MS detection was enhanced for urine metabolites using a mixture of standards representing a wide range of urine chemical classes (i.e. artificial urine). Also, a type of MS total useful signal (MSTUS) normalisation (Warrack et al., 2009) was used for LC-MS and direct ESI-MS to compensate for urine variability as it has proven to reduce the possible xenobiotics and artefact interferences associated with urine analysis. Moreover, several attempts in the literature have been reported to improve the performance of LC-MS and direct

ESI-MS by using advanced ESI sources such as nanoelectrospray ionisation (nanoESI) (Trunzer et al., 2007, Froesch et al., 2004, Zamfir et al., 2004, Chetwynd et al., 2015, Chetwynd et al., 2014). In this study, the versatility of nanoESI has been tested for urine analysis by developing and validating a chip-based infusion and LESA-MS methods for untargeted metabolomics.

One particularly challenging aspect of the method development was performing validation. Most of the available reports of direct ESI-MS analysis of urine lack proper validation (Hasegawa et al., 2010, Hasegawa et al., 2007) or evaluation against conventional methods such as LC-MS for untargeted metabolomics. Therefore, before commencement of clinical urine analysis, the developed direct ESI-MS methods were validated using a small-scale dietary intervention with green tea. The results were compared with the developed LC-MS method. Although fewer metabolites were detected by direct ESI-MS compared to LC-MS, the direct ESI-MS datasets still generated a robust model as compared with LC-MS. In direct ESI-MS, the run time is less than 1.0 min compared with several 10s min for LC-MS, hence, direct ESI-MS undoubtedly provides a much higher analysis throughput. Chip-based infusion and LESA-MS use a disposable chip with nanoESI nozzles provide analysis without carryover effects. Although this adds an extra cost to the analysis it results in improved the throughput as well. In terms of overall performance of the developed methods, FIE-MS provided the best throughput solution compared with chip-based infusion and LESA-MS. It had the lowest analysis cost/sample, lower variability and detects more metabolites in urine samples. However, LESA-MS with no sample preparation step gives the method the advantage of being the highest throughput by reducing pre-analysis time significantly.

The strengths of the developed methods in this thesis are: (1) the use of rigorously optimised urine dilution protocol to ensure wide metabolic coverage, (2) the use of validated HILIC LC-MS, FIE-MS, chip-based infusion and LESA-MS for ensured a high specificity and sensitivity, (3) the use of high throughput direct ESI-MS methods to detect most of the metabolites in shorter analytical runs, hence minimising the risks of invalid results due to sample instability during the analysis, (4) exploring the versatility of LESA-MS with no sample preparation step for the metabolic profiling of urine with its potential ability as a simple and rapid diagnostic and screening tool, and (5) the rigorous validation methods used in the data analysis to minimise the risks of false positive

biomarkers and on the same time retaining the true biomarkers from being missed. The ability of the developed methods to differentiate and classify urine samples after ingestion of green tea from controls (reported in chapter 2) clarify that these methods have the potential to differentiate disease from control of various clinical studies using urine samples to identify potential disease biomarkers and study metabolic perturbations during health and disease state.

In this thesis, certain caveats and method limitations should also be addressed, including, direct ESI-MS lacks the ability to resolve isobaric compounds and fewer peaks were detected compared to LC-MS. However, the high throughput power of the developed methods (i.e. they require ~6.7% of the LC-MS analysis time) and yet provide about 40-50% information compared to LC-MS. Also, the overall batch analysis time for both LC-MS and direct ESI-MS should be less than 48 h as significant changes in urine composition were reported when urine samples kept for the analysis for more than 48 h (Gika et al., 2008c). Considering such constraints, direct ESI-MS methods have the potential to analyse larger batch of samples compared to LC-MS, hence more information could be gained. In addition, LC-MS/MS method was developed for confirming the identity of the reported urinary biomarkers, but the lack of available suitable standards to generate an MS/MS spectral database, meant that the discovered biomarkers could only be putatively identified by accurate mass alone. Some biomarkers were definitively identified with reference to pure authentic standards, however, funds and time placed constraints to confirm the identity of all reported biomarkers in this thesis. Also, in clinical applications of the developed methods, some of the biomarkers remained unknown due to the limitation of the available databases. Therefore, further investigation to reveal the identity of these metabolites may provide additional information; however, structure elucidation of these metabolites is beyond the scope of this thesis.

6.2 The clinical applications of the developed LC-MS and direct ESI-MS for urine metabolomics

Metabolomics studies have identified different biomarkers of osteoarthritis (OA) in serum, synovial fluid and plasma which helped in studying the pathophysiology, progression and diagnosis of OA (Rousseau and Delmas, 2007, Mobasheri, 2012, Loeuille, 2012a, Lotz et al., 2013, Goldring and Berenbaum, 2015). However, there are few studies in the literature related to the use of urine metabolomics in OA (Li et al., 2010, Nepple et al., 2015, Lamers et al., 2003a, Lamers et al., 2005). In this thesis, the validated LC-MS and direct ESI-MS methods were applied to study urinary metabolic signature of OA. The developed methods successfully differentiated between the metabolic profiles of OA patients in the active stage and healthy controls as detailed in chapter 3. A novel multivariate design and SUS plot was used to balance the number of samples in each group in the study for multivariate analysis. FIE-MS, chip-based infusion and LESA-MS were able to provide 33%, 44% and 44%, respectively, of the LC-MS information, indicating their great potential for diagnostic application in OA, while LC-MS is necessary when comprehensive biomarker screening is required. The altered level of metabolites detected in OA patients suggesting a perturbed activity in TCA cycle, pyruvate metabolism, β-oxidation pathway, amino acids and glycerophospholipids metabolism, which supports the previously reported literature of mitochondrial dysfunction, inflammation, oxidative stress, collagen destruction and use of lipolysis as an alternative energy source in the cartilage cells of OA patients (Gavriilidis et al., 2013, Blanco et al., 2004, Li et al., 2010, Damyanovich et al., 1999, Kosinska et al., 2014, Zhang et al., 2014b).

In chapter 4, the developed LC-MS and FIE-MS methods were applied to investigate the effect of malaria on urinary metabolic profiles and for biomarker discovery. Both methods successfully identified potential biomarkers from the urinary profiles of malaria patients and healthy controls. Some of the biomarkers were definitively identified by LC-MS/MS with reference to authentic standards. FIE-MS was able to cover 50% of the LC-MS information and detected all the potential candidates found with LC-MS. These findings indicated that FIE-MS performance was comparable to LC-MS and demonstrated its great potential as a high throughput diagnostic tool for malaria. The altered level of urinary pipecolic acid and taurine in malaria was consistent with previously reported studies (Sengupta et al., 2015, Ghosh et al., 2012, Sengupta et al., 2011b, Li et al., 2008,

Tritten et al., 2013), supporting the suggestion of their use as diagnostic biomarkers of malaria. Whereas the altered levels of acetylated polyamines such as 1,3-diacetylpropane, N-acetylspermidine and N-acetylputrescine were detected in the urine of malaria patients for the first time, which may provide surrogate biomarkers of malaria.

In chapter 5, the developed methods were also applied in nutrition. Different dietary interventions have been reported to investigate the effect of oligofructose and its fermentation products on intestinal microbiota in health and disease state (Windey et al., 2014, Vitali et al., 2010, Verhoef et al., 2011, De Preter et al., 2013, De Preter et al., 2011, De Preter et al., 2010, Dewulf et al., 2013). Most of these studies relied on the analysis of faecal samples (De Preter et al., 2010, De Preter et al., 2011, De Preter et al., 2013, Windey et al., 2014), despite the anticipated loss of metabolites in faeces due to absorption (Lewis and Cochrane, 2007). Few studies used urine for untargeted metabolomics to study the effect of oligofructose on microbiota in cats (Verbrugghe et al., 2010) and obese women (Dewulf et al., 2013). In this study, the developed LC-MS and FIE-MS methods were applied for urinary metabolic profiling of healthy volunteers to study the effect of the low FODMAP and oligofructose diet on gut microbiota. The developed methods were able to find significant differences between the urinary metabolic profiles of healthy subjects on oligofructose diet and controls. FIE-MS as a high throughput method was able to provide about 45% of the information obtained with LC-MS. An increased production of SCFAs, fatty acids and amino acids by gut microbiota as a result of oligofructose intake were observed, suggesting an increased population of the health-promoting *Bifidobacterium* and a decreased *Lactobacillus* and *Enterococcus* genera in the colon. However, further investigation of the effect of oligofructose on gut microbial composition is necessary to correlate and conclude such findings.

General parameters such as diet, life-style, genetic disposition and sample collection protocol may influence the metabolic profiles between the study groups and subsequently the biological outcome. In order to circumvent these possible confounding factors, metabolic profiling of urine samples (pooled 24 h urine sample is preferential, if possible) from cohorts of study and control subjects with a longitudinal follow-up and a controlled dietary intake may provide more specific information directly related to metabolic changes found between the study groups.

6.3 The potential applications of the developed direct ESI-MS in clinical practice and research

In conclusion, the developed direct ESI-MS methods demonstrated the ability to obtain relevant information for diagnosis and differentiation between inherent types of urine samples in disease and health state. The use of urine with all its advantages as a simple, non-invasive and requires minimal skills to handle, suggesting that the above direct ESI-MS methods can be implemented in clinical practice (e.g. hospitals) for high throughput screening and diagnosis of diseases. The identified list of urinary biomarkers from OA, malaria and oligofructose dietary intervention may provide further insight in identifying disease mechanisms and progression, as discussed with comparison to the relevant literature in chapter 3-5. However, these biomarkers require certain prerequisites (which is beyond the scope of this thesis) to be considered specific enough for diagnostic purposes. Most of the biomarkers were tentatively identified, hence absolute confirmation of identity using LC-MS/MS is desirable. Following this, a targeted analysis of these biomarkers will help in quantifying each metabolite in the samples and increase the confidence in specificity of the biomarker for the disease under question. Addressing these issues, the methodologies could be applied in two ways: the discovery and confirmation of biomarkers could be used to develop a urine 'dipstick' type of immune method to enable rapid screening with simple equipment, and with the use of cheap portable mass spectrometers coupled with ambient or direct analysis could make a mobile diagnostic screening unit with rapid feedback to patients. For instance, FIE-MS demonstrated high performance as LC-MS in detecting surrogate urinary biomarkers of *P. falciparum* infection, and yet provides a high throughput analysis. Hence, the method can be used as a portable screening tool for early diagnosis of malaria in endemic areas, where the availability of equipped laboratory and skilled personnel are limited. Furthermore, the use of chip-based infusion and LESA-MS adds further advantages for their clinical use as they demonstrated a higher throughput power, enhanced sensitivity and no carry-over effect. In particular, LESA-MS requires no sample preparation step and only involves placing urine samples directly onto a slide surface, gives the method further credit of being very simple and easy to use in clinical practice. While the developed LC-MS is necessary when comprehensive biomarker screening is required.

The developed urine dilution protocol for the direct ESI-MS demonstrated high reproducibility with minimum ion suppression on the ionisation of urinary metabolites, therefore, it may be beneficial to transfer such protocol for LC-MS analysis to improve sensitivity and coverage of urinary metabolites. Moreover, the use of the nanoESI interface for direct ESI-MS demonstrated higher sensitivity and S/N ratio compared to standard ESI interface (Table 2-4). Therefore, development of an LC-MS method combined with the use of the nanoESI interface might improve the coverage of urine metabolites for metabolomics analysis. Recently, Chetwynd et al., reported an improved MS sensitivity of low abundance compounds in urine metabolomics using nanoESI with nanoUHPLC and a wider range of metabolites was detected compared to standard LC-MS (Chetwynd et al., 2014, Chetwynd et al., 2015). Nevertheless, the use of LC-MS for comperhensive screening of large batches of urine samples (e.g. 200 samples) is limited to the stability of the urine samples during the analysis (less than 48 h) and the stability of the analytical system (Gika et al., 2008c). This problem can be overcome by using LC-MS batch analysis, however, extensive data analysis and certain techniques such as nonlinear locally estimated smoothing function (LOESS) (Cleveland, 1979) are required to correct for signal intensity drift and integration of data from multiple analytical batches (Dunn et al., 2011).

It is worth mentioning that the developed method has already been used by other researchers in the group; FIE-MS successfully applied for urinary metabolic signature after broccoli consumption, whereas chip-based infusion and LESA-MS have been adapted for targeted analysis of analgesics and their metabolites in urine. Also, LESA-MS has been used for targeted analysis of opiate analgesics in saliva, plasma and urine. The established data analysis protocol has also been used for LESA-MS metabolic profiling of human skin, and profiling of amino acids adsorbed on surfaces treated with oxygen and plasma.

6.4 Future work and recommendations

Additional work which is beyond the scope of this thesis would be beneficial and includes definitive confirmation of biomarker identity using the developed LC-MS/MS, and detailed pathway analysis and mechanisms of each biomarker in the study. In this thesis, due to time and funds constraints most of the reported biomarkers were tentatively identified and these could be better defined by using the developed LC-MS/MS method with reference to pure authentic standards to confirm the identity and to increase the confidence of the obtained results. A detailed further targeted MS method for a specific class of urinary metabolites could give better qualitative and quantitative information of the reported biomarkers. Nonetheless, the developed methods could be further adapted as a targeted method. Moreover, the developed direct ESI-MS can be further tested and adapted with a portable mass spectrometer for disease diagnosis such as malaria; however, design modification is essentially required.

In some studies of this thesis such as malaria (n = 46) and oligofructose dietary intervention (n = 37) it would be advantageous to have a larger group of subjects for untargeted metabolomics. Also some of the reported biomarkers remained unknown due to limitation of databases; it would be better to structurally elucidate the identity of those metabolites. This will provide additional information which may help in further understanding of the biological processes related to the study in question. Finally the developed methods could be applied into any study involving urine analysis and therefore, have the potential to be applied in different research areas such as drug doping, toxicology and clinical trials.

CHAPTER SEVEN

References

7. References

AFENDI, F. M., OKADA, T., YAMAZAKI, M., HIRAI-MORITA, A., NAKAMURA, Y., NAKAMURA, K., IKEDA, S., TAKAHASHI, H., ALTAF-UL-AMIN, M., DARUSMAN, L. K., SAITO, K. & KANAYA, S. 2012. KNApSAcK Family Databases: Integrated Metabolite–Plant Species Databases for Multifaceted Plant Research. *Plant and Cell Physiology,* 53, e1.

AHARONI, A., RIC DE VOS, C. H., VERHOEVEN, H. A., MALIEPAARD, C. A., KRUPPA, G., BINO, R. & GOODENOWE, D. B. 2002. Nontargeted Metabolome Analysis by Use of Fourier Transform Ion Cyclotron Mass Spectrometry. *OMICS: A Journal of Integrative Biology,* 6, 217-234.

ALBERICE, J. V., AMARAL, A. F. S., ARMITAGE, E. G., LORENTE, J. A., ALGABA, F., CARRILHO, E., MŒRQUEZ, M., GARCÍA, A., MALATS, N. & BARBAS, C. 2013. Searching for urine biomarkers of bladder cancer recurrence using a liquid chromatography–mass spectrometry and capillary electrophoresis–mass spectrometry metabolomics approach. *Journal of Chromatography A,* 1318, 163-170.

ALLEN, J., DAVEY, H. M., BROADHURST, D., HEALD, J. K., ROWLAND, J. J., OLIVER, S. G. & KELL, D. B. 2003. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nature Biotechnology,* 21, 692-696.

ALLEN, J., DAVEY, H. M., BROADHURST, D., ROWLAND, J. J., OLIVER, S. G. & KELL, D. B. 2004. Discrimination of Modes of Action of Antifungal Substances by Use of Metabolic Footprinting. *Applied and Environmental Microbiology,* 70, 6157-6165.

ALMEIDA, R., MOSOARCA, C., CHIRITA, M., UDRESCU, V., DINCA, N., VUKELIC, Z., ALLEN, M. & ZAMFIR, A. D. 2008. Coupling of fully automated chip-based electrospray ionization to high-capacity ion trap mass spectrometer for ganglioside analysis. *Analytical Biochemistry,* 378, 43-52.

ALVES, J., NETO, W. B., MITSUTAKE, H., ALVES, P. S. P. & AUGUSTI, R. 2010. Extra virgin (EV) and ordinary (ON) olive oils: distinction and detection of adulteration (EV with ON) as determined by direct infusion electrospray ionization mass spectrometry and chemometric approaches. *Rapid Communications in Mass Spectrometry,* 24, 1875-1880.

AMARAL, P. H., FERNANDES, R., EBERLIN, M. N. & HOEHR, N. F. 2011. Direct monitoring of drug degradation by easy ambient sonic-spray ionization mass spectrometry: the case of enalapril. *Journal of Mass Spectrometry,* 46, 1269-1273.

ANDERSEN, M.-B. S., RINNAN, A., MANACH, C., POULSEN, S. K., PUJOS-GUILLOT, E., LARSEN, T. M., ASTRUP, A. & DRAGSTED, L. O. 2014. Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. *Journal of Proteome Research,* 13, 1405-1418.

ANDRADE, P. B., FERRERES, F., PEREIRA, D. M. & VALENTAO, P. 2010. *Metabolomic analysis of natural products*, Bentham Science Publishers Ltd.

ANTUNES, L. C. M., ANDERSEN, S. K., MENENDEZ, A., ARENA, E. T., HAN, J., FERREIRA, R. B. R., BORCHERS, C. H. & FINLAY, B. B. 2011a. Metabolomics Reveals Phospholipids as Important Nutrient Sources during Salmonella Growth in Bile In Vitro and In Vivo. *Journal of Bacteriology,* 193**,** 4719-4725.

ANTUNES, L. C. M., ARENA, E. T., MENENDEZ, A., HAN, J., FERREIRA, R. B. R., BUCKNER, M. M. C., LOLIĆ, P., MADILAO, L. L., BOHLMANN, J., BORCHERS, C. H. & FINLAY, B. B. 2011b. Impact of Salmonella Infection on Host Hormone Metabolism Revealed by Metabolomics. *Infection and Immunity,* 79**,** 1759-1769.

ARAKI, S., SATA, F. & MURATA, K. 1990. Adjustment for urinary flow rate: an improved approach to biological monitoring. *International Archives of Occupational and Environmental Health,* 62**,** 471-477.

ARDAWI, M. S. & NEWSHOLME, E. A. 1984. Metabolism of ketone bodies, oleate and glucose in lymphocytes of the rat. *Biochemical Journal,* 221**,** 255-260.

ARNETT, F. C., EDWORTHY, S. M., BLOCH, D. A., MCSHANE, D. J., FRIES, J. F., COOPER, N. S., HEALEY, L. A., KAPLAN, S. R., LIANG, M. H., LUTHRA, H. S., MEDSGER, T. A., MITCHELL, D. M., NEUSTADT, D. H., PINALS, R. S., SCHALLER, J. G., SHARP, J. T., WILDER, R. L. & HUNDER, G. G. 1988. The american rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheumatism,* 31**,** 315-324.

ARSLANOGLU, S., MORO, G. E. & BOEHM, G. 2007. Early Supplementation of Prebiotic Oligosaccharides Protects Formula-Fed Infants against Infections during the First 6 Months of Life. *The Journal of Nutrition,* 137**,** 2420-2424.

ARTHRITIS RESEARCH UK 2013. Osteoarthritis in general practice report. Chesterfield, UK.

ASSARAF, Y. G., GOLENSER, J., SPIRA, D. T. & BACHRACH, U. 1984. Polyamine levels and the activity of their biosynthetic enzymes in human erythrocytes infected with the malarial parasite, Plasmodium falciparum. *Biochemical Journal,* 222**,** 815-819.

BAIGORRI, R., ZAMARRENO, A. M., FUENTES, M., GONZALEZ-GAITANO, G., GARCIA-MINA, J. M., ALMENDROS, G. & GONZALEZ-VILA, F. J. 2008. Multivariate Statistical Analysis of Mass Spectra as a Tool for the Classification of the Main Humic Substances According to Their Structural and Conformational Features. *Journal of Agricultural and Food Chemistry,* 56**,** 5480-5487.

BARAN R FAU - KOCHI, H., KOCHI H FAU - SAITO, N., SAITO N FAU - SUEMATSU, M., SUEMATSU M FAU - SOGA, T., SOGA T FAU - NISHIOKA, T., NISHIOKA T FAU - ROBERT, M., ROBERT M FAU

- TOMITA, M. & TOMITA, M. 2006. MathDAMP: a package for differential analysis of metabolite profiles. *BMC Bioinformatics.,* 13, 530.

BARNWELL, J. W., NICHOLS, M. E. & RUBINSTEIN, P. 1989. In vitro evaluation of the role of the Duffy blood group in erythrocyte invasion by Plasmodium vivax. *The Journal of experimental medicine,* 169, 1795-1802.

BASANT, A., REGE, M., SHARMA, S. & SONAWAT, H. 2010. Alterations in urine, serum and brain metabolomic profiles exhibit sexual dimorphism during malaria disease progression. *Malaria Journal,* 9, 110.

BAUER, D. C., HUNTER, D. J., ABRAMSON, S. B., ATTUR, M., CORR, M., FELSON, D., HEINEGARD, D., JORDAN, J. M., KEPLER, T. B., LANE, N. E., SAXNE, T., TYREE, B. & KRAUS, V. B. 2006. Classification of osteoarthritis biomarkers: a proposed approach. *Osteoarthritis and Cartilage,* 14, 723-727.

BECKMANN, M., ENOT, D. P., OVERY, D. P. & DRAPER, J. 2007. Representation, Comparison, and Interpretation of Metabolome Fingerprint Data for Total Composition Analysis and Quality Trait Investigation in Potato Cultivars. *Journal of Agricultural and Food Chemistry,* 55, 3444-3451.

BECKMANN, M., ENOT, D. P., OVERY, D. P., SCOTT, I. M., JONES, P. G., ALLAWAY, D. & DRAPER, J. 2010. Metabolite fingerprinting of urine suggests breed-specific dietary metabolism differences in domestic dogs. *British Journal of Nutrition,* 103, 1127-1138.

BECKMANN, M., PARKER, D., ENOT, D. P., DUVAL, E. & DRAPER, J. 2008. High-throughput, nontargeted metabolite fingerprinting using nominal mass flow injection electrospray mass spectrometry. *Nature Protocols,* 3, 486-504.

BENTON, H. P., WONG, D. M., TRAUGER, S. A. & SIUZDAK, G. 2008. XCMS2: Processing Tandem Mass Spectrometry Data for Metabolite Identification and Structural Characterization. *Analytical Chemistry,* 80, 6382-6389.

BERGMEYER, H. U. 1984. *Methods of enzymatic analysis,* Weinheim, Verlag Chemie.

BERRIDGE, M. J. & IRVINE, R. F. 1989. Inositol phosphates and cell signalling. *Nature,* 341, 197-205.

BHATTACHARYA, M., FUHRMAN, L., INGRAM, A., NICKERSON, K. W. & CONWAY, T. 1995. Single-Run Separation and Detection of Multiple Metabolic Intermediates by Anion-Exchange High-Performance Liquid Chromatography and Application to Cell Pool Extracts Prepared from Escherichia coli. *Analytical Biochemistry,* 232, 98-106.

BIJLSMA, S., BOBELDIJK, I., VERHEIJ, E. R., RAMAKER, R., KOCHHAR, S., MACDONALD, I. A., VAN OMMEN, B. & SMILDE, A. K. 2006. Large-Scale Human Metabolomics Studies: A Strategy for Data (Pre-) Processing and Validation. *Analytical Chemistry,* 78, 567-574.

BJORKMAN, A. & MARTENSSON, A. 2010. Risks and Benefits of Targeted Malaria Treatment Based on Rapid Diagnostic Test Results. *Clinical Infectious Diseases,* 51, 512-514.

BLANCO, F. J., LÓPEZ-ARMADA, M. J. & MANEIRO, E. 2004. Mitochondrial dysfunction in osteoarthritis. *Mitochondrion,* 4, 715-728.

BLAUT, M. 2002. Relationship of prebiotics and food to intestinal microflora. *European Journal of Nutrition,* 41, i11-i16.

BLIDDAL, H. & CHRISTENSEN, R. 2006. The management of osteoarthritis in the obese patient: practical considerations and guidelines for therapy. *Obesity Reviews,* 7, 323-331.

BLUM, R. A., COMSTOCK, T. J., SICA, D. A., SCHULTZ, R. W., KELLER, E., REETZE, P., BOCKBRADER, H., TUERCK, D., BUSCH, J. A., REECE, P. A. & SEDMAN, A. J. 1994. Pharmacokinetics of gabapentin in subjects with various degrees of renal function. *Clinical Pharmacology & Therapeutics,* 56, 154-159.

BLYDT-HANSEN, T. D., SHARMA, A., GIBSON, I. W., MANDAL, R. & WISHART, D. S. 2014. Urinary Metabolomics for Noninvasive Detection of Borderline and Acute T Cell–Mediated Rejection in Children After Kidney Transplantation. *American Journal of Transplantation,* 14, 2339-2349.

BOLLARD, M. E., STANLEY, E. G., LINDON, J. C., NICHOLSON, J. K. & HOLMES, E. 2005. NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. *NMR in Biomedicine,* 18, 143-162.

BONFIGLIO, R., KING, R. C., OLAH, T. V. & MERKLE, K. 1999. The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds. *Rapid Communications in Mass Spectrometry,* 13, 1175-1185.

BORNER, J., BUCHINGER, S. & SCHOMBURG, D. 2007. A high-throughput method for microbial metabolome analysis using gas chromatography/mass spectrometry. *Analytical Biochemistry,* 367, 143-151.

BOUATRA, S., AZIAT, F., MANDAL, R., GUO, A. C., WILSON, M. R., KNOX, C., BJORNDAHL, T. C., KRISHNAMURTHY, R., SALEEM, F., LIU, P., DAME, Z. T., POELZER, J., HUYNH, J., YALLOU, F. S., PSYCHOGIOS, N., DONG, E., BOGUMIL, R., ROEHRING, C. & WISHART, D. S. 2013. The Human Urine Metabolome. *PLoS ONE,* 8, e73076.

BOUHNIK, Y., VAHEDI, K., ACHOUR, L., ATTAR, A., SALFATI, J., POCHART, P., MARTEAU, P., FLOURIÉ, B., BORNET, F. & RAMBAUD, J.-C. 1999. Short-Chain Fructo-Oligosaccharide Administration Dose-Dependently Increases Fecal Bifidobacteria in Healthy Humans. *The Journal of Nutrition,* 129, 113-116.

BRADFORD, B. U., O'CONNELL, T. M., HAN, J., KOSYK, O., SHYMONYAK, S., ROSS, P. K., WINNIKE, J., KONO, H. & RUSYN, I. 2008. Metabolomic profiling of a modified alcohol liquid diet model for liver injury in the mouse uncovers new markers of disease. *Toxicology and Applied Pharmacology,* 232, 236-243.

BROADHURST, D. & KELL, D. 2006. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics,* 2, 171-196.

BROECKLING, C. D., REDDY, I. R., DURAN, A. L., ZHAO, X. & SUMNER, L. W. 2006. MET-IDEA: Data Extraction Tool for Mass Spectrometry-Based Metabolomics. *Analytical Chemistry,* 78, 4334-4341.

BROOKS, T. & KEEVIL, C. W. 1997. A simple artificial urine for the growth of urinary pathogens. *Letters in Applied Microbiology,* 24, 203-206.

BUCHHOLZ, A., HURLEBAUS, J., WANDREY, C. & TAKORS, R. 2002. Metabolomics: quantification of intracellular metabolite dynamics. *Biomolecular Engineering,* 19, 5-15.

BUNK, B., KUCKLICK, M., JONAS, R., MUNCH, R., SCHOBERT, M., JAHN, D. & HILLER, K. 2006. MetaQuant: a tool for the automatic quantification of GC/MS-based metabolome data. *Bioinformatics,* 22, 2962-2965.

BUSZEWSKI, B. & NOGA, S. 2012. Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Analytical and Bioanalytical Chemistry,* 402, 231-247.

BYLESJO, M., RANTALAINEN, M., CLOAREC, O., NICHOLSON, J. K., HOLMES, E. & TRYGG, J. 2006. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics,* 20, 341-351.

CAI, H.-L., LI, H.-D., YAN, X.-Z., SUN, B., ZHANG, Q., YAN, M., ZHANG, W.-Y., JIANG, P., ZHU, R.-H., LIU, Y.-P., FANG, P.-F., XU, P., YUAN, H.-Y., ZHANG, X.-H., HU, L., YANG, W. & YE, H.-S. 2012. Metabolomic Analysis of Biochemical Changes in the Plasma and Urine of First-Episode Neuroleptic-Naïve Schizophrenia Patients after Treatment with Risperidone. *Journal of Proteome Research,* 11, 4338-4350.

CANI, P. D., JOLY, E., HORSMANS, Y. & DELZENNE, N. M. 2005. Oligofructose promotes satiety in healthy human: a pilot study. *The American Journal of Clinical Nutrition,* 60, 567-572.

CANI, P. D., LECOURT, E., DEWULF, E. M., SOHET, F. M., PACHIKIAN, B. D., NASLAIN, D., DE BACKER, F., NEYRINCK, A. M. & DELZENNE, N. M. 2009. Gut microbiota fermentation of prebiotics increases satietogenic and incretin gut peptide production with consequences for appetite sensation and glucose response after a meal. *The American Journal of Clinical Nutrition,* 90, 1236-1243.

CANTAREL, B. L., WAUBANT, E., CHEHOUD, C., KUCZYNSKI, J., DESANTIS, T. Z., WARRINGTON, J., VENKATESAN, A., FRASER, C. M. & MOWRY, E. M. 2015. Gut Microbiota in Multiple Sclerosis: Possible Influence of Immunomodulators. *Journal of Investigative Medicine,* 63, 729-734.

CAO, M., KOULMAN, A., JOHNSON, L. J., LANE, G. A. & RASMUSSEN, S. 2008. Advanced Data-Mining Strategies for the Analysis of Direct-Infusion Ion Trap Mass Spectrometry Data from the Association of Perennial Ryegrass with Its Endophytic Fungus, Neotyphodium lolii. *Plant Physiology,* 146, 1501-1514.

CARPITA, N. C., KANABUS, J. & HOUSLEY, T. L. 1989. Linkage Structure of Fructans and Fructan Oligomers from Triticum aestivum and Festuca arundinacea Leaves. *Journal of Plant Physiology,* 134, 162-168.

CARRIERI, M., TREVISAN, A. & BARTOLUCCI, G. B. 2000. Adjustment to concentration-dilution of spot urine samples: correlation between specific gravity and creatinine. *International Archives of Occupational and Environmental Health,* 74, 63-67.

CASTRILLO, J. I., HAYES, A., MOHAMMED, S., GASKELL, S. J. & OLIVER, S. G. 2003. An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. *Phytochemistry,* 62, 929-937.

CATCHPOLE, G. S., BECKMANN, M., ENOT, D. P., MONDHE, M., ZYWICKI, B., TAYLOR, J., HARDY, N., SMITH, A., KING, R. D., KELL, D. B., FIEHN, O. & DRAPER, J. 2005. Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. *Proceedings of the National Academy of Sciences of the United States of America,* 102, 14458-14462.

CATHARINO, R. R., MILAGRE, H. M. S., SARAIVA, S. A., GARCIA, C. M., SCHUCHARDT, U., EBERLIN, M. N., AUGUSTI, R., PEREIRA, R. C. L., GUIMAR-ES, M. J. R., DE SŒ, G. F., CAIXEIRO, J. M. R. & DE SOUZA, V. 2007. Biodiesel Typification and Quality Control by Direct Infusion Electrospray Ionization Mass Spectrometry Fingerprinting. *Energy & Fuels,* 21, 3698-3701.

CDC. 2015. *Malaria* [Online]. Centres for Disease Control and Prevention. Available: http://www.cdc.gov/malaria/about/disease.html [Accessed 24/5 2015].

CEGLAREK, U., LEICHTLE, A., BRÜGEL, M., KORTZ, L., BRAUER, R., BRESLER, K., THIERY, J. & FIEDLER, G. M. 2009. Challenges and developments in tandem mass spectrometry based clinical metabolomics. *Molecular and Cellular Endocrinology,* 301, 266-271.

CHADHA, V., GARG, U. & ALON, U. S. 2001. Measurement of urinary concentration: a critical appraisal of methodologies. *Pediatric nephrology (Berlin, Germany),* 16, 374-382.

CHAIMBAULT, P., ALBERIC, P., ELFAKIR, C. & LAFOSSE, M. 2004. Development of an LC–MS–MS method for the quantification of taurine derivatives in marine invertebrates. *Analytical Biochemistry,* 332, 215-225.

CHEN, G.-Y., LIAO, H.-W., TSENG, Y. J., TSAI, I. L. & KUO, C.-H. 2015a. A matrix-induced ion suppression method to normalize concentration in urinary metabolomics studies using flow injection analysis electrospray ionization mass spectrometry. *Analytica Chimica Acta,* 864, 21-29.

CHEN, H., YUAN, B., MIAO, H., TAN, Y., BAI, X., ZHAO, Y.-Y. & WANG, Y. 2015b. Urine metabolomics reveals new insights into hyperlipidemia and the therapeutic effect of rhubarb. *Analytical Methods*.

CHEN, J., ZHOU, L., ZHANG, X., LU, X., CAO, R., XU, C. & XU, G. 2012. Urinary hydrophilic and hydrophobic metabolic profiling based on liquid chromatography-mass spectrometry methods: Differential metabolite discovery specific to ovarian cancer. *ELECTROPHORESIS,* 33, 3361-3369.

CHEN, P., HARNLY, J. M. & LESTER, G. E. 2010a. Flow Injection Mass Spectral Fingerprints Demonstrate Chemical Differences in Rio Red Grapefruit with Respect to Year, Harvest Time, and Conventional versus Organic Farming. *Journal of Agricultural and Food Chemistry,* 58, 4545-4553.

CHEN, P., LIN, L.-Z. & HARNLY, J. M. 2010b. Mass spectroscopic fingerprinting method for differentiation between Scutellaria lateriflora and the germander (Teucrium canadense and T. chamaedrys) species. *Journal of AOAC International,* 93, 1148-1154.

CHEN, Y., SHEN, G., ZHANG, R., HE, J., ZHANG, Y., XU, J., YANG, W., CHEN, X., SONG, Y. & ABLIZ, Z. 2013a. Combination of Injection Volume Calibration by Creatinine and MS Signals' Normalization to Overcome Urine Variability in LC-MS-Based Metabolomics Studies. *Analytical Chemistry,* 85, 7659-7665.

CHEN, Y., XU, J., ZHANG, R., SHEN, G., SONG, Y., SUN, J., HE, J., ZHAN, Q. & ABLIZ, Z. 2013b. Assessment of data pre-processing methods for LC-MS/MS-based metabolomics of uterine cervix cancer. *Analyst,* 138, 2669-2677.

CHETWYND, A. J., ABDUL-SADA, A. & HILL, E. M. 2015. Solid-Phase Extraction and Nanoflow Liquid Chromatography-Nanoelectrospray Ionization Mass Spectrometry for Improved Global Urine Metabolomics. *Analytical Chemistry,* 87, 1158-1165.

CHETWYND, A. J., DAVID, A., HILL, E. M. & ABDUL-SADA, A. 2014. Evaluation of analytical performance and reliability of direct nanoLC-nanoESI-high resolution mass spectrometry for profiling the (xeno)metabolome. *Journal of Mass Spectrometry,* 49, 1063-1069.

CHIODINI, P. L., BOWERS, K., JORGENSEN, P., BARNWELL, J. W., GRADY, K. K., LUCHAVEZ, J., MOODY, A. H., CENIZAL, A. & BELL, D. 2007. The heat stability of Plasmodium lactate dehydrogenase-

based and histidine-rich protein 2-based malaria rapid diagnostic tests. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 101, 331-337.

CHUMPITAZI, B. P., COPE, J. L., HOLLISTER, E. B., TSAI, C. M., MCMEANS, A. R., LUNA, R. A., VERSALOVIC, J. & SHULMAN, R. J. 2015. Randomised clinical trial: gut microbiome biomarkers are associated with clinical response to a low FODMAP diet in children with the irritable bowel syndrome. *Alimentary Pharmacology & Therapeutics,* 42, 418-427.

CIBERE, J. 2006. Do we need radiographs to diagnose osteoarthritis? *Best Practice & Research Clinical Rheumatology,* 20, 27-38.

CLEVELAND, W. S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association,* 74, 829-836.

COLLINS, W. E. & JEFFERY, G. M. 2007. Plasmodium malariae: Parasite and Disease. *Clinical Microbiology Reviews,* 20, 579-592.

CONE, E. J., CAPLAN, Y. H., MOSER, F., ROBERT, T., SHELBY, M. K. & BLACK, D. L. 2009. Normalization of Urinary Drug Concentrations with Specific Gravity and Creatinine. *Journal of Analytical Toxicology,* 33, 1-7.

COOPER, H. J. & MARSHALL, A. G. 2001. Electrospray Ionization Fourier Transform Mass Spectrometric Analysis of Wine. *Journal of Agricultural and Food Chemistry,* 49, 5710-5718.

COUTO, A. S., KIMURA, E. A., PERES, V. J., UHRIG, M. L. & KATZIN, A. M. 1999. Active isoprenoid pathway in the intra-erythrocytic stages of Plasmodium falciparum: presence of dolichols of 11 and 12 isoprene units. *Biochemical Journal,* 341, 629-637.

COX-SINGH, J., DAVIS, T. M. E., LEE, K.-S., SHAMSUL, S. S. G., MATUSOP, A., RATNAM, S., RAHMAN, H. A., CONWAY, D. J. & SINGH, B. 2008. Plasmodium knowlesi Malaria in Humans Is Widely Distributed and Potentially Life Threatening. *Clinical Infectious Diseases,* 46, 165-171.

COXAM, V. 2007. Current Data with Inulin-Type Fructans and Calcium, Targeting Bone Health in Adults. *The Journal of Nutrition,* 137, 2527S-2533S.

CUBBON, S., BRADBURY, T., WILSON, J. & THOMAS-OATES, J. 2007. Hydrophilic Interaction Chromatography for Mass Spectrometric Metabonomic Studies of Urine. *Analytical Chemistry,* 79, 8911-8918.

CUMMINGS, J. H. 1981. Short chain fatty acids in the human colon. *Gut,* 22, 763-779.

DA SILVA, L., GODEJOHANN, M., MARTIN, F.-P. J., COLLINO, S., BURKLE, A., MORENO-VILLANUEVA, M., BERNHARDT, J., TOUSSAINT, O., GRUBECK-LOEBENSTEIN, B., GONOS, E. S., SIKORA, E., GRUNE, T., BREUSING, N., FRANCESCHI, C.,

HERVONEN, A., SPRAUL, M. & MOCO, S. 2013. High-Resolution Quantitative Metabolome Analysis of Urine by Automated Flow Injection NMR. *Analytical Chemistry,* 85, 5801-5809.

DAI, W., YIN, P., ZENG, Z., KONG, H., TONG, H., XU, Z., LU, X., LEHMANN, R. & XU, G. 2014. Nontargeted Modification-Specific Metabolomics Study Based on Liquid Chromatography–High-Resolution Mass Spectrometry. *Analytical Chemistry,* 86, 9146-9153.

DAMYANOVICH, A. Z., STAPLES, J. R., CHAN, A. D. M. & MARSHALL, K. W. 1999. Comparative study of normal and osteoarthritic canine synovial fluid using 500 MHz 1H magnetic resonance spectroscopy. *Journal of Orthopaedic Research,* 17, 223-231.

DANKS, D. M., TIPPETT, P., ADAMS, C. & CAMPBELL, P. 1975. Cerebro-hepato-renal syndrome of Zellweger: A report of eight cases with comments upon the incidence, the liver lesion, and a fault in pipecolic acid metabolism. *The Journal of Pediatrics,* 86, 382-387.

DAVEY, M. P., BURRELL, M. M., WOODWARD, F. I. & QUICK, W. P. 2008. Population-specific metabolic phenotypes of Arabidopsis lyrata ssp. petraea. *New Phytologist,* 177, 380-388.

DAY, N. P. J., HIEN, T. T., SCHOLLAARDT, T., LOC, P. P., CHUONG, L. V., HONG CHAU, T. T., HOANG MAI, N. T., PHU, N. H., SINH, D. X., WHITE, N. J. & HO, M. 1999. The Prognostic and Pathophysiologic Role of Pro- and Antiinflammatory Cytokines in Severe Malaria. *Journal of Infectious Diseases,* 180, 1288-1297.

DE CEUNINCK, F. & BERENBAUM, F. 2009. Proteomics: addressing the challenges of osteoarthritis. *Drug Discovery Today,* 14, 661-667.

DE CEUNINCK, F., SABATINI, M. & PASTOUREAU, P. 2011. Recent progress toward biomarker identification in osteoarthritis. *Drug Discovery Today,* 16, 443-449.

DE GRAAF, A. A., STRIEGEL, K., WITTIG, R. M., LAUFER, B., SCHMITZ, G., WIECHERT, W., SPRENGER, G. A. & SAHM, H. 1999. Metabolic state of Zymomonas mobilis in glucose-, fructose-, and xylose-fed continuous cultures as analysed by 13C- and 31P-NMR spectroscopy. *Archives of Microbiology,* 171, 371-385.

DE LA PUERTA, C., ARRIETA, F. J., BALSA, J. A., BOTELLA-CARRETERO, J. I., ZAMARRON, I. & VAZQUEZ, C. 2010. Taurine and glucose metabolism: a review. *Nutricion Hospitalaria,* 25, 910-919.

DE PRETER, V., FALONY, G., WINDEY, K., HAMER, H. M., DE VUYST, L. & VERBEKE, K. 2010. The prebiotic, oligofructose-enriched inulin modulates the faecal metabolite profile: An in vitro analysis. *Molecular Nutrition & Food Research,* 54, 1791-1801.

DE PRETER, V., GHEBRETINSAE, A. H., ABRAHANTES, J. C., WINDEY, K., RUTGEERTS, P. & VERBEKE, K. 2011. Impact of the synbiotic combination of Lactobacillus casei shirota and oligofructose-enriched

inulin on the fecal volatile metabolite profile in healthy subjects. *Molecular Nutrition & Food Research,* 55, 714-722.

DE PRETER, V., JOOSSENS, M., BALLET, V., SHKEDY, Z., RUTGEERTS, P., VERMEIRE, S. & VERBEKE PHD, K. 2013. Metabolic profiling of the impact of oligofructose-enriched inulin in Crohn's disease patients: a double-blinded randomized controlled trial. *Clinical and translational gastroenterology,* 4, e30.

DE VOS, R. C. H., MOCO, S., LOMMEN, A., KEURENTJES, J. J. B., BINO, R. J. & HALL, R. D. 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols,* 2, 778-791.

DELIC, D., WARSKULAT, U., BORSCH, E., AL-QAHTANI, S., AL-QURAISHI, S., H¨USSINGER, D. & WUNDERLICH, F. 2010. Loss of Ability To Self-Heal Malaria upon Taurine Transporter Deletion. *Infection and Immunity,* 78, 1642-1649.

DELZENNE, N. M. & CANI, P. D. 2011. Interaction Between Obesity and the Gut Microbiota: Relevance in Nutrition. *Annual Review of Nutrition,* 31, 15-31.

DELZENNE, N. M., NEYRINCK, A. M., BACKHED, F. & CANI, P. D. 2011a. Targeting gut microbiota in obesity: effects of prebiotics and probiotics. *Nature Reviews Endocrinology,* 7, 639-646.

DELZENNE, N. M., NEYRINCK, A. M. & CANI, P. D. 2011b. Modulation of the gut microbiota by nutrients with prebiotic properties: consequences for host health in the context of obesity and metabolic syndrome. *Microbial cell factories,* 10 Suppl 1, S10.

DEMIREV, P. A., FELDMAN, A. B., KONGKASURIYACHAI, D., SCHOLL, P., SULLIVAN, D. & KUMAR, N. 2002. Detection of Malaria Parasites in Blood by Laser Desorption Mass Spectrometry. *Analytical Chemistry,* 74, 3262-3266.

DETHY, J.-M., ACKERMANN, B. L., DELATOUR, C., HENION, J. D. & SCHULTZ, G. A. 2003. Demonstration of Direct Bioanalysis of Drugs in Plasma Using Nanoelectrospray Infusion from a Silicon Chip Coupled with Tandem Mass Spectrometry. *Analytical Chemistry,* 75, 805-811.

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

DEWULF, E. M., CANI, P. D., CLAUS, S. P., FUENTES, S., PUYLAERT, P. G., NEYRINCK, A. M., BINDELS, L. B., DE VOS, W. M., GIBSON, G. R., THISSEN, J.-P. & DELZENNE, N. M. 2013. Insight into the prebiotic concept: lessons from an exploratory, double blind intervention study with inulin-type fructans in obese women. *Gut,* 62, 1112-1121.

DINARELLO, C. A. 1984. Interleukin-1 and the Pathogenesis of the Acute-Phase Response. *New England Journal of Medicine,* 311, 1413-1418.

DOLE, M., MACK, L. L., HINES, R. L., MOBLEY, R. C., FERGUSON, L. D. & ALICE, M. B. 1968. Molecular Beams of Macroions. *The Journal of Chemical Physics,* 49, 2240-2249.

DONDORP, A. M., KAGER, P. A., VREEKEN, J. & WHITE, N. J. 2000. Abnormal Blood Flow and Red Blood Cell Deformability in Severe Malaria. *Parasitology Today,* 16, 228-232.

DONDORP, A. M., PONGPONRATN, E. & WHITE, N. J. 2004. Reduced microcirculatory flow in severe falciparum malaria: pathophysiology and electron-microscopic pathology. *Acta Tropica,* 89, 309-317.

DRAPER, J., LLOYD, A., GOODACRE, R. & BECKMANN, M. 2013. Flow infusion electrospray ionisation mass spectrometry for high throughput, non-targeted metabolite fingerprinting: a review. *Metabolomics,* 9, 4-29.

DUBBELBOER, I. R., POHANKA, A., SAID, R., ROSENBORG, S. & BECK, O. 2012. Quantification of Tacrolimus and Three Demethylated Metabolites in Human Whole Blood Using LC-ESI-MS/MS. *Therapeutic Drug Monitoring,* 34, 134-142.

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

DUNN, W. B., BROADHURST, D., BEGLEY, P., ZELENA, E., FRANCIS-MCINTYRE, S., ANDERSON, N., BROWN, M., KNOWLES, J. D., HALSALL, A., HASELDEN, J. N., NICHOLLS, A. W., WILSON, I. D., KELL, D. B. & GOODACRE, R. 2011. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols,* 6, 1060-1083.

EDWARDS, J. L., CHISOLM, C. N., SHACKMAN, J. G. & KENNEDY, R. T. 2006. Negative mode sheathless capillary electrophoresis electrospray ionization-mass spectrometry for metabolite analysis of prokaryotes. *Journal of Chromatography A,* 1106, 80-88.

EGGINS, B. R. 2008. *Chemical Sensors and Biosensors*, Wiley.

EISEN, D. P. & SAUL, A. 2000. Disappearance of pan-malarial antigen reactivity using the ICT Malaria P.f/ P.vTM kit parallels decline of patent parasitaemia as shown by microscopy. *Transactions of The Royal Society of Tropical Medicine and Hygiene,* 94, 169-170.

ENOT, D. P., BECKMANN, M., OVERY, D. & DRAPER, J. 2006. Predicting interpretability of metabolome models based on behavior, putative identity, and biological relevance of explanatory signals. *Proceedings of the National Academy of Sciences,* 103, 14865-14870.

ERIKSSON, L., JOHANSSON, E., KETTANEH-WOLD, N., J.TRYGG, C. W. & WOLD, S. 2006a. *Multi- and Megavariate Data Analysis: Advanced Applications and Method Extensions,* Umea, Umetrics AB.

ERIKSSON, L., JOHANSSON, E., KETTANEH-WOLD, N., J.TRYGG, C. W. & WOLD, S. 2006b. *Multi- and Megavariate Data Analysis: Basic Principles and Applications,* Umea, Umetrics AB.

EULER, A. R., MITCHELL, D. K., KLINE, R. & PICKERING, L. K. 2005. Prebiotic Effect Of Fructo-Oligosaccharide Supplemented Term Infant Formula at Two Concentrations Compared with Unsupplemented Formula and Human Milk. *Journal of Pediatric Gastroenterology and Nutrition,* 40, 157-164.

FALONY, G., VERSCHAEREN, A., DE BRUYCKER, F., DE PRETER, V., VERBEKE, K., LEROY, F. & DE VUYST, L. 2009. In Vitro Kinetics of Prebiotic Inulin-Type Fructan Fermentation by Butyrate-Producing Colon Bacteria: Implementation of Online Gas Chromatography for Quantitative Analysis of Carbon Dioxide and Hydrogen Gas Production. *Applied and Environmental Microbiology,* 75, 5884-5892.

FAN, J., HONG, J., HU, J.-D. & CHEN, J.-L. 2012. Ion Chromatography Based Urine Amino Acid Profiling Applied for Diagnosis of Gastric Cancer. *Gastroenterology Research and Practice,* 2012, 8.

FARD, A. M., TURNER, A. G. & WILLETT, G. D. 2003. High-Resolution Electrospray-Ionization FourierTransform Ion Cyclotron Resonance and Gas Chromatography–Mass Spectrometry of Macadamia Nut Oil. *Australian Journal of Chemistry,* 56, 499-508.

FAROOQUI, A. A., HORROCKS, L. A. & FAROOQUI, T. 2000. Glycerophospholipids in brain: their metabolism, incorporation into membranes, functions, and involvement in neurological disorders. *Chemistry and Physics of Lipids,* 106, 1-29.

FAVE, G., BECKMANN, M., LLOYD, A., ZHOU, S., HAROLD, G., LIN, W., TAILLIART, K., XIE, L., DRAPER, J. & MATHERS, J. 2011. Development and validation of a standardized protocol to monitor human dietary exposure by metabolite fingerprinting of urine samples. *Metabolomics,* 7, 469-484.

FAVRETTO, D. & FLAMINI, R. 2000. Application of Electrospray Ionization Mass Spectrometry to the Study of Grape Anthocyanins. *American Journal of Enology and Viticulture,* 51, 55-64.

FAVRETTO, D., PIOVAN, A., FILIPPINI, R. & CANIATO, R. 2001. Monitoring the production yields of vincristine and vinblastine in Catharanthus roseus from somatic embryogenesis. Semiquantitative determination by flow-injection electrospray ionization mass spectrometry. *Rapid Communications in Mass Spectrometry,* 15, 364-369.

FDA 2013. FDA Guidance for Industry, Bioanalytical method Validation. Food and Drug Administration, Centre for Drug Evaluation and Research (CDER).

FELSON, D. T. 2010. Arthroscopy as a treatment for knee osteoarthritis. *Clinical rheumatology,* 24, 47-50.

FELSON, D. T., CHAISSON, C. E., HILL, C. L., TOTTERMAN, S. M. S., GALE, M. E., SKINNER, K. M., KAZIS, L. & GALE, D. R. 2001. The Association of Bone Marrow Lesions with Pain in Knee Osteoarthritis. *Annals of Internal Medicine,* 134**,** 541-549.

FENN, J. B., MANN, M., CHIN-KAI, M., SHEK-FU, W. & WHITEHOUSE, C. M. 1989. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science (Washington),* 246**,** 64.

FERNANDEZ-PERALBO, M. A. & LUQUE DE CASTRO, M. D. 2012. Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. *TrAC Trends in Analytical Chemistry,* 41**,** 75-85.

FITZGERALD, R. L., GRIFFIN, T. L., YUN, Y.-M., GODFREY, R. A., WEST, R., PESCE, A. J. & HEROLD, D. A. 2012. Dilute and Shoot: Analysis of Drugs of Abuse Using Selected Reaction Monitoring for Quantification and Full Scan Product Ion Spectra for Identification. *Journal of Analytical Toxicology,* 36**,** 106-111.

FLANGEA, C., SERB, A., SISU, E. & ZAMFIR, A. D. 2011. Chip-based nanoelectrospray mass spectrometry of brain gangliosides. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids,* 1811**,** 513-535.

FROESCH, M., BINDILA, L. M., BAYKUT, G., ALLEN, M., PETER-KATALINIC, J. & ZAMFIR, A. D. 2004. Coupling of fully automated chip electrospray to Fourier transform ion cyclotron resonance mass spectrometry for high-performance glycoscreening and sequencing. *Rapid Communications in Mass Spectrometry,* 18**,** 3084-3092.

FUHRER, T., HEER, D., BEGEMANN, B. & ZAMBONI, N. 2011. High-Throughput, Accurate Mass Metabolome Profiling of Cellular Extracts by Flow Injection–Time-of-Flight Mass Spectrometry. *Analytical Chemistry,* 83**,** 7074-7080.

FUJITA, T., AMURO, Y., HADA, T. & HIGASHINO, K. 1999a. Plasma levels of pipecolic acid, both l- and d-enantiomers, in patients with chronic liver diseases, especially hepatic encephalopathy. *Clinica Chimica Acta,* 287**,** 99-109.

FUJITA, T., HADA, T. & HIGASHINO, K. 1999b. Origin of d- and l-pipecolic acid in human physiological fluids: a study of the catabolic mechanism to pipecolic acid using the lysine loading test. *Clinica Chimica Acta,* 287**,** 145-156.

FURUYA, T., NISHI, T., SHIBATA, D., SUZUKI, H., OHTA, D. & KINO, K. 2008. Characterization of Orphan Monooxygenases by Rapid Substrate Screening Using FT-ICR Mass Spectrometry. *Chemistry & Biology,* 15**,** 563-572.

FUZFAI, Z., KATONA, Z. F., KOVACS, E. & MOLNAR-PERL, I. 2004. Simultaneous Identification and Quantification of the Sugar, Sugar Alcohol, and Carboxylic Acid Contents of Sour Cherry, Apple, and Ber Fruits, as Their Trimethylsilyl Derivatives, by Gas

Chromatography–Mass Spectrometry. *Journal of Agricultural and Food Chemistry,* 52, 7444-7452.

GABRIELSSON, J., LINDBERG, N.-O. & LUNDSTEDT, T. 2002. Multivariate methods in pharmaceutical applications. *Journal of Chemometrics,* 16, 141-160.

GAMACHE, P. H., MEYER, D. F., GRANGER, M. C. & ACWORTH, I. N. 2004. Metabolomic applications of electrochemistry/Mass spectrometry. *Journal of the American Society for Mass Spectrometry,* 15, 1717-1726.

GANTI, S., TAYLOR, S. L., KIM, K., HOPPEL, C. L., GUO, L., YANG, J., EVANS, C. & WEISS, R. H. 2012. Urinary acylcarnitines are altered in human kidney cancer. *International Journal of Cancer,* 130, 2791-2800.

GARDNER, M. J., HALL, N., FUNG, E., WHITE, O., BERRIMAN, M., HYMAN, R. W., CARLTON, J. M., PAIN, A., NELSON, K. E., BOWMAN, S., PAULSEN, I. T., JAMES, K., EISEN, J. A., RUTHERFORD, K., SALZBERG, S. L., CRAIG, A., KYES, S., CHAN, M.-S., NENE, V., SHALLOM, S. J., SUH, B., PETERSON, J., ANGIUOLI, S., PERTEA, M., ALLEN, J., SELENGUT, J., HAFT, D., MATHER, M. W., VAIDYA, A. B., MARTIN, D. M. A., FAIRLAMB, A. H., FRAUNHOLZ, M. J., ROOS, D. S., RALPH, S. A., MCFADDEN, G. I., CUMMINGS, L. M., SUBRAMANIAN, G. M., MUNGALL, C., VENTER, J. C., CARUCCI, D. J., HOFFMAN, S. L., NEWBOLD, C., DAVIS, R. W., FRASER, C. M. & BARRELL, B. 2002. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature,* 419, 498-511.

GARNER, C. E., SMITH, S., DE LACY COSTELLO, B., WHITE, P., SPENCER, R., PROBERT, C. S. J. & RATCLIFFE, N. M. 2007. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *The FASEB Journal,* 21, 1675-1688.

GAUT, J. P., BYUN, J., TRAN, H. D. & HEINECKE, J. W. 2002. Artifact-Free Quantification of Free 3-Chlorotyrosine, 3-Bromotyrosine, and 3-Nitrotyrosine in Human Plasma by Electron Capture–Negative Chemical Ionization Gas Chromatography Mass Spectrometry and Liquid Chromatography–Electrospray Ionization Tandem Mass Spectrometry. *Analytical Biochemistry,* 300, 252-259.

GAVAGHAN, C. L., NICHOLSON, J. K., CONNOR, S. C., WILSON, I. D., WRIGHT, B. & HOLMES, E. 2001. Directly Coupled High-Performance Liquid Chromatography and Nuclear Magnetic Resonance Spectroscopic with Chemometric Studies on Metabolic Variation in Sprague–Dawley Rats. *Analytical Biochemistry,* 291, 245-252.

GAVRIILIDIS, C., MIWA, S., VON ZGLINICKI, T., TAYLOR, R. W. & YOUNG, D. A. 2013. Mitochondrial dysfunction in osteoarthritis is associated with down-regulation of superoxide dismutase 2. *Arthritis & Rheumatism,* 65, 378-387.

GERMAN, J. B., ROBERTS, M.-A. & WATKINS, S. M. 2003. Personal Metabolomics as a Next Generation Nutritional Assessment. *The Journal of Nutrition,* 133, 4260-4266.

GETHING, P., PATIL, A., SMITH, D., GUERRA, C., ELYAZAR, I., JOHNSTON, G., TATEM, A. & HAY, S. 2011. A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malaria Journal,* 10, 378.

GHOSH, S., SENGUPTA, A., SHARMA, S. & SONAWAT, H. M. 2012. Metabolic Fingerprints of Serum, Brain, and Liver Are Distinct for Mice with Cerebral and Noncerebral Malaria: A 1H NMR Spectroscopy-Based Metabonomic Study. *Journal of Proteome Research,* 11, 4992-5004.

GIARDI, M. T., REA, G. & BERRA, B. 2010. *Bio-farms for nutraceuticals : functional food and safety control by biosensors,* Austin, Tex.; New York, Springer.

GIAVALISCO, P., HUMMEL, J., LISEC, J., INOSTROZA, A. C., CATCHPOLE, G. & WILLMITZER, L. 2008. High-Resolution Direct Infusion-Based Mass Spectrometry in Combination with Whole 13C Metabolome Isotope Labeling Allows Unambiguous Assignment of Chemical Sum Formulas. *Analytical Chemistry,* 80, 9417-9425.

GIBSON, G. R., BEATTY, E. R., WANG, X. & CUMMINGS, J. H. 1995. Selective stimulation of bifidobacteria in the human colon by oligofructose and inulin. *Gastroenterology,* 108, 975-982.

GIBSON, G. R. & ROBERFROID, M. B. 1995. Dietary Modulation of the Human Colonic Microbiota: Introducing the Concept of Prebiotics. *The Journal of Nutrition,* 125, 1401-1412.

GIBSON, P. R. & SHEPHERD, S. J. 2005. Personal view: food for thought – western lifestyle and susceptibility to Crohn's disease. The FODMAP hypothesis. *Alimentary Pharmacology & Therapeutics,* 21, 1399-1409.

GIBSON, P. R. & SHEPHERD, S. J. 2010. Evidence-based dietary management of functional gastrointestinal symptoms: The FODMAP approach. *Journal of Gastroenterology and Hepatology,* 25, 252-258.

GIEMSA, G. 1904. Eine Vereinfachung und Vervollkommnung meiner Methylenblau-Eosin-Farbemethode zur Erzielung der Romanowsky-Nocht'schen Chromatinfarbung. *Centralblatt fur Bakteriologie I Abteilung* 32, 307–313.

GIKA, H. G., MACPHERSON, E., THEODORIDIS, G. A. & WILSON, I. D. 2008a. Evaluation of the repeatability of ultra-performance liquid chromatography–TOF-MS for global metabolic profiling of human urine samples. *Journal of Chromatography B,* 871, 299-305.

GIKA, H. G., THEODORIDIS, G. A., EARLL, M., SNYDER, R. W., SUMNER, S. J. & WILSON, I. D. 2010. Does the Mass Spectrometer Define the Marker? A Comparison of Global Metabolite Profiling Data Generated Simultaneously via UPLC-MS on Two Different Mass Spectrometers. *Analytical Chemistry,* 82, 8226-8234.

GIKA, H. G., THEODORIDIS, G. A., EARLL, M. & WILSON, I. D. 2012. A QC approach to the determination of day-to-day reproducibility and

robustness of LC–MS methods for global metabolite profiling in metabonomics/metabolomics. *Bioanalysis,* 4, 2239-2247.

GIKA, H. G., THEODORIDIS, G. A., PLUMB, R. S. & WILSON, I. D. 2014. Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. *Journal of Pharmaceutical and Biomedical Analysis,* 87, 12-25.

GIKA, H. G., THEODORIDIS, G. A. & WILSON, I. D. 2008b. Hydrophilic interaction and reversed-phase ultra-performance liquid chromatography TOF-MS for metabonomic analysis of Zucker rat urine. *Journal of Separation Science,* 31, 1598-1608.

GIKA, H. G., THEODORIDIS, G. A. & WILSON, I. D. 2008c. Liquid chromatography and ultra-performance liquid chromatography–mass spectrometry fingerprinting of human urine: Sample stability under different handling and storage conditions for metabonomics studies. *Journal of Chromatography A,* 1189, 314-322.

GIKA, H. G., THEODORIDIS, G. A., WINGATE, J. E. & WILSON, I. D. 2007. Within-Day Reproducibility of an HPLC−MS-Based Method for Metabonomic Analysis: Application to Human Urine. *Journal of Proteome Research,* 6, 3291-3303.

GLAUSER, G., VEYRAT, N., ROCHAT, B., WOLFENDER, J.-L. & TURLINGS, T. C. J. 2013. Ultra-high pressure liquid chromatography–mass spectrometry for plant metabolomics: A systematic comparison of high-resolution quadrupole-time-of-flight and single stage Orbitrap mass spectrometers. *Journal of Chromatography A,* 1292, 151-159.

GOLDRING, M. B. & BERENBAUM, F. 2015. Emerging targets in osteoarthritis therapy. *Current Opinion in Pharmacology,* 22, 51-63.

GONCALVES, L., SUBTIL, A., DE OLIVEIRA, M. R., DO ROSŒRIO, V., LEE, P.-W. & SHAIO, M.-F. 2012. Bayesian Latent Class Models in Malaria Diagnosis. *PLoS ONE,* 7, e40633.

GONZALEZ-DOMINGUEZ, R., CASTILLA-QUINTERO, R., GARCIA-BARRERA, T. & GOMEZ-ARIZA, J. L. 2014. Development of a metabolomic approach based on urine samples and direct infusion mass spectrometry. *Analytical Biochemistry,* 465, 20-27.

GOODACRE, R., BROADHURST, D., SMILDE, A., KRISTAL, B., BAKER, J. D., BEGER, R., BESSANT, C., CONNOR, S., CAPUANI, G., CRAIG, A., EBBELS, T., KELL, D., MANETTI, C., NEWTON, J., PATERNOSTRO, G., SOMORJAI, R., SJÖSTRÖM, M., TRYGG, J. & WULFERT, F. 2007. Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics,* 3, 231-241.

GOODACRE, R., HEALD, J. K. & KELL, D. B. 1999. Characterisation of intact microorganisms using electrospray ionisation mass spectrometry. *FEMS Microbiology Letters,* 176, 17-24.

GOODACRE, R., VAIDYANATHAN, S., BIANCHI, G. & KELL, D. B. 2002. Metabolic profiling using direct infusion electrospray ionisation mass spectrometry for the characterisation of olive oils. *Analyst,* 127**,** 1457-1462.

GOODACRE, R., YORK, E. V., HEALD, J. K. & SCOTT, I. M. 2003. Chemometric discrimination of unfractionated plant extracts analyzed by electrospray mass spectrometry. *Phytochemistry,* 62**,** 859-863.

GOUGEON, R. D., LUCIO, M., FROMMBERGER, M., PEYRON, D., CHASSAGNE, D., ALEXANDRE, H., FEUILLAT, F., VOILLEY, A., CAYOT, P., GEBEFÜGI, I., HERTKORN, N. & SCHMITT-KOPPLIN, P. 2009. The chemodiversity of wines can reveal a metabologeography expression of cooperage oak wood. *Proceedings of the National Academy of Sciences,* 106**,** 9174-9179.

GRATA, E., BOCCARD, J., GLAUSER, G., CARRUPT, P.-A., FARMER, E. E., WOLFENDER, J.-L. & RUDAZ, S. 2007. Development of a two-step screening ESI-TOF-MS method for rapid determination of significant stress-induced metabolome modifications in plant leaf extracts: The wound response in Arabidopsis thaliana as a case study. *Journal of Separation Science,* 30**,** 2268-2278.

GRAY, G. R. & HEATH, D. 2005. A global reorganization of the metabolome in Arabidopsis during cold acclimation is revealed by metabolic fingerprinting. *Physiologia Plantarum,* 124**,** 236-248.

GRIFFITH, K. S., LEWIS, L. S., MALI, S. & PARISE, M. E. 2007. Treatment of malaria in the united states: A systematic review. *JAMA,* 297**,** 2264-2277.

GROSS, J. H. 2011. *Mass spectrometry,* Berlin, Springer.

GUARNER, F. & MALAGELADA, J.-R. 2003. Gut flora in health and disease. *The Lancet,* 361**,** 512-519.

GYAMLANI, G. G., BERGSTRALH, E. J., SLEZAK, J. M. & LARSON, T. S. 2003. Urinary albumin to osmolality ratio predicts 24-hour urine albumin excretion in diabetes mellitus. *American Journal of Kidney Diseases,* 42**,** 685-692.

HALDAR, K., MURPHY, S. C., MILNER, D. A. & TAYLOR, T. E. 2007. Malaria: Mechanisms of Erythrocytic Infection and Pathological Correlates of Severe Disease. *Annual Review of Pathology: Mechanisms of Disease,* 2**,** 217-249.

HALMOS, E., POWER, V., SHEPHERD, S., GIBSON, P. & MUIR, J. 2014. A Diet Low in FODMAPs Reduces Symptoms of Irritable Bowel Syndrome. *Gastroenterology,* 146**,** 67-75.e5.

HAMER, H. M., DE PRETER, V., WINDEY, K. & VERBEKE, K. 2011. Functional analysis of colonic bacterial metabolism: relevant to health? *American Journal of Physiology - Gastrointestinal and Liver Physiology,* 302**,** G1-G9.

HAN, J., DANELL, R. M., PATEL, J. R., GUMEROV, D. R., SCARLETT, C. O., SPEIR, J. P., PARKER, C. E., RUSYN, I., ZEISEL, S. & BORCHERS, C. H. 2008a. Towards high-throughput metabolomics using ultrahigh-field Fourier transform ion cyclotron resonance mass spectrometry. *Metabolomics : Official journal of the Metabolomic Society,* 4, 128-140.

HAN, J., DATLA, R., CHAN, S. & BORCHERS, C. H. 2009. Mass spectrometry-based technologies for high-throughput metabolomics. *Bioanalysis,* 1, 1665-1684.

HAN, X., YANG, K. & GROSS, R. W. 2008b. Microfluidics-based electrospray ionization enhances the intrasource separation of lipid classes and extends identification of individual molecular species through multi-dimensional mass spectrometry: development of an automated high-throughput platform for shotgun lipidomics. *Rapid Communications in Mass Spectrometry,* 22, 2115-2124.

HANDLEY, C. J., SPEIGHT, G., LEYDEN, K. M. & LOWTHER, D. A. 1980. Extracellular matrix metabolism by chondrocytes 7. Evidence that L-glutamine is an essential amino acid for chondrocytes and other connective tissue cells. *Biochimica et Biophysica Acta (BBA) - General Subjects,* 627, 324-331.

HANSCHEID, T. 2003. Current strategies to avoid misdiagnosis of malaria. *Clinical Microbiology and Infection,* 9, 497-504.

HARPER, J. D., CHARIPAR, N. A., MULLIGAN, C. C., ZHANG, X., COOKS, R. G. & OUYANG, Z. 2008. Low-Temperature Plasma Probe for Ambient Desorption Ionization. *Analytical Chemistry,* 80, 9097-9104.

HARVEY, R. A., HARVEY, R. A. & FERRIER, D. R. 2011. *Biochemistry*, Wolters Kluwer Health.

HASEGAWA, M., IDE, M., KUWAMURA, M., YAMATE, J. & TAKENAKA, S. 2010. Metabolic Fingerprinting in Toxicological Assessment Using FT-ICR MS. *Journal of Toxicologic Pathology,* 23, 67-74.

HASEGAWA, M., TAKENAKA, S., KUWAMURA, M., YAMATE, J. & TSUYAMA, S. 2007. Urinary metabolic fingerprinting for amiodarone-induced phospholipidosis in rats using FT-ICR MS. *Experimental and Toxicologic Pathology,* 59, 115-120.

HEAVNER, D. L., MORGAN, W. T., SEARS, S. B., RICHARDSON, J. D., BYRD, G. D. & OGDEN, M. W. 2006. Effect of creatinine and specific gravity normalization techniques on xenobiotic biomarkers in smokers' spot and 24-h urines. *Journal of Pharmaceutical and Biomedical Analysis,* 40, 928-942.

HELM, R. F., JERVIS, J., RAY, W. K., WILLOUGHBY, N., IRVIN, B., HASTIE, J., SCHELL, D. J. & NAGLE, N. 2010. Mass Spectral Analyses of Corn Stover Prehydrolysates To Assess Conditioning Processes. *Journal of Agricultural and Food Chemistry,* 58, 12642-12649.

HIGGS, R. E., ZAHN, J. A., GYGI, J. D. & HILTON, M. D. 2001. Rapid Method To Estimate the Presence of Secondary Metabolites in Microbial Extracts. *Applied and Environmental Microbiology,* 67, 371-376.

HILLER, K., HANGEBRAUK, J., JAGER, C., SPURA, J., SCHREIBER, K. & SCHOMBURG, D. 2009. MetaboliteDetector: Comprehensive Analysis Tool for Targeted and Nontargeted GC/MS Based Metabolome Analysis. *Analytical Chemistry,* 81, 3429-3439.

HILLS, B. A. 2002. Surface-active phospholipid: a Pandora' s box of clinical applications. Part II. Barrier and lubricating properties. *Internal Medicine Journal,* 32, 242-251.

HINMAN, R. S. & CROSSLEY, K. M. 2007. Patellofemoral joint osteoarthritis: an important subgroup of knee osteoarthritis. *Rheumatology,* 46, 1057-1062.

HO, C. S., LAM, C. W. K., CHAN, M. H. M., CHEUNG, R. C. K., LAW, L. K., LIT, L. C. W., NG, K. F., SUEN, M. W. M. & TAI, H. L. 2003. Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *The Clinical Biochemist Reviews,* 24, 3-12.

HOAD, C. L., MARCIANI, L., FOLEY, S., TOTMAN, J. J., WRIGHT, J., BUSH, D., COX, E. F., CAMPBELL, E., SPILLER, R. C. & GOWLAND, P. A. 2007. Non-invasive quantification of small bowel water content by MRI: a validation study. *Physics in Medicine and Biology,* 52, 6909.

HOJER-PEDERSEN, J., SMEDSGAARD, J. & NIELSEN, J. 2008. The yeast metabolome addressed by electrospray ionization mass spectrometry: Initiation of a mass spectral library and its applications for metabolic footprinting by direct infusion mass spectrometry. *Metabolomics,* 4, 393-405.

HOLMES, E., LOO, R. L., STAMLER, J., BICTASH, M., YAP, I. K. S., CHAN, Q., EBBELS, T., DE IORIO, M., BROWN, I. J., VESELKOV, K. A., DAVIGLUS, M. L., KESTELOOT, H., UESHIMA, H., ZHAO, L., NICHOLSON, J. K. & ELLIOTT, P. 2008. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature,* 453, 396-400.

HORAI, H., ARITA, M., KANAYA, S., NIHEI, Y., IKEDA, T., SUWA, K., OJIMA, Y., TANAKA, K., TANAKA, S., AOSHIMA, K., ODA, Y., KAKAZU, Y., KUSANO, M., TOHGE, T., MATSUDA, F., SAWADA, Y., HIRAI, M. Y., NAKANISHI, H., IKEDA, K., AKIMOTO, N., MAOKA, T., TAKAHASHI, H., ARA, T., SAKURAI, N., SUZUKI, H., SHIBATA, D., NEUMANN, S., IIDA, T., TANAKA, K., FUNATSU, K., MATSUURA, F., SOGA, T., TAGUCHI, R., SAITO, K. & NISHIOKA, T. 2010. MassBank: a public repository for sharing mass spectral data for life sciences. *Journal of Mass Spectrometry,* 45, 703-714.

HRYDZIUSZKO, O., SILVA, M. A., P. R. PERERA, M. T., RICHARDS, D. A., MURPHY, N., MIRZA, D. & VIANT, M. R. 2010. Application of Metabolomics to Investigate the Process of Human Orthotopic Liver

Transplantation: A Proof-of-Principle Study. *OMICS: A Journal of Integrative Biology,* 14, 143-150.

HUANG, Z., CHEN, Y., HANG, W., GAO, Y., LIN, L., LI, D., XING, J. & YAN, X. 2013. Holistic metabonomic profiling of urine affords potential early diagnosis for bladder and kidney cancers. *Metabolomics,* 9, 119-129.

IAVARONE, A., JURCHEN, J. & WILLIAMS, E. 2000. Effects of solvent on the maximum charge state and charge state distribution of protein ions produced by electrospray ionization. *Journal of the American Society for Mass Spectrometry,* 11, 976-985.

IDBORG-BJORKMAN, H., EDLUND, P.-O., KVALHEIM, O. M., SCHUPPE-KOISTINEN, I. & JACOBSSON, S. P. 2003. Screening of Biomarkers in Rat Urine Using LC/Electrospray Ionization-MS and Two-Way Data Analysis. *Analytical Chemistry,* 75, 4784-4792.

ISSAQ, H. J., NATIV, O., WAYBRIGHT, T., LUKE, B., VEENSTRA, T. D., ISSAQ, E. J., KRAVSTOV, A. & MULLERAD, M. 2008. Detection of Bladder Cancer in Human Urine by Metabolomic Profiling Using High Performance Liquid Chromatography/Mass Spectrometry. *The Journal of Urology,* 179, 2422-2426.

IUB-IUPAC 1982. Abbreviated terminology of oligosaccharide chains. Recommendations 1980. IUB-IUPAC Joint Commission on Biochemical Nomenclature (JCBN). *Journal of Biological Chemistry,* 257, 3347-51.

JAIN, P., CHAKMA, B., PATRA, S. & GOSWAMI, P. 2014. Potential Biomarkers and Their Applications for Rapid and Reliable Detection of Malaria. *BioMed Research International,* 2014, 20.

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.

JEANNERET, F., BOCCARD, J., BADOUD, F., SORG, O., TONOLI, D., PELCLOVA, D., VLCKOVA, S., RUTLEDGE, D. N., SAMER, C. F., HOCHSTRASSER, D., SAURAT, J.-H. & RUDAZ, S. 2014. Human urinary biomarkers of dioxin exposure: Analysis by metabolomics and biologically driven data dimensionality reduction. *Toxicology Letters,* 230, 234-243.

JIMENEZ GIRON, A., DEVENTER, K., ROELS, K. & VAN EENOO, P. 2012. Development and validation of an open screening method for diuretics, stimulants and selected compounds in human urine by UHPLC–HRMS for doping control. *Analytica Chimica Acta,* 721, 137-146.

JIN, X., YUN, S. J., JEONG, P., KIM, I. Y., KIM, W.-J. & PARK, S. 2014. Diagnosis of bladder cancer and prediction of survival by urinary metabolomics. *Oncotarget,* 5, 1635-1645.

JOHNSON, C. H., SLANAR, O., KRAUSZ, K. W., KANG, D. W., PATTERSON, A. D., KIM, J.-H., LUECKE, H., GONZALEZ, F. J. & IDLE, J. R. 2012. Novel metabolites and roles for α-tocopherol in

humans and mice discovered by mass spectrometry – based metabolomics. *The American Journal of Clinical Nutrition,* 96, 818-830.

JOHNSON, C. L. & VERSALOVIC, J. 2012. The Human Microbiome and Its Potential Importance to Pediatrics. *Pediatrics,* 129, 950-960.

JOHNSON, H., LLOYD, A., MUR, L. J., SMITH, A. & CAUSTON, D. 2007. The application of MANOVA to analyse Arabidopsis thaliana metabolomic data from factorially designed experiments. *Metabolomics,* 3, 517-530.

JONES, M. C. 2008. The family and their close and distant relations. *Journal of the Korean Statistical Society,* 37, 293-302.

JORDAN, K., ARDEN, N., DOHERTY, M., BANNWARTH, B., BIJLSMA, J., DIEPPE, P., GUNTHER, K., HAUSELMANN, H., HERRERO-BEAUMONT, G., KAKLAMANIS, P., LOHMANDER, S., LEEB, B., LEQUESNE, M., MAZIERES, B., MARTIN-MOLA, E., PAVELKA, K., PENDLETON, A., PUNZI, L., SERNI, U., SWOBODA, B., VERBRUGGEN, G., ZIMMERMAN-GORSKA, I. & DOUGADOS, M. 2003. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Annals of the Rheumatic Diseases,* 62, 1145-1155.

JORGENSEN, P., CHANTHAP, L., REBUENO, A., TSUYUOKA, R. & BELL, D. 2006. Malaria rapid diagnostic tests in tropical climates: the need for a cool chain. *The American Journal of Tropical Medicine and Hygiene,* 74, 750-754.

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.

KAI, K., HASHIDZUME, H., YOSHIMURA, K., SUZUKI, H., SAKURAI, N., SHIBATA, D. & OHTA, D. 2009. Metabolomics for the characterization of cytochromes P450-dependent fatty acid hydroxylation reactions in Arabidopsis. *Plant Biotechnology,* 26, 175-182.

KATAJAMAA, M., MIETTINEN, J. & ORESIC, M. 2006. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics,* 22, 634-636.

KATAJAMAA, M. & ORESIC, M. 2005. Processing methods for differential analysis of LC/MS profile data. *BMC Bioinformatics,* 6, 179.

KATZUNG, B., MASTERS, S. & TREVOR, A. 2012. *Basic and Clinical Pharmacology*, McGraw-Hill Education.

KAYSER, O. & WARZECHA, H. 2012. *Pharmaceutical Biotechnology: Drug Discovery and Clinical Applications*, John Wiley & Sons.

KELLER, M. D., POLLITT, C. C. & MARX, U. C. 2011. Nuclear magnetic resonance-based metabonomic study of early time point laminitis in an oligofructose-overload model. *Equine Veterinary Journal,* 43, 737-743.

KERTESZ, V. & VAN BERKEL, G. J. 2010. Fully automated liquid extraction-based surface sampling and ionization using a chip-based robotic nanoelectrospray platform. *Journal of Mass Spectrometry,* 45, 252-260.

KHYMENETS, O., ANDRES-LACUEVA, C., URPI-SARDA, M., VAZQUEZ-FRESNO, R., MART, M. M., REGLERO, G., TORRES, M. & LLORACH, R. 2015. Metabolic fingerprint after acute and under sustained consumption of a functional beverage based on grape skin extract in healthy human subjects. *Food & Function,* 6, 1288-1298.

KIM, H., AHN, E. & MOON, M. H. 2008. Profiling of human urinary phospholipids by nanoflow liquid chromatography/tandem mass spectrometry. *Analyst,* 133, 1656-1663.

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.

KING, R., BONFIGLIO, R., FERNANDEZ-METZLER, C., MILLER-STEIN, C. & OLAH, T. 2000. Mechanistic investigation of ionization suppression in electrospray ionization. *Journal of the American Society for Mass Spectrometry,* 11, 942-950.

KISS, A., BORDES, C., BUISSON, C., LASNE, F., LANTERI, P. & CREN-OLIVE, C. 2014. Data-handling strategies for metabonomic studies: example of the UHPLC-ESI/ToF urinary signature of tetrahydrocannabinol in humans. *Analytical and Bioanalytical Chemistry,* 406, 1209-1219.

KISS, A., LUCIO, M., FILDIER, A., BUISSON, C., SCHMITT-KOPPLIN, P. & CREN-OLIVE, C. 2013. Doping Control Using High and Ultra-High Resolution Mass Spectrometry Based Non-Targeted Metabolomics-A Case Study of Salbutamol and Budesonide Abuse. *PLoS ONE,* 8, e74584.

KOLIDA, S., TUOHY, K. & GIBSON, G. R. 2002. Prebiotic effects of inulin and oligofructose. *British Journal of Nutrition,* 87, S193-S197.

KOSINSKA, M. K., LIEBISCH, G., LOCHNIT, G., WILHELM, J., KLEIN, H., KAESSER, U., LASCZKOWSKI, G., RICKERT, M., SCHMITZ, G. & STEINMEYER, J. 2014. Sphingolipids in Human Synovial Fluid - A Lipidomic Study. *PLoS ONE,* 9, e91769.

KOULMAN, A., TAPPER, B. A., FRASER, K., CAO, M., LANE, G. A. & RASMUSSEN, S. 2007. High-throughput direct-infusion ion trap mass spectrometry: a new method for metabolomics. *Rapid Communications in Mass Spectrometry,* 21, 421-428.

KOULMAN, A., WOFFENDIN, G., NARAYANA, V. K., WELCHMAN, H., CRONE, C. & VOLMER, D. A. 2009. High-resolution extracted ion chromatography, a new tool for metabolomics and lipidomics using a

second-generation orbitrap mass spectrometer. *Rapid Communications in Mass Spectrometry,* 23, 1411-1418.

KOUSKOUMVEKAKI, I. & PANAGIOTOU, G. 2011. Navigating the Human Metabolome for Biomarker Identification and Design of Pharmaceutical Molecules. *Journal of Biomedicine and Biotechnology,* 2011.

KRISHNAPPA, R. 2011. Molecular expression profiling with respect to KEGG hsa05219 pathway. *Ecancermedicalscience,* 5, 189.

KUSANO, T., BERBERICH, T., TATEDA, C. & TAKAHASHI, Y. 2008. Polyamines: essential factors for growth and survival. *Planta,* 228, 367-381.

KUSHNER, I. 1988. The acute phase response: An overview. *In:* GIOVANNI DI, S. (ed.) *Methods in Enzymology.* Academic Press.

LACOMBE, A., LI, R. W., KLIMIS-ZACAS, D., KRISTO, A. S., TADEPALLI, S., KRAUSS, E., YOUNG, R. & WU, V. C. H. 2013. Lowbush wild blueberries have the potential to modify gut microbiota and xenobiotic metabolism in the rat colon. *PloS ONE,* 8, e67497.

LAIAKIS, E. C., MAK, T. D., ANIZAN, S., AMUNDSON, S. A., BARKER, C. A., WOLDEN, S. L., BRENNER, D. J. & FORNACE, A. J. 2014. Development of a Metabolomic Radiation Signature in Urine from Patients Undergoing Total Body Irradiation. *Radiation Research,* 181, 350-361.

LAKSHMANAN, V., RHEE, K., WANG, W., YU, Y., KHAFIZOV, K., FISER, A., WU, P., NDIR, O., MBOUP, S., NDIAYE, D. & DAILY, J. 2012. Metabolomic analysis of patient plasma yields evidence of plant-like α-linolenic acid metabolism in Plasmodium falciparum. *Journal of Infectious Diseases*.

LAKSHMANAN, V., RHEE, K. Y. & DAILY, J. P. 2011. Metabolomics and malaria biology. *Molecular and Biochemical Parasitology,* 175, 104-111.

LAMERS, R.-J. A. N., DEGROOT, J., SPIES-FABER, E. J., JELLEMA, R. H., KRAUS, V. B., VERZIJL, N., TEKOPPELE, J. M., SPIJKSMA, G. K., VOGELS, J. T. W. E., VAN DER GREEF, J. & VAN NESSELROOIJ, J. H. J. 2003a. Identification of Disease- and Nutrient-Related Metabolic Fingerprints in Osteoarthritic Guinea Pigs. *The Journal of Nutrition,* 133, 1776-1780.

LAMERS, R.-J. A. N., WESSELS, E. C. H. H., VAN DE SANDT, J. J. M., VENEMA, K., SCHAAFSMA, G., VAN DER GREEF, J. & VAN NESSELROOIJ, J. H. J. 2003b. A Pilot Study to Investigate Effects of Inulin on Caco-2 Cells through In Vitro Metabolic Fingerprinting. *The Journal of Nutrition,* 133, 3080-3084.

LAMERS, R. J. A. N., VAN NESSELROOIJ, J. H. J., KRAUS, V. B., JORDAN, J. M., RENNER, J. B., DRAGOMIR, A. D., LUTA, G., VAN DER GREEF, J. & DEGROOT, J. 2005. Identification of an urinary metabolite profile associated with osteoarthritis. *Osteoarthritis and Cartilage,* 13, 762-768.

LAMOND-LAB. 2014. *Cell Biologist's Guide to Proteomics* [Online]. Available: http://www.lamondlab.com/MSResource/LCMS/MassSpectrometry/electrosprayIonisation.php [Accessed 16/12 2014].

LAU, Y.-L., FONG, M.-Y., MAHMUD, R., CHANG, P.-Y., PALAEYA, V., CHEONG, F.-W., CHIN, L.-C., ANTHONY, C. N., AL-MEKHLAFI, A. M. & CHEN, Y. 2011. Specific, sensitive and rapid detection of human plasmodium knowlesi infection by loop-mediated isothermal amplification (LAMP) in blood samples. *Malaria Journal,* 10, 197-197.

LAW, W. S., HUANG, P. Y., ONG, E. S., ONG, C. N., LI, S. F. Y., PASIKANTI, K. K. & CHAN, E. C. Y. 2008. Metabonomics investigation of human urine after ingestion of green tea with gas chromatography/mass spectrometry, liquid chromatography/mass spectrometry and 1H NMR spectroscopy. *Rapid Communications in Mass Spectrometry,* 22, 2436-2446.

LE GAC, S. & VAN DEN BERG, A. 2009. *Miniaturization and Mass Spectrometry,* Cambridge, The Royal Society of Chemistry.

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.

LE GRAVERAND-GASTINEAU, M. P. 2010. Disease modifying osteoarthritis drugs: facing development challenges and choosing molecular targets. *Curr Drug Targets,* 11, 528-35.

LE, J., PERIER, C., PEYROCHE, S., RASCLE, F., BLANCHON, M. A., GONTHIER, R., FREY, J. & CHAMSON, A. 1999. Urine glycyl-L-proline increase and skin trophicity. *Amino Acids,* 17, 315-322.

LEE, S., SONG, K.-M., JEON, W., JO, H., SHIM, Y.-B. & BAN, C. 2012. A highly sensitive aptasensor towards Plasmodium lactate dehydrogenase for the diagnosis of malaria. *Biosensors and Bioelectronics,* 35, 291-296.

LENZ, E. M., BRIGHT, J., WILSON, I. D., HUGHES, A., MORRISSON, J., LINDBERG, H. & LOCKTON, A. 2004. Metabonomics, dietary influences and cultural differences: a 1H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *Journal of Pharmaceutical and Biomedical Analysis,* 36, 841-849.

LENZ, E. M., BRIGHT, J., WILSON, I. D., MORGAN, S. R. & NASH, A. F. P. 2003. A 1H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. *Journal of Pharmaceutical and Biomedical Analysis,* 33, 1103-1115.

LEWIS, S. & COCHRANE, S. 2007. Alteration of Sulfate and Hydrogen Metabolism in the Human Colon by Changing Intestinal Transit Rate. *The American Journal of Gastroenterology,* 102, 624-633.

LI, F., PATTERSON, A. D., HÖFER, C. C., KRAUSZ, K. W., GONZALEZ, F. J. & IDLE, J. R. 2011. A comprehensive understanding of thioTEPA

metabolism in the mouse using UPLC–ESI-QTOFMS-based metabolomics. *Biochemical Pharmacology,* 81, 1043-1053.

LI, J., WANG, Y., SARIC, J., NICHOLSON, J., DIRNHOFER, S., SINGER, B., TANNER, M., WITTLIN, S., HOLMES, E. & UTZINGER, J. 2008. Global metabolic responses of NMRI mice to an experimental Plasmodium berghei infection. *J Proteome Res,* 7, 3948 - 3956.

LI, P., ZHAO, Z., WANG, Y., XING, H., PARKER, D. M., YANG, Z., BAUM, E., LI, W., SATTABONGKOT, J., SIRICHAISINTHOP, J., LI, S., YAN, G., CUI, L. & FAN, Q. 2014. Nested PCR detection of malaria directly using blood filter paper samples from epidemiological surveys. *Malaria Journal,* 13, 175-175.

LI, X., LU, X., TIAN, J., GAO, P., KONG, H. & XU, G. 2009. Application of Fuzzy c-Means Clustering in Data Analysis of Metabolomics. *Analytical Chemistry,* 81, 4468-4475.

LI, X., YANG, S., QIU, Y., ZHAO, T., CHEN, T., SU, M., CHU, L., LV, A., LIU, P. & JIA, W. 2010. Urinary metabolomics as a potentially novel diagnostic and stratification tool for knee osteoarthritis. *Metabolomics,* 6, 109-118.

LI, Y.-J., WANG, Z.-Z., BI, Y.-A., DING, G., SHENG, L.-S., QIN, J.-P., XIAO, W., LI, J.-C., WANG, Y.-X. & WANG, X. 2012. The evaluation and implementation of direct analysis in real time quadrupole time-of-flight tandem mass spectrometry for characterization and quantification of geniposide in Re Du Ning Injections. *Rapid Communications in Mass Spectrometry,* 26, 1377-1384.

LIAN, L.-Y., AL-HELAL, M., ROSLAINI, A., FISHER, N., BRAY, P., WARD, S. & BIAGINI, G. 2009. Glycerol: An unexpected major metabolite of energy metabolism by the human malaria parasite. *Malaria Journal,* 8, 38.

LIANG, Q., YU, Q., WU, H., ZHU, Y.-Z. & ZHANG, A.-H. 2014. Metabolite fingerprint analysis of cervical cancer using LC-QTOF/MS and multivariate data analysis. *Analytical Methods,* 6, 3937-3942.

LIGER-BELAIR, G., CILINDRE, C., GOUGEON, R. D., LUCIO, M., GEBEFUGI, I., JEANDET, P. & SCHMITT-KOPPLIN, P. 2009. Unraveling different chemical fingerprints between a champagne wine and its aerosols. *Proceedings of the National Academy of Sciences of the United States of America,* 106, 16545-16549.

LIN, L., YU, Q., YAN, X., HANG, W., ZHENG, J., XING, J. & HUANG, B. 2010. Direct infusion mass spectrometry or liquid chromatography mass spectrometry for human metabonomics? A serum metabonomic study of kidney cancer. *Analyst,* 135, 2970-2978.

LINUSSON, A., GOTTFRIES, J., OLSSON, T., ÖRNSKOV, E., FOLESTAD, S., NORDÉN, B. & WOLD, S. 2001. Statistical Molecular Design, Parallel Synthesis, and Biological Evaluation of a Library of Thrombin Inhibitors. *Journal of Medicinal Chemistry,* 44, 3424-3439.

LIU, P., DUAN, J., WANG, P., QIAN, D., GUO, J., SHANG, E., SU, S., TANG, Y. & TANG, Z. 2013. Biomarkers of primary dysmenorrhea and herbal formula intervention: an exploratory metabonomics study of blood plasma and urine. *Molecular BioSystems,* 9**,** 77-87.

LLORACH, R., MEDINA, S., GARCÍA-VIGUERA, C., ZAFRILLA, P., ABELLŒN, J., JAUREGUI, O., TOMŒS-BARBERŒN, F. A., GIL-IZQUIERDO, A. & ANDRÉS-LACUEVA, C. 2014. Discovery of human urinary biomarkers of aronia-citrus juice intake by HPLC-q-TOF-based metabolomic approach. *Electrophoresis,* 35**,** 1599-1606.

LLORACH, R., URPI-SARDA, M., TULIPANI, S., GARCIA-ALOY, M., MONAGAS, M. & ANDRES-LACUEVA, C. 2013. Metabolomic fingerprint in patients at high risk of cardiovascular disease by cocoa intervention. *Molecular Nutrition & Food Research,* 57**,** 962-973.

LLOYD, A. J., BECKMANN, M., FAVE, G., MATHERS, J. C. & DRAPER, J. 2011a. Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. *British Journal of Nutrition,* 106**,** 812-824.

LLOYD, A. J., FAVE, G., BECKMANN, M., LIN, W., TAILLIART, K., XIE, L., MATHERS, J. C. & DRAPER, J. 2011b. Use of mass spectrometry fingerprinting to identify urinary metabolites after consumption of specific foods. *The American Journal of Clinical Nutrition,* 94**,** 981-991.

LLOYD, A. J., WILLIAM ALLWOOD, J., WINDER, C. L., DUNN, W. B., HEALD, J. K., CRISTESCU, S. M., SIVAKUMARAN, A., HARREN, F. J. M., MULEMA, J., DENBY, K., GOODACRE, R., SMITH, A. R. & MUR, L. A. J. 2011c. Metabolomic approaches reveal that cell wall modifications play a major role in ethylene-mediated resistance against Botrytis cinerea. *The Plant Journal,* 67**,** 852-868.

LOEUILLE, D. 2012a. When should MRI for knee or hip osteoarthritis should be performed. *La Revue du praticien,* 62**,** 625-629.

LOEUILLE, D. 2012b. [When should MRI for knee or hip osteoarthritis should be performed?]. *Rev Prat,* 62**,** 625-9.

LOKHOV, P. & ARCHAKOV, A. 2009. Mass spectrometry methods in metabolomics. *Biochemistry (Moscow) Supplemental Series B: Biomedical Chemistry,* 3**,** 1-9.

LOKHOV, P. G., KHARYBIN, O. N. & ARCHAKOV, A. I. 2012. Diagnosis of lung cancer based on direct-infusion electrospray mass spectrometry of blood plasma metabolites. *International Journal of Mass Spectrometry,* 309**,** 200-205.

LOTZ, M., MARTEL-PELLETIER, J., CHRISTIANSEN, C., BRANDI, M.-L., BRUYÈRE, O., CHAPURLAT, R., COLLETTE, J., COOPER, C., GIACOVELLI, G., KANIS, J. A., KARSDAL, M. A., KRAUS, V., LEMS, W. F., MEULENBELT, I., PELLETIER, J.-P., RAYNAULD, J.-P., REITER-NIESERT, S., RIZZOLI, R., SANDELL, L. J., VAN SPIL, W. E. & REGINSTER, J.-Y. 2013. Value of biomarkers in osteoarthritis:

current status and perspectives. *Annals of the Rheumatic Diseases,* 72, 1756-1763.

LU, C., DENG, J., LI, L., WANG, D. & LI, G. 2014. Application of metabolomics on diagnosis and treatment of patients with psoriasis in traditional Chinese medicine. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* 1844, 280-288.

LUA, I. A., LIN, S.-L., LIN, H. R. & LUA, A. C. 2012. Replacing Immunoassays for Mephedrone, Ketamines and Six Amphetamine-Type Stimulants with Flow Injection Analysis Tandem Mass Spectrometry. *Journal of Analytical Toxicology,* 36, 575-581.

LUAN, H., LIU, L.-F., MENG, N., TANG, Z., CHUA, K.-K., CHEN, L.-L., SONG, J.-X., MOK, V. C. T., XIE, L.-X., LI, M. & CAI, Z. 2015. LC–MS-Based Urinary Metabolite Signatures in Idiopathic Parkinson's Disease. *Journal of Proteome Research,* 14, 467-478.

LUCIO, M., FEKETE, A., WEIGERT, C., W¨GELE, B., ZHAO, X., CHEN, J., FRITSCHE, A., H¨RING, H.-U., SCHLEICHER, E. D., XU, G., SCHMITT-KOPPLIN, P. & LEHMANN, R. 2010. Insulin Sensitivity Is Reflected by Characteristic Metabolic Fingerprints - A Fourier Transform Mass Spectrometric Non-Targeted Metabolomics Approach. *PLoS ONE,* 5, e13317.

LUEDEMANN, A., STRASSBURG, K., ERBAN, A. & KOPKA, J. 2008. TagFinder for the quantitative analysis of gas chromatography—mass spectrometry (GC-MS)-based metabolite profiling experiments. *Bioinformatics,* 24, 732-737.

LUO, J., LAMB, J. H. & LIM, C. K. 1997. Analysis of urinary and faecal porphyrin excretion patterns in human porphyrias by fast atom bombardment mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis,* 15, 1289-1294.

LUTHRIA, D. L., LIN, L.-Z., ROBBINS, R. J., FINLEY, J. W., BANUELOS, G. S. & HARNLY, J. M. 2008. Discriminating between Cultivars and Treatments of Broccoli Using Mass Spectral Fingerprinting and Analysis of Variance−Principal Component Analysis. *Journal of Agricultural and Food Chemistry,* 56, 9819-9827.

MACFARLANE, G. T., ALLISON, C. & GIBSON, G. R. 1988. Effect of pH on protease activities in the large intestine. *Letters in Applied Microbiology,* 7, 161-164.

MACFARLANE, G. T. & MACFARLANE, S. 2012. Bacteria, Colonic Fermentation, and Gastrointestinal Health. *Journal of AOAC International,* 95, 50-60.

MACPHERSON, G. G., WARRELL, M. J., WHITE, N. J., LOOAREESUWAN, S. & WARRELL, D. A. 1985. Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. *The American Journal of Pathology,* 119, 385-401.

MACRAE, J., DIXON, M., DEARNLEY, M., CHUA, H., CHAMBERS, J., KENNY, S., BOTTOVA, I., TILLEY, L. & MCCONVILLE, M. 2013. Mitochondrial metabolism of sexual and asexual blood stages of the malaria parasite Plasmodium falciparum. *BMC Biology,* 11, 67.

MACZULAK, A. E., DAWSON, K. A. & BAKER, J. P. 1985. Nitrogen utilization in bacterial isolates from the equine cecum. *Applied and Environmental Microbiology,* 50, 1439-1443.

MADALINSKI, G., GODAT, E., ALVES, S., LESAGE, D., GENIN, E., LEVI, P., LABARRE, J., TABET, J.-C., EZAN, E. & JUNOT, C. 2008. Direct Introduction of Biological Samples into a LTQ-Orbitrap Hybrid Mass Spectrometer as a Tool for Fast Metabolome Analysis. *Analytical Chemistry,* 80, 3291-3303.

MAGGE, S. & LEMBO, A. 2012. Low-FODMAP Diet for Treatment of Irritable Bowel Syndrome. *Gastroenterology & Hepatology,* 8, 739-745.

MAHAPATRA, S., HESS, A., JOHNSON, J., EISENACH, K., DEGROOTE, M., GITTA, P., JOLOBA, M., KAPLAN, G., WALZL, G., BOOM, W. & BELISLE, J. 2014. A metabolic biosignature of early response to anti-tuberculosis treatment. *BMC Infectious Diseases,* 14, 53.

MALLERET, B., CLASER, C., ONG, A. S. M., SUWANARUSK, R., SRIPRAWAT, K., HOWLAND, S. W., RUSSELL, B., NOSTEN, F. & RÉNIA, L. 2011. A rapid and robust tri-color flow cytometry assay for monitoring malaria parasite development. *Scientific Reports,* 1, 118.

MALTHA, J., GILLET, P., BOTTIEAU, E., CNOPS, L., VAN ESBROECK, M. & JACOBS, J. 2010. Evaluation of a rapid diagnostic test (CareStartTM Malaria HRP-2/pLDH (Pf/pan) Combo Test) for the diagnosis of malaria in a reference setting. *Malaria Journal,* 9, 171.

MARTIN, S. K., RAJASEKARIAH, G.-H., AWINDA, G., WAITUMBI, J. & KIFUDE, C. 2009. Unified Parasite Lactate Dehydrogenase and Histidine-Rich Protein ELISA for Quantification of Plasmodium falciparum. *The American Journal of Tropical Medicine and Hygiene,* 80, 516-522.

MAS, S., VILLAS-BÔAS, S. G., EDBERG HANSEN, M., ÅKESSON, M. & NIELSEN, J. 2007. A comparison of direct infusion MS and GC-MS for metabolic footprinting of yeast mutants. *Biotechnology and Bioengineering,* 96, 1014-1022.

MATTARUCCHI, E., BARALDI, E. & GUILLOU, C. 2012. Metabolomics applied to urine samples in childhood asthma; differentiation between asthma phenotypes and identification of relevant metabolites. *Biomedical Chromatography,* 26, 89-94.

MATTARUCCHI, E. & GUILLOU, C. 2012. Critical aspects of urine profiling for the selection of potential biomarkers using UPLC-TOF-MS. *Biomedical Chromatography,* 26, 512-517.

MATTOLI, L., CANGI, F., GHIARA, C., BURICO, M., MAIDECCHI, A., BIANCHI, E., RAGAZZI, E., BELLOTTO, L., SERAGLIA, R. &

TRALDI, P. 2011. A metabolite fingerprinting for the characterization of commercial botanical dietary supplements. *Metabolomics,* 7, 437-445.

MATUSZEWSKI, B. K., CONSTANZER, M. L. & CHAVEZ-ENG, C. M. 1998. Matrix Effect in Quantitative LC/MS/MS Analyses of Biological Fluids: A Method for Determination of Finasteride in Human Plasma at Picogram Per Milliliter Concentrations. *Analytical Chemistry,* 70, 882-889.

MAUGHAN, R. J. & GRIFFIN, J. 2003. Caffeine ingestion and fluid balance: a review. *Journal of Human Nutrition and Dietetics,* 16, 411-420.

MAURI, P., MIGLIAZZA, B. & PIETTA, P. 1999. Liquid chromatography/electrospray mass spectrometry of bioactive terpenoids in Ginkgo biloba L. *Journal of Mass Spectrometry,* 34, 1361-1367.

MAURI, P. & PIETTA, P. 2000a. Electrospray characterization of selected medicinal plant extracts. *Journal of Pharmaceutical and Biomedical Analysis,* 23, 61-68.

MAURI, P. & PIETTA, P. 2000b. High performance liquid chromatography/electrospray mass spectrometry of Hypericum perforatum extracts. *Rapid Communications in Mass Spectrometry,* 14, 95-99.

MAY, D. H., NAVARRO, S. L., RUCZINSKI, I., HOGAN, J., OGATA, Y., SCHWARZ, Y., LEVY, L., HOLZMAN, T., MCINTOSH, M. W. & LAMPE, J. W. 2013. Metabolomic profiling of urine: response to a randomised, controlled feeding study of select fruits and vegetables, and application to an observational study. *British Journal of Nutrition,* 110, 1760-1770.

MCDOUGALL, G., MARTINUSSEN, I. & STEWART, D. 2008. Towards fruitful metabolomics: High throughput analyses of polyphenol composition in berries using direct infusion mass spectrometry. *Journal of Chromatography B,* 871, 362-369.

MEDINA, S., FERRERES, F., GARCÍA-VIGUERA, C., HORCAJADA, M. N., ORDUNA, J., SAVIRÓN, M., ZUREK, G., MARTÍNEZ-SANZ, J. M., GIL, J. I. & GIL-IZQUIERDO, A. 2013. Non-targeted metabolomic approach reveals urinary metabolites linked to steroid biosynthesis pathway after ingestion of citrus juice. *Food Chemistry,* 136, 938-946.

MEHTA, M., SONAWAT, H. & SHARMA, S. 2005. Malaria parasite-infected erythrocytes inhibit glucose utilization in uninfected red cells. *FEBS Letters,* 579, 6151 - 6158.

MEHTA, M., SONAWAT, H. M. & SHARMA, S. 2006. Glycolysis in Plasmodium falciparum results in modulation of host enzyme activities. *Journal of Vector Borne Diseases.*

MEIJERS, B. K. I., DE PRETER, V., VERBEKE, K., VANRENTERGHEM, Y. & EVENEPOEL, P. 2010. p-Cresyl sulfate serum concentrations in haemodialysis patients are reduced by the prebiotic oligofructose-enriched inulin. *Nephrology Dialysis Transplantation,* 25, 219-224.

MENASHE, L., HIRKO, K., LOSINA, E., KLOPPENBURG, M., ZHANG, W., LI, L. & HUNTER, D. J. 2012. The diagnostic performance of MRI in osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis Cartilage,* 20**,** 13-21.

MEYER, D. & STASSE-WOLTHUIS, M. 2009. The bifidogenic effect of inulin and oligofructose and its consequences for gut health. *European Journal of Clinical Nutrition,* 63**,** 1277-1289.

MICKIEWICZ, B., HEARD, B. J., CHAU, J. K., CHUNG, M., HART, D. A., SHRIVE, N. G., FRANK, C. B. & VOGEL, H. J. 2015. Metabolic profiling of synovial fluid in a unilateral ovine model of anterior cruciate ligament reconstruction of the knee suggests biomarkers for early osteoarthritis. *Journal of Orthopaedic Research,* 33**,** 71-77.

MILLER, J. N. & MILLER, J. C. 2010. *Statistics and chemometrics for analytical chemistry* Harlow, : Prentice Hall, .

MILLER, R. C., BRINDLE, E., HOLMAN, D. J., SHOFER, J., KLEIN, N. A., SOULES, M. R. & O'CONNOR, K. A. 2004. Comparison of Specific Gravity and Creatinine for Normalizing Urinary Reproductive Hormone Concentrations. *Clinical Chemistry,* 50**,** 924-932.

MOBASHERI, A. 2012. Osteoarthritis year 2012 in review: biomarkers. *Osteoarthritis and Cartilage,* 20**,** 1451-1464.

MONAGAS, M., KHAN, N., ANDRES-LACUEVA, C., URPI-SARDA, M., VAZQUEZ-AGELL, M., LAMUELA-RAVENTOS, R. M. & ESTRUCH, R. 2009. Dihydroxylated phenolic acids derived from microbial metabolism reduce lipopolysaccharide-stimulated cytokine secretion by human peripheral blood mononuclear cells. *British Journal of Nutrition,* 102**,** 201-206.

MORASSIN, B., FABRE, R., BERRY, A. & MAGNAVAL, J. F. 2002. One year's experience with the polymerase chain reaction as a routine method for the diagnosis of imported malaria. *The American Journal of Tropical Medicine and Hygiene,* 66**,** 503-8.

MOUGOUS, J. D., LEAVELL, M. D., SENARATNE, R. H., LEIGH, C. D., WILLIAMS, S. J., RILEY, L. W., LEARY, J. A. & BERTOZZI, C. R. 2002. Discovery of sulfated metabolites in mycobacteria with a genetic and mass spectrometric approach. *Proceedings of the National Academy of Sciences,* 99**,** 17037-17042.

MULLER, E. & KOLKER, S. 2004. Reduction of lysine intake while avoiding malnutrition—Major goals and major problems in dietary treatment of glutaryl-CoA dehydrogenase deficiency. *Journal of Inherited Metabolic Disease,* 27**,** 903-910.

MULLER, S., COOMBS, G. H. & WALTER, R. D. 2001. Targeting polyamines of parasitic protozoa in chemotherapy. *Trends in Parasitology,* 17**,** 242-249.

MULLER, S., DA'DARA, A., LUERSEN, K., WRENGER, C., DAS GUPTA, R., MADHUBALA, R. & WALTER, R. D. 2000. In the Human Malaria

Parasite Plasmodium falciparum,Polyamines Are Synthesized by a Bifunctional Ornithine Decarboxylase,S-Adenosylmethionine Decarboxylase. *Journal of Biological Chemistry,* 275, 8097-8102.

MULLER, V. 2001. *Bacterial Fermentation*, John Wiley & Sons, Ltd.

MUNDT, L. A. & SHANAHAN, K. 2010. *Graff's Textbook of Routine Urinalysis and Body Fluids*, Wolters Kluwer/Lippincott Williams & Wilkins Health.

MUNGUR, R., GLASS, A. D. M., GOODENOW, D. B. & LIGHTFOOT, D. A. 2005. Metabolite Fingerprinting in Transgenic Nicotiana tabacum Altered by the Escherichia coli Glutamate Dehydrogenase Gene. *Journal of Biomedicine and Biotechnology,* 2005, 198-214.

MURRAY, K., WILKINSON-SMITH, V., HOAD, C., COSTIGAN, C., COX, E., LAM, C., MARCIANI, L., GOWLAND, P. & SPILLER, R. C. 2014. Differential Effects of FODMAPs (Fermentable Oligo-, Di-, Mono-Saccharides and Polyols) on Small and Large Intestinal Contents in Healthy Subjects Shown by MRI. *American Journal of Gastroenterology,* 109, 110-119.

MUSSO, G., GAMBINO, R. & CASSADER, M. 2010. Obesity, Diabetes, and Gut Microbiota: The hygiene hypothesis expanded? *Diabetes Care,* 33, 2277-2284.

NAKAMURA, Y., KIMURA, A., SAGA, H., OIKAWA, A., SHINBO, Y., KAI, K., SAKURAI, N., SUZUKI, H., KITAYAMA, M., SHIBATA, D., KANAYA, S. & OHTA, D. 2007. Differential metabolomics unraveling light/dark regulation of metabolic activities in Arabidopsis cell culture. *Planta,* 227, 57-66.

NAM, S., JOO, S., KIM, S., BAEK, N.-I., CHOI, H.-K. & PARK, S. 2008. Induced metabolite changes inMyriophyllum spicatum during co-existence experiment with the cyanobacteriumMicrocystis aeruginosa. *Journal of Plant Biology,* 51, 373-378.

NANKABIRWA, J., ZUROVAC, D., NJOGU, J., RWAKIMARI, J., COUNIHAN, H., SNOW, R. & TIBENDERANA, J. 2009. Malaria misdiagnosis in Uganda - implications for policy change. *Malaria Journal,* 8, 66.

NAVARRO-MUNOZ, M., IBERNON, M., BONET, J., PEREZ, V., PASTOR, M. C., BAYES, B., CASADO-VELA, J., NAVARRO, M., ARA, J., ESPINAL, A., FLUVIA, L., SERRA, A., LOPEZ, D. & ROMERO, R. 2012. Uromodulin and α(1)-antitrypsin urinary peptide analysis to differentiate glomerular kidney diseases. *Kidney and Blood Pressure Research,* 35, 314-325.

NDAGIJIMANA, M., LAGHI, L., VITALI, B., PLACUCCI, G., BRIGIDI, P. & GUERZONI, M. E. 2009. Effect of a synbiotic food consumption on human gut metabolic profiles evaluated by 1H Nuclear Magnetic Resonance spectroscopy. *International Journal of Food Microbiology,* 134, 147-153.

NEPPLE, J., THOMASON, K., AN, T., HARRIS-HAYES, M. & CLOHISY, J. 2015. What Is the Utility of Biomarkers for Assessing the Pathophysiology of Hip Osteoarthritis? A Systematic Review. *Clinical Orthopaedics and Related Research,* 473**,** 1683-1701.

NFMCPA. 2015. *Osteoarthritis (OA)* [Online]. NFMCPA. Available: http://www.fmcpaware.org/o-p/osteoarthritis [Accessed 19.04 2015].

NICHOLSON, J. K., HOLMES, E., KINROSS, J., BURCELIN, R., GIBSON, G., JIA, W. & PETTERSSON, S. 2012. Host-Gut Microbiota Metabolic Interactions. *Science,* 336**,** 1262-1267.

NIELSEN, J. & OLIVER, S. 2005. The next wave in metabolome analysis. *Trends in Biotechnology,* 23**,** 544-546.

NIELSEN, K. F., SMEDSGAARD, J., NIELSEN, K. F. & SMEDSGAARD, J. 2003. Fungal metabolite screening: database of 474 mycotoxins and fungal metabolites for dereplication by standardised liquid chromatography-UV-mass spectrometry methodology. *Journal of Chromatography A,* 1002**,** 111-136.

NIESSER, M., KOLETZKO, B. & PEISSNER, W. 2012. Determination of Creatinine in Human Urine with Flow Injection Tandem Mass Spectrometry. *Annals of Nutrition and Metabolism,* 61**,** 314-321.

NINESS, K. R. 1999. Inulin and Oligofructose: What Are They? *The Journal of Nutrition,* 129**,** 1402S-1406s.

NOBEL MEDIA AB. 2014. *Press Release: The Nobel Prize in Chemistry 2002* [Online]. Nobelprize.org. Available: http://www.nobelprize.org/nobel˙prizes/chemistry/laureates/2002/press.html [Accessed 21/12 2014].

NOMURA, Y., OKUMA, Y., SEGAWA, T., SCHMIDT-GLENEWINKEL, T. & GIACOBINI, E. 1981. Comparison of synaptosomal and glial uptake of pipecolic acid and GABA in rat brain. *Neurochemical Research,* 6**,** 391-400.

NONLINEAR-DYNAMICS. 2014. *Progenesis QI* [Online]. Available: http://www.nonlinear.com/progenesis/qi/ [Accessed 5/02/ 2014].

NORDSTROM, A., WANT, E., NORTHEN, T., LEHTIO, J. & SIUZDAK, G. 2007. Multiple Ionization Mass Spectrometry Strategy Used To Reveal the Complexity of Metabolomics. *Analytical Chemistry,* 80**,** 421-429.

NUDNOVA, M. M., ZHU, L. & ZENOBI, R. 2012. Active capillary plasma source for ambient mass spectrometry. *Rapid Communications in Mass Spectrometry,* 26**,** 1447-1452.

OHTA, D., SHIBATA, D. & KANAYA, S. 2007. Metabolic profiling using Fourier-transform ion-cyclotron-resonance mass spectrometry. *Analytical and Bioanalytical Chemistry,* 389**,** 1469-1475.

OIKAWA, A., NAKAMURA, Y., OGURA, T., KIMURA, A., SUZUKI, H., SAKURAI, N., SHINBO, Y., SHIBATA, D., KANAYA, S. & OHTA, D.

2006. Clarification of Pathway-Specific Inhibition by Fourier Transform Ion Cyclotron Resonance/Mass Spectrometry-Based Metabolic Phenotyping Studies. *Plant Physiology,* 142, 398-413.

OLDIGES, M., LUTZ, S., PFLUG, S., SCHROER, K., STEIN, N. & WIENDAHL, C. 2007. Metabolomics: current state and evolving methodologies and tools. *Applied Microbiology and Biotechnology,* 76, 495-511.

OLSZEWSKI, K. L. & LLINAS, M. 2011. Central carbon metabolism of Plasmodium parasites. *Molecular and Biochemical Parasitology,* 175, 95-103.

OLSZEWSKI, K. L., MORRISEY, J. M., WILINSKI, D., BURNS, J. M., VAIDYA, A. B., RABINOWITZ, J. D. & LLINŒS, M. 2009. Host-Parasite Interactions Revealed by Plasmodium falciparum Metabolomics. *Cell Host & Microbe,* 5, 191-199.

OVERY, D. P., ENOT, D. P., TAILLIART, K., JENKINS, H., PARKER, D., BECKMANN, M. & DRAPER, J. 2008. Explanatory signal interpretation and metabolite identification strategies for nominal mass FIE-MS metabolite fingerprints. *Nature Protocols,* 3, 471-485.

OVERY, S. A., WALKER, H. J., MALONE, S., HOWARD, T. P., BAXTER, C. J., SWEETLOVE, L. J., HILL, S. A. & QUICK, W. P. 2005. Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. *Journal of Experimental Botany,* 56, 287-296.

PACHER, P., BECKMAN, J. S. & LIAUDET, L. 2007. Nitric Oxide and Peroxynitrite in Health and Disease. 87, 315-424.

PAIN, A. & HERTZ-FOWLER, C. 2009. Plasmodium genomics: latest milestone. *Nature Reviews Microbiology,* 7, 180-181.

PARKER, D., BECKMANN, M., ENOT, D. P., OVERY, D. P., RIOS, Z. C., GILBERT, M., TALBOT, N. & DRAPER, J. 2008. Rice blast infection of Brachypodium distachyon as a model system to study dynamic host/pathogen interactions. *Nature Protocols,* 3, 435-445.

PARKER, D., BECKMANN, M., ZUBAIR, H., ENOT, D. P., CARACUEL-RIOS, Z., OVERY, D. P., SNOWDON, S., TALBOT, N. J. & DRAPER, J. 2009. Metabolomic analysis reveals a common pattern of metabolic re-programming during invasion of three host plant species by Magnaporthe grisea. *The Plant Journal,* 59, 723-737.

PATOCKA, J. & KUEHN, G. D. 2000. Natural polyamines and their biological consequence in mammals. *Current Opinion in Pharmacology,* 43, 119-124.

PAYNE, T. G., SOUTHAM, A. D., ARVANITIS, T. N. & VIANT, M. R. 2009. A Signal Filtering Method for Improved Quantification and Noise Discrimination in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry-Based Metabolomics Data. *Journal of the American Society for Mass Spectrometry,* 20, 1087-1095.

PENG, J., CHEN, Y.-T., CHEN, C.-L. & LI, L. 2014. Development of a Universal Metabolome-Standard Method for Long-Term LC–MS Metabolome Profiling and Its Application for Bladder Cancer Urine-Metabolite-Biomarker Discovery. *Analytical Chemistry,* 86, 6540-6547.

PERETZ, H., WATSON, D., BLACKBURN, G., ZHANG, T., LAGZIEL, A., SHTAUBER-NAAMATI, M., MORAD, T., KEREN-TARDAI, E., GREENSHPUN, V., USHER, S., SHALEV, H., LANDAU, D. & LEVARTOVSKY, D. 2012. Urine metabolomics reveals novel physiologic functions of human aldehyde oxidase and provides biomarkers for typing xanthinuria. *Metabolomics,* 8, 951-959.

PHIPPS, A. N., STEWART, J., WRIGHT, B. & WILSON, I. D. 1998. Effect of diet on the urinary excretion of hippuric acid and other dietary-derived aromatics in rat. A complex interaction between diet, gut microflora and substrate specificity. *Xenobiotica,* 28, 527-537.

PI, Z., YUE, H., MA, L., DING, L., LIU, Z. & LIU, S. 2011. Differentiation of various kinds of Fructus schisandrae by surface desorption atmospheric pressure chemical ionization mass spectrometry combined with Principal Component Analysis. *Analytica Chimica Acta,* 706, 285-290.

PIOVAN, A., FILIPPINI, R., CANIATO, R., BORSARINI, A., BINI MALECI, L. & CAPPELLETTI, E. M. 2004. Detection of hypericins in the "red glands" of Hypericum elodes by ESI–MS/MS. *Phytochemistry,* 65, 411-414.

PIOVAN, A., FILIPPINI, R. & FAVRETTO, D. 1998. Characterization of the anthocyanins of Catharanthus roseus (L.) G. Don in vivo and in vitro by electrospray ionization ion trap mass spectrometry. *Rapid Communications in Mass Spectrometry,* 12, 361-367.

PIPER, R., LEBRAS, J., WENTWORTH, L., HUNT-COOKE, A., HOUZE, S., CHIODINI, P. & MAKLER, M. 1999. Immunocapture diagnostic assays for malaria using Plasmodium lactate dehydrogenase (pLDH). *The American Journal of Tropical Medicine and Hygiene,* 60, 109-18.

PITT, J., WILLIS, F., TZANAKOS, N., BELOSTOTSKY, R. & FRISHBERG, Y. 2015. 4-Hydroxyglutamate Is a Biomarker for Primary Hyperoxaluria Type 3. *JIMD Reports, Volume 15,* 15, 1-6.

PLANCHE, T. & KRISHNA, S. 2006. Severe Malaria: Metabolic Complications. *Current Molecular Medicine,* 6, 141-153.

PLECKO, B., STÖCKLER-IPSIROGLU, S., PASCHKE, E., ERWA, W., STRUYS, E. A. & JAKOBS, C. 2000. Pipecolic acid elevation in plasma and cerebrospinal fluid of two patients with pyridoxine-dependent epilepsy. *Annals of Neurology,* 48, 121-125.

POPE, G. A., MACKENZIE, D. A., DEFERNEZ, M., AROSO, M. A. M. M., FULLER, L. J., MELLON, F. A., DUNN, W. B., BROWN, M., GOODACRE, R., KELL, D. B., MARVIN, M. E., LOUIS, E. J. & ROBERTS, I. N. 2007. Metabolic footprinting as a tool for discriminating between brewing yeasts. *Yeast,* 24, 667-679.

PROSCHOGO, N. W., ALBERTSON, P. L., BURSLE, J., MCCONCHIE, C. A., TURNER, A. G. & WILLETT, G. D. 2012. Aging Effects on Macadamia Nut Oil Studied by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Journal of Agricultural and Food Chemistry,* 60**,** 1973-1980.

PUJOS-GUILLOT, E., HUBERT, J., MARTIN, J.-F., LYAN, B., QUINTANA, M., CLAUDE, S., CHABANAS, B., ROTHWELL, J. A., BENNETAU-PELISSERO, C., SCALBERT, A., COMTE, B., HERCBERG, S., MORAND, C., GALAN, P. & MANACH, C. 2013. Mass Spectrometry-based Metabolomics for the Discovery of Biomarkers of Fruit and Vegetable Intake: Citrus Fruit as a Case Study. *Journal of Proteome Research,* 12**,** 1645-1659.

PURICH, D. L. & ALLISON, R. D. 1999. *Handbook of Biochemical Kinetics: A Guide to Dynamic Processes in the Molecular Life Sciences*, Elsevier Science.

QIN, J., LI, R., RAES, J., ARUMUGAM, M., BURGDORF, K. S., MANICHANH, C., NIELSEN, T., PONS, N., LEVENEZ, F., YAMADA, T., MENDE, D. R., LI, J., XU, J., LI, S., LI, D., CAO, J., WANG, B., LIANG, H., ZHENG, H., XIE, Y., TAP, J., LEPAGE, P., BERTALAN, M., BATTO, J.-M., HANSEN, T., LE PASLIER, D., LINNEBERG, A., NIELSEN, H. B., PELLETIER, E., RENAULT, P., SICHERITZ-PONTEN, T., TURNER, K., ZHU, H., YU, C., LI, S., JIAN, M., ZHOU, Y., LI, Y., ZHANG, X., LI, S., QIN, N., YANG, H., WANG, J., BRUNAK, S., DORE, J., GUARNER, F., KRISTIANSEN, K., PEDERSEN, O., PARKHILL, J., WEISSENBACH, J., BORK, P., EHRLICH, S. D. & WANG, J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature,* 464**,** 59-65.

RANASINGHE, A., RAMANATHAN, R., JEMAL, M., D'ARIENZO, C. J., HUMPHREYS, W. G. & OLAH, T. V. 2012. Integrated quantitative and qualitative workflow for in vivo bioanalytical support in drug discovery using hybrid Q-TOF-MS. *Bioanalysis,* 4**,** 511-528.

RANE, H. S., BERNARDO, S. M., HOWELL, A. B. & LEE, S. A. 2013. Cranberry-derived proanthocyanidins prevent formation of Candida albicans biofilms in artificial urine through biofilm- and adherence-specific mechanisms. *Journal of Antimicrobial Chemotherapy*.

RATHAHAO-PARIS, E., PARIS, A., BURSZTYKA, J., JAEG, J.-P., CRAVEDI, J.-P. & DEBRAUWER, L. 2014. Identification of xenobiotic metabolites from biological fluids using flow injection analysis high-resolution mass spectrometry and post-acquisition data filtering. *Rapid Communications in Mass Spectrometry,* 28**,** 2713-2722.

RAUHA, J.-P., VUORELA, H. & KOSTIAINEN, R. 2001. Effect of eluent on the ionization efficiency of flavonoids by ion spray, atmospheric pressure chemical ionization, and atmospheric pressure photoionization mass spectrometry. *Journal of Mass Spectrometry,* 36**,** 1269-1280.

RIDDER, L., VAN DER HOOFT, J. J. J., VERHOEVEN, S., DE VOS, R. C. H., VERVOORT, J. & BINO, R. J. 2014. In Silico Prediction and

Automatic LC–MSn Annotation of Green Tea Metabolites in Urine. *Analytical Chemistry,* 86, 4767-4774.

RIDLEY, R. G. 2003. Malaria: To kill a parasite. *Nature,* 424, 887-889.

RIST, V. T. S., WEISS, E., EKLUND, M. & MOSENTHIN, R. 2013. Impact of dietary protein on microbiota composition and activity in the gastrointestinal tract of piglets in relation to gut health: a review. *animal,* 7, 1067-1078.

ROBERFROID, M. 2007. Prebiotics: The Concept Revisited. *The Journal of Nutrition,* 137, 830S-837S.

ROBERFROID, M., GIBSON, G. R., HOYLES, L., MCCARTNEY, A. L., RASTALL, R., ROWLAND, I., WOLVERS, D., WATZL, B., SZAJEWSKA, H., STAHL, B., GUARNER, F., RESPONDEK, F., WHELAN, K., COXAM, V., DAVICCO, M.-J., LEOTOING, L., WITTRANT, Y., DELZENNE, N. M., CANI, P. D., NEYRINCK, A. M. & MEHEUST, A. 2010. Prebiotic effects: metabolic and health benefits. *British Journal of Nutrition,* 104, S1-S63.

ROSSELLO-MORA, R., LUCIO, M., PENA, A., BRITO-ECHEVERRIA, J., LOPEZ-LOPEZ, A., VALENS-VADELL, M., FROMMBERGER, M., ANTON, J. & SCHMITT-KOPPLIN, P. 2008. Metabolic evidence for biogeographic isolation of the extremophilic bacterium Salinibacter ruber. *The ISME Journal,* 2, 242-253.

ROTHWELL, J. A., FILLØTRE, Y., MARTIN, J.-F., LYAN, B., PUJOS-GUILLOT, E., FEZEU, L., HERCBERG, S., COMTE, B., GALAN, P., TOUVIER, M. & MANACH, C. 2014. New Biomarkers of Coffee Consumption Identified by the Non-Targeted Metabolomic Profiling of Cohort Study Subjects. *PLoS ONE,* 9, e93474.

ROUSSEAU, J.-C. & DELMAS, P. D. 2007. Biological markers in osteoarthritis. *Nature Clinical Practice Rheumatology,* 3, 346-356.

RYBI–SZUMINSKA, A., MICHALUK-SKUTNIK, J., OSIPIUK-REMZA, B., KOSSAKOWSKA, A. & WASILEWSKA, A. 2014. Normal values for urine renalase excretion in children. *Pediatric Nephrology (Berlin, Germany),* 29, 2191-2195.

RYCROFT, C. E., FOOKS, L. J. & GIBSON, G. R. 1999. Methods for assessing the potential of prebiotics and probiotics. *Current Opinion in Clinical Nutrition & Metabolic Care,* 2, 481-484.

SABATER-MOLINA, M., LARQUE, E., TORRELLA, F. & ZAMORA, S. 2009. Dietary fructooligosaccharides and potential benefits on health. *Journal of Physiology and Biochemistry,* 65, 315-328.

SACCENTI, E., HOEFSLOOT, H. J., SMILDE, A., WESTERHUIS, J. & HENDRIKS, M. W. B. 2014. Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics,* 10, 361-374.

SANA, T. R., GORDON, D. B., FISCHER, S. M., TICHY, S. E., KITAGAWA, N., LAI, C., GOSNELL, W. L. & CHANG, S. P. 2013. Global Mass

Spectrometry Based Metabolomics Profiling of Erythrocytes Infected with Plasmodium falciparum. *PLoS ONE,* 8, e60840.

SANGSTER, T., MAJOR, H., PLUMB, R., WILSON, A. J. & WILSON, I. D. 2006. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis. *Analyst,* 131, 1075-1078.

SAUDE, E., ADAMKO, D., ROWE, B., MARRIE, T. & SYKES, B. 2007. Variation of metabolites in normal human urine. *Metabolomics,* 3, 439-451.

SAWADA, Y., AKIYAMA, K., SAKATA, A., KUWAHARA, A., OTSUKI, H., SAKURAI, T., SAITO, K. & HIRAI, M. Y. 2009. Widely Targeted Metabolomics Based on Large-Scale MS/MS Data for Elucidating Metabolite Accumulation Patterns in Plants. *Plant and Cell Physiology,* 50, 37-47.

SAWAYA, A. C. H. F., CATHARINO, R. R., FACCO, E. M. P., FOGACA, A., GODOY, H. T., DAUDT, C. E. & EBERLIN, M. N. 2011. Monitoring of wine aging process by electrospray ionization mass spectrometry. *Food Science and Technology (Campinas),* 31, 730-734.

SCANZELLO, C. R. & GOLDRING, S. R. 2012. The role of synovitis in osteoarthritis pathogenesis. *Bone,* 51, 249-257.

SCHMIDT, J., BOETTCHER, C., KUHNT, C., KUTCHAN, T. M. & ZENK, M. H. 2007. Poppy alkaloid profiling by electrospray tandem mass spectrometry and electrospray FT-ICR mass spectrometry after [ring-13C6]-tyramine feeding. *Phytochemistry,* 68, 189-202.

SCHOLTENS, P. A. M. J., OOZEER, R., MARTIN, R., AMOR, K. B. & KNOL, J. 2012. The Early Settlers: Intestinal Microbiology in Early Life. *Annual Review of Food Science and Technology,* 3, 425-447.

SCHOLZ-AHRENS, K. E., SCHAAFSMA, G., VAN DEN HEUVEL, E. G. & SCHREZENMEIR, J. 2001. Effects of prebiotics on mineral metabolism. *The American Journal of Clinical Nutrition,* 73, 459s-464s.

SCOTT, I. M., VERMEER, C. P., LIAKATA, M., COROL, D. I., WARD, J. L., LIN, W., JOHNSON, H. E., WHITEHEAD, L., KULAR, B., BAKER, J. M., WALSH, S., DAVE, A., LARSON, T. R., GRAHAM, I. A., WANG, T. L., KING, R. D., DRAPER, J. & BEALE, M. H. 2010. Enhancement of Plant Metabolite Fingerprinting by Machine Learning. *Plant Physiology,* 153, 1506-1520.

SEDIC, M., GETHINGS, L. A., VISSERS, J. P. C., SHOCKCOR, J. P., MCDONALD, S., VASIEVA, O., LEMAC, M., LANGRIDGE, J. I., BATINIĆ, D. & PAVELIĆ, S. K. 2014. Label-free mass spectrometric profiling of urinary proteins and metabolites from paediatric idiopathic nephrotic syndrome. *Biochemical and Biophysical Research Communications,* 452, 21-26.

SEN, A., WANG, Y., CHIU, K., WHILEY, L., COWAN, D., CHANG, R. C.-C. & LEGIDO-QUIGLEY, C. 2013. Metabolic Phenotype of the Healthy

Rodent Model Using In-Vial Extraction of Dried Serum, Urine, and Cerebrospinal Fluid Spots. *Analytical Chemistry,* 85, 7257-7263.

SENGUPTA, A., BASANT, A., GHOSH, S., SHARMA, S. & SONAWAT, H. M. 2011a. Liver Metabolic Alterations and Changes in Host Intercompartmental Metabolic Correlation during Progression of Malaria. *Journal of Parasitology Research,* 2011.

SENGUPTA, A., GHOSH, S., BASANT, A., MALUSARE, S., JOHRI, P., PATHAK, S., SHARMA, S. & SONAWAT, H. 2011b. Global host metabolic response to Plasmodium vivax infection: a 1H NMR based urinary metabonomic study. *Malaria Journal,* 10, 384.

SENGUPTA, A., GHOSH, S., PATHAK, S., GOGTAY, N., THATTE, U., DOSHI, M., SHARMA, S. & SONAWAT, H. 2015. Metabolomic analysis of urine samples of vivax malaria in-patients for biomarker identification. *Metabolomics*, 1-12.

SHANE ANDERSON, A. & LOESER, R. F. 2010. Why is osteoarthritis an age-related disease? *Best practice & research. Clinical rheumatology,* 24, 15-26.

SHARMA, M. K., RAO, V. K., AGARWAL, G. S., RAI, G. P., GOPALAN, N., PRAKASH, S., SHARMA, S. K. & VIJAYARAGHAVAN, R. 2008. Highly Sensitive Amperometric Immunosensor for Detection of Plasmodium falciparum Histidine-Rich Protein 2 in Serum of Humans with Malaria: Comparison with a Commercial Kit. *Journal of Clinical Microbiology,* 46, 3759-3765.

SHARMA, M. K., RAO, V. K., MERWYN, S., AGARWAL, G. S., UPADHYAY, S. & VIJAYARAGHAVAN, R. 2011. A novel piezoelectric immunosensor for the detection of malarial Plasmodium falciparum histidine rich protein-2 antigen. *Talanta,* 85, 1812-1817.

SHOCKCOR, J. P., UNGER, S. E., WILSON, I. D., FOXALL, P. J. D., NICHOLSON, J. K. & LINDON, J. C. 1996. Combined HPLC, NMR Spectroscopy, and Ion-Trap Mass Spectrometry with Application to the Detection and Characterization of Xenobiotic and Endogenous Metabolites in Human Urine. *Analytical Chemistry,* 68, 4431-4435.

SIRICH, T. L., ARONOV, P. A., PLUMMER, N. S., HOSTETTER, T. H. & MEYER, T. W. 2013. Numerous protein-bound solutes are cleared by the kidney with high efficiency. *Kidney International,* 84, 585-590.

SMEDSGAARD, J. 1997. Terverticillate Penicillia Studied by Direct Electrospray Mass Spectrometric Profiling of Crude Extracts. II. Database and Identification. *Biochemical Systematics and Ecology,* 25, 65-71.

SMEDSGAARD, J. & FRISVAD, J. C. 1996. Using direct electrospray mass spectrometry in taxonomy and secondary metabolite profiling of crude fungal extracts. *Journal of Microbiological Methods,* 25, 5-17.

SMEDSGAARD, J., HANSEN, M. E., FRISVAD, J. C., SMEDSGAARD, J., HANSEN, M. E. & FRISVAD, J. C. 2004. Classification of terverticillate

Penicillia by electrospray mass spectrometric profiling. *Studies in Mycology,* 49, 243-251.

SMITH, C. A., WANT, E. J., O'MAILLE, G., ABAGYAN, R. & SIUZDAK, G. 2006. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Analytical Chemistry,* 78, 779-787.

SMRS 2005. Summary recommendations for standardization and reporting of metabolic analyses. *Nature Biotechnology,* 23, 833-838.

SOMMER, F. & BACKHED, F. 2013. The gut microbiota [mdash] masters of host development and physiology. *Nature Reviews Microbiology,* 11, 227-238.

SOUTHAM, A. D., PAYNE, T. G., COOPER, H. J., ARVANITIS, T. N. & VIANT, M. R. 2007. Dynamic range and mass accuracy of wide-scan direct infusion nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Analytical Chemistry,* 79, 4595-4602.

SPENCER, J. P. E. 2003. Metabolism of Tea Flavonoids in the Gastrointestinal Tract *The Journal of Nutrition,* 133, 3255S-3261S.

STERNER, J. L., JOHNSTON, M. V., NICOL, G. R. & RIDGE, D. P. 2000. Signal suppression in electrospray ionization Fourier transform mass spectrometry of multi-component samples. *Journal of Mass Spectrometry,* 35, 385-391.

STOLL, N., SCHMIDT, E. & THUROW, K. 2006. Isotope Pattern Evaluation for the Reduction of Elemental Compositions Assigned to High-Resolution Mass Spectral Data from Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Journal of the American Society for Mass Spectrometry,* 17, 1692-1699.

STORM, J., SETHIA, S., BLACKBURN, G. J., CHOKKATHUKALAM, A., WATSON, D. G., BREITLING, R., COOMBS, G. H. & MULLER, S. 2014. Phosphoenolpyruvate Carboxylase Identified as a Key Enzyme in Erythrocytic Plasmodium falciparum Carbon Metabolism. *PLoS Pathogens,* 10, e1003876.

STRASINGER, S. K. & DI-LORENZO, M. S. 2008. *Urinalysis and Body Fluids,* Philadelphia, F. A. Davis Company.

STRUCK, W., SILUK, D., YUMBA-MPANGA, A., MARKUSZEWSKI, M., KALISZAN, R. & MARKUSZEWSKI, M. J. 2013. Liquid chromatography tandem mass spectrometry study of urinary nucleosides as potential cancer markers. *Journal of Chromatography A,* 1283, 122-131.

SU, S., DUAN, J., WANG, P., LIU, P., GUO, J., SHANG, E., QIAN, D., TANG, Y. & TANG, Z. 2013. Metabolomic Study of Biochemical Changes in the Plasma and Urine of Primary Dysmenorrhea Patients Using UPLC–MS Coupled with a Pattern Recognition Approach. *Journal of Proteome Research,* 12, 852-865.

SUMNER, L., AMBERG, A., BARRETT, D., BEALE, M., BEGER, R., DAYKIN, C., FAN, T. M., FIEHN, O., GOODACRE, R., GRIFFIN, J., HANKEMEIER, T., HARDY, N., HARNLY, J., HIGASHI, R., KOPKA, J., LANE, A., LINDON, J., MARRIOTT, P., NICHOLLS, A., REILY, M., THADEN, J. & VIANT, M. 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics,* 3**,** 211-221.

SUN, J. & CHEN, P. 2011. A flow-injection mass spectrometry fingerprinting method for authentication and quality assessment of Scutellaria lateriflora-based dietary supplements. *Analytical and Bioanalytical Chemistry,* 401**,** 1577-1584.

SURAPANENI, K. M. & VENKATARAMANA, G. 2007. Status of lipid peroxidation, glutathione, ascorbic acid, vitamin E and antioxidant enzymes in patients with osteoarthritis. *Indian Journal of Medical Sciences,* 61**,** 9-14.

SUROWIEC, I., ORIKIIRIZA, J., KARLSSON, E., NELSON, M., BONDE, M., KYAMANWA, P., KARENZI, B., BERGSTRÖM, S., TRYGG, J. & NORMARK, J. 2015. Metabolic signature profiling as a diagnostic and prognostic tool in paediatric Plasmodium falciparum malaria. *Open Forum Infectious Diseases*.

SUTHERLAND, C. J., TANOMSING, N., NOLDER, D., OGUIKE, M., JENNISON, C., PUKRITTAYAKAMEE, S., DOLECEK, C., HIEN, T. T., DO ROSŒRIO, V. E., AREZ, A. P., PINTO, J., MICHON, P., ESCALANTE, A. A., NOSTEN, F., BURKE, M., LEE, R., BLAZE, M., OTTO, T. D., BARNWELL, J. W., PAIN, A., WILLIAMS, J., WHITE, N. J., DAY, N. P. J., SNOUNOU, G., LOCKHART, P. J., CHIODINI, P. L., IMWONG, M. & POLLEY, S. D. 2010. Two Nonrecombining Sympatric Forms of the Human Malaria Parasite Plasmodium ovale Occur Globally. *Journal of Infectious Diseases,* 201**,** 1544-1550.

SZABO, C., ISCHIROPOULOS, H. & RADI, R. 2007. Peroxynitrite: biochemistry, pathophysiology and development of therapeutics. *Nature Reviews Drug Discovery,* 6**,** 662-680.

TAKAHASHI, H., KAI, K., SHINBO, Y., TANAKA, K., OHTA, D., OSHIMA, T., ALTAF-UL-AMIN, M., KUROKAWA, K., OGASAWARA, N. & KANAYA, S. 2008. Metabolomics approach for determining growth-specific metabolites based on Fourier transform ion cyclotron resonance mass spectrometry. *Analytical and Bioanalytical Chemistry,* 391**,** 2769-2782.

TAKATS, Z., WISEMAN, J. M., GOLOGAN, B. & COOKS, R. G. 2004. Mass Spectrometry Sampling Under Ambient Conditions with Desorption Electrospray Ionization. *Science (Washington),* 306**,** 471-473.

TANG, D.-Q., BIAN, T.-T., ZHENG, X.-X., LI, Y., WU, X.-W., LI, Y.-J., DU, Q. & JIANG, S.-S. 2014. LC-MS/MS methods for the determination of edaravone and/or taurine in rat plasma and its application to a pharmacokinetic study. *Biomedical Chromatography,* 28**,** 1173-1182.

TAYLOR, N., WEBER, R. M., SOUTHAM, A., PAYNE, T., HRYDZIUSZKO, O., ARVANITIS, T. & VIANT, M. 2009. A new approach to toxicity

testing in Daphnia magna: application of high throughput FT-ICR mass spectrometry metabolomics. *Metabolomics,* 5, 44-58.

TENG, R., JUNANKAR, P. R., BUBB, W. A., RAE, C., MERCIER, P. & KIRK, K. 2009. Metabolite profiling of the intraerythrocytic malaria parasite Plasmodium falciparum by 1H NMR spectroscopy. *NMR in Biomedicine,* 22, 292-302.

THOMSON, B. A. & IRIBARNE, J. V. 1979. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *The Journal of Chemical Physics,* 71, 4451-4463.

TJITRA, E., SUPRIANTO, S., DYER, M., CURRIE, B. J. & ANSTEY, N. M. 1999. Field Evaluation of the ICT Malaria P.f/P.v Immunochromatographic Test for Detection of Plasmodium falciparum andPlasmodium vivax in Patients with a Presumptive Clinical Diagnosis of Malaria in Eastern Indonesia. *Journal of Clinical Microbiology,* 37, 2412-2417.

TOMLINSON, L., FUCHSER, J., FÜTTERER, A., BAUMERT, M., HASSALL, D. G., WEST, A. & MARSHALL, P. S. 2014. Using a single, high mass resolution mass spectrometry platform to investigate ion suppression effects observed during tissue imaging. *Rapid Communications in Mass Spectrometry,* 28, 995-1003.

TONHOSOLO, R., D'ALEXANDRI, F. L., DE ROSSO, V. V., GAZARINI, M. L., MATSUMURA, M. Y., PERES, V. J., MERINO, E. F., CARLTON, J. M., WUNDERLICH, G., MERCADANTE, A. Z., KIMURA, E. A. & KATZIN, A. M. 2009. Carotenoid Biosynthesis in Intraerythrocytic Stages of Plasmodium falciparum. *Journal of Biological Chemistry,* 284, 9974-9985.

TOPPING, D. L., FUKUSHIMA, M. & BIRD, A. R. 2003. Resistant starch as a prebiotic and synbiotic: state of the art. *Proceedings of the Nutrition Society,* 62, 171-176.

TRAMPUZ, A., JEREB, M., MUZLOVIC, I. & PRABHU, R. M. 2003. Clinical review: Severe malaria. *Critical Care,* 7, 315-323.

TRITTEN, L., KEISER, J., GODEJOHANN, M., UTZINGER, J., VARGAS, M., BECKONERT, O., HOLMES, E. & SARIC, J. 2013. Metabolic Profiling Framework for Discovery of Candidate Diagnostic Markers of Malaria. *Scientific Reports,* 3.

TRUNZER, M., GRAF, D. & KIFFE, M. 2007. Comparison of a two-dimensional liquid chromatography/mass spectrometry approach with a chip-based nanoelectrospray device for structural elucidation of metabolites in a human ADME study using a quadrupole time-of-flight mass spectrometer. *Rapid Communications in Mass Spectrometry,* 21, 937-944.

TRYGG, J. & WOLD, S. 2002. Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics,* 16, 119-128.

TSENG, C.-L. & LI, L. 2014. High-performance isotope-labeling liquid chromatography mass spectrometry for investigating the effect of

drinking Goji tea on urine metabolome profiling. *Science China Chemistry,* 57, 678-685.

TSIKAS, D., MITSCHKE, A. & GUTZKI, F.-M. 2012. Measurement of 3-Nitro-Tyrosine in Human Plasma and Urine by Gas Chromatography-Tandem Mass Spectrometry. *Amino Acid Analysis,* 828, 255-270.

TSUCHIYA, Y., TAKAHASHI, Y., JINDO, T., FURUHAMA, K. & SUZUKI, K. T. 2003. Comprehensive evaluation of canine renal papillary necrosis induced by nefiracetam, a neurotransmission enhancer. *European Journal of Pharmacology,* 475, 119-128.

TWEEDDALE, H., NOTLEY-MCROBB, L. & FERENCI, T. 1998. Effect of slow growth on metabolism of Escherichia coli, as revealed by global metabolite pool ("Metabolome") analysis. *Journal of Bacteriology,* 180, 5109-5116.

VAIDYANATHAN, S., KELL, D. B. & GOODACRE, R. 2002. Flow-injection electrospray ionization mass spectrometry of crude cell extracts for high-throughput bacterial identification. *Journal of the American Society for Mass Spectrometry,* 13, 118-128.

VAIDYANATHAN, S., ROWLAND, J. J., KELL, D. B. & GOODACRE, R. 2001. Discrimination of Aerobic Endospore-forming Bacteria via Electrospray-Ionization Mass Spectrometry of Whole Cell Suspensions. *Analytical Chemistry,* 73, 4134-4144.

VAN DEN BERG, R., HOEFSLOOT, H., WESTERHUIS, J., SMILDE, A. & VAN DER WERF, M. 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics,* 7, 142.

VAN DER HOOFT, J. J. J., DE VOS, R. C. H., MIHALEVA, V., BINO, R. J., RIDDER, L., DE ROO, N., JACOBS, D. M., VAN DUYNHOVEN, J. P. M. & VERVOORT, J. 2012. Structural Elucidation and Quantification of Phenolic Conjugates Present in Human Urine after Tea Intake. *Analytical Chemistry,* 84, 7263-7271.

VAN HOUT, M. W. J., NIEDERL¨NDER, H. A. G., DE ZEEUW, R. A. & DE JONG, G. J. 2003. Ion suppression in the determination of clenbuterol in urine by solid-phase extraction atmospheric pressure chemical ionisation ion-trap mass spectrometry. *Rapid Communications in Mass Spectrometry,* 17, 245-250.

VAN WIETMARSCHEN, H. A., DAI, W., VAN DER KOOIJ, A. J., REIJMERS, T. H., SCHROËN, Y., WANG, M., XU, Z., WANG, X., KONG, H., XU, G., HANKEMEIER, T., MEULMAN, J. J. & VAN DER GREEF, J. 2012. Characterization of Rheumatoid Arthritis Subtypes Using Symptom Profiles, Clinical Chemistry and Metabolomics Measurements. *PLoS ONE,* 7, e44331.

VERBRUGGHE, A., HESTA, M., GOMMEREN, K., DAMINET, S., WUYTS, B., BUYSE, J. & JANSSENS, G. P. J. 2009. Oligofructose and inulin modulate glucose and amino acid metabolism through propionate

production in normal-weight and obese cats. *British Journal of Nutrition,* 102, 694-702.

VERBRUGGHE, A., JANSSENS, G. P. J., MEININGER, E., DAMINET, S., PIRON, K., VANHAECKE, L., WUYTS, B., BUYSE, J. & HESTA, M. 2010. Intestinal fermentation modulates postprandial acylcarnitine profile and nitrogen metabolism in a true carnivore: the domestic cat (Felis catus). *British Journal of Nutrition,* 104, 972-979.

VERHOEF, S. P. M., MEYER, D. & WESTERTERP, K. R. 2011. Effects of oligofructose on appetite profile, glucagon-like peptide 1 and peptide YY3-36 concentrations and energy intake. *British Journal of Nutrition,* 106, 1757-1762.

VILLAGRA, E., SANTOS, L. S., VAZ, B. G., EBERLIN, M. N. & FELIPE LAURIE, V. 2012. Varietal discrimination of Chilean wines by direct injection mass spectrometry analysis combined with multivariate statistics. *Food Chemistry,* 131, 692-697.

VINAIXA, M., SAMINO, S., SAEZ, I., DURAN, J., GUINOVART, J. J. & YANES, O. 2012. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites,* 2, 775-795.

VISWANATHAN, C. T., BANSAL, S., BOOTH, B., DESTEFANO, A., ROSE, M., SAILSTAD, J., SHAH, V., SKELLY, J., SWANN, P. & WEINER, R. 2007. Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays. *Pharmaceutical Research,* 24, 1962-1973.

VITALI, B., NDAGIJIMANA, M., CRUCIANI, F., CARNEVALI, P., CANDELA, M., GUERZONI, M. & BRIGIDI, P. 2010. Impact of a synbiotic food on the gut microbial ecology and metabolic profiles. *BMC Microbiology,* 10, 4.

WAGNER, B. D., ACCURSO, F. J. & LAGUNA, T. A. 2010. The applicability of urinary creatinine as a method of specimen normalization in the cystic fibrosis population. *Journal of Cystic Fibrosis,* 9, 212-216.

WAIKAR, S. S., SABBISETTI, V. S. & BONVENTRE, J. V. 2010. Normalization of urinary biomarkers to creatinine during changes in glomerular filtration rate. *Kidney International,* 78, 486-494.

WALLNER-LIEBMANN, S., GRALKA, E., TENORI, L., KONRAD, M., HOFMANN, P., DIEBER-ROTHENEDER, M., TURANO, P., LUCHINAT, C. & ZATLOUKAL, K. 2014. The impact of free or standardized lifestyle and urine sampling protocol on metabolome recognition accuracy. *Genes & Nutrition,* 10, 1-9.

WALSH, M. C., BRENNAN, L., MALTHOUSE, J. P. G., ROCHE, H. M. & GIBNEY, M. J. 2006. Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *The American Journal of Clinical Nutrition,* 84, 531-539.

WANG, X., ZHANG, A., HAN, Y., WANG, P., SUN, H., SONG, G., DONG, T., YUAN, Y., YUAN, X., ZHANG, M., XIE, N., ZHANG, H., DONG, H. & DONG, W. 2012a. Urine Metabolomics Analysis for Biomarker Discovery and Detection of Jaundice Syndrome in Patients With Liver Disease. *Molecular & Cellular Proteomics,* 11**,** 370-380.

WANG, Y., XIAO, J., SUZEK, T. O., ZHANG, J., WANG, J., ZHOU, Z., HAN, L., KARAPETYAN, K., DRACHEVA, S., SHOEMAKER, B. A., BOLTON, E., GINDULYTE, A. & BRYANT, S. H. 2012b. PubChem's BioAssay Database. *Nucleic Acids Research,* 40**,** D400-D412.

WANT, E. J., WILSON, I. D., GIKA, H., THEODORIDIS, G., PLUMB, R. S., SHOCKCOR, J., HOLMES, E. & NICHOLSON, J. K. 2010. Global metabolic profiling procedures for urine using UPLC-MS. *Nature Protocols,* 5**,** 1005-1018.

WARD, J. L., BAKER, J. M. & BEALE, M. H. 2007. Recent applications of NMR spectroscopy in plant metabolomics. *FEBS Journal,* 274**,** 1126-1131.

WARD, J. L., FORCAT, S., BECKMANN, M., BENNETT, M., MILLER, S. J., BAKER, J. M., HAWKINS, N. D., VERMEER, C. P., LU, C., LIN, W., TRUMAN, W. M., BEALE, M. H., DRAPER, J., MANSFIELD, J. W. & GRANT, M. 2010. The metabolic transition during disease following infection of Arabidopsis thaliana by Pseudomonas syringae pv. tomato. *The Plant Journal,* 63**,** 443-457.

WARRACK, B. M., HNATYSHYN, S., OTT, K.-H., REILY, M. D., SANDERS, M., ZHANG, H. & DREXLER, D. M. 2009. Normalization strategies for metabonomic analysis of urine samples. *Journal of Chromatography B,* 877**,** 547-552.

WATERS 2005. Waters Micromass Q-TOF Premier Mass Spectrometer Operator's Guide. Milford: Waters Corporation.

WHO 2014. World Malaria Report: 2013. .

WHO. 2015a. *Chronic rheumatic conditions* [Online]. World health organisation Available: http://www.who.int/chp/topics/rheumatic/en/ [Accessed 15.04 2015].

WHO. 2015b. *World Malaria Day: call to close gaps in prevention and treatment to defeat malaria* [Online]. World health organisation Available: http://www.who.int/mediacentre/news/releases/2015/world-malaria-day-2015/en/ [Accessed 04/06 2015].

WICKREMSINHE, E. R., ACKERMANN, B. L. & CHAUDHARY, A. K. 2005. Validating regulatory-compliant wide dynamic range bioanalytical assays using chip-based nanoelectrospray tandem mass spectrometry. *Rapid Communications in Mass Spectrometry,* 19**,** 47-56.

WIKLUND, S., JOHANSSON, E., SJOSTROM, L., MELLEROWICZ, E. J., EDLUND, U., SHOCKCOR, J. P., GOTTFRIES, J., MORITZ, T. & TRYGG, J. 2008. Visualization of GC/TOF-MS-Based Metabolomics

Data for Identification of Biochemically Interesting Compounds Using OPLS Class Models. *Analytical Chemistry,* 80, 115-122.

WILLIAM ALLWOOD, J., ELLIS, D. I., HEALD, J. K., GOODACRE, R. & MUR, L. A. J. 2006. Metabolomic approaches reveal that phosphatidic and phosphatidyl glycerol phospholipids are major discriminatory non-polar metabolites in responses by Brachypodium distachyon to challenge by Magnaporthe grisea. *The Plant Journal,* 46, 351-368.

WILLIAMS, H. R. T., COX, I. J., WALKER, D. G., NORTH, B. V., PATEL, V. M., MARSHALL, S. E., JEWELL, D. P., GHOSH, S., THOMAS, H. J. W., TEARE, J. P., JAKOBOVITS, S., ZEKI, S., WELSH, K. I., TAYLOR-ROBINSON, S. D. & ORCHARD, T. R. 2009. Characterization of Inflammatory Bowel Disease With Urinary Metabolic Profiling. *American Journal of Gastroenterology,* 104, 1435-1444.

WILSON, I. D. 2011. High-Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS)-Based Drug Metabolite Profiling. *Metabolic Profiling,* 708, 173-190.

WINDEY, K., FRANÇOIS, I., BROEKAERT, W., DE PRETER, V., DELCOUR, J. A., LOUAT, T., HERMAN, J. & VERBEKE, K. 2014. High dose of prebiotics reduces fecal water cytotoxicity in healthy subjects. *Molecular Nutrition & Food Research,* 58, 2206-2218.

WINNIKE, J. H., BUSBY, M. G., WATKINS, P. B. & O'CONNELL, T. M. 2009. Effects of a prolonged standardized diet on normalizing the human metabolome. *The American Journal of Clinical Nutrition,* 90, 1496-1501.

WISHART, D. S., JEWISON, T., GUO, A. C., WILSON, M., KNOX, C., LIU, Y., DJOUMBOU, Y., MANDAL, R., AZIAT, F., DONG, E., BOUATRA, S., SINELNIKOV, I., ARNDT, D., XIA, J., LIU, P., YALLOU, F., BJORNDAHL, T., PEREZ-PINEIRO, R., EISNER, R., ALLEN, F., NEVEU, V., GREINER, R. & SCALBERT, A. 2013. HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Research,* 41, D801-D807.

WOLD, S., JOSEFSON, M., GOTTFRIES, J. & LINUSSON, A. 2004. The utility of multivariate design in PLS modeling. *Journal of Chemometrics,* 18, 156-165.

WOMERSLEY, C. 1981. A micromethod for the extraction and quantitative analysis of "free" carbohydrates in nematode tissue. *Analytical Biochemistry,* 112, 182-189.

WONG, J. W. H., CAGNEY, G. & CARTWRIGHT, H. M. 2005. SpecAlign—processing and alignment of mass spectra datasets. *Bioinformatics,* 21, 2088-2090.

WOOLF, A. D. & PFLEGER, B. 2003. Burden of major musculoskeletal conditions. *Bulletin of the World Health Organization,* 81, 646-656.

WU, Q., WANG, Y., GU, X., ZHOU, J., ZHANG, H., LV, W., CHEN, Z. & YAN, C. 2014. Urinary metabolomic study of non-small cell lung carcinoma based on ultra high performance liquid chromatography

coupled with quadrupole time-of-flight mass spectrometry. *Journal of Separation Science,* 37, 1728-1735.

WU, Y. & LI, L. 2012. Determination of Total Concentration of Chemically Labeled Metabolites as a Means of Metabolome Sample Normalization and Sample Loading Optimization in Mass Spectrometry-Based Metabolomics. *Analytical Chemistry,* 84, 10723-10731.

WUNDERLICHOVA, L., BUNKOVA, L., KOUTNY, M., JANCOVA, P. & BUNKA, F. 2014. Formation, Degradation, and Detoxification of Putrescine by Foodborne Bacteria: A Review. *Comprehensive Reviews in Food Science and Food Safety,* 13, 1012-1030.

XIA, J., BROADHURST, D., WILSON, M. & WISHART, D. 2013. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics,* 9, 280-299.

XIA, J., MANDAL, R., SINELNIKOV, I. V., BROADHURST, D. & WISHART, D. S. 2012. MetaboAnalyst 2.0—a comprehensive server for metabolomic data analysis. *Nucleic Acids Research*, 1–7.

XIAYAN, L. & LEGIDO-QUIGLEY, C. 2008. Advances in separation science applied to metabonomics. *Electrophoresis,* 29, 3724-3736.

XIE, G., ZHAO, A., ZHAO, L., CHEN, T., CHEN, H., QI, X., ZHENG, X., NI, Y., CHENG, Y., LAN, K., YAO, C., QIU, M. & JIA, W. 2012. Metabolic Fate of Tea Polyphenols in Humans. *Journal of Proteome Research,* 11, 3449-3457.

YAMASHITA, M. & FENN, J. B. 1984. Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry,* 88, 4451-4459.

YANG, L., BENNETT, R., STRUM, J., ELLSWORTH, B., HAMILTON, D., TOMLINSON, M., WOLF, R., HOUSLEY, M., ROBERTS, B., WELSH, J., JACKSON, B., WOOD, S., BANKA, C., THULIN, C. & LINFORD, M. 2009. Screening phosphatidylcholine biomarkers in mouse liver extracts from a hypercholesterolemia study using ESI-MS and chemometrics. *Analytical and Bioanalytical Chemistry,* 393, 643-654.

YANG, W., WANG, Y., ZHOU, Q. & TANG, H. 2008. Analysis of human urine metabolites using SPE and NMR spectroscopy. *Science in China Series B: Chemistry,* 51, 218-225.

YANG, Y., KAMEOKA, J., WACHS, T., HENION, J. D. & CRAIGHEAD, H. G. 2004. Quantitative Mass Spectrometric Determination of Methylphenidate Concentration in Urine Using an Electrospray Ionization Source Integrated with a Polymer Microchip. *Analytical Chemistry,* 76, 2568-2574.

YIN, P., WAN, D., ZHAO, C., CHEN, J., ZHAO, X., WANG, W., LU, X., YANG, S., GU, J. & XU, G. 2009. A metabonomic study of hepatitis B-induced liver cirrhosis and hepatocellular carcinoma by using RP-LC and HILIC coupled with mass spectrometry. *Molecular BioSystems,* 5, 868-876.

YU, S. P.-C. & HUNTER, D. J. 2015. Emerging drugs for the treatment of knee osteoarthritis. *Expert Opinion on Emerging Drugs,* 0, 1-18.

YU, Z.-R., NING, Y., YU, H. & TANG, N.-J. 2014. A HPLC-Q-TOF-MS-based urinary metabolomic approach to identification of potential biomarkers of metabolic syndrome. *Journal of Huazhong University of Science and Technology,* 34, 276-283.

YUE, M.-E., XU, J., LI, Q.-Q. & HOU, W.-G. 2011. Identification of coumarins in traditional Chinese medicine by direct-injection electrospray ionisation tandem mass spectrometry and high-performance liquid chromatography-mass spectrometry. *Journal of Analytical Chemistry,* 66, 290-295.

YUSUF, E., NELISSEN, R. G., IOAN-FACSINAY, A., STOJANOVIC-SUSULIC, V., DEGROOT, J., VAN OSCH, G., MIDDELDORP, S., HUIZINGA, T. W. J. & KLOPPENBURG, M. 2010. Association between weight or body mass index and hand osteoarthritis: a systematic review. *Annals of the Rheumatic Diseases,* 69, 761-765.

ZAMBONI, N. & FISHER, S. M. 2012. High-Throughput, High-Efficiency Metabolome Profiling Using the Agilent 6550 iFunnel Q-TOF LC/MS System. *Agilent Technologies*, 1-7.

ZAMFIR, A., VAKHRUSHEV, S., STERLING, A., NIEBEL, H. J., ALLEN, M. & PETER-KATALINIĆ, J. 2004. Fully Automated Chip-Based Mass Spectrometry for Complex Carbohydrate System Analysis. *Analytical Chemistry,* 76, 2046-2054.

ZAMPA, A., SILVI, S., FABIANI, R., MOROZZI, G., ORPIANESI, C. & CRESCI, A. 2004. Effects of different digestible carbohydrates on bile acid metabolism and SCFA production by human gut micro-flora grown in an in vitro semi-continuous culture. *Anaerobe,* 10, 19-26.

ZELENA, E., DUNN, W. B., BROADHURST, D., FRANCIS-MCINTYRE, S., CARROLL, K. M., BEGLEY, P., O'HAGAN, S., KNOWLES, J. D., HALSALL, A., WILSON, I. D. & KELL, D. B. 2009. Development of a Robust and Repeatable UPLC−MS Method for the Long-Term Metabolomic Study of Human Serum. *Analytical Chemistry,* 81, 1357-1364.

ZHANG, A.-H., WANG, P., SUN, H., YAN, G.-L., HAN, Y. & WANG, X.-J. 2013a. High-throughput ultra-performance liquid chromatography-mass spectrometry characterization of metabolites guided by a bioinformatics program. *Molecular BioSystems,* 9, 2259-2265.

ZHANG, A., SUN, H., YAN, G., HAN, Y., YE, Y. & WANG, X. 2013b. Urinary metabolic profiling identifies a key role for glycocholic acid in human liver cancer by ultra-performance liquid-chromatography coupled with high-definition mass spectrometry. *Clinica Chimica Acta,* 418, 86-90.

ZHANG, J., HUANG, Z., CHEN, M., XIA, Y., MARTIN, F. L., HANG, W. & SHEN, H. 2014a. Urinary metabolome identifies signatures of oligozoospermic infertile men. *Fertility and Sterility,* 102, 44-53.e12.

ZHANG, T., WATSON, D. G., WANG, L., ABBAS, M., MURDOCH, L., BASHFORD, L., AHMAD, I., LAM, N.-Y., NG, A. C. F. & LEUNG, H. Y. 2013c. Application of Holistic Liquid Chromatography-High Resolution Mass Spectrometry Based Urinary Metabolomics for Prostate Cancer Detection and Biomarker Discovery. *PLoS ONE,* 8**,** e65880.

ZHANG, W., LIKHODII, S., ZHANG, Y., AREF-ESHGHI, E., HARPER, P. E., RANDELL, E., GREEN, R., MARTIN, G., FUREY, A., SUN, G., RAHMAN, P. & ZHAI, G. 2014b. Classification of osteoarthritis phenotypes by metabolomics analysis. *BMJ Open,* 4.

ZHANG, Y., REN, H., JIANG, Y., GAO, Y.-F. & LIU, S.-Y. 2013d. Urinary metabonomics of stomach cancer assessed by rapid resolution liquid chromatography/time-of-flight mass spectrometry. *Chinese Medical Journal,* 10**,** 1930-1933.

ZHANG, Y., TOBIAS, H. J., AUCHUS, R. J. & BRENNA, J. T. 2011. Comprehensive Two Dimensional Gas Chromatography Fast Quadrupole Mass Spectrometry (GC×GC-qMS) for Urinary Steroid Profiling. Mass Spectral Characteristics with Chemical Ionization. *Drug testing and analysis,* 3**,** 857-867.

ZHOU, H., TAI, Y., SUN, C. & PAN, Y. 2005. Rapid identification of vinca alkaloids by direct-injection electrospray ionisation tandem mass spectrometry and confirmation by high-performance liquid chromatography–mass spectrometry. *Phytochemical Analysis,* 16**,** 328-333.

ZUBAREV, R. A. & MAKAROV, A. 2013. Orbitrap Mass Spectrometry. *Analytical Chemistry,* 85**,** 5288-5296.

APPENDICES

Appendices

Appendix A Artificial urine composition and concentrations

Compound	MW (Da)	Conc. (µg/mL)	$[M+H]^+$	$[M-H]^-$
Adrenaline	183.2044	83.5	184.0974	182.0817
Alanine	89.0932	38.7	90.0555	88.0399
AM4-Aminohippuric acid	194.1873	12.8	195.0770	193.0613
Aspartic acid	133.1027	29.0	134.0453	132.0297
Citric acid	192.1235	28.8	193.0348	191.0192
Creatine	131.1332	29.6	132.0773	130.0617
Creatinine	113.1179	22.6	114.0667	112.0511
Cytidine	243.2166	75.9	244.0933	242.0777
3,4-Dihydroxymandelic acid	184.1461	17.7	185.0450	183.0293
3,4-Dihydroxyphenylacetic acid (DOPAC)	168.1467	60.5	169.0501	167.0344
3,4-Dihydroxyphenylalanine (L-DOPA)	197.1879	42.9	198.0766	196.0610
Glutamic acid	147.1293	21.1	148.0610	146.0453
Hippuric acid	179.1727	92.1	180.0661	178.0504
Histidine	155.1546	31.6	156.0773	154.0617
4-Hydroxy-3-methoxymandelic acid	198.1727	5.5	199.0606	197.0450
Homovanillic acid	182.1733	18.0	183.0657	181.0501
4-Hydroxyphenylacetic acid	152.1473	22.5	153.0552	151.0395
5-Hydroxyindole-3-acetic acid	191.1834	11.1	192.0661	190.0504
2-Hydroxyisobutyric acid	104.1045	60.4	105.0552	103.0395
Hypoxanthine	136.1115	12.7	137.0463	135.0307
Lactic acid	90.0779	60.2	91.0395	89.0239
4-Methyl-2-oxovaleric acid	130.1418	13.5	131.0708	129.0552
Metanephrine	197.231	2.7	198.1130	196.0974
3-Methyl-L-histidine	169.1811	51.1	170.0930	168.0773
MHPG sulphate	264.252	1.0	265.0382	263.0225
Normetanephrine	183.2044	14.9	184.0974	182.0817
Phenylalanine	165.1891	21.1	166.0868	164.0712
Serotonin	176.2151	13.3	177.1028	175.0871
trans-Aconitic acid	174.1082	23.0	175.0243	173.0086
Threonine	119.1192	34.8	120.0661	118.0504
Tryptophan	204.2252	31.9	205.0977	203.0821
Tyrosine	181.1885	54.7	182.0817	180.0661
Urea	60.0553	67.8	61.0402	59.0245
Uric acid	168.1103	62.5	169.0362	167.0205
Uridine	244.2014	48.5	245.0774	243.0617

Appendix B Artificial urine database (attached CD)

Appendix C Performance of the developed direct ESI-MS and LC-MS for urine analysis

			HILIC LC-MS					FIE-MS		Chip-based infusion		LESA-MS	
Artificial urine compounds	*m/z*	Ions	RT(min)	RT shift (min)	RSD%	Peak area	RSD%	Peak counts	RSD%	Peak counts	RSD%	Peak counts	RSD%
Adrenaline	166.0868	$[M+H-H2O]^+$	4.49	0.004	0.1%	2,931,350	2%	18,182	13%	104,768	12%	286,945	22%
Alanine	90.0555	$[M+H]^+$	-	-	-	Not detected	-	25,905	8%	130,257	19%	312,207	26%
4-Aminohippuric acid	193.0613	$[M-H]^-$	4.58	0.004	0.1%	257,993	3%	1,651	21%	27,287	13%	18,143	7%
Aspartic acid	132.0293	$[M-H]^-$	5.36	0.022	0.2%	227,452	7%	4,016	11%	64,435	10%	94,664	20%
Citric acid	173.0086	$[M-H-H2O]^-$	6.56	0.003	0.0%	21,647,473	1%	29,917	6%	4,374,650	9%	5,217,463	11%
Creatine	132.0773	$[M+H]^+$	6.56	0.015	0.2%	14,732,724	1%	55,737	10%	426,022	11%	808,210	23%
Creatinine	114.0667	$[M+H]^+$	3.15	0.012	0.2%	1,076,748,825	1%	3,382,380	3%	40,858,691	25%	32,433,272	25%
Cytidine	244.0933	$[M+H]^+$	3.10	0.014	0.3%	5,090,960	3%	3,096	21%	61,508	16%	150,504	5%
3,4-Dihydroxymandelic acid	183.0293	$[M-H]^-$	-	-	-	Not detected	-	1,414	9%	31,857	10%	57,274	20%
DOPAC	167.0342	$[M-H]^-$	0.90	0.001	0.2%	30,447	2%	Not detected	-	95,973	10%	218,381	16%
L-DOPA	198.0766	$[M+H]^+$	-	-	-	Not detected	-	2,001	18%	Not detected	-	29,288	21%
Glutamic acid	146.0453	$[M-H]^-$	-	-	-	Not detected	-	4,007	4%	50,162	14%	59,127	27%
Hippuric acid	178.0504	$[M-H]^-$	3.98	0.014	0.3%	167,721,038	1%	1,011,798	5%	13,421,248	8%	12,974,603	16%
Histidine	156.0773	$[M+H]^+$	1.36	0.002	0.1%	350,174	2%	32,875	6%	466,699	20%	456,689	12%
Homovanillic acid	181.0501	$[M-H]^-$	3.08	0.014	0.4%	7,178,010	2%	3,026	14%	52,233	5%	79,383	28%
5-Hydroxyindole-3-acetic acid	190.0504	$[M-H]^-$	2.74	0.016	0.5%	1,451,876	5%	1,434	18%	Not detected	-	9,628	6%
2-Hydroxyisobutyric acid	103.0395	$[M-H]^-$	2.92	0.020	0.4%	1,512,983	2%	20,510	4%	468,202	8%	706,803	7%
4-Hydroxy-3-methoxymandelic acid	197.0450	$[M-H]^-$	2.73	0.027	0.7%	1,332,275	4%	2,973	11%	50,801	14%	83,016	13%
4-Hydroxyphenylacetic acid	151.0395	$[M-H]^-$	2.84	0.012	0.4%	9,708,292	1%	2,997	7%	72,934	11%	117,690	21%
Hypoxanthine	137.0463	$[M+H]^+$	1.95	0.017	0.4%	35,143,665	2%	4,839	14%	46,839	22%	41,088	19%
Lactic acid	89.0239	$[M-H]^-$	2.28	0.003	0.1%	4,955,643	1%	11,751	9%	334,980	8%	852,108	9%
Metanephrine	180.1025	$[M+H-H2O]^+$	4.71	0.004	0.0%	5,055,099	2%	2,937	16%	29,334	12%	51,344	11%
MHPG sulphate	263.0225	$[M-H]^-$	0.76	0.002	0.2%	2,957,966	3%	Not detected	-	Not detected	-	17,244	22%
3-Methyl-L-histidine	170.0930	$[M+H]^+$	8.11	0.028	0.2%	24,290,882	2%	109,984	5%	1,254,331	16%	1,142,199	12%
4-Methyl-2-oxovaleric acid	129.0552	$[M-H]^-$	0.95	0.002	0.1%	1,910,989	3%	6,552	13%	177,531	7%	224,956	11%
Normetanephrine	184.0974	$[M+H-H2O]^+$	4.43	0.014	0.3%	64,347	7%	1,964	16%	29,523	18%	76,585	11%
Phenylalanine	164.0712	$[M-H]^-$	5.34	0.021	0.2%	815,668	3%	Not detected	-	16,235	9%	67,259	14%
Serotonin	177.1028	$[M+H]^+$	4.65	0.018	0.3%	280,219	4%	2,586	19%	48,753	9%	85,426	16%
t-Aconitic acid	173.0086	$[M-H]^-$	6.59	0.140	1.0%	21,653,714	2%	29,917	6%	4,374,650	9%	5,217,463	11%
Threonine	120.0661	$[M+H]^+$	6.58	0.013	0.2%	999,418	8%	11,024	7%	130,288	13%	124,375	21%
Tryptophan	203.0821	$[M-H]^-$	5.27	0.017	0.2%	748,628	3%	Not detected	-	Not detected	-	Not detected	-
Tyrosine	180.0660	$[M-H]^-$	4.11	0.018	0.3%	322,222	1%	122,507	6%	35,934	8%	31,907	19%
Urea	61.0402	$[M+H]^+$	1.44	0.002	0.1%	57,787,213	1%	2,586,231	4%	13,146,281	21%	9,608,980	30%
Uric acid	167.0233	$[M-H]^-$	4.48	0.023	0.3%	7,283,733	5%	8,186	10%	Not detected	-	Not detected	-
Uridine	243.0617	$[M-H]^-$	1.69	0.020	0.6%	34,794,114	1%	Not detected	-	34,902	12%	35,444	22%
Max RSD%	-	-	-	-	1.0%	-	8%	-	21%		25%	-	30%
Artificial urine compounds detected	-	-	31	-	-	31	-	30	-	30	-	33	-

RSD% of retention time (RT), Peak area/count of the artificial urine compounds were calculated from the analysis of urine samples spiked with artificial urine (n = 6).

Appendix D Validation of LC-MS, FIE-MS, chip-based infusion and LESA-MS analyses for green tea intervention metabolomics study using selected urine metabolites peaks from pooled QCs injected throughout the run

		Compounds present in	LC-MS (BEH-HILIC)					FIE-MS		Chip-based infusion		LESA-MS	
m/z	Polarity	artificial urine	RT(min)	Shift(min)	RSD%	Peak Area	RSD%	Counts	RSD%	Counts	RSD%	Counts	RSD%
61.0402	ESI+	Urea	1.438	0.014	0.65%	43,670,703	13%	2,552,632	6%	15,295,067	23%	6,492,819	16%
89.0239	ESI-	Lactic acid	2.262	0.016	0.40%	2,949,002	1%	1,904	18%	78,197	19%	379,888	24%
90.0555	ESI+	Alanine	-	-	-	-	-	12,620	6%	91,571	21%	58,451	25%
96.0682	ESI+	-	6.637	0.020	0.17%	291,764	6%	-	-	-	-	-	-
96.9600	ESI-	-	6.590	0.019	0.18%	31,097,261	1%	161,779	15%	1,948,150	11%	90,909	20%
103.0395	ESI-	2-Hydroxyisobutyric acid	-	-	-	-	-	6,531	14%	149,040	10%	211,104	32%
104.9923	ESI+	-	5.556	0.017	0.19%	20,591,625	3%	-	-	-	-	-	-
112.0515	ESI-	-	3.151	0.026	0.41%	173,519,038	1%	3,587	24%	110,811	24%	542,017	28%
114.0667	ESI+	Creatinine	3.117	0.029	0.55%	726,511,481	5%	3,044,014	6%	37,980,925	30%	29,040,752	29%
118.0862	ESI+	-	5.947	0.018	0.20%	4,871,985	2%	118,245	7%	1,587,559	22%	1,551,406	16%
120.0661	ESI+	Threonine	6.574	0.025	0.21%	467,339	3%	6,992	14%	126,616	29%	65,497	13%
129.0552	ESI-	4-Methyl-2-oxovaleric acid	-	-	-	-	-	1,584	16%	48,442	12%	39,423	16%
132.0293	ESI-	Aspartic acid	5.387	0.022	0.22%	92,553	8%	2,428	13%	47,491	9%	62,025	27%
132.0773	ESI+	Creatine	6.541	0.014	0.13%	8,506,068	3%	36,888	5%	402,164	18%	267,558	21%
136.0482	ESI+	-	3.187	0.019	0.41%	133,881,969	2%	139,545	13%	1,261,793	25%	770,216	28%
137.0463	ESI+	Hypoxanthine	1.939	0.014	0.38%	16,483,991	3%	3,086	14%	36,361	18%	-	-
144.1018	ESI+	-	6.058	0.010	0.14%	17,262,621	2%	323,422	8%	5,722,468	24%	4,588,623	11%
146.0453	ESI-	Glutamic acid	-	-	-	-	-	2,546	6%	39,763	15%	43,485	27%
151.0395	ESI-	4-Hydroxyphenylacetic acid	2.829	0.018	0.44%	4,630,459	4%	2,648	12%	71,454	10%	97,127	31%
152.0706	ESI+	-	4.344	0.017	0.22%	4,754,329	2%	2,425	20%	-	-	-	-
156.0773	ESI+	Histidine	1.324	0.007	0.27%	183,271	3%	24,198	5%	384,774	28%	278,000	22%
156.9909	ESI-	-	5.706	0.012	0.11%	60,208,713	0%	-	-	-	-	-	-
164.0712	ESI-	Phenylalanine	-	-	-	-	-	-	-	7,296	9%	16,341	14%
166.0868	ESI+	Adrenaline	-	-	-	-	-	5,816	9%	30,158	26%	-	-
167.0233	ESI-	Uric acid	4.446	0.020	0.28%	1,792,755	15%	-	-	-	-	-	-
170.0930	ESI+	3-Methyl-L-histidine	8.299	0.021	0.17%	23,299,696	4%	80,664	6%	1,023,530	23%	719,589	24%
172.991	ESI-	-	0.646	0.002	0.27%	33,875,833	3%	13,719	5%	523,654	9%	455,870	14%
173.0086	ESI-	Citric acid	6.620	0.027	0.24%	12,417,391	2%	26,699	7%	3,132,752	31%	4,359,845	32%
173.0086	ESI-	t-Aconitic acid	-	-	-	-	-	26,699	7%	3,132,752	31%	4,359,845	32%
174.0408	ESI-	-	6.845	0.016	0.17%	541,325	5%	2,137	15%	47,618	11%	55,010	34%
177.1028	ESI+	Serotonin	-	-	-	-	-	-	-	29,234	12%	-	-
178.0504	ESI-	Hippuric acid	3.967	0.011	0.19%	88,291,205	2%	737,726	3%	12,107,797	10%	11,838,764	10%

		Compounds present in	LC-MS (BEH-HILIC)					FIE-MS		Chip-based infusion		LESA-MS	
m/z	Polarity	artificial urine	RT(min)	Shift(min)	RSD%	Peak Area	RSD%	Counts	RSD%	Counts	RSD%	Counts	RSD%
180.0660	ESI-	Tyrosine	4.100	0.010	0.15%	159,759	4%	95,486	11%	29,472	5%	25,559	18%
180.1025	ESI+	Metanephrine	4.718	0.017	0.21%	26,157	9%	-	-	-	-	-	-
181.0501	ESI-	Homovanillic acid	3.062	0.017	0.31%	2,641,782	5%	2,668	8%	54,775	9%	46,693	11%
183.0293	ESI-	3,4-Dihydroxymandelic acid	-	-	-	-	-	-	-	10,792	9%	9,179	16%
184.0974	ESI+	Normetanephrine	-	-	-	-	-	1,427	19%	-	-	-	-
186.9953	ESI+	-	5.556	0.017	0.19%	14,363,621	2%	-	-	-	-	-	-
187.0070	ESI-	-	0.646	0.002	0.27%	201,375,264	3%	29,358	8%	2,875,052	12%	2,466,036	9%
190.0504	ESI-	5-Hydroxyindole-3-acetic acid	2.869	0.027	0.52%	228,654	4%	-	-	-	-	-	-
193.0613	ESI-	4-Aminohippuric acid	-	-	-	-	-	-	-	19,109	7%	12,743	15%
194.9275	ESI-	-	6.597	0.015	0.13%	17,316,364	2%	4,216	16%	8,451	28%	-	-
195.0521	ESI-	-	3.124	0.027	0.56%	5,493,711	4%	20,556	6%	168,079	4%	173,190	24%
197.0450	ESI-	4-Hydroxy-3-methoxymandelic acid	-	-	-	-	-	1,286	13%	27,417	17%	28,966	30%
203.0821	ESI-	Tryptophan	5.291	0.025	0.28%	170,484	7%	-	-	-	-	-	-
203.1502	ESI+	-	9.358	0.024	0.17%	18,479,597	4%	11,110	7%	139,469	11%	100,786	16%
216.9810	ESI-	-	1.356	0.007	0.29%	11,012,194	2%	1,411	13%	52,290	8%	52,727	16%
225.0628	ESI-	-	1.722	0.013	0.39%	8,202,141	2%	22,430	5%	150,481	12%	179,836	26%
229.1544	ESI+	-	8.141	0.019	0.14%	19,678,192	6%	76,580	5%	1,920,355	32%	983,268	9%
232.0273	ESI+	-	0.702	0.003	0.28%	9,722,750	3%	1,240	19%	-	-	-	-
243.0617	ESI-	Uridine	1.678	0.012	0.39%	23,815,846	2%	6,764	12%	35,640	19%	36,959	28%
245.0123	ESI-	-	1.457	0.014	0.49%	1,952,090	4%	-	-	19,542	6%	18,151	12%
249.1070	ESI+	-	3.187	0.023	0.52%	20,940,966	3%	41,411	17%	69,941	21%	41,659	15%
254.9814	ESI-	-	-	-	-	-	-	3,444	15%	85,333	19%	86,737	31%
263.0225	ESI-	MHPG sulphate	0.754	0.004	0.30%	1,684,117	4%	-	-	-	-	-	-
263.1036	ESI-	-	4.754	0.017	0.21%	15,618,321	1%	80,200	3%	1,280,472	9%	940,453	15%
265.1180	ESI+	-	4.750	0.013	0.18%	9,580,283	2%	19,738	13%	184,849	7%	92,354	9%
RSD% (maximum)					0.65%		15%		24%		32%		34%

RSD% of retention time (RT), Peak area/count of the selected peaks were calculated from the analysis of the pooled QC urine sample (n = 7).

Appendix E Validation of LC-MS, FIE-MS, chip-based infusion and LESA-MS metabolomics analyses of osteoarthritis urine samples using selected metabolites peaks from pooled QC injections interspaced throughout the analytical run

			LC-MS (BEH-HILIC)				FIE-MS		Chip-based infusion		LESA-MS	
m/z	Polarity	Compounds present in artificial urine	RT(min)	RSD(%)	Peak Area	RSD(%)	Counts	RSD(%)	Counts	RSD(%)	Counts	RSD(%)
61.0399	ESI+	Urea	1.42	0.89%	18,726,739	15.5%	90,504	19.4%	2,710,008	11.2%	1,266,953	27.9%
76.0758	ESI+	-	7.77	0.22%	8,071,358	11.9%	183,442	17.4%	6,696,967	11.7%	8,039,761	21.6%
88.0399	ESI-	L-alanine	6.15	0.00	284818.23	18.0%	176	13.3%	10,535	21.0%	19,615	29.2%
94.0653	ESI+	-	5.90	0.28%	3,045,676	11.7%	5625.031	7.5%	196081.067	12.7%	244362.667	13.1%
96.9600	ESI-	-	6.61	0.18%	8,332,453	16.4%	8,959	8.8%	387,847	19.5%	317,509	34.5%
103.0395	ESI-	2-Hydroxyisobutyric acid	4.10	0.01	373351.77	13.1%	1,146	17.4%	97,855	12.2%	50,868	28.7%
104.9923	ESI+	-	5.63	0.16%	5,962,086	8.9%	-	-	-	-	-	-
112.0507	ESI+	-	2.31	1.31%	1,931,229	17.4%	3,370	19.0%	-	-	-	-
112.0515	ESI-	-	3.21	0.89%	36,022,537	12.1%	3,370	19.0%	-	-	190,516	26.2%
114.0664	ESI+	Creatinine	3.16	1.13%	219,528,745	15.0%	114,433	11.6%	12,955,540	24.3%	13,768,035	15.6%
115.0697	ESI+	-	3.16	1.09%	9,432,436	15.5%	4,400	10.7%	508,848	19.9%	577,182	14.1%
118.0864	ESI+	-	6.06	0.00	1983323.97	15.1%	3,932	14.0%	232,031	14.7%	297,976	23.6%
120.0661	ESI+	L-threonine	6.55	0.55%	83,885	15.9%	386	17.9%	19,117	18.7%	16,380	17.4%
130.1089	ESI+	-	5.32	0.20%	4,951,975	16.3%	22,186	5.7%	2,458,381	12.1%	3,484,016	19.6%
132.0769	ESI+	Creatine	6.59	0.15%	2,183,826	14.3%	1,326	10.9%	76,982	13.8%	83,701	18.9%
137.0459	ESI+	Hypoxanthine	1.95	0.70%	4,615,790	16.3%	213	15.3%	7,388	22.8%	9099.625	26.8%
141.0167	ESI-	-	5.63	0.26%	24,499,130	6.6%	-	-	-	-	-	-
142.0864	ESI+	-	3.97	0.00	2089146.02	17.8%	204	14.8%	10,000	16.8%	11,369	23.8%
143.1180	ESI+	-	8.68	0.23%	2,077,661	14.6%	2,085	17.2%	51,544	7.8%	72,523	18.9%
144.1019	ESI+	-	6.18	0.31%	4,946,851	17.0%	13,420	16.5%	631819.200	24.9%	769863.467	29.5%
152.0706	ESI+	-	4.40	0.37%	2,231,276	12.3%	392	20.3%	-	-	-	-
153.0660	ESI+	-	1.38	1.24%	5,054,141	14.8%	290.705	12.9%	9039.700	10.8%	9573.000	
160.1333	ESI+	-	8.89	0.18%	2,751,757	15.5%	1666.263	16.2%	96,333	9.3%	111,769	16.0%
162.1125	ESI+	-	7.68	0.00	3323420.84	18.3%	1,919	10.9%	186,749	22.2%	239070.667	22.4%
166.0725	ESI+	-	2.78	0.83%	5,790,744	17.5%	457.979	18.7%	28281.133	16.4%	28538.333	18.3%
166.0868	ESI+	Phenylalanine	5.25	0.27%	112,671	17.3%	307	17.8%	-	-	-	-
170.0925	ESI+	3-Methyl-L-Histidine	7.77	0.50%	3,781,756	18.2%	4,297	16.1%	127,254	13.4%	132,873	18.2%
172.9914	ESI-	-	0.61	0.13%	8,521,617	13.9%	11,081	17.4%	209,592	7.9%	226,687	11.6%
174.0406	ESI-	-	6.89	0.00	206454.27	17.7%	293	13.6%	13,131	15.2%	15,393	13.5%
178.0509	ESI-	Hippuric acid	4.00	0.51%	21,442,240	17.5%	46,475	14.6%	1,885,018	16.8%	2,003,627	19.7%
180.0660	ESI-	Tyrosine	5.50	0.00	72516.90	17.6%	2791.639	13.9%	6,901	11.1%	6541.667	11.1%
186.9954	ESI+	-	5.63	0.18%	3,938,743	10.9%	-	-	-	-	-	-

			LC-MS (BEH-HILIC)				FIE-MS		Chip-based infusion		LESA-MS	
m/z	Polarity	Compounds present in artificial urine	RT(min)	RSD(%)	Peak Area	RSD(%)	Counts	RSD(%)	Counts	RSD(%)	Counts	RSD(%)
187.0070	ESI-	-	0.61	0.13%	50,119,797	15.6%	61,605	20.2%	991,281	10.5%	994,878	17.1%
190.0504	ESI-	5-Hydroxyindole-3-acetic acid	2.95	0.97%	127,512	17.8%	-	-	-	-	-	-
193.0613	ESI-	4-Aminohippuric acid	3.89	0.41%	92,201	18.0%	-	-	5,224	16.7%	2,548	15.1%
194.9273	ESI-	-	6.61	0.00	3590258.40	11.3%	229.968	14.5%	-	-	-	-
197.0450	ESI-	4-Hydroxy-3-methoxymandelic acid	2.55	0.01	294092.52	18.3%	167	12.4%	7682.786	11.0%	7915.200	20.2%
203.0821	ESI-	L-Tryptophan	5.17	0.22%	97,002	13.6%	-	-	-	-	-	-
203.1504	ESI+	-	9.16	0.11%	6,092,154	17.2%	1,392	13.4%	41,043	8.8%	47,097	14.4%
212.0022	ESI-	-	0.63	0.14%	15,997,383	14.6%	3,461	20.2%	102099.923	14.1%	111675.733	22.8%
223.0199	ESI-	-	5.62	0.00	3172090.29	7.4%	-	-	-	-	-	-
225.0627	ESI-	-	1.75	1.07%	3,505,031	13.2%	5,672	16.8%	85,136	24.2%	79485.667	32.2%
226.0828	ESI+	Cytidine	2.27	1.03%	596,664	17.3%	-	-	-	-	-	-
229.1549	ESI+	-	7.50	0.01	6015114.50	16.7%	5,802	19.2%	230,568	24.5%	282,101	24.5%
232.0275	ESI+	-	0.67	0.16%	4,322,180	10.3%	-	-	-	-	-	-
243.0621	ESI-	Uridine	1.65	0.81%	5,991,699	12.9%	431	17.0%	13,039	19.9%	13,710	22.7%
245.0123	ESI-	-	1.45	0.97%	1,177,384	15.7%	239	15.4%	8,574	11.3%	8,267	10.4%
254.9815	ESI-	-	0.74	1.14%	1,069,467	18.3%	703	18.5%	13,358	24.2%	5,920	30.3%
263.1037	ESI-	-	4.77	0.58%	7,618,696	17.5%	7,965	13.0%	334,091	16.3%	335,456	23.1%
265.1182	ESI+	-	4.77	0.65%	3,658,075	15.1%	1,303	17.1%	23629.933	9.7%	11359.800	19.1%
268.9984	ESI+	-	5.63	0.18%	6,826,822	12.5%	-	-	-	-	-	-
305.0335	ESI-	-	1.46	1.21%	1,397,431	15.5%	-	-	4,589	17.4%	5,192	16.6%
357.1093	ESI-	-	4.01	0.39%	1,399,247	18.1%	216	13.1%	2,800	17.7%	4,000	34.3%
515.0079	ESI+	-	5.63	0.18%	1,287,547	12.0%	-	-	-	-	-	-

RSD% of retention time (RT), Peak area/count of the selected peaks were calculated from the analysis of the pooled QC urine sample (n = 15).

Appendix F Validation of LC-MS and FIE-MS analyses for metabolomics profiling of malaria using selected urine metabolites peaks from pooled QC samples

			LC-MS (BEH-HILIC)				FIE-MS	
m/z	Polarity	Compounds present in artificial urine	RT(min)	RSD(%)	Peak area	RSD(%)	Counts	RSD(%)
61.0399	ESI+	Urea	1.45	0.69%	47,187,222	3%	654,684	4%
76.0758	ESI+	-	7.79	0.25%	24,251,989	5%	498,887	7%
88.0399	ESI-	L-alanine	6.19	0.21%	541,218	8%	1,676	11%
89.0268	ESI-	Lactic acid	2.16	0.66%	32,393,711	3%	-	-
94.0653	ESI+	-	5.93	0.41%	18,454,689	5%	77,197	10%
96.9600	ESI-	-	6.62	0.29%	52,033,613	3%	184,566	12%
103.0395	ESI-	2-Hydroxyisobutyric acid	4.24	0.33%	7,063,654	3%	6,200	21%
104.9923	ESI+	-	5.64	0.38%	33,746,354	3%	-	-
105.0336	ESI+	-	4.01	0.40%	2,458,530	8%	6,204	8%
112.0507	ESI+	-	2.29	0.41%	22,255,228	3%	2,390	15%
112.0515	ESI-	-	3.18	0.61%	169,499,037	2%	1,860	18%
114.0664	ESI+	Creatinine	3.12	0.42%	910,770,147	3%	1,247,896	4%
115.0697	ESI+	-	3.12	0.39%	39,539,573	2%	53,256	5%
118.0864	ESI+	-	6.06	0.34%	5,356,289	7%	32,226	6%
120.0661	ESI+	L-Threonine	6.53	0.24%	627,755	12%	4,326	14%
130.1089	ESI+	-	5.36	0.44%	156,338	8%	-	-
132.0769	ESI+	Creatine	6.60	0.24%	9,889,928	3%	9,430	6%
137.0459	ESI+	Hypoxanthine	1.95	0.51%	23,083,594	5%	1,496	20%
141.0167	ESI-	-	5.64	0.34%	144,825,883	1%	-	-
142.0864	ESI+	-	3.97	0.52%	12,521,078	5%	1,961	20%
143.1180	ESI+	-	8.72	0.18%	11,237,483	4%	14,364	9%
144.1019	ESI+	-	6.20	0.30%	6,106,461	4%	28,465	7%
152.0706	ESI+	-	4.40	0.43%	764,484	4%	-	-
153.0660	ESI+	-	1.41	0.66%	7,554,338	4%	-	-
160.1333	ESI+	-	8.89	0.13%	3,975,983	6%	1,980	23%
162.1125	ESI+	-	7.69	0.20%	19,556,764	5%	34,971	5%
166.0725	ESI+	-	2.70	0.59%	35,449,941	3%	3,948	13%
166.0868	ESI+	Phenylalanine	5.31	0.27%	1,087,605	3%	5,867	13%
167.0205	ESI-	-	1.45	0.64%	656,995	6%	64,270	4%
172.9914	ESI-	-	0.64	0.56%	69,253,890	3%	24,856	8%
174.0406	ESI-	-	6.88	0.14%	1,592,452	8%	1,421	23%
178.0509	ESI-	Hippuric acid	4.00	0.46%	102,229,693	1%	266,231	9%
180.0660	ESI-	Tyrosine	5.52	0.28%	293,726	12%	28,774	9%
180.0881	ESI+	-	2.23	0.75%	17,198,686	2%	-	-
181.0501	ESI-	Homovanillic acid	3.15	0.46%	10,650,934	2%	5,884	7%
186.9954	ESI+	-	5.64	0.34%	26,186,276	4%	-	-
187.0070	ESI-	-	0.64	0.56%	319,751,678	3%	95,750	10%
190.0504	ESI-	5-Hydroxyindole-3-acetic acid	2.80	1.37%	449,565	10%	-	-
193.0613	ESI-	4-Aminohippuric acid	3.88	0.40%	463,402	3%	-	-
194.9273	ESI-	-	6.62	0.29%	26,678,056	2%	4,087	19%
195.0522	ESI-	-	3.13	0.81%	14,782,718	6%	12,156	7%
197.0450	ESI-	4-Hydroxy-3-methoxymandelic acid	2.51	0.69%	1,651,895	5%	1,167	21%
203.0821	ESI-	L-Tryptophan	5.24	0.24%	585,848	3%	-	-

		Compounds present in	LC-MS (BEH-HILIC)				FIE-MS	
m/z	Polarity	artificial urine	RT(min)	RSD(%)	Peak area	RSD(%)	Counts	RSD(%)
203.1504	ESI+	-	9.20	0.17%	35,846,006	6%	9,836	8%
204.1231	ESI+	-	7.35	0.25%	15,814,618	4%	36,489	7%
212.0022	ESI-	-	0.66	1.14%	83,944,045	4%	6,474	10%
216.9810	ESI-	-	1.35	0.84%	14,492,929	6%	-	-
223.0199	ESI-	-	5.64	0.32%	27,260,765	2%	-	-
225.0627	ESI-	-	1.72	0.80%	3,075,832	9%	3,744	10%
226.0828	ESI+	Cytidine	2.24	0.52%	15,315,227	5%	-	-
232.0275	ESI+	-	0.69	0.52%	2,255,168	3%	-	-
243.0621	ESI+	Uridine	1.64	0.42%	45,620,277	5%	5,141	9%
254.9815	ESI-	-	0.74	1.34%	2,754,107	5%	1,248	6%
263.0225	ESI-	MHPG sulphate	0.74	0.52%	3,524,896	3%	-	-
263.1037	ESI-	-	4.76	0.30%	31,252,508	1%	57,466	5%
265.1182	ESI-	-	4.75	0.29%	21,164,752	5%	12,045	12%
268.9984	ESI-	-	5.64	0.39%	47,467,723	2%	-	-
357.1093	ESI+	-	4.01	0.46%	8,660,359	2%	-	-
391.1119	ESI+	-	3.13	0.76%	2,101,625	10%	-	-
515.0079	ESI-	-	5.64	0.39%	9,504,267	2%	-	-

RSD% of retention time (RT), Peak area/count of the selected peaks were calculated from the analysis of the pooled QC urine sample (n = 10).

Appendix G Validation of LC-MS and FIE-MS analyses for FOG dietary intervention using selected urine metabolites peaks from pooled QCs.

			Spot urine analysis						24 h urine analysis					
			LC-MS (BEH-HILIC)				FIE-MS		LC-MS (BEH-HILIC)				FIE-MS	
m/z	Polarity	Compounds present in artificial urine	RT(min)	RSD(%)	Peak Area	RSD(%)	Counts	RSD(%)	RT(min)	RSD(%)	Peak Area	RSD(%)	Counts	RSD(%)
61.0399	ESI+	Urea	1.46	0.32%	78,487,199	8%	4,048,442	7%	1.44	0.81%	22,498,379	8%	5,896,627	4%
76.0758	ESI+		7.78	0.06%	40,620,025	3%	2,380,957	3%	7.72	0.36%	10,375,509	12%	1,261,813	4%
88.0399	ESI-	L-alanine	6.65	0.09%	1,406,798	8%	2,573	16%	6.62	0.13%	632,362	13%	3,018	10%
89.0242	ESI-	Lactic acid	2.33	0.25%	4,454,704	7%	-	-	2.15	0.27%	686,491	8%	-	-
94.0653	ESI+		5.92	0.13%	13,743,248	5%	66,432	2%	5.87	0.41%	5,844,364	6%	64,846	6%
96.9600	ESI-		6.70	0.09%	52,861,048	7%	184,042	18%	6.64	0.34%	17,379,699	9%	141,815	20%
103.0395	ESI-	2-Hydroxyisobutyric acid	3.26	0.26%	10,921,025	9%	13,220	6%	3.15	0.42%	3,572,236	7%	3,720	14%
104.9923	ESI+		5.68	0.07%	25,962,199	5%	-	-	5.61	0.19%	12,249,793	4%	-	-
105.0336	ESI+		4.06	0.24%	9,934,866	14%	18,327	7%	3.97	0.52%	1,185,691	6%	8,148	10%
112.0507	ESI+		2.46	0.49%	15,051,628	9%	4,573	13%	2.34	0.52%	1,210,474	9%	3,186	11%
112.0515	ESI-		3.20	0.31%	269,537,096	3%	4,573	13%	3.08	0.54%	72,215,808	4%	3,161	10%
114.0664	ESI+	Creatinine	3.15	0.55%	1,539,222,543	6%	4,844,998	5%	3.05	0.64%	377,792,257	11%	1,556,672	6%
115.0697	ESI+		3.15	0.55%	65,816,944	6%	206,649	5%	3.05	0.63%	16,218,782	11%	66,407	4%
118.0864	ESI+		6.08	0.14%	1,670,379	4%	55,496	10%	5.97	0.50%	1,172,480	7%	45,861	6%
120.0661	ESI+	L-threonine	-	-	-	-	10,275	16%	6.03	0.35%	28,096	8%	8,393	8%
132.0769	ESI+	Creatine	6.65	0.09%	11,026,123	8%	31,173	10%	6.55	0.32%	4,924,340	14%	52,347	7%
137.0459	ESI+	Hypoxanthine	1.98	0.26%	57,963,049	6%	6,540	11%	1.88	0.50%	4,231,559	10%	-	-
142.0864	ESI+		4.14	0.20%	9,374,639	20%	6,068	13%	4.13	0.47%	2,030,220	13%	7,919	21%
143.1180	ESI+		8.93	0.19%	6,144,224	5%	26,980	12%	8.49	0.31%	256,005	8%	13,791	14%
144.1019	ESI+		6.19	0.13%	10,539,684	14%	246,980	7%	6.12	0.39%	7,843,285	10%	424,716	5%
152.0706	ESI+		4.47	0.13%	3,667,164	10%	1,782	20%	4.32	0.67%	552,228	8%	2,272	21%
153.0660	ESI+		1.40	0.34%	34,294,375	4%	6,898	20%	1.40	0.72%	6,729,831	8%	7,467	12%
160.1333	ESI+		8.91	0.07%	16,266,337	6%	15,952	9%	8.76	0.22%	4,407,289	14%	13,224	10%
162.1125	ESI+		7.69	0.06%	7,211,681	8%	50,831	3%	7.59	0.19%	2,034,656	10%	36,335	9%
166.0725	ESI+		2.78	0.41%	51,009,736	7%	14,872	11%	2.65	0.32%	4,832,693	4%	5,064	15%
166.0868	ESI+	Phenylalanine	5.47	0.23%	331,421	6%	7,749	16%	5.46	0.48%	46,785	11%	5,652	13%
167.0205	ESI-	Uric acid	1.49	0.51%	693,238	5%	172,431	2%	1.44	0.80%	66,913	12%	131,502	3%
172.9914	ESI-		0.64	1.36%	93,513,639	6%	34,072	6%	0.63	0.17%	13,914,906	11%	11,148	10%
174.0406	ESI-		6.98	0.09%	4,758,727	10%	4,289	9%	6.82	0.35%	139,226	7%	3,825	11%

			Spot urine analysis						24 h urine analysis					
			LC-MS (BEH-HILIC)				FIE-MS		LC-MS (BEH-HILIC)				FIE-MS	
m/z	Polarity	Compounds present in artificial urine	RT(min)	RSD(%)	Peak Area	RSD(%)	Counts	RSD(%)	RT(min)	RSD(%)	Peak Area	RSD(%)	Counts	RSD(%)
178.0509	ESI-	Hippuric acid	4.06	0.20%	147,690,552	6%	607,252	5%	3.96	0.68%	22,086,726	4%	402,930	3%
180.0881	ESI+		2.05	0.94%	3,912,320	7%	-	-	2.39	0.57%	1,142,146	8%	-	-
181.0501	ESI-	Homovanillic acid	3.15	0.37%	19,318,117	9%	11,150	9%	3.08	0.43%	2,700,425	6%	7,454	9%
186.9954	ESI+		5.68	0.07%	17,475,857	3%	-	-	5.60	0.39%	8,401,089	6%	-	-
187.0070	ESI-		0.63	1.24%	304,339,281	6%	36,311	13%	0.63	0.17%	55,380,796	11%	17,868	17%
190.0504	ESI-	5-Hydroxyindole-3-acetic acid	3.02	0.30%	638,604	7%	2,524	22%	-	-	-	-	1,788	21%
193.0613	ESI-	4-Aminohippuric acid	3.96	0.21%	388,245	9%	1,111	3%	3.85	0.60%	85,841	6%	1,154	11%
194.9273	ESI-		6.72	0.08%	30,544,804	6%	4,561	19%	6.62	0.32%	8,904,405	9%	3,435	21%
195.0522	ESI-		3.21	0.13%	40,897,158	12%	39,373	4%	3.01	0.39%	1,811,722	5%	17,220	7%
197.0450	ESI-	4-Hydroxy-3-methoxymandelic acid	2.80	0.69%	1,102,000	12%	2,044	12%	2.69	0.35%	154,465	14%	2,070	14%
203.0821	ESI-	L-Tryptophan	5.39	0.14%	284,990	8%	-	-	5.37	0.41%	13,882	8%	-	-
204.1231	ESI+		7.36	0.08%	19,761,392	6%	38,009	4%	7.27	0.32%	3,302,924	12%	29,949	6%
212.0022	ESI-		0.65	0.59%	118,900,289	6%	6,788	21%	0.66	0.11%	15,723,519	12%	1,803	22%
216.9810	ESI-		1.36	0.84%	25,570,263	4%	1,276	12%	1.32	0.78%	3,752,584	4%	-	-
223.0199	ESI-		5.68	0.10%	19,345,106	4%	-	-	5.59	0.47%	11,190,849	3%	-	-
225.0627	ESI-		1.76	0.40%	17,185,846	3%	30,522	6%	1.65	0.38%	1,903,889	8%	8,229	3%
226.0828	ESI+	Cytidine	2.40	0.82%	5,335,680	8%	-	-	2.28	0.57%	245,016	15%		
229.1549	ESI+		7.38	0.13%	40,332,389	7%	105,095	5%	7.35	0.31%	3,654,667	10%	50,233	11%
232.0275	ESI+		0.70	0.48%	8,137,240	5%	-	-	0.69	0.18%	2,625,077	5%	-	-
243.0621	ESI-	Uridine	1.71	0.34%	50,806,572	4%	8,782	9%	1.64	0.46%	9,393,648	9%	6,081	6%
245.0123	ESI-		1.48	0.60%	17,916,898	5%	-	-	1.32	0.97%	1,529,179	7%	1,064	3%
254.9815	ESI-		0.74	1.15%	6,033,463	9%	3,928	16%	0.74	0.66%	1,085,442	13%	2,722	7%
263.0225	ESI-	MHPG sulphate	0.75	0.98%	2,695,715	5%	-	-	0.75	0.65%	455,316	14%	-	-
263.1037	ESI-		4.85	0.10%	28,379,398	5%	73,575	5%	4.77	0.47%	3,917,168	4%	61,834	3%
265.1182	ESI+		4.85	0.42%	20,616,618	5%	25,173	11%	4.76	0.66%	4,003,646	11%	22,436	7%
268.9984	ESI+		5.68	0.07%	29,194,542	3%	-	-	5.61	0.30%	14,317,585	6%	-	-
305.0335	ESI-		1.50	0.56%	15,937,110	6%	-	-	1.33	0.87%	287,322	14%	-	-
357.1093	ESI-		4.06	0.20%	13,975,776	4%	1,338	6%	3.98	0.78%	986,373	8%	-	-
515.0079	ESI+		5.68	0.07%	6,257,022	3%	-	-	5.59	0.49%	3,654,763	3%	-	-

RSD% of retention time (RT), Peak area/count of the selected peaks were calculated from the analysis of the pooled QC urine sample interspaced throughout the run (n = 9).