2009

Metabolomic investigation of a new rat model of autosomal recessive polycystic kidney disease

Hayley White
Edith Cowan University

Edith Cowan University

Copyright Warning

EDITH COWAN UNIVERSITY

School of Natural Sciences

Metabolomic Investigation of a New Rat Model of Autosomal Recessive Polycystic Kidney Disease

By

Hayley White BSc.

This thesis is presented for the degree Bachelor of Science (Applied and Analytical Chemistry) Honours from the School of Natural Sciences; Faculty of Computing Health and Science; Edith Cowan University, Western Australia

Principal Supervisor: Dr Mary Boyce (Edith Cowan University)

Associate Supervisor: Associate Professor Robert Trengove (Murdoch University)

Date of Submission: October 30th, 2009

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

i. incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;

ii. contain any material previously published or written by another person except where due reference is made in the text; or

iii. contain any defamatory material

I also grant permission for the Library at Edith Cowan University to make duplicate copies of this thesis as required.

Signature: ___

Date: 09.12.09

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my deepest gratitude to a vast number of people who made this project possible. First and foremost to Dr Mary Boyce, thank you for agreeing to supervise me whilst embarking on my first research experience. Your support, particularly in the development of this thesis, is greatly appreciated. To Associate Professor Robert Trengove, thank you for steering this research and opening my eyes to a wealth of innovative projects in metabolomics. To Professor Jacqueline Phillips, Dr Garth Maker, Mr Joel Gummer, Dr Joanne Harrison, Mr Mark Bannister and Dr Matthew Sharman; for assisting me with the logistics of the project, in particular, Garth and Joel who were always helping me in the laboratory, with data analysis and the thesis. In fact, the entire metabolomics group at Murdoch deserve mention here. I have felt very welcome this year, thanks to you all. To Professor Rudolf Appels, Professor Richard Oliver and Dr Peter Solomon thank you for reviewing my work particularly Richard for reviewing both proposal and thesis stages. To the various institutions involved in this project, School of Natural Sciences; Edith Cowan University, Research and Development; Murdoch University, Animal Resources Centre; Western Australia, thank you for the funding and infrastructure that has made this project possible. To my family and friends; Mum, Dad, Kevin and Scott, thank you for encouraging me, for your understanding and for your support this year. To Chris; I have much to thank you for. For inspiring me, motivating me, making me laugh, but most of all for understanding and helping me make it through this stressful year.

ABSTRACT

Metabolomics is complementary to genomics, transcriptomics and proteomics; however, it has the capacity to reflect the activities of the organism at a functional level. Metabolomics can therefore be used as a diagnostic tool by identifying the up- or down-regulation of metabolites in the cell, tissue, plasma, serum or urine. Specifically, these are, but are not limited to, identifying biomarkers of disease, monitoring drug treatments, and monitoring surgical procedures such as organ transplant. Autosomal recessive polycystic kidney disease (ARPKD) makes up 5-8% of patients requiring kidney dialysis and/or transplantation and of these, an estimated 50% of patients progress to end-stage renal disease (ESRD) by the age of 10 years, resulting in renal and liver-related morbidity and mortality. The purpose of this research was to utilise the Lewis Polycystic Kidney (LPK) rat to investigate the ARPKD phenotype using metabolomics. Spot urine samples were collected from 7 male Lewis; 8 male LPK; and 6 female LPK rats aged 5 weeks. Metabolites were extracted from urine and derivatised for GC/MS analysis. The peak area of target components was normalised to the internal standard ribitol and then to creatinine. Principal component analysis (PCA) was used to visualise sample grouping and the loadings plot of the PCA was used to determine key metabolites attributed to the variance between sample groups. α-ketoglutarate, uric acid and allantoin were proposed as potential biomarkers for ARPKD in the 5-week old male LPK rat. The findings of this study, particularly the development of a GC/MS method to analyse Lewis and LPK rat urine, demonstrate the potential of metabolomics to further investigate ARPKD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure 10. Principal component analysis of male 5-week old Lewis and LPK rat urine samples. This output was generated using log transformed peak area data normalised to ribitol then creatinine.

Figure 11. PCA X-loadings plot of male 5-week old Lewis and LPK rat urine samples. This output was generated using log transformed peak area data normalised to ribitol then creatinine and shows the key metabolites which influence the model described in Figure 10.

CHAPTER ONE

INTRODUCTION

1.1 Metabolomics

The relatively new field of metabolomics is defined as the study of the metabolome (i.e. all metabolites in the cell, tissue and/or biofluid(s)) [1-5]. Metabolomics is complementary to genomics, transcriptomics and proteomics, however, it has the capacity to reflect the activities of the organism at a functional level [6, 7]. Metabolites are products of enzyme-catalysed reactions [8] and, although it is known that changes in the quantities of enzymes may have little effect on metabolic processes, significant effects can be seen in the concentration of individual metabolites [6]. As this suggests, metabolites are integral to the function and survival of the cell [3], most commonly participating in the production and release of energy or as a precursor to such transfer processes [3]. Metabolomics can therefore be used as a diagnostic tool by identifying the up- or down-regulation of metabolites in the cell, tissue, plasma, serum or urine [9]. Specifically, these are, but are not limited to, identifying biomarkers of disease [9], monitoring drug treatments [9, 10] and monitoring surgical procedures such as organ transplant [4, 5].

Also termed metabonomics or metabolic profiling [4], the advent of metabolomics has been in light of recent technological advancements in the separation and identification of small, low molecular weight molecules [4, 11], and also in software for processing the resultant large quantities of data [5]. As the properties of metabolites such as polarity, volatility and solubility vary so greatly, metabolomic analysis requires the use of several complimentary analytical techniques [2]. The most

common approaches in metabolomics have been capillary electrophoresis [12, 13], gas chromatography mass spectrometry (GC/MS) [2, 8, 14, 15], liquid chromatography mass spectrometry (LC/MS) [2, 14, 16] and proton nuclear magnetic resonance spectroscopy (^{1}H NMR) [1, 14, 16]. GC/MS is the method of choice for the screening of non-targeted metabolites due to its ability to resolve hundreds of metabolites in a single injection [2]. Coupled with the sensitivity and ability of MS to provide structural information, the detection, separation and identification of metabolites is now possible [2].

Gas chromatography has been widely used in metabolomics studies [3, 6, 8-10, 15, 17-20]. GC consists of an injection system, mobile phase, the column (stationary phase) and a detector [3, 21]. A sample (gas or liquid) is injected through a septum into an injection port and is rapidly volatilised [21]. Analytes are swept onto the column and are separated according to volatility and diffusion between the mobile and stationary phases [21]. Separated analytes then flow through to the detector which records a response [21].

The mobile phase in gas chromatography is known as the carrier gas, and is typically helium, nitrogen, or hydrogen [21]. The carrier gas used is largely dependent upon the detector employed however, where the detector is not an issue, significant differences can be found in separation efficiency and the speed of GC analysis by selecting the appropriate carrier gas [21]. For example, hydrogen will yield the fastest separation time followed by helium and then nitrogen [21]. In metabolomics, however, mass spectrometers are generally used as the method of detection in gas chromatography [3]. For this reason, helium is employed as the carrier gas. Helium and

hydrogen show superior resolution as analytes diffuse between phases more rapidly through these carrier gases than through nitrogen [21]. Hydrogen may protonate analytes and can potentially interfere with the MS vacuum pump [21]. For these reasons, helium is the obvious choice of carrier gas in metabolomics separations using GC/MS.

The method of injection in gas chromatography is critical as it is at this point that the sample is transferred to the mobile phase and focused at the beginning of the column [3]. Villas-Boas *et al.* [3] have identified this as one of the major causes of error in gas chromatography, being particularly attributed to slow and/or incomplete volatilisation of analytes. To prevent this, metabolites must be converted to a volatile form. This process is known as derivatisation and is a considerable limitation of GC/MS due to its arduous and time consuming nature. Splitless injections are widely used in non-targeted metabolic profiling [9, 15, 17, 19] as these are typically classed as trace analyses. A split injection is commonly used to achieve baseline separation when the concentration of a sample is known to be high or is unknown. It is a preventative measure to minimise overloading the column and this is especially the case when using capillary columns as micrograms of analyte can cause overloading [21].

There is a wealth of column technology available to GC practitioners. Essentially, there are two types of GC columns; packed and open. Open tubular columns are the most commonly used in GC analyses. They are long and narrow in structure and are made of polyimide-coated fused silica. Open tubular columns offer greater sensitivity, higher resolution, and shorter analysis time [21]. These features make the application of open tubular columns ideal for research in the field of

metabolomics. The common coupling of GC to MS requires Low bleed columns. These are typically non-polar capillary columns which amplify the need for derivatisation (i.e. to reduce the polarity of metabolites) for successful high resolution chromatography.

Mass spectrometry is commonly employed as the method of detection in metabolomics research [7, 11, 22]. In fact, Villas-Boas *et al.* [3] have identified mass spectrometry as the driving force behind the recent emergence of metabolomics. To briefly describe MS, molecules are ionised by either electron or chemical ionisation processes, are accelerated by an electric field and then separated according to their mass-to-charge (*m/z*) ratio [21]. There are many different types of mass spectrometers however, for the purpose of this research, only two will be discussed here. These are the linear quadrupole mass spectrometer and the ion trap quadrupole mass spectrometer.

In linear quadrupole mass spectrometry, four rods are assembled parallel to one another. A constant voltage and a radio frequency oscillating voltage are applied to the rods to stabilise the ion trajectory for a given *m/z* species as it migrates to the detector. This can, however, result in the loss of sensitivity for destabilised ions. In ion trap quadrupole mass spectrometry, ions are trapped by a ring electrode and two end caps. Ion trajectories are destabilised in order to reach the detector in a uniform stream which results in superior sensitivity in comparison to the linear quadrupole mass spectrometer.

1.2 Polycystic Kidney Disease

Polycystic kidney diseases (PKD) are inherited diseases characterised by renal tubular defects [23]. Advances in molecular genetics have allowed the characterisation of the three PKD genes; PKD1, PKD2 and PKHD1 [24]. By combining these findings with the descriptive data obtained from previous mapping of genetic cystic diseases, the molecular and cellular mechanisms of the diseases can now be systematically and thoroughly investigated [24]. Two modes of inheritance have been characterised for the polycystic kidney diseases. The first and most common is autosomal dominant polycystic kidney disease (ADPKD) which is caused by mutations in the PKD1 (polycystic kidney disease 1) or PKD2 (polycystic kidney disease 2) genes [25] and results in focal cyst development [23]. The second is autosomal recessive polycystic kidney disease (ARPKD) which is caused by mutations in the PKHD1 (polycystic kidney and hepatic disease 1) gene [25], commonly resulting in tubule dilation [23, 26, 27].

Autosomal recessive polycystic kidney disease (ARPKD), far less common than ADPKD, has an incidence of 1 in 20 000 [24, 28] which makes up 5-8% of patients requiring kidney dialysis and/or transplantation [24]. Unlike ADPKD, ARPKD is most commonly evident *in utero* or at birth but has also been characterised in later childhood and in patients over 20 years of age [24, 28]. An estimated 50% of patients progress to end-stage renal disease (ESRD) by the age of 10 years, resulting in renal and liver-related morbidity and mortality despite recent advancements in neonatal and intensive care [24].

Symptoms of ARPKD include fusiform dilation of renal collecting tubules or ducts and/or focal cyst development [23] as well as dysgenesis of the hepatic portal triad (congenital hepatic fibrosis) [24]. As this suggests, ARPKD is a member of a group of congenital hepatorenal fibrocystic diseases [29] characterised by both renal and hepatic involvement [24]. Dilation of the collecting tubules can be evident in 10-90% of tubules, accounting for the vast variability in renal dysfunction and kidney enlargement [24]. The manifestation of liver disease consisting of biliary dysgenesis and congenital hepatic fibrosis varies from patient to patient according to age at presentation, however liver involvement is a symptom of all ARPKD patients [24]. The most common cause of morbidity and mortality in new born patients is pulmonary hypoplasia, often resulting in respiratory insufficiency [24]. However, for patients who survive the neonatal and perinatal periods, ARPKD can present with systemic hypertension, renal failure, portal hypertension, and fibrosis of both the kidneys and liver also resulting in morbidity and mortality [24].

The different phenotypes of the disease were originally thought to be caused by different mutant genes [24]. Since then, all phenotypes of ARPKD have been attributed to a mutation at a single locus; the PKHD1 (polycystic kidney and hepatic disease 1) gene, which is located on chromosome 6p21.1–p12 [24, 30]. PKHD1 is a large gene, spanning approximately 470 kb of genomic DNA and producing cDNA of 9–16 kb [24, 28]. A minimum of 86 exons is assembled into a variety of alternatively spliced transcripts [24, 28]. The longest continuous open reading frame (ORF) is predicted to yield a 4074-amino acid protein called fibrocystin or polyductin which is believed to be involved in collecting tubule and biliary duct formation, however the

exact function is unknown [24, 31]. It is anticipated that multiple variations of the protein are produced by the complicated transcription of the PKHD1 gene [24].

Recent research in ARPKD has centred on determining the localisation and function of the PKHD1 protein fibrocystin/polyductin [23, 27, 28, 32, 33]; the molecular/cellular pathogenesis and cystogenesis of the disease [24, 28]; diagnostic applications [34]; and therapeutic interventions for patients suffering from ARPKD [24, 28]. This research has encompassed the use of many laboratory animal models of ARPKD including the $Pkhd1^{del2/del2}$ mice [27], the polycystic kidney (PCK) rat [33] and the lewis polycystic kidney (LPK) rat [35]. There is little evidence to show that metabolomics has been identified as a tool to investigate ARPKD. Therefore, the purpose of this research is to utilise the LPK rat model to investigate the ARPKD phenotype using a metabolomics approach.

1.3 Research Questions

1. Is it possible to accurately and reliably detect metabolites in Lew/LPK rat urine using GC/MS?
2. Can this method be optimised to see a greater profile of metabolites at lower detection limits?
3. Are there any significant differences between the metabolic profiles of Lewis and LPK rats?

1.4 Definitions of Selected Terms

ADPKD	Autosomal dominant polycystic kidney disease
AEC	Animal Ethics Committee
ANOVA	Analysis of variance
ARC	Animal Resources Centre
ARPKD	Autosomal recessive polycystic kidney disease
ESRD	End stage renal disease
eV	Electron volt
GC	Gas chromatography
GC/MS	Gas chromatography mass spectrometry
IS	Internal standard
LC	Liquid chromatography
LC/MS	Liquid chromatography mass spectrometry
LEW	Lewis
LPK	Lewis polycystic kidney
m/z	Mass-to-charge ratio
M	Moles per litre
mg	Milligram
mL	Millilitre
MeOH	Methanol
MS	Mass spectrometry
MSTFA	N-methyl-N-(trimethylsilyl)trifluroacetamide
ng	Nanogram
$ng.\mu L^{-1}$	Nanograms per microlitre
NIST	National Institute of Standards and Technology

^{1}H NMR	Proton nuclear magnetic resonance
P	Probability
PC1	Principal component 1
PC2	Principal component 2
PCA	Principal component analysis
PCK	Polycystic kidney
PKD	Polycystic kidney disease
PKD1	Polycystic kidney disease 1 gene
PKD2	Polycystic kidney disease 2 gene
PKHD1	Polycystic kidney and hepatic disease 1 gene
U	Unit
$U.\mu L^{-1}$	Units per microlitre
µg	microgram
µL	microlitre

CHAPTER TWO

METHODS

2.1 Chemicals

D-mannitol (>99%), ribitol (>99%), methoxyamine-HCl (98%), N-methyl-N-(trimethylsilyl)trifluroacetamide (MSTFA) (>98.5%), pyridine (>99%), *n*-alkanes C_{10}, C_{12}, C_{15}, C_{19}, C_{22}, C_{28}, C_{32} and C_{36} (>98%), type III jack bean urease (40100 $U.g^{-1}$) and creatinine (>98%) were all purchased from Sigma Aldrich; Australia. *n*-hexane (>95%) and methanol (HPLC grade >99%) were purchased from LabScan; Australia. Helium cylinders (UHP) were purchased from BOC Gases; Australia.

2.2 Standards

Seven concentrations (0.01, 0.05, 0.1, 0.5, 1, 5 and 10 $ng.\mu L^{-1}$) of D-mannitol were prepared to calculate limit of detection (LOD) and precision of the GC/MS. A stock solution of D-mannitol (6.5 $mg.L^{-1}$) was prepared from which 0.1, 0.5, 1, 5, 10 and 50 mL aliquots were taken and made up to 100 mL with milli-Q water (the stock was used to prepare the 10 $ng.\mu L^{-1}$ standard). The internal standard, ribitol, was added to each at a desired final concentration of 50 $ng.\mu L^{-1}$. Multiple 100 µL aliquots of each of the standards were dried in an Eppendorf Concentrator Plus vacuum concentrator and were either derivatised for GC/MS analysis or frozen at -80°C until further use. A similar approach was taken to add ribitol to experimental samples, however the desired final concentration was 15 $ng.\mu L^{-1}$. *n*-alkanes were used for Kovats Index calculation. Individual solutions of the each of the alkanes were made up to 1 mL in hexane at a concentration of 0.5 $mg.mL^{-1}$. An alkane mixture was prepared from the individual solutions at a concentration of 0.0625 $mg.mL^{-1}$ in 8 mL.

2.3 Laboratory Animals

Lewis and LPK rat urine samples were purchased from the Animal Resources Centre (ARC), Canning Vale, Western Australia. The animals were housed together in cages with steam cut wheaten chaff, at ambient temperature (20-25°C). The animals were fed a vegetarian diet of Standard Laboratory Rat and Mouse Cubes (Specialty Feeds, Glen Forrest; Western Australia) and had water available at all times. Phillips *et al.* [35] outlines the method by which the disease was transmitted to the laboratory animals and also the method by which the mode of inheritance (autosomal recessive) was characterised. Spot urine samples (50-500 µL) were collected from 3 male Lewis; 2 female Lewis; and 5 male LPK rats aged approximately 11 weeks. These samples were pooled and used for method development. For experimental work, urine was collected from 7 male Lewis; 8 male LPK; and 6 female LPK rats aged 5 weeks. Registration as an animal researcher was sought and approved by Murdoch University's Animal Ethics Committee (AEC) and ethical approval for this study was sought and approved by Edith Cowan University's AEC. Approval number: Project 2144 MARTINS; Investigating Beta-amyloid clearance *in vivo* 06-A9.

2.4 Metabolite Extraction

Urine from Lewis rats (control) was used to optimise the extraction procedure. Samples from 11-week old male and female rats were pooled (50 µL unless otherwise stated) and urease (1 unit per microlitre of urine unless otherwise stated) was added. The samples were incubated at 37°C (30 minutes unless otherwise stated). The samples were then centrifuged at 4000 rpm for 10 minutes. Following this, -40°C methanol containing the ribitol internal standard was added to the sample. The calculation for the concentration of ribitol in the extraction solvent was based on

achieving a final ribitol concentration of 15 ng.μL^{-1} or 15 ng 'on column'. The samples were then agitated at 1200 rpm for 10 minutes and following this were centrifuged at 4°C and 12000 rpm for 10 minutes. Aliquots (50 µL) of the sample were dried in an Eppendorf Concentrator Plus vacuum concentrator and frozen at -80°C until further use.

2.5 Metabolite Derivatisation

Extracted metabolites were treated with 20 µL of 20 mg.mL^{-1} methoxylamine-HCl in pyridine and agitated at 1200 rpm at a temperature of 30°C in an Eppendorf Thermomixer Comfort (NSW, Australia) for 90 minutes. Following this, 40 µL of N-methyl-N-(trimethylsilyl)trifluroacetamide (MSTFA) was added and the metabolites were incubated at 37°C for 30 minutes in the thermomixer. The metabolites were agitated for the first 1-2 minutes of this incubation period to mix the MSTFA. The 60 µL sample was then transferred to a 31 × 5 mm 0.1 mL clear GC vial insert inside an Alltech GC vial containing 5 µL of an alkane mix in hexane.

2.6 Instrumentation

2.6.1 Quadrupole GC/MS (Murdoch University)

An Agilent 6890 series Gas Chromatograph equipped with an Agilent 5973N mass selective detector was used. The column employed was a Varian Factor Four fused silica capillary column VF 5MS (30 m × ID = 0.25 mm × DF = 0.25 µm + 10 m EZ-Guard). The derivatised sample (1 µL) was injected splitless into the inlet at a temperature of 230°C. The initial column temperature was 70°C and the temperature ramp was set to 5.63°C.min^{-1} to a final temperature of 330°C. The carrier gas used was helium at a flow rate of 1 mL.min^{-1} (constant flow). The transfer line was set at 330°C and the ion source was set to 230°C. Ionisation was achieved with a 70 eV electron

beam and the mass spectrometer was set to scan ion masses in the range of 45 - 600 *m/z*.

2.6.2 Ion trap GC/MS (Murdoch University)

A Varian CP 3800 Gas Chromatograph (Australia) equipped with a Varian 1177 split/splitless capillary injector, Varian CP 8400 autosampler (Australia) and a Varian Saturn 2200 GC/MS/MS ion trap mass spectrometer (Australia) was used. The column employed was a Varian Factor Four fused silica capillary column VF 5MS (30 m × ID = 0.25 mm × DF = 0.25 µm + 10 m EZ-Guard). The derivatised sample (1 µL) was injected splitless into the inlet at a temperature of 230°C. The initial column temperature was 70°C and the temperature ramp was set to 5.63°C.min^{-1} to a final temperature of 330°C. The carrier gas used was helium at a flow rate of 1 mL.min^{-1} (constant flow). The transfer line was set at 330°C and the ion source was set to 230°C. Ionisation was achieved by a 70 eV electron beam and the mass spectrometer was set to scan ion masses in the range of 45 - 600 *m/z*.

2.6.3 Ion trap GC/MS (Edith Cowan University)

A Varian Star 3400 CX gas chromatography system (Australia) equipped with a Varian Saturn 2000 GC/MS/MS ion trap mass spectrometer (Australia) was used. The column employed was a Varian Factor Four fused silica capillary column VF 5MS (30 m × ID = 0.25 mm × DF = 0.25 µm + 10 m EZ-Guard). The derivatised sample (1 µL) was injected splitless into the inlet at a temperature of 230°C. The initial column temperature was 50°C and was ramped to 70°C in the first 2 minutes of the run time. Thereafter the temperature ramp was set to 5.60°C.min^{-1} for 45.35 minutes to a final temperature of 330°C. The carrier gas was helium at a flow rate of 1 mL.min^{-1} (constant pressure). The transfer line was set at 330°C and the ion source was set to

230°C. Ionisation was achieved by a 70 eV electron beam and the mass spectrometer was set to scan ion masses in the range of 45 - 600 *m/z*.

2.7 Data Analysis

GC chromatograms and mass spectra were viewed using AnalyzerPro v2.2.0.7 (SpectralWorks, UK). AnalyzerPro was used to identify all unique components (i.e. peaks) of the chromatogram and also to identify target components, which were stored in a user-generated library (Metabolomics Australia Node; Murdoch University). Target components were added to the library according to retention time, Kovats Index and deconvoluted mass spectrum (base peak). Analytes in the sample chromatograms were resultantly matched to these parameters. In addition to being compared with the user-generated library, spectra were compared with the National Institute of Standards and Technology (NIST) Mass Spectral library to aid in the identification of significant unknown metabolites. The identification generated by the NIST library was given a probability score (%) which was based on base peak and fragmentation pattern identification. AnalyzerPro was used to generate a Microsoft Office Excel spreadsheet of the metabolites and their peak area in each of the samples. This output was used to determine the precision and reproducibility of the results and was also used to run univariate statistical and multivariate analyses. The peak area of each metabolite was normalised to the internal standard, ribitol, and then to creatinine.

2.8 Statistical Analysis

Univariate statistical analyses were conducted using SPSS v17.0. Where there were two groups of data to compare, a paired samples t-test was used with a confidence interval of 0.95 and significance level of P=0.05. Where there were more than two groups of data to compare, a one-way analysis of variance (ANOVA) was used with Bonferroni *post hoc* multiple comparisons. A confidence interval of 0.95 and significance level of $P<0.05$ were adopted. Where it was appropriate to test for correlations in the data, particularly between the sums of the peak areas of the components in a sample to the peak area of the component used to normalise the data, a Pearson correlation was employed. In this instance a confidence interval of 0.99 and significance level of $P<0.01$ were used. Multivariate principal component analysis (PCA) was conducted using The Unscrambler® v9.8 software package (CAMO, Norway). The data was log transformed using the equation $x = \log(x + 1)$.

CHAPTER THREE

RESULTS AND DISCUSSION

3.1 Extraction Optimisation

Urine, containing the highest number of polar metabolites in comparison to other biofluids, offers a rich source of metabolic information and is commonly used for metabolomics research [9, 10, 15, 16, 25, 36-39]. Urine was of particular interest in this research as the focus was on kidney disease. There is currently no agreed guideline for sample preparation and analysis of urine for metabolomics [40]. For this reason, it was necessary to undertake an extraction optimisation study to suit the operating conditions of the various GC/MS systems. This involved optimising urease incubation (concentration and incubation time), the urine volume extracted, and centrifugation procedures. The results and discussion of these parameters are detailed below.

3.1.1 Urease incubation

Urea is present in high concentrations in urine and has the potential to obscure closely eluting analytes and overload the mass spectrometer [15]. This issue is commonly resolved by adding urease [9, 15, 18, 19] which catalyses the hydrolysis of urea ($CO(NH_2)_2$) to carbon dioxide and ammonia. Initially, urease was added to a 200 µL pooled urine sample (11-week Lewis male and female) at a concentration of 0.3 $U.\mu L^{-1}$ of urine and was incubated at 37°C for 10 minutes. The sample was then dried and derivatised with MSTFA as previously described in section *2.3*. The resulting chromatogram (Figure 1) shows an overloaded peak in the region of RT 16.7-19.8 minutes which was attributed (by retention time and mass spectral analysis) to urea. Zhang *et al.* [15] reported that 0.15 units of urease per microlitre of urine and an incubation period of 15 minutes was adequate to degrade urea from rat urine,

however the results here clearly showed the need for a greater concentration of urease and an extended incubation period for the Lewis/LPK rat urine. It can also be seen that the intensity of some of the other analytes was 1.2×10^8, suggesting overloading, which was attributed to the volume of urine used (200 µL). Therefore, the experiment was repeated with three modifications: 1. the units of urease added per microlitre of urine was increased to 1; 2. the incubation period was increased to 30 minutes; and 3. the volume of urine treated with urease was decreased to 50 µL. The resulting chromatogram (Figure 2) showed that the urea peak, located at RT 16.7 minutes, did not exhibit excessive overloading. By comparing Figures 1 and 2 it can be seen that there are a greater number of peaks in Figure 1. This is due to the greater volume of urine used to extract metabolites (i.e. 200 µL compared to 50 µL). This comparison is summarised in Table 1, along with the results of experiments using two other extraction volumes (25 µL and 10 µL).

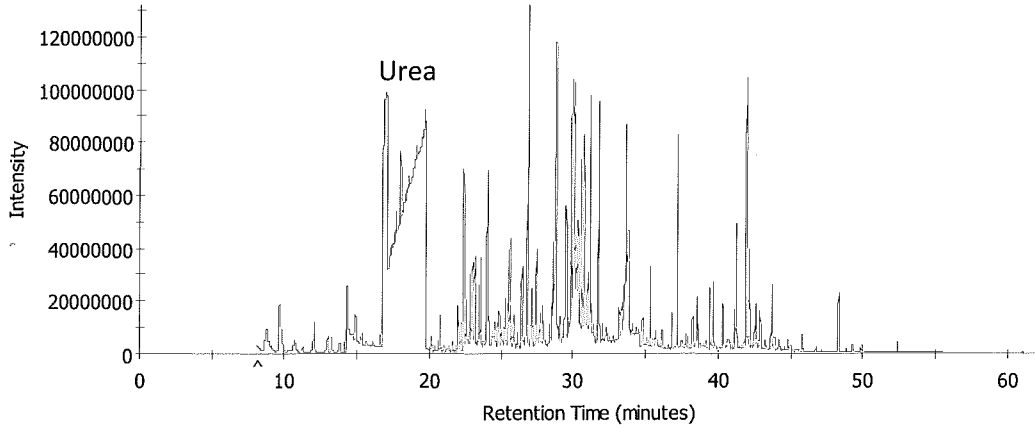

Figure 1. GC/MS chromatogram of a 200 µL urine extract. 0.3 units of urease were added per microlitre of urine which was then incubated for 10 minutes.

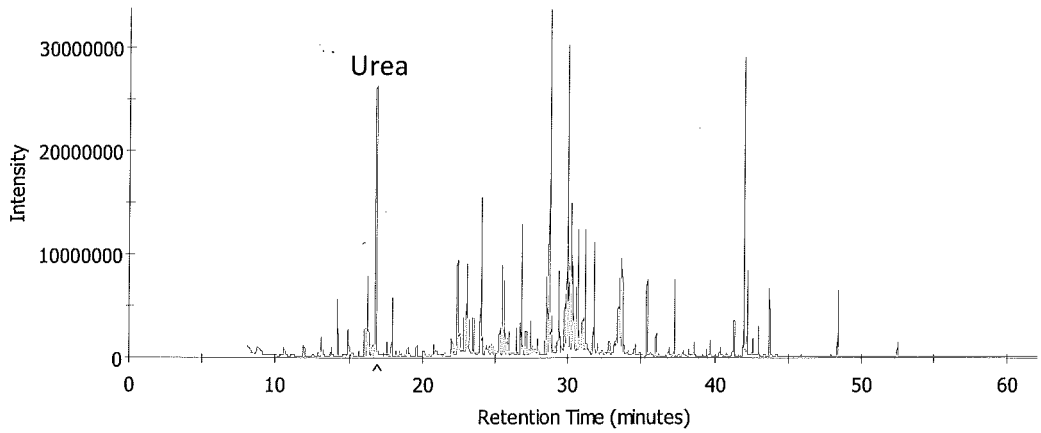

Figure 2. GC/MS chromatogram of a 50 µL urine extract. 1 unit of urease was added per microlitre of urine and was incubated for 30 minutes.

3.1.2 Urine volume

Initial experiments involving catalysed urea hydrolysis saw the need to investigate the extraction volume of urine. Table 1 shows the number of components (mean ± SD) and target components (mean ± SD) detected in four different extraction volumes of pooled 11-week old male and female Lewis rat urine. It can be seen here that there is a significant difference ($P<0.05$) between each of the volumes tested and there is a significant correlation ($P<0.01$) between the number of components found and the urine volume, demonstrating good linearity of the assay method. There is no significant difference between the number of target components found for the volumes 25 µL, 50 µL and 200 µL. At an earlier stage, it was concluded that 200 µL of urine contributed an overload of material on the GC column (Figure 1), however it also yielded the greatest total number of components (Table 1). Extracted urine volumes of 10, 25 and 50 µL showed significantly less components than 200 µL ($P<0.01$), however there was much less evidence of overloading with these smaller extraction volumes. A 50 µL extract is shown in Figure 2 with a maximum ion intensity of 3×10^7. Although

there was no significant difference between the number of target components detected with the 25 and 50 µL volumes, it was concluded that 50 µL was the optimum volume to proceed with due to the presence of a significantly higher total number of components detected (Table 1).

Table 1. Number of components (mean ± SD) and target components (mean ± SD) detected in 10, 25, 50 and 200 µL extraction volumes of urine.

Urine Volume (µL)	Samples (*n*)	Number of Components	Number of Target Components
10	3	$71 \pm 4^{\dagger\#¥}$	50 ± 5
25	3	$164 \pm 25^{*\#¥}$	93 ± 6*
50	3	$230 \pm 9^{*\dagger¥}$	108 ± 3*
200	5	$580 \pm 26^{*\dagger\#}$	100 ± 8*

*significantly different from 10 µL ($P<0.05$), †significantly different from 25 µL ($P<0.05$), $^{\#}$significantly different from 50 µL ($P<0.05$), $^{¥}$significantly different from 200 µL ($P<0.05$).

3.1.3 Centrifugation

Some debate has centred on the effect of centrifuging urine prior to metabolomic analysis, specifically, the potential loss of metabolites in the sediment of the urine during centrifugation [9]. Urine samples (in triplicate) with and without an initial centrifugation step (see section *2.2*) were prepared for extraction. The number of components (mean ± SD) and target components (mean ± SD) that were detected are shown in Table 2. Centrifugation resulted in 473 ± 7 components and 44 ± 1 target components being detected and the non-centrifuged samples resulted in 436 ± 33 components and 43 ± 4 components being detected. There was no significant difference between the number of components and target components found

between the centrifuged and non-centrifuged groups; which show that initial centrifugation does not affect the number of metabolites detectable in Lewis and LPK rat urine by GC/MS. A *et al.* [9] found that metabolites such as myo-inositol, urate, hippurate, glycerate, p-hydroxyphenylacetate and a few other unidentified compounds were found in greater concentrations in a non-centrifuged group of rat urine. However, A *et al.* [9] also found that centrifugation improved the response of metabolites such as creatinine, phosphorate and allantoin and predicted that this was due to improved trimethylsilylation for the lower concentrations of analytes in the centrifuged supernatants. Since one of the main aims of this study was to develop a method for non-targeted urinary metabolic profiling, it was decided that the centrifuged method should be adopted for effective trimethylsilylation, particularly because creatinine was to be used as a normalisation technique.

Table 2. Number of components (Mean ± SD) and target compounds (Mean ± SD) detected with and without the initial centrifugation step in the extraction method (n = 3).

Treatment	Number of Components	Number of Target Components
Centrifuged	473 ± 7	44 ± 1
Non-centrifuged	436 ± 33	43 ± 4

3.2 Instrumentation

3.2.1 Limit of detection

Limit of detection is determined by a signal (detector response) that is three times the baseline noise. Low detection limits are important for the detection of low abundance metabolites in non-targeted metabolomics [11]. Trimethylsilylated D-mannitol was used to determine the LOD of three different GC/MS systems at two sites (Table 3). Seven different concentrations (0.01, 0.05, 0.1, 0.5, 1, 5 and 10 $ng.\mu L^{-1}$) of D-mannitol were analysed with splitless injections (1 µL) on each of the instruments to obtain the limit of detection. The most sensitive instrument was the ion trap GC/MS at Murdoch University (with an LOD of 0.10 ng on column) followed by the quadrupole GC/MS at Murdoch University (with an LOD of 0.14 ng on column). These results were expected as ion trap mass spectrometers trap all ions and destabilise their trajectories in order that they reach the detector in a uniform stream. Linear quadrupole mass spectrometers, however, stabilise the ion trajectory for a given *m/z* species, but this can result in the loss of sensitivity for the destabilised ions. An LOD could not be obtained for the ion trap GC/MS at ECU due to the apparent absence of D-mannitol. It was proposed that D-mannitol was co-eluting with *n*-nonadecane which resulted from the operation of the instrument in constant pressure mode, the only mode available due to the age of the instrument. To rectify this problem, the pressure was reduced slightly (from 15 to 14 psi) and the temperature ramp was increased in the second segment of the run (from 1.0 to 1.1 $°C.min^{-1}$). This saw a small peak eluting after *n*-nonadecane even though the concentration of D-mannitol was 100 ng on column and should have been an intense peak in comparison to the alkane. This resulted in the exclusion of the GC/MS at ECU due to the inferior sensitivity it exhibited for a

compound which is typical of metabolite properties. This finding was supported in section *3.2.2.*

Table 3. Comparison of the limit of detection (ng on column) of a quadrupole GC/MS system and two ion trap GC/MS/MS systems at two different sites.

Instrument	Institution	Limit of Detection
Quadrupole GC/MS	Murdoch	0.14
Ion Trap GC/MS	Murdoch	0.10
Ion Trap GC/MS	ECU	ND

3.2.2 Detectable components

As the limit of detection was not particularly different between the quadrupole GC/MS and the ion trap GC/MS at Murdoch University, urine samples were prepared using the optimised extraction method and run on both instruments as well as the ion trap GC/MS at ECU. Surprisingly, the quadrupole GC/MS returned a significantly higher ($P<0.05$) number of components than the ion trap GC/MS at Murdoch University in both the centrifuged and non-centrifuged groups (Table 4). The ion trap GC/MS at ECU showed significantly less ($P<0.05$) components than both of the GC/MS systems at Murdoch, however, these results were expected due to the inferior limit of detection. The number of target components was not reported here as the AnalyzerPro user-generated library was created using the quadrupole GC/MS at Murdoch University with D-mannitol retention time lock based retention times. For this reason, library matched components would not be returned for either of the ion trap instruments. This could not be corrected with retention time locking as Varian instruments do not

have this functionality. As a result of these findings, it was decided to proceed with urine analysis using the linear quadrupole GC/MS.

Table 4. Comparison of the number of components (mean ± SD) detected in pooled 11 week old male and female Lewis rats by the quadrupole GC/MS system and two ion trap GC/MS/MS systems (n = 3).

Instrument	Number of Components	
	Centrifuged	Non-centrifuged
Quadrupole GC/MS (Murdoch)	$473 \pm 7^{\dagger\#}$	$436 \pm 33^{\dagger\#}$
Ion Trap GC/MS (Murdoch)	$160 \pm 7^{*\#}$	$152 \pm 13^{*\#}$
Ion Trap GC/MS (ECU)	$56 \pm 0^{*\dagger}$	$56 \pm 4^{*\dagger}$

*Significantly different from Quadrupole GC/MS (Murdoch), †significantly different from Ion Trap GC/MS/MS (Murdoch), #significantly different from Ion Trap GC/MS/MS ECU.

3.2.3 Precision and reproducibility

A standard curve generated for quantitative analysis is also a useful measure of precision for both sample preparation and instrumentation steps. Figure 3 shows the peak area of five different concentrations (0.1, 0.5, 1, 5, 10 $ng.\mu L^{-1}$) of trimethylsilylated D-mannitol splitless injections (1 µL) on the quadrupole GC/MS at Murdoch. The R^2 value of 0.9994 indicates a precise and reproducible method and system for the concentrations between 0.1 to 10 ng of trimethylsilylated D-mannitol on column. D-mannitol was used to obtain LOD, precision and reproducibility data as the majority of the high throughput metabolomics work in this laboratory has been carried out on fungi, for which D-mannitol was the most abundant compound detected. Additionally, D-mannitol has been utilised by Metabolomics Australia to lock retention times of instruments across Australia. D-mannitol is cheap, and elutes at

approximately half of the total GC/MS analysis time, making it useful for these experiments and also for retention time locking. Internal standards are also used as a measure of method reproducibility [6, 15]. Figure 4 shows the peak area of 50 ng.μL^{-1} of the internal standard trimethylsilylated ribitol in each of the injected D-mannitol standards. The relative standard deviation (RSD) of the ribitol peak area was 2.68% which indicated that the assay (sample preparation and instrumental analysis) was highly reproducible. Ribitol was used as an internal standard because it is not present in fungi (nor is it present in mammals) and elutes before the 30 minute 'sugar overload' in fungal samples, and is therefore readily resolved. There are many other internal standards that have proven to be stable and reliable such as [$^{2}H_3$]myristic acid [15], succinic d_4 acid, malonic d_2 acid, glycine d_5 and glucose $^{13}C_6$ [6]. However, due to the ready availability, the low cost, and the recent success using D-mannitol and ribitol, these procedures were continued.

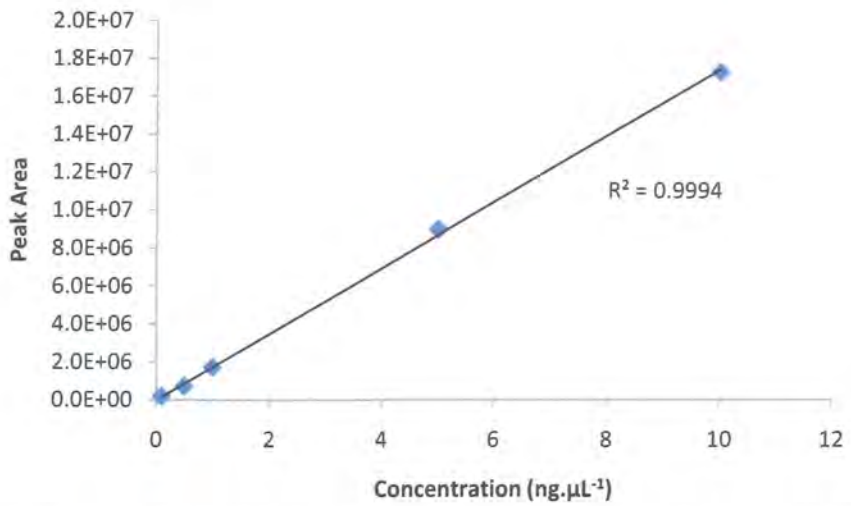

Figure 3. Peak area of 0.1, 0.5, 1, 5 and 10 ng.μL^{-1} injections of trimethylsilylated D-mannitol (RT = 30.6 minutes).

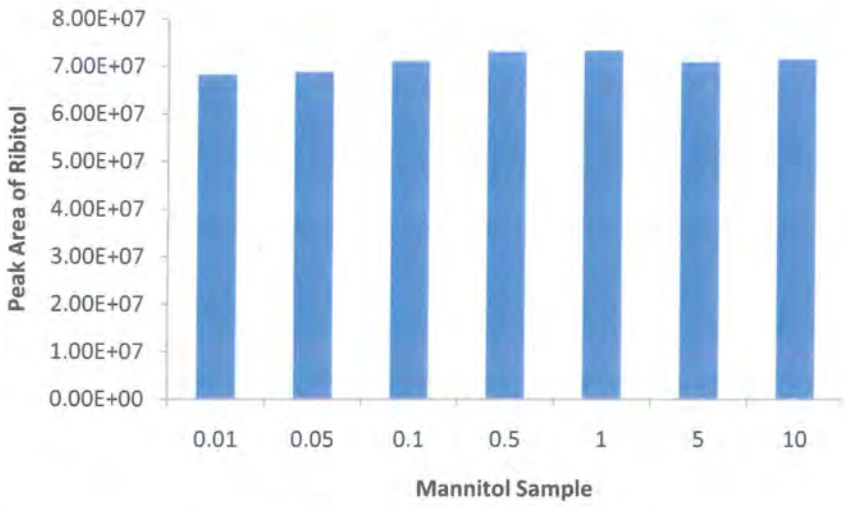

Figure 4. Peak area of 50 ng.μL^{-1} ribitol in each of the D-mannitol standards (%RSD = 2.68).

3.2.4 Injection mode

A split injection is commonly used to achieve baseline separation when the concentration of a sample is known to be high or is unknown. It is a preventative measure to minimise overloading. This is especially the case when using capillary columns as tens of micrograms of analyte can cause overloading. Although splitless injections are widely used in non-targeted urinary metabolic profiling [9, 15, 17, 19], a split injection was tested in an attempt to improve signal-to-noise ratio and resolution, and to minimise overloading. Whilst the chromatography was improved, the number of components detectable was reduced significantly ($P<0.05$). Table 5 shows the number of components (mean ± SD) and target components detected using splitless and split (20:1) injections. The splitless method resulted in the detection of 473 ± 7 components and 44 ± 1 target component library matches. The 20:1 split method, however, resulted in the detection of only 179 ± 7 components and 18 ± 0 target component library matches. Due to these findings, subsequent injections were performed in splitless mode.

Table 5. Number of components (mean ± SD) and target components (mean ± SD) detected using splitless and split (20:1) GC methods (n = 3).

Treatment	Number of Components	Number of Target Components
Splitless	473 ± 7	44 ± 1
Split (20:1)	179 ± 7*	18 ± 0*

*significantly different from splitless.

3.3 Data Analysis

From the 584 ± 44 components detected in the male Lewis rat urine (n = 7) only 92 ± 5 were library matched (target components) (Table 6). For female LPK rat urine (n = 4) 554 ± 31 components were detected of which 96 ± 5 were library matched. Analysis of male LPK rat urine resulted in the detection of 587 ± 47 components of which 94 ± 5 were library matched. There was no significant difference between the number of components and target components found in the two strains and sexes.

Table 6. Number of components (mean ± SD) and target components found in Lewis male and LPK male and female rat urine.

Strain	Sex	Replicates	Number of Components	Number of Target Components
Lewis	Male	7	585 ± 44	92 ± 5
LPK	Female	4	554 ± 31	96 ± 5
LPK	Male	6	587 ± 47	94 ± 5

3.3.1 Normalisation

As mentioned previously, internal standards are a measure of methodology reproducibility and can highlight if the signal-to-noise ratio changes between analyses. Ribitol was added to each of the urine samples at a desired final concentration of 15 $ng.\mu L^{-1}$ or 15 ng on column. By doing so, each of the peaks in a particular sample were normalised to the internal standard to account for any variability in both the sample preparation and the instrument. Trimethylsilylated ribitol is a target component in the AnalyzerPro user-generated library. In addition to this, the mass spectrum of the compound was obtained, as was the deconvoluted mass spectrum, to ensure the library match was accurate and to further ensure data was being normalised to the correct component. There was a significant correlation ($P<0.01$) between the sum of the peak area of all components in the sample to the respective ribitol peak area. Based on this result it was assumed that if ribitol was affected by sample preparation and instrument variability, all peaks were subsequently affected and normalisation to account for this was therefore warranted. Figure 5 shows the peak area of ribitol in each of the urine samples, showing a relative standard deviation (RSD) of 14.02%. A desirable RSD is <10%, however, due to the limited volume of sample obtained from the rats, this experiment could not be repeated.

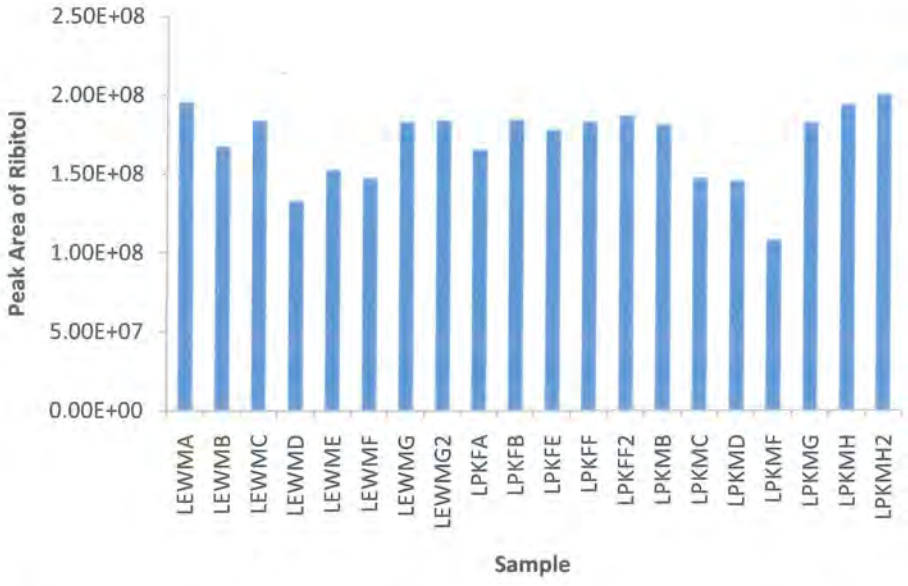

Figure 5. Peak area of trimethylsilylated ribitol in each of the rat urine samples (%RSD = 14.02). Samples are labelled strain (LEW/LPK), sex (M/F) and from A-H for individual identification.

Biofluids such as plasma, serum and cerebrospinal fluid are physiologically controlled, however, urine can be particularly varied, mainly due to water consumption, but also as a result of other factors [39]. Although food and water were available to the animals at all times, it was expected that hydration state would affect the concentration of metabolites in the urine and potentially cause significant variability in the metabolic profiles, masking any differences that may have been seen as a result of disease state. To correct for this, creatinine was identified as a potential normalisation strategy. This is common practice in clinical chemistry and metabolomics as creatinine is an indicator of the concentration of urine, assuming a constant secretion of creatinine into urine [36, 39]. The focus of this study was ARPKD. Therefore, there was concern that creatinine levels would be significantly lower in the LPK animals than in the Lewis. The result of a one-way ANOVA found that there was no significant difference in creatinine attributed to sex and/or strain. This finding was supported by Phillips *et al.* [35] who found that serum creatinine was not significantly

increased in the LPK rat until age 12-weeks. In order to positively identify creatinine in the samples, a standard was derivatised and run on the quadrupole GC/MS at Murdoch University. The retention time of creatinine was found to be 23.4 minutes (Figure 6) and the base peak *m/z* 115 (Figures 7 and 8). In order to add creatinine to the existing AnalyzerPro library, a deconvoluted mass spectrum was obtained (Figure 8) which excludes interferences such as *m/z* 73 which is attributed to trimethylsilyl groups. The deconvoluted mass spectrum aids in the identification of the correct base peak. After the addition of creatinine to the library, the data was re-processed to include the identification of creatinine. To ensure that it was correct to assume creatinine levels would increase and decrease with changing metabolite concentration, the sum of all peak areas in the sample were correlated to the respective creatinine peak area. This was found to be significant ($P<0.01$) after ribitol normalisation.

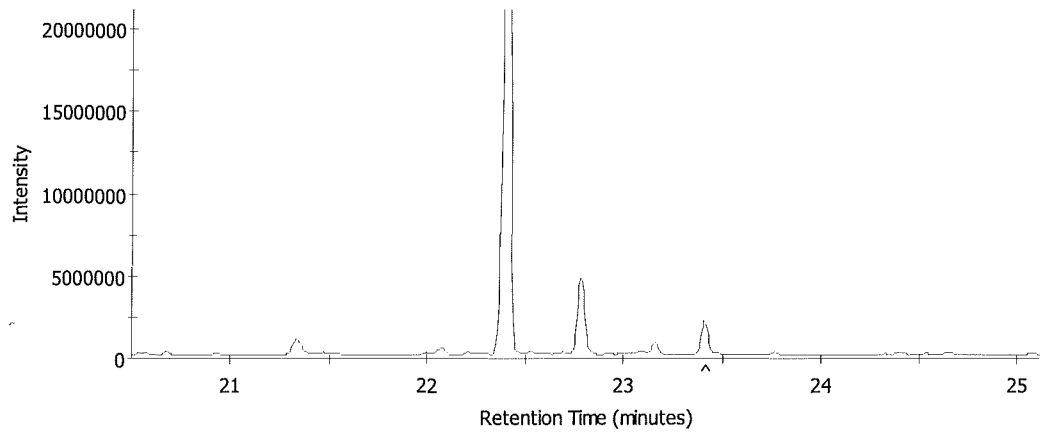

Figure 6. GC Chromatogram of trimethylsilyl creatinine (RT 23.4 minutes).

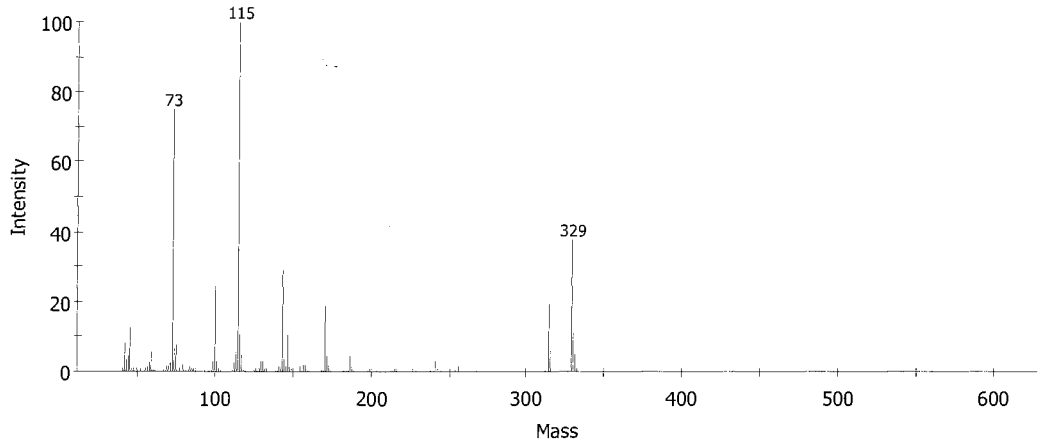

Figure 7. Mass spectrum of trimethylsilyl creatinine.

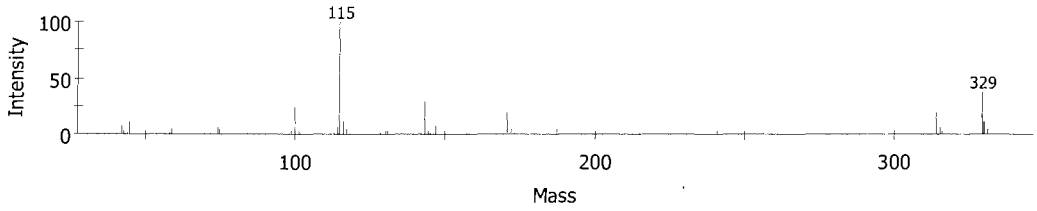

Figure 8. Deconvoluted mass spectrum of trimethylsilyl creatinine.

3.3.2 Principal component analysis

Principal component analysis has been employed in metabolomics studies using GC [8, 9, 15, 17], LC [2, 41-44], MS [11, 22, 39] and ^{1}H NMR [16, 36, 38, 45, 46]. It is a method of data transformation that assigns the data a coordinate system whereby in any projection of the data (2D or 3D), the greatest variance (%) lies on the first coordinate (principal component 1). Principal component 2 shows the next greatest variance and so on. Figure 9a shows the principal component analysis of male and female 5-week old Lewis and LPK rats. This 2-D projection of data was generated using the log-transformed raw peak area of target components. At this point, the target components had not been normalised to ribitol or creatinine. 18% of the variance between the samples was attributed to PC1 and 14% was attributed to PC2. The remaining principal

components 3-15 inclusive (not shown) explain ≤9% of variance. With only 32% of variance explained in PC1 and PC2 and clear overlapping of the sample groups, this model provides little information regarding variability in metabolites as a function of disease state.

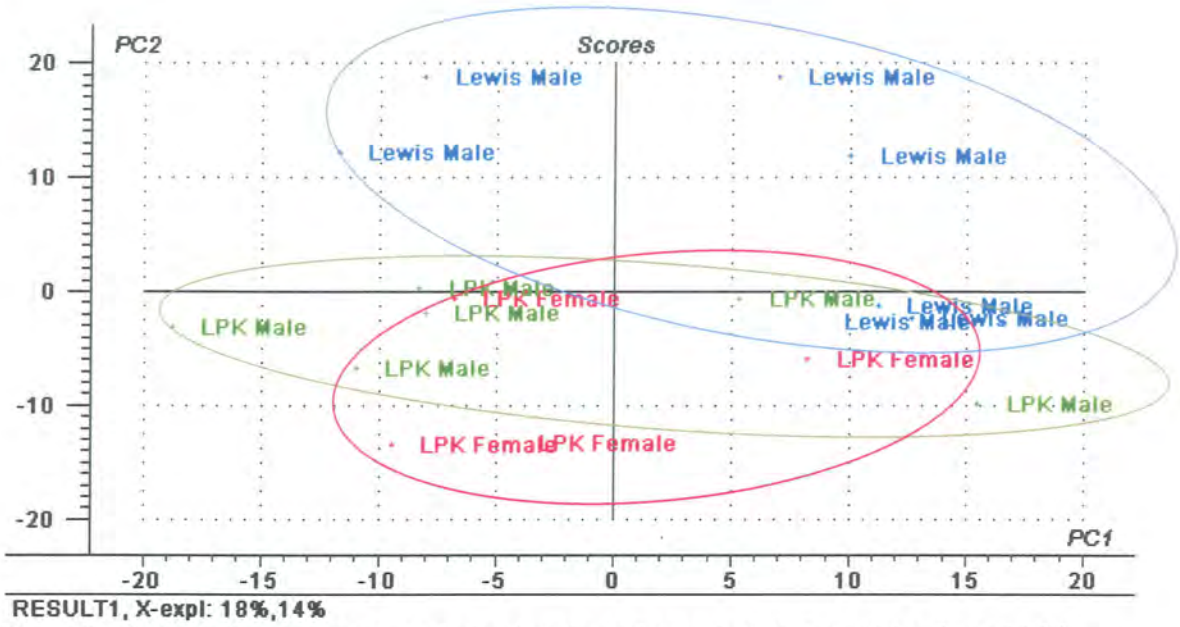

Figure 9a. Principal component analysis of male and female 5-week old Lewis and LPK rat urine samples. This output was generated using log-transformed peak area data.

Figure 9b shows the principal component analysis of male and female 5-week old Lewis and LPK rats. This projection of data was generated using the log-transformed peak area of target components normalised to the internal standard ribitol. PC1 explained 28% of the variance between samples and PC2 explained 17% of the variance. PC3-15 (not shown) explained ≤14% of variance. With an overall explained variance of 45% in PC1 and PC2, it was concluded that 13% (i.e. the difference in explained variance between Figures 9a and 9b) of variance between samples was attributed to the concentration of metabolites in the individual samples (i.e. overloading) and not to sample preparation and instrumental analysis. Results

reported previously in this section (*3.2.3*) indicate that the sample preparation and instrumental analysis were highly reproducible with a relative standard deviation of 2.68% for the internal standard ribitol. Section *3.3.1,* however, reports the relative standard deviation of 14.02% for the internal standard, ribitol. This result (Figure 9b) was therefore expected as the signal-to-noise ratio changed between biological replicates while the signal-to-noise ratio between technical replicates stayed relatively stable. Figure 9b once again shows overlapping of sample groups showing little discrimination between diseased and non-diseased animals.

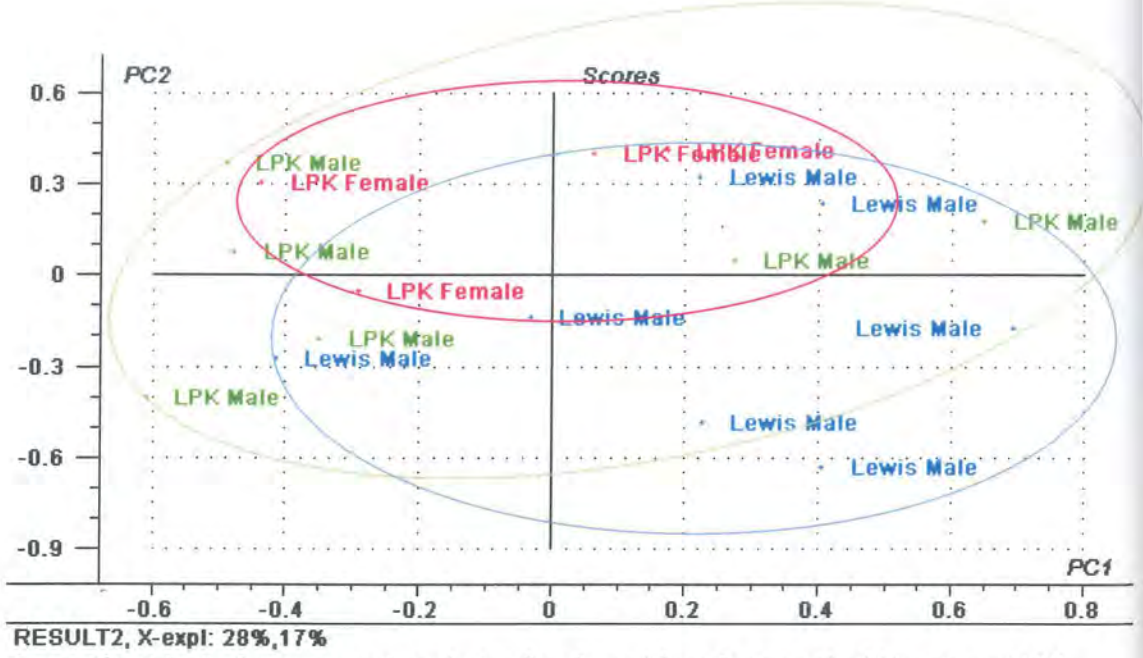

Figure 9b. Principal component analysis of male and female 5-week old Lewis and LPK rat urine samples. This output was generated using log transformed peak area data normalised to the internal standard ribitol.

Figure 9c shows the principal component analysis of male and female 5-week old Lewis and LPK rats. This projection of data was generated using the log-transformed peak area of target components normalised to ribitol and then creatinine. 40% of the variance between samples is explained by PC1 and 16% by PC2. PC3-15 (not shown) each explained ≤ 13%. Therefore, 56% of the total variance was explained by the model (PC1 and PC2), of which, 11% was attributed to variation in analyte concentration due to the hydration state of the animal. Figures 9a-c shows the importance of data transformation and normalisation. These processes minimise the impact of variability of high intensity peaks [11] and further illustrate the need for high inter-sample reproducibility. Even more so than Figures 9a and 9b, there was a clear overlap in sample groups. Having accounted for signal-to-noise variability and hydration state, it was expected that this model represented variability attributed to individual target components in the samples. Although there is an overlap in sample groups, it was seen here that the 5 male animals on the right are the key influence attributing to the variance explained by PC1 (40%). Given that 80% of the animals influencing PC1 were Lewis males, it was noticed that male diseased and non-diseased animals began to show discrimination (Figure 10) after the various normalisation techniques.

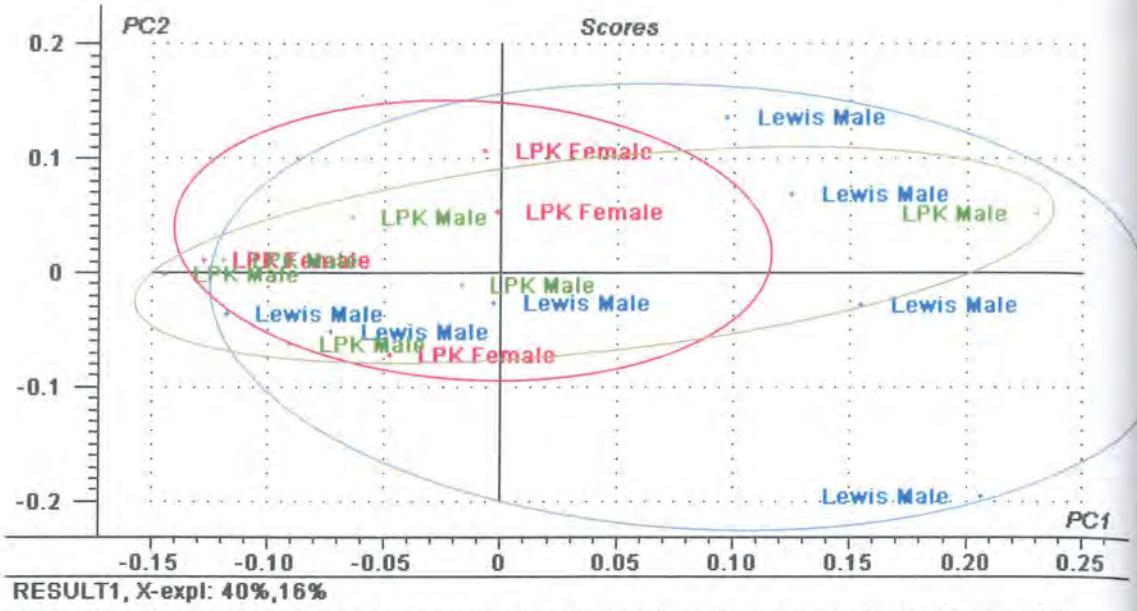

Figure 9c. Principal component analysis of male and female 5-week old Lewis and LPK rat urine samples. This output was generated using log transformed peak area data normalised to ribitol then creatinine.

These results (Figures 9a-10) show that PCA is a powerful tool for visualising metabolomics data. These models (Figures 9a-c) are representative of approximately 100 metabolites for 17 individual animals. PCA is commonly used to identify specific metabolites which are attributed to variations between disease states [8, 36, 44, 46], however, it can also be used as a method for data mining to identify interferences due to assay method [15]. In this case, it has been used for both. The X-loadings plot for the PCA in Figure 10 (Figure 11) revealed that the metabolites attributing to the variance between sample groups were, but were not limited to, D-mannitol, α-ketoglutarate, uric acid, allantoin and phosphoric acid. The presence of D-mannitol is an indication that the chromatographic separations were not sound for the samples that attributed to the variance. In particular, the presence of D-mannitol is an indication that there was overloading in the samples as higher eukaryotes do not synthesise or catabolise D-mannitol [47]. This finding is indicative of a shift in retention time due to overloading and resultant misidentification of peaks. It was postulated

that the presence of D-mannitol was the result of the presence of a closely eluting sugar with a similar fragmentation pattern. For this reason, D-mannitol was not excluded from the data analysis as it may represent an important sugar which should be investigated further in the future. The presence of α-ketoglutarate, uric acid and allantoin, however, are indicative that the groups are likely to be separated due to disease state. α-ketoglutarate has been linked to hypertension by Akira *et al.* [36], and hypertension has been identified as a clinical symptom of ARPKD in the LPK rat by Phillips *et al.* [35]. Further, it is postulated that uric acid and allantoin may be linked to renal and/or liver dysgenesis. The reason for the presence of phosphoric acid in the X-loadings plot remains unknown. It has been postulated that it is an artefact of sample preparation. Whilst these findings show a solid foundation for investigating ARPKD using metabolomics, food and water intake and urine collection (i.e. spot urine) cannot be ruled out as factors influencing variance. Walsh *et al.* [40] have identified urinary metabolomics as an effective method of determining and assessing dietary intake which suggests that food and water should be withheld from the animals to eliminate this as a factor attributing to variance. Similarly, Walsh *et al.* have identified the need to investigate the difference between 24-hour samples and spot samples. Akira *et al.* [36] were able to explain 76% of variance between young spontaneously hypertensive rats and healthy controls using a PCA model after creatinine normalisation. Similarly, Vallejo *et al.* were able to explain 94% of variance between atherosclerosis patients and healthy controls using a PCA model. The total explained variance of the PCA in Figure 10 is 68%. This comparatively low explained variance may be attributed to the limitations of the study, in particular, food and water intake, and limited sample size. In this analysis (figure 10) principal components 3-15 (not shown) explain ≤ 15% of variance. This suggests that there is further interpretation needed for this data set with

respect to principal components beyond PC1 and PC2. In order to do so, all of the limitations of the study need to be identified to make accurate and reliable interpretations regarding the influence of individual metabolites.

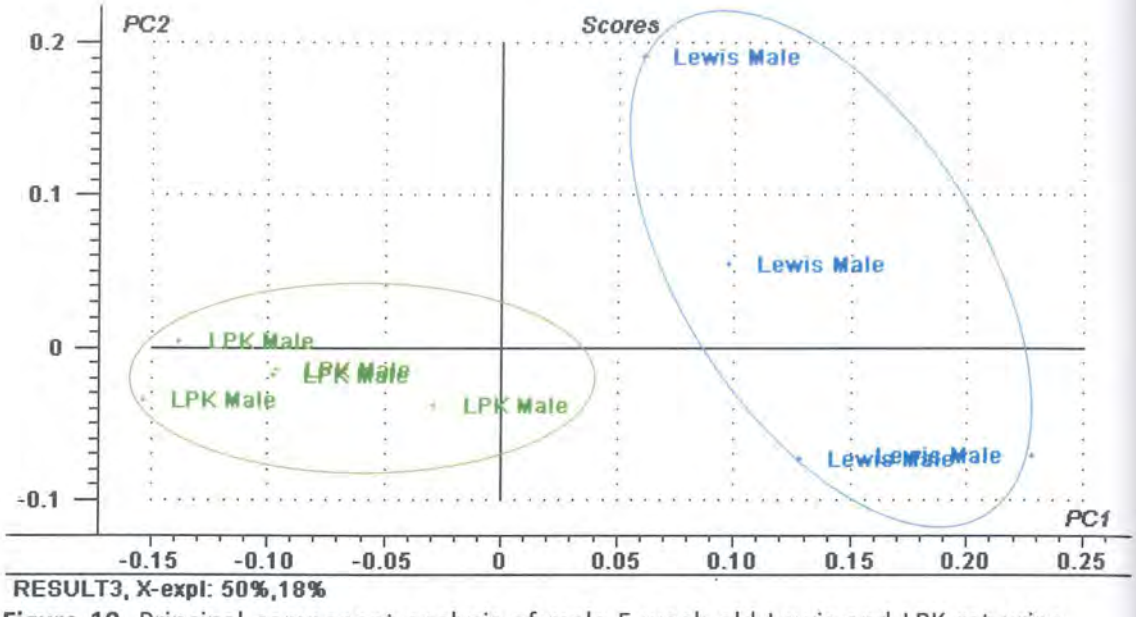

Figure 10. Principal component analysis of male 5-week old Lewis and LPK rat urine samples. This output was generated using log transformed peak area data normalised to ribitol then creatinine.

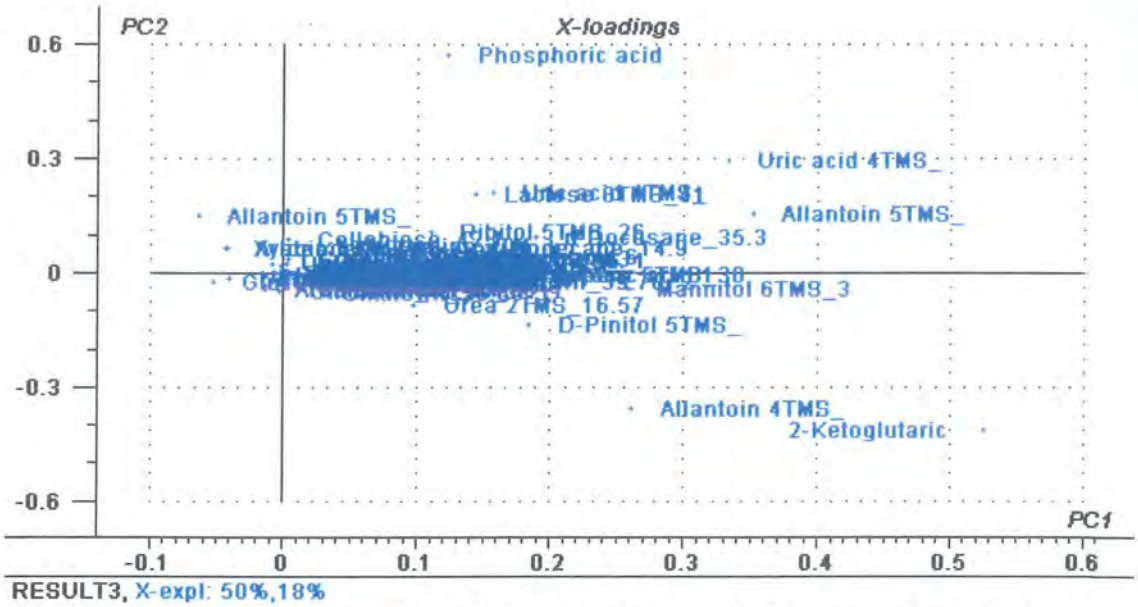

Figure 11. PCA X-loadings plot of male 5-week old Lewis and LPK rat urine samples. This output was generated using log transformed peak area data normalised to ribitol then creatinine and shows the key metabolites which influence the model described in Figure 10.

CHAPTER FOUR

CONCLUSION

4.1 Conclusion

There are some interesting findings in the present research, the overall outcome of this work being the identification of the need for a well-controlled experiment with a greater sample size. With the development of an optimised method for non-targeted urinary metabolic profiling of the LPK rat, it is hoped that a metabolomics approach may lead to the identification of biomarkers for ARPKD in the future. Once the phenotype of the LPK rat model of ARPKD has been thoroughly investigated, it is hoped that the findings will contribute to research in human ARPKD. The findings of this study, particularly the development of a GC/MS method to analyse Lewis and LPK rat urine, demonstrate the potential of metabolomics to investigate ARPKD. The work in this project has provided preliminary data demonstrating the ability to differentiate urine samples on the basis of sex and disease state. Future detailed temporal studies are hoped to elucidate key metabolites to identify the onset of kidney dysgenesis and aid in determining the function of the ARPKD protein fibrocystin/polyductin.

4.2 Recommendations

The user generated library at the Metabolomics Australia Node; Murdoch University, primarily consists of polar plant metabolites and for further work; mammalian metabolites and non-polar metabolites, including lipids, need to be added to the library. This can be carried out by complementary analytical techniques such as ^{1}H NMR to elucidate the structures of the unknown analytes in the samples. These results would then be validated with pure standards. Controlled sample collection, including food and water prior to sampling, and further optimisation of sample size will improve reproducibility of data in the future. Furthermore, the role of creatinine as an appropriate urine normalisation technique as a function of kidney dysgenesis needs to be investigated. This can be undertaken as a temporal animal based study with adequate biological replicates.

CHAPTER FIVE

REFERENCES

1. Assfalg, M., et al., *Evidence of different metabolic phenotypes in humans.* Proceedings of the National Academy of Sciences, 2008. **105**(5): p. 1420-1424.
2. Ceglarek, U., et al., *Challenges and developments in tandem mass spectrometry based clinical metabolomics.* Molecular and Cellular Endocrinology, 2009. **301**: p. 266-271.
3. Villas-Boas, S.G., et al., *Metabolome Analysis*. Wiley-Interscience Series in Mass Spectrometry, ed. D.M. Desiderio and N.M.M. Nibbering. 2007, New Jersey: John Wiley & Sons, Inc. 311.
4. Wishart, D.S., *Metabolomics in monitoring kidney transplants.* Current Opinion in Nephrology and Hypertension, 2006. **15**: p. 637-642.
5. Wishart, D.S., *Metabolomics: The Principles and Potential Applications to Transplantation.* American Journal of Transplantation, 2005. **5**: p. 2814-2820.
6. Kenny, L.C., et al., *Novel biomarkers for pre-eclampsia detected using metabolomics and machine learning.* Metabolomics, 2005. **1**(3): p. 227-234.
7. Feng, X., et al., *Mass Spectrometry in Systems Biology: An Overview.* Mass Spectrometry Reviews, 2008. **27**: p. 635-660.
8. Vallejo, M., et al., *Plasma fingerprinting with GC-MS in acute coronary syndrome.* Analytical and Bioanalytical Chemistry, 2008. **Epub Ahead of Print**.
9. A, J., et al., *Global analysis of metabolites in rat and human urine based on gas chromatography/time-of-flight mass spectrometry.* Analytical Biochemistry, 2008. **379**: p. 20-26.

10. Lin, L., et al., *Identification of 2,5-dimethoxy-4-ethylthiophenethylamine and its metabolites in the urine of rats by gas chromatography-mass spectrometry.* Journal of Chromatography B, 2003. **798**: p. 241-247.

11. Dettmer, K., P.A. Aronov, and B.D. Hammock, *Mass Spectrometry-Based Metabolomics.* Mass Spectrometry Reviews, 2007. **26**: p. 51-78.

12. Ramautar, R., A. Demirci, and G.J. de Jong, *Capillary electrophoresis in metabolomics.* Trends in Analytical Chemistry 2006. **25**(5): p. 455-466.

13. Barbas, C., et al., *Capillary electrophoresis as a metabolomic tool in antioxidant therapy studies.* Journal of Pharmaceutical and Biomedical Analysis 2008. **47**: p. 388-398.

14. Lindon, J.C. and J.K. Nicholson, *Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery.* Trends in Analytical Chemistry, 2008. **27**(3): p. 194-204.

15. Zhang, Q., et al., *GC/MS analysis of the rat urine for metabonomic research.* Journal of Chromatography B, 2007. **854**: p. 20-25.

16. Beckonert, O., et al., *Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts.* Nature Protocols, 2007. **2**(11): p. 2692-2703.

17. A, J., et al., *Extraction and GC/MS Analysis of the Human Blood Plasma Metabolome.* Analytical Chemistry, 2005. **77**: p. 8086-8094.

18. Kuhara, T., *Gas Chromatographic-Mass Spectrometric Urinary Metabolome Analysis to Study Mutations of Inborn Errors of Metabolism.* Mass Spectrometry Reviews, 2005. **24**: p. 814-827.

19. Lin, H.M., et al., *Non-targeted urinary metabolite profiling of a mouse model of Crohn's disease.* Journal of Proteome Research, 2009. **In press**.

20. Matsumoto, I. and T. Kuhara, *A New Chemical Diagnostic Method for Inborn Errors of Metabolism by Mass Spectrometry- Rapid, Practical, and Simultaneous Urinary Metabolites Analysis* Mass Spectrometry Reviews, 1996. **15**: p. 43-57.

21. Harris, D.C., *Quantitative Chemical Analysis*. 2nd ed. 2007, New York: W.H. Freeman and Company.

22. Vaidyanathan, S., et al., *A laser desorption ionisation mass spectrometry approach for high throughput metabolomics.* Metabolomics, 2005. **1**(3): p. 243-250.

23. Ward, C.J., et al., *Cellular and subcellular localization of the ARPKD protein; fibrocystin is expressed on primary cilia.* Human Molecular Genetics, 2003. **12**(20): p. 2703-2710.

24. Sweeney, W.E. and E.D. Avner, *Molecular and cellular pathophysiology of autosomal recessive polycystic kidney disease (ARPKD).* Cell Tissue Research, 2006. **326**: p. 671-685.

25. Bukanov, N.O., et al., *Long-lasting arrest of murine polycystic kidney disease with CDK inhibitor roscovitine.* Nature 2006. **444**: p. 949-952.

26. Fischer, D.C., et al., *Activation of the AKT/mTOR pathway in autosomal recessive polycystic kidney disease (ARPKD).* Nephrology Dialysis Transplantation, 2009. **24**: p. 1819-1827.

27. Woollard, J.R., et al., *A mouse model of autosmal recessive polycystic kidney disease with biliary duct and proximal tubule dilatation.* Kidney International, 2007. **72**: p. 328-336.

28. Menezes, L.F.C. and L.F. Onuchic, *Molecular and cellular pathogenesis of autosomal recessive polycystic kidney disease* Brazilian Journal of Medical and Biological Research, 2006. **39**: p. 1537-1548.

29. Turkbey, B., et al., *Autosomal recessive polycystic kidney disease and congenital hepatic fibrosis (ARPKD/CHF).* Pediatric Radiology, 2009. **39**: p. 100-111.

30. Mucher, G., et al., *Fine Mapping of the Autosomal Recessive Polycystic Kidney Disease Locus (PKHD1) and the Genes MUT, RDS, CSNK2β, and GSTA1 at 6p21.1-p12.* Genomics, 1998. **48**: p. 40-45.

31. Herman, T.E. and M.J. Siegel, *Neonatal autosomal recessive polycystic kidney disease.* Journal of Perinatology, 2008. **28**: p. 584-585.

32. Kim, I., et al., *Fibrocystin/Polyductin Modulates Renal Tubular Formation by Regulating Polycystin-2 Expression and Function.* Journal of the American Society of Nephrology, 2008. **19**: p. 455-468.

33. Muff, M.A., et al., *Development and characterization of a cholangiocyte cell line from the PCK rat, an animal model of Autosomal Recessive Polycystic Kidney Disease.* Laboratory Investigation, 2006. **86**: p. 940-950.

34. Gigarel, N., et al., *Preimplantation genetic diagnosis for autosomal recessive polycystic kidney disease.* Reproductive BioMedicine Online, 2008. **16**(1).

35. Phillips, J.K., et al., *Temporal relationship between renal cyst development, hypertension and cardiac hypertrophy in a new rat model of autosomal recessive polycystic kidney disease.* Kidney and Blood Pressure Research, 2007. **30**: p. 129-144.

36. Akira, K., et al., *^{1}H NMR-based metabonomic analysis of urine from young spontaneously hypertensive rats.* Journal of Pharmaceutical and Biomedical Analysis, 2008. **46**: p. 550-556.

37. Alzweiri, M., et al., *Comparison of different water-miscible solvents for the preparation of plasma and urine samples in metabolic profiling studies.* Talanta, 2008. **74**: p. 1060-1065.

38. Potts, B.C.M., et al., *NMR of biofluids and pattern recognition: assessing the impact of NMR parameters on the principal component analysis of urine from rat and mouse.* Journal of Pharmaceutical and Biomedical Analysis, 2001. **26**: p. 463-476.

39. Warrack, B.M., et al., *Normalization strategies for metabonomic analysis of urine samples.* Journal of Chromatography B, 2009. **877**: p. 547-552.

40. Walsh, M.C., et al., *Understanding the metabolome – challenges for metabolomics.* Nutrition Bulletin, 2008. **33**: p. 316-323.

41. Bruce, S.J., et al., *Investigation of Human Blood Plasma Sample Preparation for Performing Metabolomics Using Ultrahigh Performance Liquid Chromatography/Mass Spectrometry.* Analytical Chemistry, 2009. **Epub Ahead of Print**.

42. Hodson, M.P., et al., *An approach for the development and selection of chromatographic methods for high throughput metabolomic screening of urine by ultra pressure LC-ESI-ToF-MS.* Metabolomics, 2009. **5**: p. 166-182.

43. Michopoulos, F., et al., *UPLC-MS-Based Analysis of Human Plasma for Metabonomics Using Solvent Precipitation or Solid Phase Extraction.* Journal of Proteome Research, 2009. **In press**.

44. Whitfield, P.D., et al., *Metabolomics as a diagnostic tool for hepatology: validation in a naturally occurring canine model.* Metabolomics, 2005. **1**(3): p. 215-225.

45. Webb-Robertson, B.J.M., et al., *A Study of spectral integration and normalization in NMR-based metabonomic analyses.* Journal of Pharmaceutical and Biomedical Analysis, 2005. **39**: p. 830-836.

46. Whitehead, T.L., B. Monzavi-Karbassi, and T. Kieber-Emmons, *1H-NMR metabonomics analysis of sera differentiates between mammary tumor-bearing mice and healthy controls.* Metabolomics, 2005. **1**(3): p. 269-278.

47. Liberator, P., et al., *Molecular Cloning and Functional Expression of Mannitol-1-phosphatase from the Apicomplexan Parasite Eimeria tenella.* The Journal of Biological Chemistry, 1998. **273**(7): p. 4237-4244.