



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.

Keele



U N I V E R S I T Y

A urinary proteomics investigation of pregnancy-related complications

Thesis submitted for Doctor of Philosophy in
Clinical Biochemistry

Wafaa Flaifel Al-Jasim

October 2019

Faculty of Medicine and Health Sciences

Keele University

Abstract

Pre-eclampsia (PE) and intrauterine growth restriction (IUGR) are common disorders associated with placental implantation abnormalities in early pregnancy. Both conditions are heterogeneous, having complex and multiple aetiologies. PE complicates 5-8 % of pregnancies globally, and is characterised by new onset hypertension and elevated urine protein concentration. PE is a major cause of maternal and fetal morbidity and mortality. IUGR may co-occur with PE or separately, and affects babies born at or below the bottom decile for weight. Monitoring for these conditions has a tremendous cost burden for healthcare providers, and as yet we have no robust clinical predictors to identify those women who need extra monitoring. Mass spectrometry and Fourier transform infrared spectroscopy were used to analyse proteins within urine samples from pregnant women. Proteomic and bioinformatics analyses of urine samples have demonstrated that proteins, subjected to ultrafiltration and subsequent proteolysis and peptide labelling, are readily detected in urine samples, with differential detection of proteins showing disparate levels in women whose pregnancies are complicated by IUGR and PE. Quantitative changes in urinary protein expression are investigated in order to identify proteins showing differential expression in early pregnancy (15 weeks gestation) in the urine of patients who go on to develop pregnancy complications. A number of candidate proteins showing altered expression levels have been identified for future follow-up analysis. This project suggests new methods of clinical diagnostics for these conditions, which could aid improved targeting of intensive monitoring in prenatal care.

Acknowledgements

First and foremost, I would like to thank God for his never ending grace, mercy and provision during what ended up being one of the toughest times of my life.

I would like to express my sincere gratitude to my supervisor Dr. Sarah Hart for her expertise, support and her helpful suggestions. Her training and guidance helped me in all the time of research and writing of this thesis, she also encouraged me to attend at various workshops and conferences.

I express my gratitude to my co-supervisor Dr Paul Roach for his support, fruitful discussion during my scientific research, as well as helping in writing of thesis. Paul Roach not only advised me but also supported me in my experimental work at Keele University and Loughborough University.

I owe a huge thanks to my country, Iraq and particularly to the Basra University, Iraqi Ministry of Higher Education and Scientific Research for giving me this opportunity to study my PhD in the United Kingdom and for the financial support. Special thanks go to all staff at the Iraqi Cultural Attaché for their moral support and for their responses to my questions and queries.

Also, I express my gratitude to the staff at SCreening fOr Pregnancy Endpoints (SCOPE) centre at University College Cork in Republic of Ireland for their collaboration and support for clinical sample collection. I am grateful to Professor Louise C. Kenny, Professor Philip N. Baker, Dr Pensee Wu and Dr Hamid Al-Jasim for their kindly advice. In addition, I would like to thank Dr Philip Brownridge, Dr Deborah Simpson and Dr Catarina De Matos Ferraz Franco respectively for their help with proteomics analyses in Liverpool University.

I express my gratitude to the following friends and colleagues for their nicely friendship and scientific help who created a wonderful working environment during the project work: Dr Zainab Al-Mnaseer, Dr Rawaa Almyahi, Dr Arwa Al-Shakli, Dr Ibrahim Ali, Dr. Christopher Adams, Dr Munya Kamudzandu, Lekia Kumbe, Jaksha Chandrathas, Dan Merryweather, Ahmad Al-Shallawi, James Kinsella, Dr George Joseph, Yolanda Gómez, Emma Whittle and Ieuan Smith. My especially warm gratitude and respect to Dr Abigail Victoria Rutter for her help with software and her lovely friendship. My thanks also to Hussain Matoq' family, they were kind enough to accommodate me at their home when I attended proteomic workshop in London.

I owe big thanks and gratefulness to all staff at Keele University; ISTM, School of Life Sciences, special thanks to Dr Alan Harper, Professor Trevor Greenhough, Professor Paul Horrocks, Professor Divya Chari, Dr Clare Hoskins, Lisa Cartlidge and Zara Richards and Chris Bain for providing the facilities throughout the academic study. My thankfully to Stefanie Jones, Dave Bosworth, Zoe Bosworth and Debbie Adams at the animal facility; also, I extend thanks to Phil and Chris, who kept our lab clean.

There are no proper words to convey my deep gratitude to my parents, my constant source of inspiration. Massive thanks to my sisters and my brothers for their long distance support and unconditional love. **This thesis is dedicated** to humanity and to all researchers who interested in proteomic biomarkers study.

Output

Oral presentations:

- Urinary Proteomics Identification of Biomarkers for Intrauterine Growth Restriction and Pre-eclampsia, ISTM Postgraduate Symposium 2015, Keele University.

Posters:

- Detecting Pre-eclampsia Using Protein Signatures, ILAS Postgraduate Conference Crossing Paths in Keele hall, Keele University, 28 April 2017.
- Detecting Pre-eclampsia Using Protein Signatures, The 12th annual ISTM Postgraduate Research Symposium North Staffordshire Medical Institute (NSMI) - Hartshill Road, Stoke-on-Trent, Keele University, 9 May 2017.

Articles in preparation:

- Proteomic analysis for pregnancy complications using iTRAQ labelling coupled with mass spectrometry techniques
- Fourier transform infrared analysis of urine – optimisation method using animal model
- Study of spectral profile for pregnancy complications using Fourier transform infrared

Abbreviations

AGT: Angiotensinogen

as: Asymmetric

BSA: Bovine Serum Albumin

CaF₂: Calcium fluoride

CE-MS: Capillary Electrophoresis-Mass Spectrometry

CESDI: Confidential Enquiry into Stillbirths and Deaths in Infancy

DTGS: Deuterated triglycerine sulphate

EMR: Electromagnetic Radiation

ESI: Electrospray Ionization

FDR: False Discovery Rate

FIR: Far Infrared

FLDA: Fisher's Linear Discriminant Analysis

FPA: Focal Plane Array

FTIR: Fourier Transform Infrared

FT-IRMS: Fourier Transform Infrared Micro-spectroscopy

HCD: High-energy Collision Induced Dissociation

HDL: High Density Lipoprotein

HPLC: High Performance Pressure Liquid Chromatography

HTA: Human Tissue Authority

IATA/UN: International Air Transport Association/ United Nation

ICAT: Isotope Coded Affinity Tags

IR: Infrared

iTRAQ: Isobaric Tags for Relative and Absolute Quantification

IUGR: Intrauterine Growth Restriction

LC-MS or LC-MS/MS: Liquid Chromatography coupled to Mass Spectrometry (or tandem mass spectrometry)

LTQ: Linear Ion Trap Mass Spectrometer

MALDI: Matrix-assisted Laser Desorption Ionisation

MCT: Mercury Cadmium Telluride (detector)

MMTS: Methylmethanethiosulfonate

MRM: Multiple Reaction Monitoring

mRNA: Messenger RNA (Ribonucleic Acid)

MS: Mass Spectrometry

MS/MS: Tandem Mass Spectrometry

MTA: Material Transfer Agreement

MWCO: Molecular Weight Cut-Off

m/z: Mass/charge ratio

NIR: Near-Infrared

NMWCO: Nominal Molecular Weight Cut-Off

PCA: Principal Component Analysis

PCs: Principal Components

PE: Preeclampsia

PIP: Proteomics In Pre-eclampsia

PIGF: Placental Growth Factor

PPSU: Proteomic Score of Urine

PTM: Post-translational Modification

Q-TOF: Quadrupole Time-of-Flight

RF: Radio Frequency

REC: Research Ethics Committee

RNA: Ribonucleic acid

ROC: Receiver Operating Characteristic

ROS: Reactive Oxygen Species

SCOPE: SCreening fOr Pregnancy Endpoints

SCX: Strong Cation Exchange

SDS: Sodium dodecyl sulfate

SELDI- TOF: Surface Enhanced Laser Desorption and Ionization Time-of-Flight

SERPINA1: SERine Protease INhibitor A1

sFlt-1: Soluble fmslike Tyrosine Kinase 1

SOP: Standard Operating Protocol

SPS: Synchronous Precursor Selection

TCEP: Tris-(2-carboxyethyl)-Phosphine

TEAB: Triethyl ammoniumbicarbonate

TMT: Tandem Mass Tag

TOF: Time-of-flight

UCC: University College Cork

UV-vis: Ultraviolet-visible

VEGF: Vascular Endothelial Growth Factor

WHO: World Health Organization

XIC: Extracted Ion Chromatogram

δ : Bending Vibrations

ν : Stretching Vibrations

Contents:

Abstract.....	i
Acknowledgements	ii
Output	iv
Abbreviations	v
Contents.....	ix
List of Figures:	xv
List of Tables:.....	xix
1. Chapter 1 General introduction	1
1.1. Introduction	2
1.2. Preeclampsia	4
1.2.1. Preeclampsia and its complications.....	5
1.2.2. Mechanisms underlying preeclampsia and known risk factors.....	6
1.3. Proteomics	7
1.3.1. Application of proteomics methodologies to the study of preeclampsia	9
1.3.2. Urinary studies of preeclampsia	9
1.4. Aims and project objectives	15
2. Chapter 2 Proteomic analysis for pregnancy complications using iTRAQ labelling coupled with mass spectrometry techniques	18

2.1.	Introduction isobaric labelling method (iTRAQ)	19
2.2.	Proteome analysis by mass spectrometry	20
2.3.	Basics of mass spectrometry	21
2.3.1.	Ionisation	22
2.3.1.1.	Electrospray ionisation (ESI)	23
2.3.2.	Instrumentation used in biological mass spectrometry	24
2.3.2.1.	Time-of-Flight.....	24
2.3.2.2.	Quadrupole analysers	26
2.3.2.3.	Hybrid orbitrap mass spectrometers.....	29
2.4.	Application of mass spectrometry in PE and IUGR studies.....	30
2.5.	Chapter aims.....	34
2.6.	Materials and methods	35
2.6.1	Chemicals	35
2.6.2.	Sample acquisition and preparation	36
2.6.2.1.	Acquisition of human urine samples.....	36
2.6.2.2.	Transportation of human tissue and storage	36
2.6.3.	Samples preparation process	37
2.6.3.1.	Centrifugation, ultrafiltration and protein purification	37
2.6.3.2.	Verification of protein concentration	38
2.6.3.3.	Tryptic digest and iTRAQ labelling	39

2.6.3.5.	Mass spectrometry set up	40
2.6.3.6.	Data processing	42
2.6.3.7.	Statistical analysis.....	43
2.7.	Methods outline workflow	43
2.8.	Results.....	44
2.8.1	Proteomic analysis for PE at 15 and 20 weeks gestational age	44
2.8.1.1.	Highly dysregulated proteins for PE at 15 and 20 weeks	49
2.8.2.	Proteomic analysis for IUGR at 15 and 20 weeks gestational age	55
2.8.2.1	Highly dysregulated proteins for IUGR at 15 and 20 weeks	60
2.9.	Discussion.....	67
2.10.	Conclusions	76
3.	Chapter 3 Fourier transform infrared analysis of urine – optimisation method using animal model	78
3.1	Introduction Infrared Spectroscopy	79
3.2	FTIR spectroscopy and Michelson interferometer	81
3.3	Spectroscopy and Beer-Lambert law relationship.....	82
3.4	FTIR spectroscopy to study biological molecules	83
3.4.1	Protein FTIR.....	84
3.4.2	Lipid FTIR.....	86
3.4.3	Carbohydrate and nucleic acid FTIR	87

3.5	FTIR spectroscopy studies of biofluids.....	87
3.6	Chapter aims	90
3.7	Materials and methods.....	90
3.7.1	Urine sample collection	90
3.7.2	Sample preparation	91
3.7.3	Ultrafiltration method	91
3.7.4	FTIR instrument set-up	92
3.7.5	Statistical analysis	93
3.7.6	Principal Component Analysis.....	93
3.7.7	Summary workflow for methods	94
3.8	Results	94
3.8.1.	Volume sensitivity investigation of FTIR analysis using whole rat urine	96
3.8.2.	Influence of ultrafiltration preparation procedure on FTIR spectra using rat urine sample.....	98
3.8.3.	Mean differences of spectral profiles for pregnant and non-pregnant rat urine at mid-FTIR	100
3.8.4.	Spectral profile of protein fingerprint region of urine from pregnant and non-pregnant rats	101
3.8.5.	Spectral profile of lipid region for pregnant and non-pregnant rat urine ...	104
3.9	Discussion	106

3.10	Conclusions	107
4.	Chapter 4 Study of spectral profile for pregnancy complications using Fourier transform infrared spectroscopy	109
4.1.	Application of FTIR in biological studies	110
4.2.	Chapter aims	112
4.3.	Materials and methods.....	112
4.3.1.	Urine sample collection	112
4.3.2.	Sample preparation	113
4.3.2.1.	Centrifugal filtrate and protein ultrafiltration protocol	113
4.3.3.	FTIR experiment setup	113
4.3.4.	Spectral transformation.....	115
4.3.5.	Data analysis	116
4.4.	Results	116
4.4.1.	FTIR spectral profile of PE and control at 15 week gestational age	116
4.4.2.	Spectral profile study of protein fingerprint region for PE and control at 15 weeks gestational age	118
4.4.3.	Spectral profile study of lipid region for PE and control at 15 weeks gestational age	120
4.4.4.	Spectral profile study for PE and control subjects at 20 weeks' gestation..	122

4.4.5.	Spectral profile study of protein fingerprint region for PE and control at 20 weeks gestational age	123
4.4.6.	Spectral profile study of lipid region for PE and control at 20 weeks gestational age	125
4.4.7.	FTIR spectral profile of IUGR and control at 15 weeks' gestation	127
4.4.8.	Spectral profile investigation of protein fingerprint region for IUGR and control urine samples at 15 weeks' gestation	128
4.4.9.	Spectral profile investigation of lipid region for IUGR and control urine samples at 15 weeks' gestation	131
4.4.10.	Spectral profiling of urine samples collected from IUGR and control subjects at 20 weeks' gestation	133
4.4.11.	Spectral profile investigation of protein fingerprint region for urine samples taken from IUGR cases and controls at 20 weeks' gestation	134
4.4.12.	Spectral profile investigation of lipid region for IUGR and control at 20 weeks' gestation	136
4.5.	Discussion	138
4.6.	Conclusions	141
5.	Chapter 5. General discussion, conclusions and future work	143
5.1	Aim and direction of work	144
5.2	FTIR optimisation for urine-based samples	144
5.3	FTIR spectral profile for PE and IUGR	147

5.4	MS process methodology	149
5.4.1	Proteomic biomarkers of PE.....	150
5.4.2	Proteomic biomarkers of IUGR	154
5.5	Conclusions	157
5.6	Future work.....	158
6.1	Appendix 1 : Material transfer agreement	179
6.2	Appendix.2: Ethical review	183
6.3	Appendix 3: Dataset of all protein expression data for PE and IUGR with their controls.....	190
6.4	Appendix 4: Overlay of all FTIR spectra for PE and control at 15 week of pregnancy	247
6.5	Appendix 5: Overlay of all FTIR spectra for IUGR and control at 15 week gestational age	248

List of Figures:

Figure 1.1:	Schematic diagram of putative preeclampsia mechanism.....	6
Figure 1.2:	Schematic image describing the major ‘-omics’ sub-disciplines.....	8
Figure 1. 3:	Schematic workflow summarising sample process of project	15
Figure 2. 1:	Generalised structure of N-hydroxysuccinimide (NHS-) iTRAQ tag and its derivatisation activity.	20

Figure 2. 2: Schematic diagram of the basic components of a mass spectrometer and their roles in MS analysis	22
Figure 2. 3: Diagram of Time Of Flight (TOF) analyser.	26
Figure 2. 4: Diagram of quadrupole mass analyser.....	27
Figure 2. 5: Diagram of reaction monitoring using a triple quadrupole analyser	28
Figure 2. 6: Schematic chart flow illustrating typical proteomic method.....	44
Figure 2. 7: Differentially expressed protein changes quantitatively for PE at 15 weeks' gestation	46
Figure 2. 8: Differentially expressed protein changes for PE cases at 20 weeks' gestation	47
Figure 2. 9: Differentially expressed protein changes for PE cases at 15 and 20 weeks of pregnancy	48
Figure 2. 10: Dysregulated proteins comparison between 15 and 20 weeks gestational age of PE	55
Figure 2. 11: Differentially expressed protein changes quantitatively for IUGR at 15 week of gestation	57
Figure 2. 12: Differentially expressed protein changes quantitatively for IUGR at 20 week of gestational age	58
Figure 2. 13: Differentially expressed protein changes quantitatively for IUGR cases at 15 and 20 weeks' gestation.....	59

Figure 2. 14: Dysregulated proteins comparison between 15 and 20 weeks gestational age of IUGR	66
Figure 3. 1: IR spectra of CO ₂ molecule and its vibrational modes	80
Figure 3. 2: Schematic diagram of a Michelson interferometer	81
Figure 3. 3: Incident light vs. transmitted	83
Figure 3. 4: Typical FTIR absorption spectrum of a biological sample	84
Figure 3. 5: Typical structure of amide bond via amino acids linking	85
Figure 3. 6: Summary workflow protocol optimisation and application.....	94
Figure 3. 7: FTIR spectra at particular range 1800-1500 cm ⁻¹ of mid region using whole rat urine	95
Figure 3. 8: Bar graphs of FTIR analysis at (1850-1500) cm ⁻¹ range.....	97
Figure 3. 9: Typical FTIR spectra comparison of sample preparation in different conditions	99
Figure 3. 10: Mean differences of FTIR spectra intensity for rat urine from pregnant and non-pregnant females	101
Figure 3. 11: PCA plot for fingerprint protein region of pregnant and non-pregnant rat urine	103
Figure 3. 12: PCA plot for lipid region of pregnant and non-pregnant rat urine.....	105
Figure 4. 1: FTIR instrument and its sampling platform.	115
Figure 4. 2: Summary workflow of methods	116

Figure 4. 3: Averaged spectral profile (means) for PE and control at 15 week gestational age at mid FTIR range. (n = 30) samples.....	117
Figure 4. 4: PCA plot of PE and control during 15 week of gestation at protein region of 1800-1000 cm-1	119
Figure 4. 5 : PCA and loadings plot of lipid region of (3100-2700 cm-1) for 15 week gestation samples in PE vs. control subjects	121
Figure 4. 6: Averaged mid FTIR spectral profiles for PE and control samples at 20 weeks' gestation	123
Figure 4. 7: PCA plot protein region (1800-1000 cm-1) for PE and control samples at 20 weeks' gestation.....	124
Figure 4. 8: PCA plot of PE and control during 20 week of gestation at lipid region of 3100-2700 cm-1	126
Figure 4. 9: Mean spectral profiles for mid-IR analysis of IUGR and control urine samples at 15 weeks' gestation	128
Figure 4. 10: PCA plot of IUGR and control urine samples collected at 15 weeks' gestation for protein region 1800-1000 cm-1.....	130
Figure 4. 11: PCA plot of IUGR and control urine samples collected at 15 weeks' gestation over lipid region of 3100-2700 cm-1.....	132
Figure 4. 12: Spectral profile comparison of mean FTIR of IUGR and control samples at 20 weeks' gestation	133
Figure 4. 13: PCA plot of IUGR and control urine samples collected at 20 weeks' gestation over protein region, 1800-1000 cm-1	135

Figure 4. 14: PCA plot of IUGR and control subjects for 20 week’s gestation using lipid region of 3100-2700 cm-1	137
Figure S 1: Overlay of FTIR Spectra for PE and control at 15 week of gestational age at wavenumber range (4000-600) cm-1.....	246
Figure S 2: Overlay of FTIR Spectra for IUGR and control urine samples at 15 weeks’ gestation, 4000-600 cm-1.....	247

List of Tables:

Table 2. 1: Chemicals and suppliers	35
Table 2.2: Labelling reagents with all samples of each PE and IUGR samples. iTRAQ® Reagent – 8plex - Buffer Kit materials (Sciex, USA), S: Case, Ctl: Control	39
Table 2.3: Dysregulated proteins dataset of PE at 15 week gestational age compared with control.....	50
Table 2.4: Dysregulated proteins dataset of PE at 20 week gestational age compared with control.....	52
Table 2.5: Dysregulated proteins dataset of IUGR at 15 week gestational age compared with control.....	61
Table 2.6: Dysregulated proteins dataset of IUGR at 20 week gestational age compared with control.....	63
Table 2.7: Proteomic biomarkers for PE.....	70
Table 2. 8: Proteomic biomarkers for IUGR.....	74

Table 3. 1: Major FTIR bands of macromolecules described in urine	98
Table 4. 1: Assignment of main FTIR absorption bands of plasma	110
Table S 1: Dataset of protein expression for PE and control at 15 weeks gestational age	190
Table S 2: Dataset of protein expression for PE and control at 20 weeks gestation.....	204
Table S 3: Dataset of protein expression for IUGR and control at 15 weeks of pregnancy	218
Table S 4: Dataset of protein expression for IUGR and control at 20 weeks of gestation	232

Chapter 1

1. General Introduction

1.1. Introduction

The most common disorders associated with placental implantation abnormality in early pregnancy are pre-eclampsia (PE) and intrauterine growth restriction (IUGR). Both PE and IUGR are multifactorial conditions. PE is a multi-system disorder of pregnancy characterised by new onset of hypertension and elevated urine protein concentration (Law *et al.*, 2015). Other types of hypertensive disorders of pregnancy, such as gestational hypertension and chronic hypertension can occur during pregnancy; the impact of these is rarely as acute or life-threatening. By contrast, PE represents a major cause of maternal and fetal morbidity and mortality (Duley, 2009). PE improves immediately after delivery, but currently no specific therapeutic treatment is available for PE. Prophylactic administration of low dose aspirin has however been demonstrated to be of clinical use, if administered to high-risk patients during their first-to-second trimester (6-16) weeks of pregnancy; this strategy reduces overall risk of PE by approximately 50 %, and achieves ~80% reduction in severe PE incidence (Bujold *et al.*, 2014). However, prediction of individual risk remains elusive.

Various studies (Blankley *et al.*, 2013; Buhimschi *et al.*, 2008) have investigated proteomic biomarkers in preeclampsia. Due to the complexity of the proteome, various separation strategies have been employed prior to MS. Various biological samples have been used to study biomarkers in preeclampsia, such as serum (Auer *et al.*, 2010), placental tissue (Wang *et al.*, 2013), and amniotic fluid (Oh *et al.*, 2012). No clinical test has, however, been developed to accurately predict preeclampsia risk for nulliparous women with high risk (Conde-Agudelo *et al.*, 2004).

Urine and blood are two major relevant sources of material for proteomic studies. The choice of analyte is dependent on a variety of factors. For example, urine samples are considerably easier to obtain in a larger quantity than blood; collection of urine is a non-invasive technique, however the reproducibility of urine sampling may be lower than that of blood (Carty *et al.*, 2011). Many studies (Schaub *et al.*, 2004; Theodorescu *et al.*, 2006) have shown that proteins in biologic fluids other than urine could degrade quickly when handled inappropriately, whereas urinary proteins have been shown to remain stable long enough to perform reliable proteome analysis. The urinary proteome did not undergo significant alterations when urine was stored at 4° C for 3 days or at room temperature for 6 h (Theodorescu *et al.*, 2006). Urinary proteins are typically resistant to hydrolysis, especially if stored in low temperature (-80° C) for years, without significant observed proteome alterations (Theodorescu *et al.*, 2006; Schaub *et al.*, 2004). This increased stability is in stark contrast to the blood proteome; further, the collection of blood for serum or plasma analysis is relatively invasive. Serum or plasma has an extreme dynamic range of proteins that may affect data quality and comparability, and necessitates pre analytical treatment such as immune-affinity depletion, which are highly costly (Kolch *et al.*, 2005; Fliser *et al.*, 2007). Developing a non-invasive, clinically-acceptable and robust urine test which enables risk prediction for pregnant women at a sufficiently early gestational stage to enable intervention is therefore a desirable goal. Standard operating protocols (SOP) are therefore necessary for collection, storage, processing, and urine sampling (Fliser *et al.*, 2007).

1.2. Preeclampsia

Preeclampsia (PE) affects around 5-8% of pregnancies globally, and is the most significant hypertensive disorder of pregnancy, impacting both maternal and neonatal health (Van Lerberghe *et al.*, 2005). The World Health Organisation has reported an annual death toll of > 50,000 women globally due to preeclampsia; the condition is also a major cause of preterm birth (Van Lerberghe *et al.*, 2005). Diagnosis of PE is characterised by new onset hypertension ($\geq 140/90$ mmHg) and proteinuria (≥ 300 mg in 24 hr urine) (Aggarwal *et al.*, 2005, Van Lerberghe *et al.*, 2005). Symptoms of PE range from mild to severe, and include headache, abdomen pain, renal and liver failure (Jeyabalan, 2013). Unfortunately, the diagnosis of PE is frequently only made in late pregnancy, with clinical signs mainly manifesting during the third trimester (Law *et al.*, 2015). The progress of PE from mild to severe can be rapid and unpredictable, with serious, potentially fatal consequences (Law *et al.*, 2015). Eclampsia and HELLP syndrome are often associated with this condition; both are potentially life-threatening complications. HELLP syndrome is characterised by haemolysis, elevated liver enzymes, and low platelet counts; eclampsia features grand mal-like seizures (Law *et al.* 2015). A number of risk factors for PE have been identified, including ethnicity and family history (Caughey *et al.*, 2005). Other risk factors are also associated with PE such as obesity, diabetes, pre-existing hypertension, renal disease and autoimmune disease (Jeyabalan, 2013). However, there remain no diagnostic tests for women with their first time pregnancy (Myers *et al.*, 2013).

1.2.1. Preeclampsia and its complications

PE is a major cause of preterm birth, *in utero* growth restriction (IUGR), and neonatal death. Intrauterine growth restriction (IUGR) is a condition of poor fetal growth, which is commonly recognized as a weight below the 10th percentile for the gestational age (small for gestational age, SGA) (Battaglia and Lubchenco, 1967). IUGR is usually diagnosed during a routine prenatal exam in the second half of pregnancy. Diverse causes are often associated with low birth weight, with some cases of IUGR being etiologically linked to preeclampsia based on defects in implantation or placentation (Villar *et al.*, 2006). Association has been found with other factors such as, maternal malnutrition, maternal smoking, gestational hypertension, or congenital malformations (Villar *et al.*, 2001, Camm *et al.*, 2010, Langley-Evans, 2009). Nevertheless many IUGR cases remain unexplained (Villar *et al.*, 2001). Preterm birth is defined as the delivery of a live infant before 37 weeks gestation (Law *et al.*, 2015). Preterm delivery is associated with elevated risk of lifelong adverse consequences, including cerebral palsy, neurological disorders, language and learning disabilities (Irgens *et al.*, 2001). Furthermore, incomplete organ development of the premature infants can lead to poor organ development in future; for example, if born growth restricted, the child will have an increased propensity for hypertension, heart disorders, and diabetes in adulthood (Irgens *et al.*, 2001; Davis *et al.*, 2012). The mortality rate for fetuses in pregnancies complicated by PE is 3 to 10-fold higher than uncomplicated pregnancies (Van Lerberghe *et al.*, 2005). PE may be responsible for 12–25 % of babies with IUGR, and around 15–20 % of all preterm births (Duley, 2009, Goldenberg and Rouse, 1998).

1.2.2. Mechanisms underlying preeclampsia and known risk factors

The exact mechanism of preeclampsia remains unknown, but a commonly held hypothesis describes two distinct phases in the development of the condition (Carty, 2012). The first, initiation, is caused by disruption to normal placentation; this is followed by a generalised inflammatory response, resulting in endothelial dysfunction, proteinuria, and multi-organ damage (Sankaralingam *et al.*, 2006). Oxidative stress, caused by placental hypoxia, leads to release of trophoblast debris and soluble factors such as the angiogenic markers soluble endoglin (sENG) and fms-like tyrosine kinase 1 (sFlt-1) into maternal circulation (see Figure 1.1) (Carty *et al.*, 2011; Marusic *et al.*, 2013).

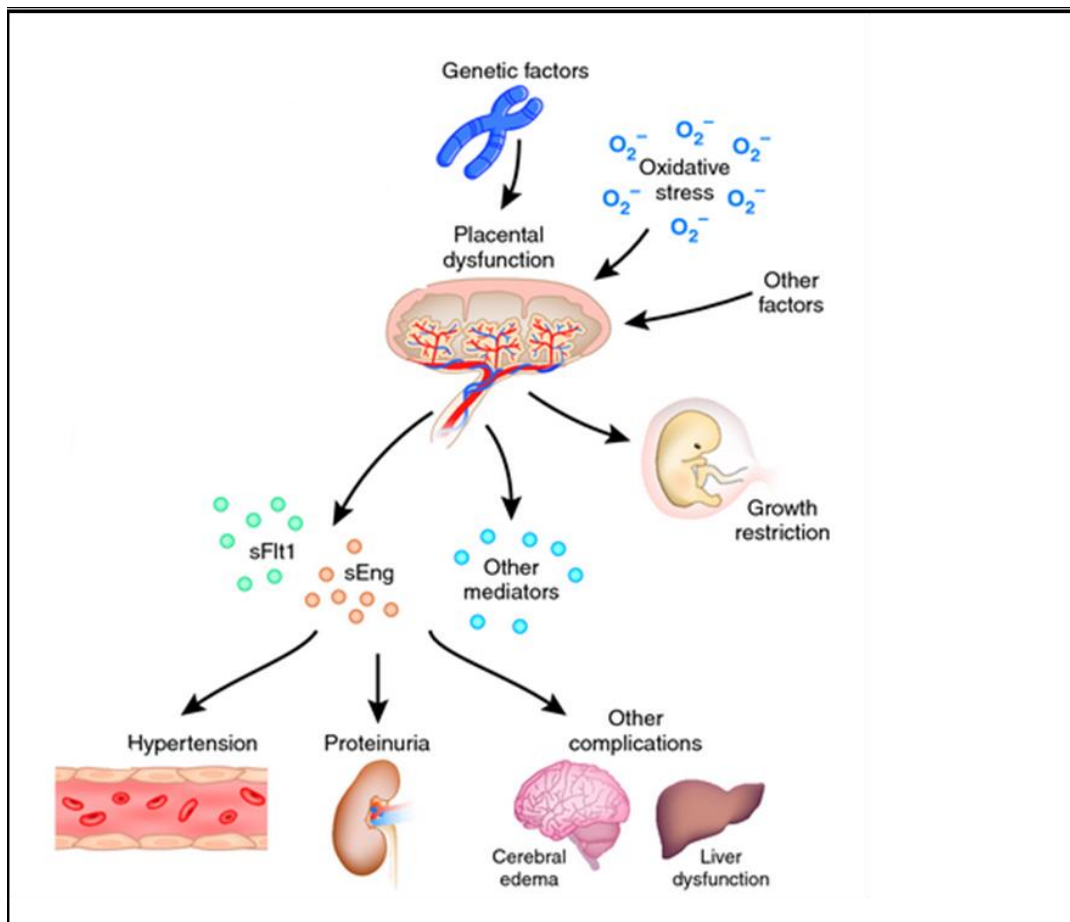


Figure 1.1: Schematic diagram of putative preeclampsia mechanism.

Disruption in placentation (implantation) leads to inflammatory response, endothelial dysfunction, growth restriction, proteinuria; placental hypoxia, results in oxidative stress, causing release of trophoblast debris and soluble factors such as the angiogenic markers soluble endoglin (sENG) and fms-like tyrosine kinase 1 (sFlt-1) into maternal circulation. These changes induce cell damage and other complications within body organs. This figure was edited from (Parikh and Karumanchi, 2008).

1.3. Proteomics

In 1995, the term '*proteome*' was described by Marc Wilkins as, "... *all proteins expressed by a genome or tissue*". This definition was later extended to include the component of time, as protein expression varies dynamically dependent upon cellular circumstances (Blackstock and Weir, 1999). The umbrella term "*-omics*" encompasses the disciplines used to discover the molecular basis, relationships, and activities of the different molecular entities which comprise the biochemistry of cells and organisms. There are four principal sub-domains within -omics: genomics, transcriptomics, proteomics and metabolomics, with "interact-omics" being used variably to ascribe functional relationships either between or within sub-domains. Genomics refers to genetic information, typically DNA sequences stored as database entries. The genome of an organism is essentially a static entity (Plebani, 2005). Transcriptomics refers to the production of gene products (transcription), i.e. the generation of messenger RNA. Many studies have confirmed that gene expression alone cannot provide full information about the behaviour of a cell or system. For instance, Lawrie and co-workers demonstrate little to no correlation between mRNA and protein levels (Lawrie *et al.*, 2001). A later study by Ghazalpour and colleagues describes the relationship between transcript (RNA) and

protein levels in mouse, with relatively little concordance being observed between RNA levels and the corresponding protein levels in response to DNA perturbations (Ghazalpour *et al.*, 2011). Metabolomics refers to the global study of small molecule levels as readouts for cellular metabolism (Barh *et al.*, 2011). Figure 1.2 illustrates general fields of study in ‘-omics’.

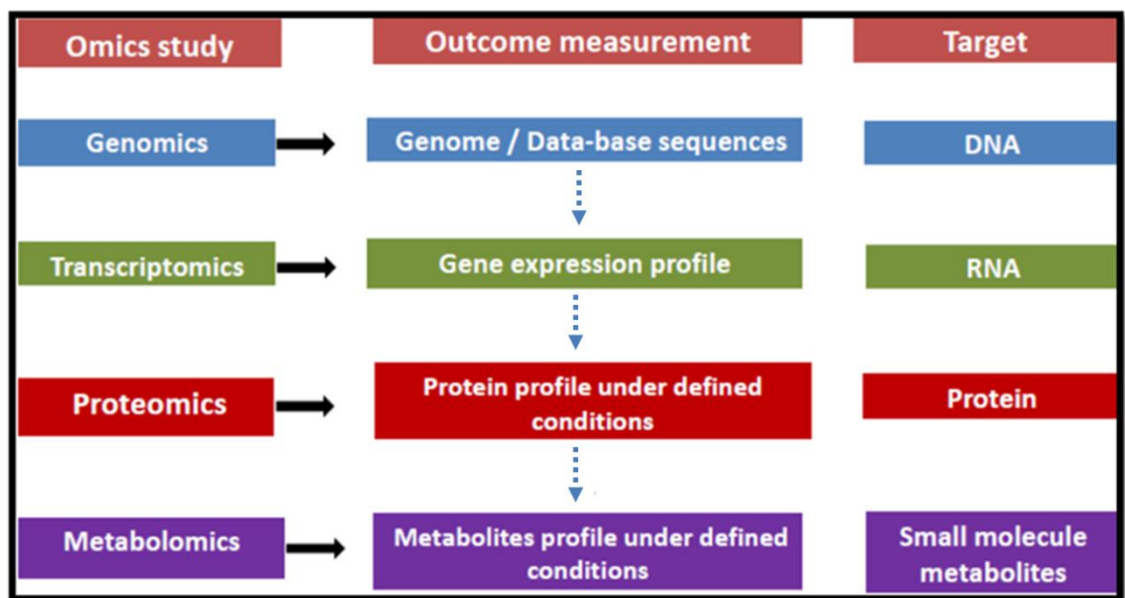


Figure 1.2: Schematic image describing the major ‘-omics’ sub-disciplines

Major ‘-omics’ terms include genomics, transcriptomics, proteomics and metabolomics. Whilst there is a broad relationship between these domains, there is also a lack of **direct** correlation between levels of analyte.

Higher organisms such as humans have multiple organs, each comprised of many cells. Each cell has various proteins, which play a vital role in physiological functions. The levels of protein show dynamic response to cellular stress via changes in expression, post-translational modification and/or protein degradation (Plebani, 2005). Post-translational modifications such as phosphorylation, glycosylation acetylation, sulphation and

methylation, and non-enzymatic chemical damage such as, glycation, oxidation and nitration increase protein heterogeneity and influence stability (Plebani, 2005). Proteins showing altered levels or changes in their modification status may therefore be considered as biomarkers for disease conditions. As such, proteomics has attracted significant interest in the wider field of clinical research. Discovery and validation of putative clinical protein biomarkers, remains challenging, with some biomarkers being seldom observed or poorly validated (Plebani, 2005).

1.3.1. Application of proteomics methodologies to the study of preeclampsia

A number of studies have used proteomic methods to identify potential biomarkers which predict preeclampsia. Various patient-derived materials (urine, serum, placenta, amniotic cervical-vaginal fluid) have been investigated. These studies have highlighted a number of candidate biomarkers that could discriminate complicated pregnancies from uncomplicated. The following section introduces the use of methodologies by presentation of different studies to detect potential biomarkers, specifically focussing on urine sampling.

1.3.2. Urinary studies of preeclampsia

A urinary proteomics study by Chen *et al.*, (2011) used iTRAQ labelling coupled with 2D LC-MS/MS to investigate complicated hypertensive pregnancies (Chen *et al.*, 2011).

Seventy-one women were enrolled in this study; 30 women in the exploratory arm and 41 women in a validation set which was used to validate potential biomarkers. Three groups of women were described in the exploratory set: preeclampsia (n=10), gestational hypertension (n=10) and normal pregnancy (n=10). 362 Non-redundant proteins were identified in total, of which, 113 were expressed differentially between complicated hypertension and normal pregnancy. 31/113 proteins showed statistically significant difference between the three study groups. Gene ontology analysis of the differentially expressed proteins identified that many of these markers had significant roles in metabolic pathogenesis processes related to preeclampsia such as blood coagulation, cell adhesion, cytoskeleton remodelling, and immune response. Hierarchical clustering demonstrated differential proteins in preeclampsia and gestational hypertension cases compared to normal pregnancy. Among these, keratin [type I cytoskeletal 19, and keratin 4], α -actinin-4, afamin, α -2-macroglobulin, serum albumin (ALB), isoform 1 of α -1-antitrypsin precursor (SERPINA1), vimentin and tubulin β -2C chain were clustered together. AGT (angiotensinogen), SERPINA1 and ALB showed significant expression differences in both preeclampsia and gestational hypertension compared to normotensive pregnancies. SERPINA1 and ALB were down-regulated in gestational hypertension, but up-regulated in preeclampsia compared to normal pregnancy. These results are in agreement with other studies (Buhimschi *et al.*, 2008). Receiver Operating Characteristic (ROC) and ELISA were used to validate urinary AGT down regulated in preeclampsia (Chen *et al.*, 2011).

In 2005, Buhimschi *et al.*, used immunoassay to study urinary protein biomarkers such as soluble fms-like tyrosine kinase 1 (sFlt-1), vascular endothelial growth factor (VEGF), and

placental growth factor (PIGF) (Buhimschi *et al.*, 2005). 68 Women were enrolled in this study and classified into four groups: healthy non-pregnant women (n=14), healthy pregnant control (n=16) at gestational age range (3-39) weeks, mild preeclampsia (n=21) at gestational age range (24-40) weeks, severe preeclampsia (n = 17) at gestational age range (16-40) weeks. In women with severe preeclampsia, the level of antiangiogenic factor sFlt-1 was increased, but PIGF level showed decreases at the time of clinical manifestation. Urinary output of VEGF was raised significantly in severe preeclampsia compared with non-pregnant women, but did not show any alteration among other pregnant groups. Moreover, the log [sFlt-1/PIGF] ratio showed 88.2% sensitivity and 100% specificity in differentiating women with severe preeclampsia from normotensive controls. Authors suggested that severe PE is associated with the increased urinary excretion of antiangiogenic factor sFlt-1 and decreased of PIGF output at the time of clinical manifestation based on sFlt/PIGF ratio may be providing a non-invasive hypertension women screening measurement for severity of the disease (Buhimschi *et al.*, 2005).

In 2008, Buhimschi, *et al.*, used tandem mass spectrometry surface enhanced laser desorption and ionization time of flight mass spectrometry (SELDI-TOF-MS) analysis to study urinary biomarkers in preeclampsia (Buhimschi *et al.*, 2008). Urine samples were collected from 284 women, classified according to having high and low risk related condition, and severity of preeclampsia: control pregnant women (n=18) at gestational age at 25 week of range (7-41) weeks, pregnant women with chronic hypertension (n=26), gestational age 33 (24-40) weeks, mild PE (n=29), gestational age 36 (24-40) weeks, severe PE (n=31), gestational age 36 (24-40) weeks, superimposed PE (n=28)

gestational age 34 (16-39) weeks, pregnancy-associated conditions unrelated to PE (n=64) at gestational age 28 of range (21-34) weeks and non-pregnant proteinuric women (n=10). The study was divided into three phases, exploratory; challenge, and translation. 59 Samples were used in the exploratory phase, where proteomic profiles were determined in severe preeclampsia that mandated delivery. Here 13 putative biomarker peaks (denoted P1-P13) were found. In the challenge phase, samples taken from 225 women were used for validation screening for both high and low risk related conditions, including preeclampsia. Here, 169 women had preeclampsia, with 86 having mandated deliveries for preeclampsia, of which 73 % occurred preterm, and in 50% of cases before 34 weeks of gestational age. Urinary [protein/creatinine and sFlt-1/PlGF] ratio measurement was used as an indicative marker in 19 women with severe preeclampsia. Peptide fragments of human SERPINA1 (SERine Protease INhibitor A1) were found to correspond to putative biomarkers: P1, P2, P3, P5, and P7; and albumin using MS/MS. The last phase (translational), identification of biomarker by tandem mass spectrometry and validation experiments in urine, serum, and placenta were employed to identify, quantify, and localize the biomarkers or related proteins (Buhimschi *et al.*, 2008).

Another study by Lee *et al.*, (2001), using SELDI-TOF-MS to analyse urinary protein profile to distinguish between the severity of preeclampsia cases, compared with uncomplicated pregnancy (Lee *et al.*, 2011). Urine samples from 26 women, categorised as having severe preeclampsia (n=11), mild preeclampsia (n = 7), or as normotensive (controls, n=8) were compared. A scoring system was constructed, designated as Preeclampsia Proteomic Score of Urine (PPSU) was used to differentiate severe preeclampsia from mild preeclampsia and uncomplicated pregnancy controls. They identified four distinctive

protein-related peaks (m/z ratio: 4155, 6044, 6663, and 7971), all of these were down-regulated in the severe preeclampsia. PPSU<2 had shown 90.9% sensitivity and 93.3% specificity (Lee *et al.*, 2011).

Proteomic analyses using ionisation methods based on laser desorption/ionization (LDI) has shown limited reproducibility, specificity and selectivity (Tiss *et al.*, 2007, Stimson *et al.*, 1997). Ultimately this led to a loss of faith by many in the clinical community for these sorts of discovery-based methods.

In 2011, Carty and co-workers used CE coupled with TOF-MS to identify proteomic biomarkers in urine samples obtained from two independent cohort studies, SCReening fOr Pregnancy Endpoints (SCOPE) and Proteomics in Preeclampsia (PIP) (Carty *et al.*, 2011). Candidate biomarkers were identified via LC-MS/MS. Outcomes showed a panel of 284 pregnancy-specific proteomic biomarkers when compared with non-pregnant controls. A reported model of 50 biomarkers from cases obtained at late stage (week 28) showed an association with PE development. The biomarkers of interest found included fibrinogen alpha chain, collagen alpha chain, and uromodulin fragments. The authors of this study suggested that the biomarkers provided a good confidence to predict preeclampsia at later gestational stage (week 28), but not reliably enough at earlier stages (weeks 12–16 and 20). The observation of different results in two cohort studies is perhaps unsurprising. The two studies (SCOPE vs PIP) had different subject gestational ages for sampling (20 vs. 28 weeks) and different study populations; for SCOPE the control subjects were healthy nulliparous women, whilst PIP recruited a mixed-parity general obstetric population (Carty *et al.*, 2011).

In 2015, Hart *et al.*, used iTRAQ method coupled with LC-MS/MS to identify proteomic biomarkers in urine samples in nulliparous women with PE (Hart *et al.*, 2015). This work was a proof-of-concept study as a potential application for the PE prediction. Samples obtained through the SCOPE consortium were collected at 15 and 20 weeks gestational age. Master pools were compared from uncomplicated pregnancies (control) to preeclampsia cases. The samples were subjected to ultracentrifugation to separate samples over 10 kDa molecular weight cut-off membranes. Bradford assay was performed before iTRAQ labelling and LC-MS/MS analysis. Around 400 proteins were observed at a 1 % false discovery rate (FDR) cut-off, of which 150 proteins had a sufficient reporter ion signal for high-confidence quantification. Some of these proteins had changes in expression in women who subsequently developed preeclampsia such as, perioplatin 1, growth and differentiation factor 15 and ICOS ligand 1. The authors here concluded that iTRAQ labelling with LC-MS/MS analysis presented a method for low-to-medium throughput for identification and quantification of proteins (Hart *et al.*, 2015). The present project was therefore developed based on this previous proof-of-concept study by (Hart *et al.*, 2015) in larger cohort of women complicated with PE and IUGR using both 15 and 20 weeks gestation samples, paired with normotensive controls at each gestational time point; again, the samples were obtained through SCOPE consortium. Methodology development is used for identification / quantification of detected proteomic biomarkers. Figure 1.3 illustrates general steps for sampling process of the project.

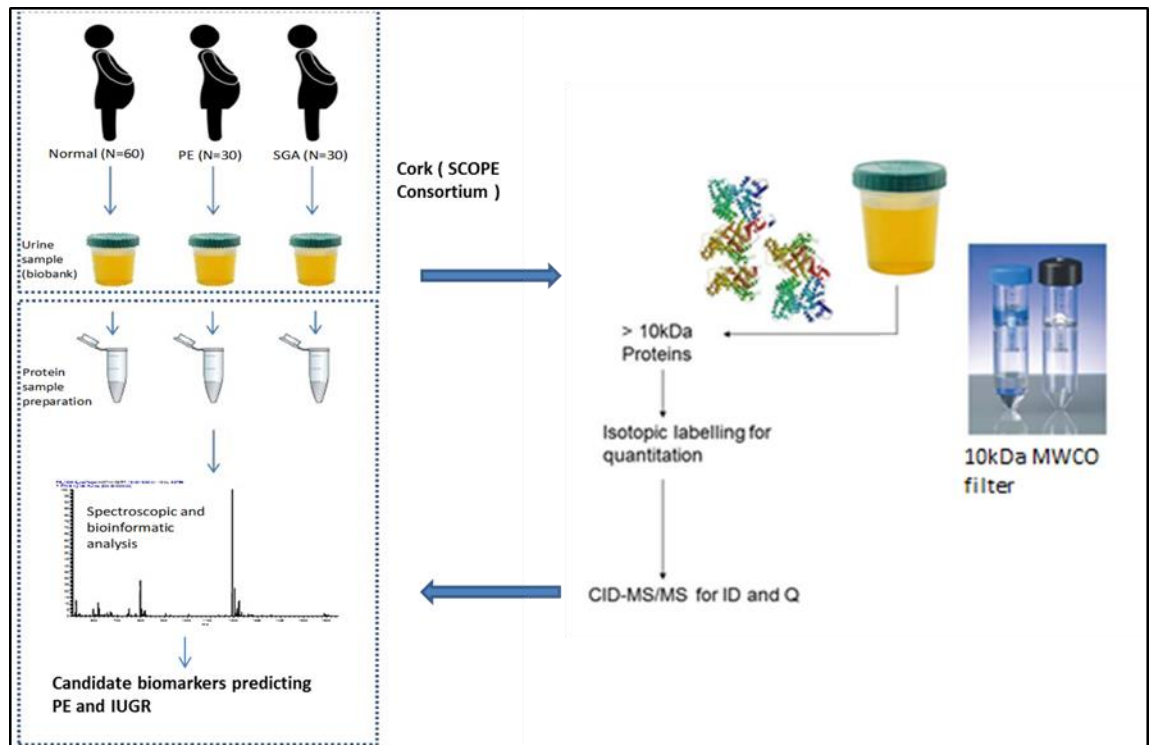


Figure 1. 3: Schematic workflow summarising sample process of project

Urine samples (240 in total) were collected from pregnant women, either uncomplicated or affected by either PE (n=30) or IUGR (n=30). Samples were obtained at 15 and 20 weeks gestation, and paired with their control at each gestational time point. All sample collection was via the SCOPE consortium (University College Cork). Sample preparation, data collection and bioinformatic analysis constitute the steps undertaken in this project. This figure was edited from (Hart *et al.*, 2015).

1.4. Aims and project objectives

Throughout this study our aims are to identify novel proteomic biomarkers that can predict in early pregnancy those who will subsequently develop late pregnancy complications such as PE or / and *in utero* growth restriction (IUGR). The hypothesis of this study is that a difference in protein presentation will be observed between control urine and that collected from women found to present clinical signs PE and IUGR. Diagnostic biomarkers could aid the development of clinical tests better predicting

maternal / fetal risk of preeclampsia. Such tests could potentially target increased monitoring frequency to at-risk first time pregnancies, and if accurate reduce rate of deaths from preeclampsia, and increase survival rate and life chances for (premature and/or small) babies. Such a test would help decrease the complex fiscal, social and emotional burden of these pregnancy related complications upon individuals and society. The aims of the series of experiments contained in this thesis are highlighted as below:

Chapter 2: Proteomic analysis for pregnancy complications using iTRAQ labelling coupled with MS techniques

The aims of this chapter are to investigate proteomic biomarkers at urine to predict PE and IUGR disorders earlier. Initial aim is to establish a Standard Operating Protocol (SOP) based on protein molecular weight cut-off membranes > 10 kDa at urine sample for proteomic analysis. Mass spectrometry-based quantitative proteomics (typically iTRAQ) strategies are used to identify robust biomarkers in PE and IUGR. These experiments aimed to generate MS / MS conditions that provide high-quality product ion spectra, enabling high-throughput sample analysis and unsupervised searching of sequences databases to ascribe product ion data to protein identity. This strategy enabled identification of peptides / proteins showing reporter ion abundance changes to investigate dysregulated urinary proteins for pregnant women with PE and IUGR at 15 and 20 weeks gestational age respectively.

Chapter 3: FTIR analysis of urine – optimisation method using animal model

FTIR spectroscopy was used to study spectral profile on rat urine samples. The main aim of this chapter is to establish a procedure for use with urine samples, using an animal

model for optimisation experiments. The use of locally available animal-derived materials allowed us to preserve the valuable and limited volume human samples. Rat urine samples were available locally and with no significant impact on the animals involved as this is their waste. The chapter aim is to investigate volume sensitivity vs spectrum detection using whole urine sample of rat using FTIR spectroscopic analysis. Moreover, methodology development and optimisation are presented for FTIR spectroscopy used to identify the spectral profile of urine samples. The processes evaluated included separation of urine into protein fractions with specific molecular weights. Optimised workflow was used for clinical urine samples collected from the international SCOPE biobank of both PE and IUGR.

Chapter 4: Study of spectral profile for pregnancy complications using FTIR spectroscopy

This chapter aims to identify spectral profile of both PE and IUGR using FTIR technique. Optimised workflow that was established in Chapter 3 was used for sample preparation protocol of clinical sample. Sample was collected from the international SCOPE biobank of both PE and IUGR. This enabled the investigation of sample spectral profiling for both PE and IUGR during 15 and 20 week of gestation, and compared with control respectively using FTIR.

Chapter 2

2. Proteomic analysis for pregnancy complications using iTRAQ labelling coupled with mass spectrometry techniques

2.1. Introduction isobaric labelling method (iTRAQ)

Mass spectrometry-based quantitative proteomics strategies can be classified in two major categories: firstly, label-free methods, based on integral MS signal of each (protein/peptide); and alternatively those approaches using stable isotope labelling, which introduce a mass difference between two proteomes or produce an internal standard for relative quantification information (Ong and Mann, 2006). Isobaric labelling methods (iTRAQ and TMT) use a common approach, labelling labile functional groups upon protein/peptide molecules, with identification and quantification being performed using various types of tandem mass spectrometer. iTRAQ and TMT employ identical target functional groups upon tryptic peptides, attacking the primary amines of peptide/protein. Both tags employ an N-hydroxysuccinimide (NHS-) activated group; this moiety reacts with N-terminal amine groups (primary amine) and with ϵ -amine groups of lysine residues to covalently attach the tags to the peptides. (Figure 2.1) illustrates the chemical structure of the 4-plex iTRAQ variant isobaric mass tagging reagent. The NHS-activated group renders most biological samples susceptible to these labelling approaches. Labelling will not occur at modified peptide N-termini, such as N-terminal cyclised glutamic acid (pyro-glutamic acid), or N-terminal acetylated peptides (Rauniyar and Yates III, 2014). The earliest experiments employed MALDI-TOF-TOF or quadrupole time-of-flight (Q-TOF) instruments (Alvarez *et al.*, 1997, Thompson *et al.*, 2003). Time-of-flight (TOF) mass spectrometry is particularly suited to iTRAQ/TMT methods, due to the ability of TOF mass analysers to detect low m/z fragment ions (the region where reporter ions are observed) (Fuller and Morris, 2012).

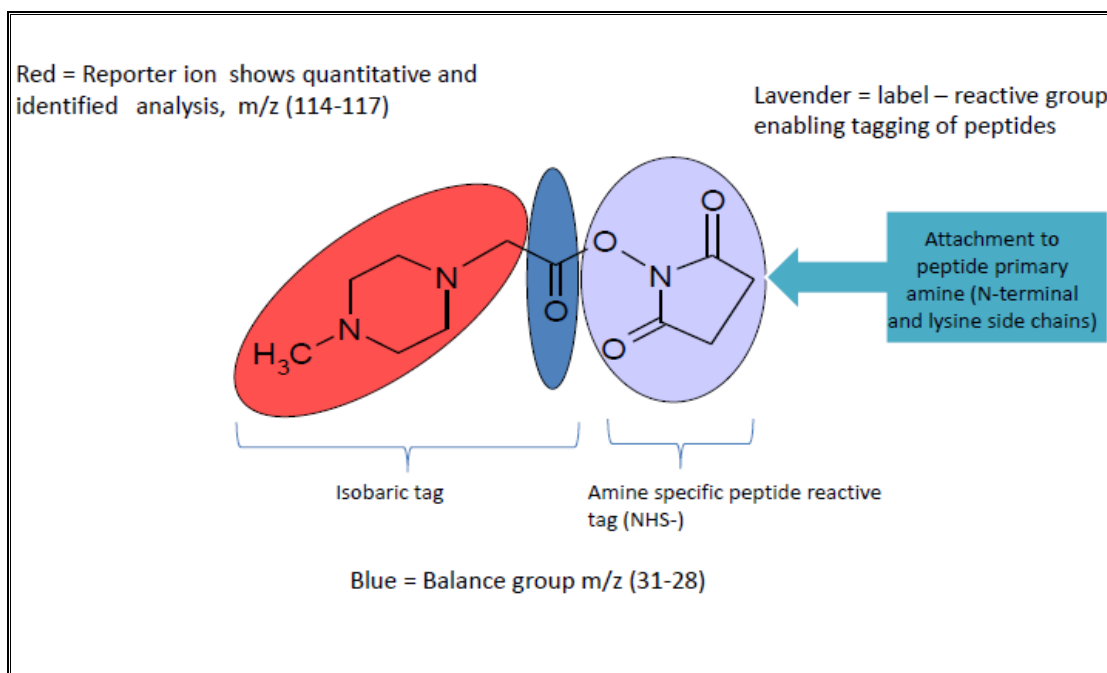


Figure 2. 1: Generalised structure of N-hydroxysuccinimide (NHS-) iTRAQ tag and its derivatisation activity.

This compound consists of an isobaric tag (reporter ion and balancer group), and NHS- tag group that reacts with free N-terminal amines and ϵ -amine of lysine residues of the peptides, allowing easy peptide identification (based on product ions resulting from backbone cleavage) and relative quantification (based upon reporter ion intensity).

2.2. Proteome analysis by mass spectrometry

Several proteomic methods can be used successfully for analysis of biomarkers (Gevaert *et al.*, 2003; Bergman and Bergquist, 2014). Mass spectrometry (MS) is a method commonly used to analyse proteins (Han *et al.*, 2008). Mass spectrometry is a primary methodology providing an effective means to chemical characterisation, with a number of desirable characteristics, such as unbiased methods, rapid, high sensitivity and high throughput (de Hoffmann, 1996; Dass and Brodbelt, 2001). Quantitative proteomics strategies highlight variations between experiment and control samples (Ong and Mann,

2006). Proteomic analyses can be broadly divided into either by '*top down*' approaches, where the intact protein is the target, or '*bottom up*' approaches where digested proteins are investigated. The latter is often used to characterise proteins (either isolated or as mixtures), which are subjected to site-specific proteolysis using enzymes (typically trypsin) (Wingert *et al.*, 2015).

2.3. Basics of mass spectrometry

Mass spectrometry (MS) plays a vital role in proteomic studies. Mass spectrometers typically consist of three essential components: ionisation source, mass analyser and detector. Figure 2.2 illustrates a generalised schematic of a mass spectrometer, describing the main components and their basic functions.

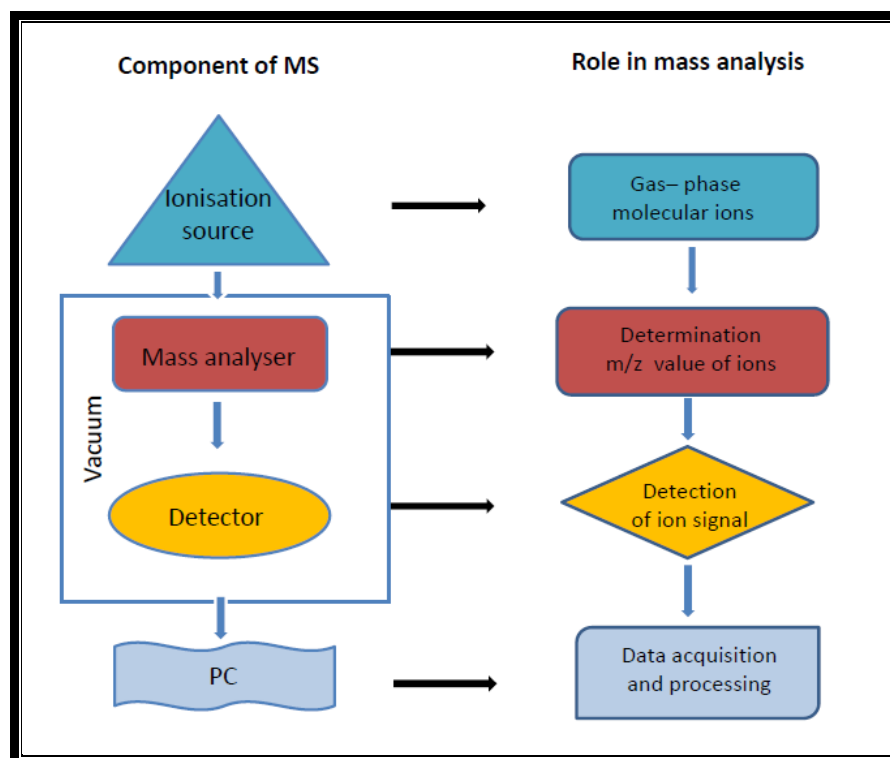


Figure 2. 2: Schematic diagram of the basic components of a mass spectrometer and their roles in MS analysis

Mass spectrometers generally consist of three major components: ion source, mass analyser and detector. Molecules are volatilised and induced into an ionised state within the ion source, allowing analysis according to m/z ratio (detected as ion signals). Acquired data are interrogated using computerised algorithms. Edited from (Yates, 2000).

2.3.1. Ionisation

In order to perform analysis of the molecular content of samples according to mass/charge (m/z), molecules must firstly undergo conversion into the gas-phase and ionisation. To be able to detect ions, they must remain intact into the mass analyser, therefore for large, in-volatile and fragile biological molecules such as proteins and peptides, low energy transfer or 'soft' ionisation techniques must be used (de Hoffmann,

1996). Two principal types of ionisation source are commonly used for this purpose, namely electrospray ionisation (ESI) (Nguyen and Schug, 2008) and Matrix Assisted Laser Desorption/Ionization (MALDI) (Korte *et al.*, 2015). ESI is the main technique which is used in this project.

2.3.1.1. Electrospray ionisation (ESI)

ESI is a method which enables the generation of intact molecular ions from liquid phase samples by application of electrical potential. In the late 1980s, Professor John Fenn's group developed electrospray ionisation (ESI) (Fenn *et al.*, 1989). This technique enables the generation of gas-phase ions via the application of an electrical charge to solution-phase sample (typically in an acid-containing organic solvent) (Gaskell, 1997). The organic solvent is volatile and thus readily evaporates by heating, increasing charge density in the resultant electrospray droplets, which causes electrostatic repulsion and thus plume expansion (Gaskell, 1997). There are a number of aspects of ESI which are attractive for characterization of biological samples; the technique has high sensitivity (femto- to attomolar, depending upon the mass analyser), it is readily coupled to high pressure liquid phase (HPLC), and further is amenable to miniaturization for further reductions in sample consumption and thus increased sensitivity. Peptide ions generated from tryptic hydrolysates typically bear multiple charges (Nguyen and Schug, 2008); this property, together with low internal energy of precursors, means that products of collisional dissociation occupy lower energy pathways, and can be readily assigned to sequence databases (Yates, 2000). ESI is not without disadvantages; it is poorly tolerant of

inorganic salts, whilst the multiplicity of charge states for larger analytes with more chargeable sites can reduce sensitivity for intact proteins and larger polypeptides (Abian et al., 1999).

2.3.2. Instrumentation used in biological mass spectrometry

There are several types of mass analyser commonly employed in biological mass spectrometry, including Time-of-flight (TOF), the quadrupole and the various types of ion trapping mass analysers (quadrupole, Fourier Transform Ion Cyclotron Resonance and orbitraps) (Wiley and McLaren, 1955, Hu *et al.*, 2005, Han *et al.*, 2008, Walther and Mann, 2010). These analysers are frequently combined into hybrid mass analysers, allowing optimum mass analysis results. ESI and MALDI are typically implemented on similar platforms, although there are some pairings which are particularly apposite. The continuous nature of ESI and finite nature of the typical mass range of ions thus generated (due to multiple charging) means that tandem quadrupoles and the various types of ion traps are commonly paired with ESI, whilst the addition of a semi-discontinuous quadrupole mass filter renders quadrupole time-of-flight (Q-q-TOF) mass analysers highly effective (Bateman and Hoyes, 2000).

2.3.2.1. Time-of-Flight

Time-of-flight mass spectrometer (TOF-MS) is a widely-used mass analyser type, which was developed in the 1950s (Balcerzak, 2003). Time-of-flight analysers work on a simple

basic principle: a set distance exists between ionisation source and ion detector. Ions are initially accelerated from the ionisation region towards the detector via application of typically electrical (field). This initial ion extraction is followed by transition of the ion population through a field-free region, where ions of differing mass-to-charge ratio (m/z) are separated and detected at differing 'time-of-flight' (Wiley and McLaren, 1955). Molecular ions of differing m/z differ in their kinetic energy (E_k). These differing kinetic energies translate to differing velocities (v) in a field-free 'drift' region with the following relationship:

$$E_k = m v^2 / 2$$

Equation 1.1

where m is mass of an ion and v is the velocity of the ion following acceleration (Balcerzak, 2003). Therefore, an ion with low mass will be transmitted faster than one with higher molecular mass. The detector is calibrated using signals of known m/z to enable conversion of time of flight to mass-to-charge (m/z) ratio for unknown signals. A schematic of this process is illustrated in Figure 2.3 (Dass and Brodbelt, 2001).

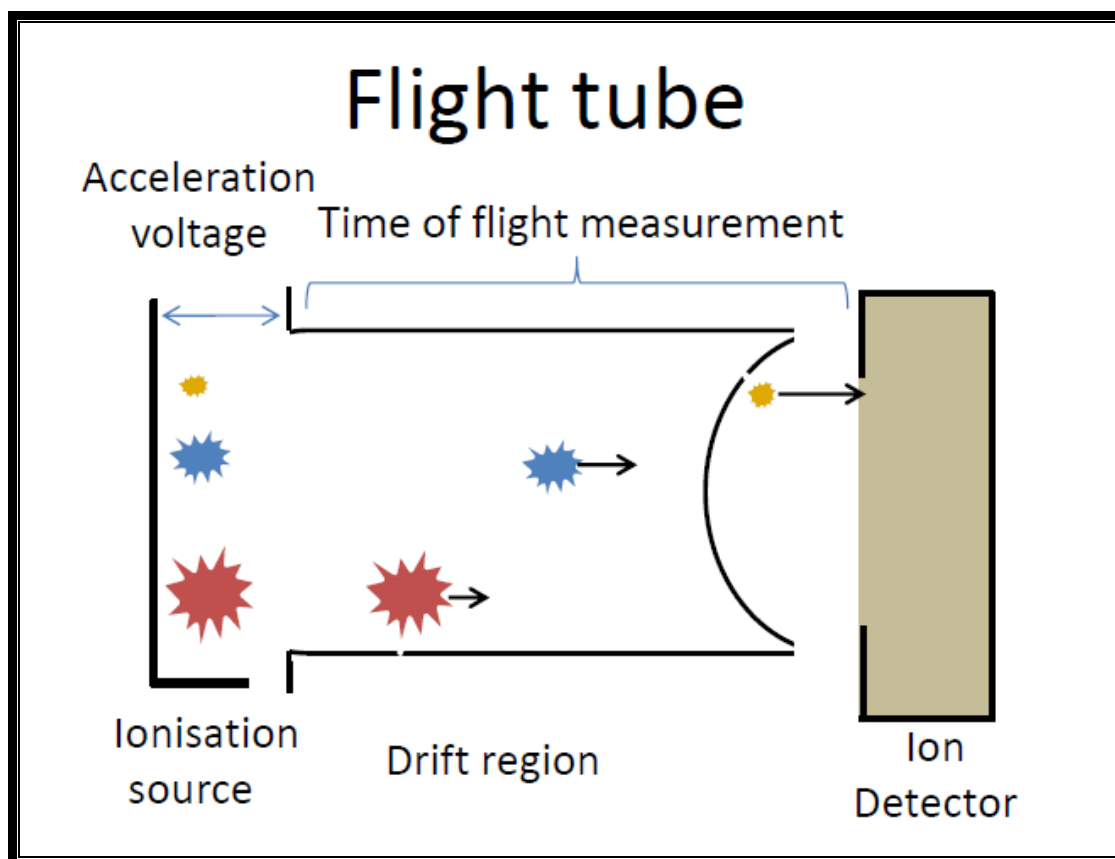


Figure 2. 3: Diagram of Time Of Flight (TOF) analyser.

Ions are subjected to mass analysis according to their time of flight, with a reference mixture being used to calibrate the relationship between time of flight and m/z ratio.

2.3.2.2. Quadrupole analysers

Quadrupole analyser devices typically consist of four parallel rods. These rods are in two pairs, with two opposite rods of positive polarity and two of negative polarity (see Figure 2.4). These rods use an oscillating electrical field to influence ions' trajectories according to their m/z ratio (Dawson, 2013). As the oscillating current changes, ions of varying m/z can be transmitted towards the detector; ions with unstable trajectories will be ejected or collide with rods (Figure 2.4). Quadrupole analysers can be employed to transmit and/or

select ions of differing m/z ratio; they have consequently been utilised in various types of tandem and hybrid mass analysers (Dawson, 2013).

Quadrupole mass analysers are well-suited to receive ions from sources which produce a constant beam, such as those created by an ESI ionisation source (Wingert *et al.*, 2015). Quadrupole mass analysers are capable of high mass accuracy, with unit mass resolution. Accordingly, quadrupoles are frequently used as ion selectors for hybrid analysers such as Q-TOF instruments. Since TOF analysers can measure multiple ions simultaneously, such analysers require an ion guide that transports all ions together at the same time over the entire mass range of the analyser (Vaz *et al.*, 2015, de Hoffmann, 1996).

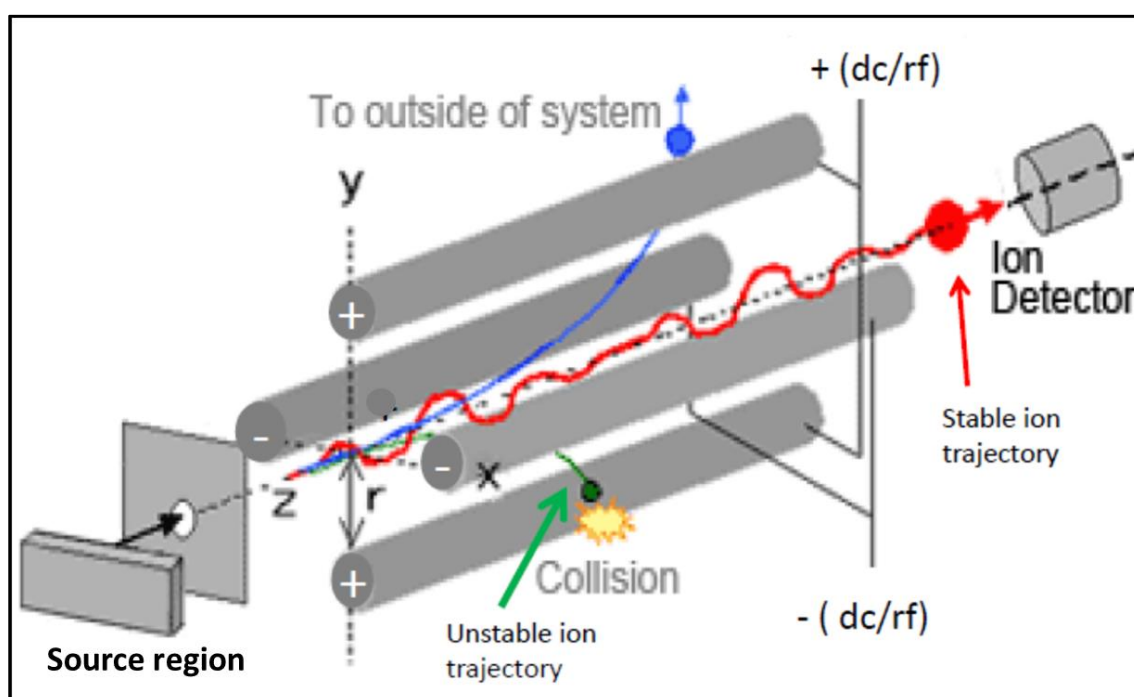


Figure 2. 4: Diagram of quadrupole mass analyser

Ions are generated in within the source region and accelerated in the z direction after applying voltage and radio frequencies to rods. Ions within a specific m/z range maintain a stable oscillatory trajectory and are thus transmitted through the quadrupole to reach the detector (red line) after applying a particular RF/DC. However, the oscillations of ions with other m/z ratio become unstable causing them either to be lost via transmission outside the mass analyser (blue

line) or collision with rods (green line). This figure was adapted from (www.Shimadzu.com, 2015).

Conventional tandem quadrupole instruments consist of three quadrupole regions; the first (Q1) can be used as a mass-selective filter or in broad pass mode. Q1 is followed by an rf-only quadrupole, typically used as a collision cell, with Q3 either being used to transmit and analyse product ions (product ion scanning), or to selectively transmit specific products (termed reaction monitoring); the latter process has an extremely high duty cycle and can be both very selective and of high sensitivity (Figure 2.5) (McLuckey, 1992).

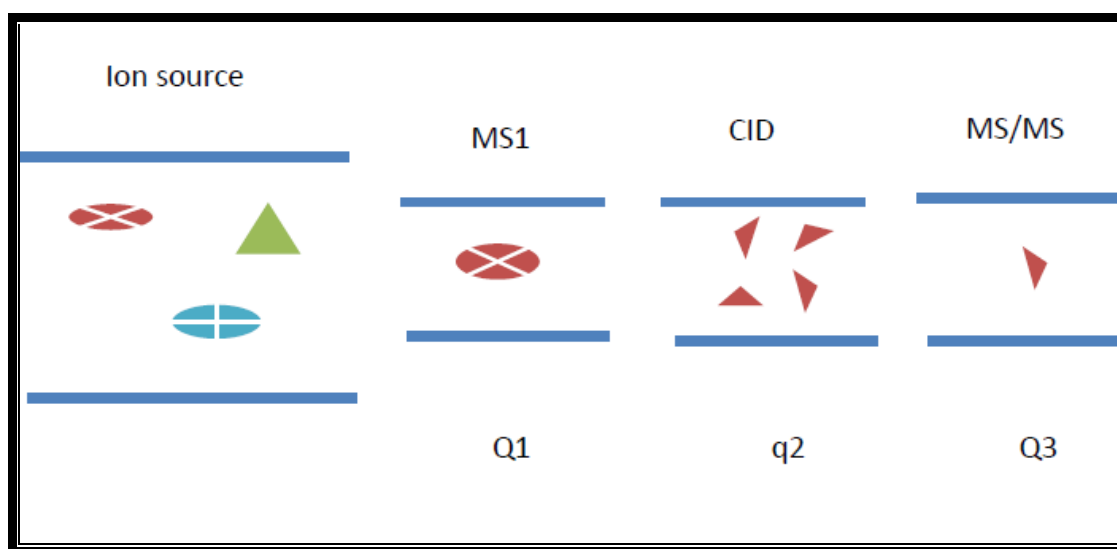


Figure 2. 5: Diagram of reaction monitoring using a triple quadrupole analyser

Initial analytical quadrupole (Q1) is used here as a mass selective filter, with the 'red' ions being selected for collisional activation in the radio frequency-only quadrupole collision cell (q2). The third quadrupole region (Q3) could be used to scan out the products (product ion scan), or as is shown here, operated as a filter to provide high-confidence quantification information for a specific product ion population generated from the selected precursor (red triangles); this is termed reaction monitoring.

2.3.2.3. Hybrid orbitrap mass spectrometers

Orbitraps are a unique class of resonance mass spectrometers which have become extremely popular due to their high resolving power and mass accuracy, combined with simple maintenance and lack of requirement for cryogenic cooling. These properties stand in direct contrast to the superconducting magnet instruments which have similar if not higher resolving power but significant ongoing maintenance requirements. The orbitrap instruments are electrostatic ion traps, with a central electrode surrounded by an outer electrode. Ions 'orbit' as oscillating packets or rings around the central electrode with different axial oscillatory frequencies depending upon the m/z of the ions (Watson and Sparkman, 2007). The ions are detected as an image current by the outer electrode according to the frequency of the oscillations (Watson and Sparkman, 2007). Orbitraps require electrodynamic 'squeezing' of ion packets to enable their coherent entry into the electrostatic field, and are therefore in practice always operated as hybrid instruments with various arrangements of quadrupole and linear ion trap acting as ion storage/filtering devices ahead of 'squeezing' in a curved linear ion trap ("C-trap") (Watson and Sparkman, 2007).

The Fusion mass spectrometer is one such hybrid instrument, in this case a "tribrid" consisting of a quadrupole mass filter, dual (low/high) pressure linear ion trap (LT), and Orbitrap mass analyser (Williamson *et al.*, 2016) . This hybrid instrument allows spatio-temporal separation of ion isolation, fragmentation, and detection. This configuration increases the efficiency of the orbitrap duty cycle to enable high-throughput, sensitive peptide identification (Williamson *et al.*, 2016). In particular, these instruments enable

sensitive and high-resolution of the low mass region for iTRAQ/TMT-labelled quantitative proteome analyses of biological samples (Wen *et al.*, 2013, McAlister *et al.*, 2014).

2.4. Application of mass spectrometry in PE and IUGR studies

Mass spectrometry-based quantitative proteomics strategies can be classified in two major categories: firstly, label-free methods, based on integral MS signal of each (protein/peptide) (Bergman and Bergquist, 2014, Blankley *et al.*, 2013); and alternatively those approaches using stable isotope labelling, such as isobaric tagging strategies (Rauniyar and Yates III, 2014). Isobaric labelling-based quantification, (e.g. iTRAQ) is a means to providing relative quantitative information, providing a set of reporter ions whose intensity profiles correlate with differences between two or more proteomes (Ong and Mann, 2006, Lehnert *et al.*, 2012, Rauniyar and Yates III, 2014).

Previously a number of studies (Auer *et al.*, 2010, Blankley *et al.*, 2013, Wen *et al.*, 2013) have been performed to study proteomic biomarkers for both PE and IUGR conditions; these studies used MS coupling with labelling technique (iTRAQ) or label free bases analysis. In 2010, Auer *et al.*, used iTRAQ labelling coupled with MALDI-TOF-MS to identified proteomic biomarkers in blood plasma from pregnant women obtained at or near term (Auer *et al.*, 2010). Women were categorised into four groups: preeclampsia (PE, n=7), isolated Intra-Uterine Growth Restriction (IUGR, n=7), PE and IUGR (n=7), and normal pregnancies (control. n=6). Before labelling with iTRAQ, the samples were immuno-depleted of major plasma proteins (albumin, IgG, antitrypsin, IgA, transferrin and haptoglobin) using multiple affinity removal spin MARS-6 cartridges. In order to

identify proteins which are changed in placental diseases, the first set samples were pooled with iTRAQ labelling from patients diagnosed with isolated PE, isolated IUGR, PE with IUGR and controls (normal pregnancies). In the second step, they used four pools categorised into two control groups. First group (6 patients from the first set and 4 others from a second independent set) and a second group included two IUGR groups (7 from the first series and 6 others from a second independent series). They selected two candidate proteins from the analysed pools to validate them on individual samples by Western blot. 166 proteins in total were identified from the first iTRAQ experiment, with 31/128 identified proteins showing expression changes in IUGR (14 up and 17 down-regulated); 54 proteins out of 130 identified were different in the serum of isolated PE, (28 up and 26 down-regulated). Finally, in (PE + IUGR) cases 28/134 proteins were significantly dysregulated (13 up and 15 down-regulated). Nine proteins were identified that showed altered expression in all pathologies investigated, for the majority of cases compared to controls; these proteins showed differential patterns of dysregulation for different conditions. For instance, fibronectin (FN1) was specifically increased in the cases of PE and PE+IUGR, but reduced in isolated IUGR. The authors concluded that the differences in the terms at which samples were collected (on average 5 weeks difference between control samples and pathological cases) may have led to many of the differences observed in these putative biomarkers, and further suggested that a larger cohort size is needed in order to gain results with increased precision (Auer *et al.*, 2010)

In 2013, Wen *et al.*, used LC-MS/MS LTQ-orbitrap to analyse serum peptidome profiles in a case control study in preeclampsia (Wen *et al.*, 2013). Sixty-two pregnant women were investigated; (n=31) had PE and (n=31) were healthy pregnant women (controls). All

subjects within the study were categorised into two groups. 2/3 of the cohort were used as a training set (n=21 preeclampsia case and n=21 control), whilst the remaining 1/3 were used as test set (n=10 case group, n=10 healthy group). They were able to identify 612 peptides in total. Two algorithms, significance analysis of microarrays (SAM) and predictive analysis of microarrays (PAM), were used for analysis. SAM identified 52 differential profiling peptides, of which 14 protein precursors were highly significant in cases compared to control. They also found nineteen peptide panels derived from 6 different protein precursors: (13 from fibrinogen alpha (FGA), 1 from alpha-1-antitrypsin (A1AT), 1 from apolipoprotein L1 (APO-L1), 1 from inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4), 2 from kininogen-1 (KNG1), and 1 from thymosin beta-4 (TMSB4). All nineteen peptides had a minimal false discovery rate significant ($q < 0.05$) value. According to trained PAM prediction analysis, all PE (n =21), were predicted correctly, whereas 3 of 21 (14.3 %) control samples were false positive. For the training set, sensitivity was 85.7 % and specificity was 100 %; these results showed around 92.9 % accuracy in the overall prediction. Test set results gave an overall prediction accuracy of 90 % (sensitivity was 80 % and specificity was 100 %). Pathway analysis (Ingenuity Pathway Analysis) was used to analyse 14 parental proteins of the 52-peptide markers, which may play vital roles in the pathophysiology of PE. The limitation of this study was that proteomic profiling data were obtained from a commercial vendor with little specific information on the clinical characteristics; the analysis can therefore only be of confirmative diagnostic rather than predictive value; the sample set lacked samples from other hypertensive disorders of pregnant women to define if the panel differentiates between these women and PE. The authors of this study also suggested using

multiplexed peptide detection coupled with multiple reaction monitoring (MRM) for confirmation of putative differential markers.

The international Screening fOr Pregnancy Endpoints (SCOPE) study enrolled 3182 pregnant women, of whom 5.6 % had preeclampsia (Blankley *et al.*, 2013). This study used iTRAQ tags coupled with a combination of MS tools (LC-MS or with MALDI-TOF-TOF) to study serum biomarkers in preeclampsia. A pool of samples (n=12) were taken at 15 weeks from women who developed PE later, early onset PE (EO-PE, delivery 34 weeks) and two control pools (uncomplicated pregnancies, n=12 each). A plasma sample was immuno-depleted before iTRAQ labelling using either the SepproIgY 14 -SuperMix Liquid Chromatography Column system or the Multiple Affinity Removal LC Column - Human 14 (MARS 14). This study performed in three experimental steps. From exploratory phase of iTRAQ experiments of first cohort, 502 proteins were identified (319 proteins in the IgY 14-SuperMix-Q-STAR data set, 331 in the IgY 14-SuperMix-5800 data set, and 189 in the MARS 14 -5800 data set), 133 proteins were modified in abundance in each one of the iTRAQ data sets, which are represent a discovery data. SRM method was used for verification of candidate proteins that were detected in the discovery data set such as, platelet basic protein (PBP, CXCL7), and pregnancy-specific beta-1-glycoprotein 9 (PSG9). PSG9 is a member of a large multigene family, which expressed in a pregnancy specific manner. In second cohort, 108 enrolled women had PE; n=16 (EP-PE), n= 42 late onset (LO-PE), and 42 control), eight technical replicates of pooled samples (after immune-depletion, proteins fractionation and digestion). LC-SRM analysis was performed on an uHPLC using QqQ-LC/MS platform with electrospray source. Six isoforms of PSG specific peptides (PSG9) were observed to show statistically significant differences between EO-

PE and controls. These changes were then verified with ROC analysis and ELISA based quantification (Blankley *et al.*, 2013).

Many studies have previously investigated metabolic profiles in PE and IUGR. No studies so far identified robust biomarkers to predict PE and IUGR disorders earlier. At the time of writing, no screening test was currently in use with predictive capacity to accurately identify women whose first time gestations will develop late pregnancy diseases. Urine sample is an ideal bodily fluid for antenatal testing due to the non-invasive nature of obtaining samples; indeed women under routine prenatal care are required to provide urine samples for testing, making this a highly-acceptable methodology. Moreover, urine proteins have more stability and low dynamic range compared to blood proteome.

2.5. Chapter aims

1. To establish Standard Operating Protocol (SOP) based on urinary protein separation over nominal molecular weight cut-off (NMWCO) membranes >10 kDa. This informed the experimental design and sample preparation for 8-plex iTRAQ labelling coupling for the LC-MS technique for proteomic analysis.
2. To undertake proteomic analysis for PE and control at both 15 and 20 weeks gestational age respectively using LC-MS analysis.
3. To perform proteomic analysis for IUGR and control at each 15 and 20 week of pregnancy respectively using LC-MS analysis

4. To determine dysregulated protein quantitatively that showing significant alteration in protein expression in PE and IUGR during 15 and 20 weeks respectively.

2.6. Materials and methods

2.6.1 Chemicals

All chemicals were used as received from the supplier. Table 2.1 presents a full list of all standard chemicals used within the work presented in this chapter.

Table 2. 1: Chemicals and suppliers

Item	Supplier
Acetonitrile	VWR Chemicals (Lutterworth, UK),
Bovine Serum Albumin (BSA)	Waters-UK
Bovine Serum Albumin digestion Standard (BSA)	(Waters-UK)
Bradford Reagent	Sigma-Aldrich (Poole, UK)
Enolase digestion standard	MassPREP (Waters-UK)
HPLC-grade water	VWR Chemicals (Lutterworth, UK)
Isopropanol	Sciex, USA
Methyl methane-thiosulfonate (MMTS) in isopropanol	Sciex, USA
SDS	Sciex, USA
Sequencing Grade Modified Trypsin:	Promega-UK
Tris thylammonium bicarbonate (TEAB)	Sigma-Aldrich, UK
Trifluoroacetic acid	Fisher Scientific (Loughborough, UK)
Tris-(2-carboxyethyl)-phosphine (TCEP).	Sciex, USA

2.6.2. Sample acquisition and preparation

2.6.2.1. Acquisition of human urine samples

Urine samples were obtained from The international SCReening fOr Pregnancy Endpoints, (SCOPE) biobank, held at University College Cork. These samples were collected prospectively at 15 and 20 week gestation from women whose pregnancies were eventually complicated by IUGR (n=30) or with PE (n=30). Samples from pregnancy complications were compared to uncomplicated pregnancies, matched for ethnicity and BMI (n=30 in each case). The acquisition and transport of the urine sample from University College Cork (UCC) were performed under a Material Transfer Agreement (MTA) (see Appendix 1). UCC undertook ethical review governing the recruitment of patients, sample and metadata collection and storage (see Appendix 2), and provided detailed assurance of their informed written consent procedures. All donor information is held by SCOPE in Cork; we have no direct access to this data; all patients are given unique references and the identifying data are held securely and accessible only by UCC. Evidence of consenting procedures was requested at the time of procurement and ethical review (see Appendix 2) arranged prior to use through Keele University's Research Ethics Committee in accordance with HTA-41.

2.6.2.2. Transportation of human tissue and storage

A total of 240 cryo-tube urine samples (450 μ L per tube) were shipped from Cork to Keele University (Huxley Building). 120 samples totally collected from each preeclampsia and IUGR respectively, 120 samples from normal pregnant (control). The samples (n=30) are matched accordance to each case-control for both PE and IUGR respectively. These

samples were collected at two period's gestational age, 15 and 20 weeks respectively. Shipping was undertaken by Marken, an authorised dry ice shipping company in accordance with the signed Material Transfer Agreement (MTA) (see Appendix 1). Cryobox temperatures were recorded and monitored during the shipment. The consignment was transported in accordance with IATA / UN 1845 standards for approved shipping containers, with double packaging to avoid leaks and the services of a registered delivery agency requiring a signature on receipt of the consignment. This group does not deal in primary patient information and thus had no access / did not store any patient identifiable information, thus loss of confidentiality was not a risk in this situation. Upon receipt, all samples were directly stored at -80° C, subject to routine temperature monitoring by University staff. Samples were clearly labelled with unique identifiers.

2.6.3. Samples preparation process

Urine samples were prepared to examine proteomic biomarkers according to protocol below:

2.6.3.1. Centrifugation, ultrafiltration and protein purification

Urine samples were selected from -80° C freezer, thawed and cleared from pelleted urinary tract cells (i.e. uroepithelial cells) using centrifuge (Eppendorf, Germany) at 3220 g for 10 min, 4° C. The remaining acellular supernatant material of each sample was subjected to isolation of proteins using >10kDa proteins ultrafiltration tubes Polyethersulfone (PES) MWCO 10000, 2 mL. (Sartorius AG,). Ultrafiltration centrifuge

devices were washed 3 X with HPLC grade water to remove PES membrane interference materials. Buffer exchange was performed on the retentate fraction of urine, spinning (3 X) with 0.5M triethylammonium bicarbonate (TEAB) (Sigma-Aldrich, UK) at 3220 g, 4° C to clear proteins from undesirable ionic materials and salts using 3220 g, 4° C (Eppendorf, Germany). Assay on protein fraction in 96-well plate using well-plate reader vs. a standard curve generated using BSA of known concentration was carried out on final concentrated pool retentate in each cohort.

2.6.3.2. Verification of protein concentration

Spectrophotometry (BioTek, E L 800, USA) was used to verify protein concentration of > 10 kDa retentate in each pool cohort samples. The amount of protein was determined using a microtiter plate format Bradford assay, standardised against bovine serum albumin. The assay was carried out as per manufacturer's protocol, monitoring spectrometric absorbance at 595 nm. 60 µg of protein was used in each labelling experiment. Protein concentration was measured using standard concentration curve (Bradford assay) at the particular wavelength to calculate their concentrations using Beer-Lambert law, Equation 2.1:

$$A = \epsilon L c \quad \text{- Equation 2.1}$$

A = absorbance, ϵ = molar extinction coefficient, c = concentration and L = light pathlength, the concentration can be calculated by given Equation 2.2:

$$c = A / \epsilon L \quad \text{- Equation 2.2}$$

2.6.3.3. Tryptic digest and iTRAQ labelling

Tryptic digestion was carried out on proteins >10 kDa at amount 60 µg. (Sigma-Aldrich, UK), Proteins were subjected to reduction using tris-(2-carboxyethyl)-phosphine (TCEP) (AB Sciex, UK), at a final concentration of 50 mM. Reduction proceeded at 60° C for 60 min. Alkylation of free cysteines was performed using methyl methanethiosulfonate (AB Sciex, UK) (final concentration 200 mM) at room temperature, 10 min. Protein was then digested overnight at 37° C using 10 µg of Promega modified trypsin (Waters, UK). A known amount at final concentration (~1 pmole) of peptides mixture derivative from yeast enolase was spiked into urine samples, these served as a convenient internal standard for quantification.

iTRAQ labelling (8-plex) was then performed as per manufacturer’s protocol. Labelling reagent (Sciex, USA) was dissolved in 100 µl of a mixture solution (0.5 M TEAB and ethanol). Reagent and digest were then mixed and incubated at room temperature for 2 hr. Table 2.2 shows descriptive method for samples labelling with iTRAQ reagents (8-plex) in each group in this study. Samples were dried down using vacuum centrifugation and dissolved in 60 µL of buffer A (5 % acetonitrile, 0.1 % formic acid) for LC-MS/MS.

Table 2.2: Labelling reagents with all samples of each PE and IUGR samples. iTRAQ® Reagent – 8plex - Buffer Kit materials (Sciex, USA), S: Case, Ctl: Control .

Condition	PE				IUGR			
	15 week (n=60)		20 week(n=60)		15 week(n=60)		20 week(n=60)	
Gestational age								
Subjects (no.)	S (30)	Ctl (30)	S (30)	Ctl (30)	S (30)	Ctl (30)	S (30)	Ctl (30)

Labelling reagents	113	114	115	116	117	118	119	121
---------------------------	-----	-----	-----	-----	-----	-----	-----	-----

2.6.3.4. Strong cation exchange clean up

Excess iTRAQ reagent was removed from the sample by performing SCX based SPE using Empore Cation (3 M) disks in a pipette tip. The disk was prepared by successive washes of acetone, methanol, water, 5% (v/v) ammonium hydroxide and water. The sample was bound to the disk and washed with 1 % (v / v) trifluoroacetic acid. The sample was eluted using 5 % ammonium hydroxide, dried and re-suspended in 3 % (v / v) methanol, 97% (v / v) HPLC water, 0.1 % (v / v) TFA. This clean-up step was carried out at Liverpool University.

2.6.3.5. Mass spectrometry set up

Experiment 1:

LC-MS was performed on a hybrid quadrupole-time-of-flight mass spectrometer Q-TOF Premier (Waters), coupled to nanoflow HPLC (Dionex Ultimate 3000, ThermoFisher Corporation, Hemel Hempstead). Tandem mass spectrometry was performed to identify tryptic peptides. Peptide mass spectra were obtained using nanoflow electrospray ionisation MS (nESI-MS) with tandem MS using MassLynx version 4.1. Approximately 1 ug of sample was analysed. Separation was performed over 120 min (90 min gradient) using a gradient from 5-90 % of acetonitrile [buffer A (5 % acetonitrile, 0.1 % formic acid and

buffer B (95 % acetonitrile, 0.1 % formic acid)] on a C 18 column (75 μm x 150 mm, poresize 200 \AA , 5 μm particle size, Dionex/ThermoFisher, San Jose, CA). Automated switching between MS (survey mode) and MS / MS (product ion mode) was used to generate sequence tag-specific product ion spectra. MS / MS was performed on the “top 3” multiply-charged ($z \geq 2$) precursors observed per survey scan. To prevent repeated MS / MS being performed upon the same precursors, dynamic exclusion was selected, with a repeat duration of 3 min. Sample analysing in duplicate injections with iTRAQ experimental set up was used. This experiment enabled verification of high labelling efficiency.

Following labelling efficiency verification, samples were transferred to Liverpool University to be analysed on the orbitrap Fusion mass spectrometer (Thermo Scientific).

Experiment 2:

The data-dependent iTRAQ analysis was performed using an Ultimate 3000 RSLCTM nano system (Thermo Scientific) coupled to a Fusion mass spectrometer (Thermo Scientific). The sample (4 μL corresponding to 400 ng of protein for human sample) was loaded onto the trapping column (Thermo Scientific, PepMap 100, C 18, 75 μm X 20 mm), using partial loop injection, for 7 minutes at a flow rate of 9 $\mu\text{L}/\text{min}$ with 0.1% (v / v) TFA. The sample was resolved on the analytical column (Easy-Spray C 18 75 μm x 500 mm 2 μm column) using a gradient of 96.2% A (0.1% formic acid) 3.8% B (79.9 % ACN, 20 % water, 0.1% formic acid) to 50% A 50% B over 90 min at a flow rate of 300 nL/min. The MS acquisition program followed the multi-notch MS 3 method described in (McAlister *et al.*, 2014). The program consisted of a 60,000 resolution full-scan MS scan (AGC set to $4e^5$ ions with a

maximum fill time of 50 ms) with MS / MS using quadrupole ion selection with a 0.7 m/z ratio window, HCD fragmentation with a normalized collision energy of 35 and LTQ analysis using the rapid scan setting and a maximum fill time of 50 msec. The 10 most intense MS 2 peaks were simultaneously selected for MS 3 analysis using SPS isolation, HCD fragmentation at collision energy 65 and analysis in the Orbitrap using a nominal resolution of 50,000 over a mass range of 100-500 m/z. The machine was set to perform as many MS/MS scans as possible, maintaining a cycle time of 3 sec. To avoid repeated selection of peptides for MS/MS the program used a 60 sec dynamic exclusion window.

2.6.3.6. Data processing

Raw product ion spectral data from experiment 1 (see Section 2.6 3.5.) were extracted as text formatted .mgf files using Mascot Distiller (Matrix Science, London), altering the default parameters to include isotopic peak extraction over the reporter ion region (m / z ratio 112-122). Mgf files were subjected to database searching using Mascot (Matrix Science). SwissProt protein sequence database was used for these searches, with the following parameters: trypsin as enzyme, precursor ion tolerance 100 ppm, product ion tolerance 0.6 Da, fixed modification methylthio cysteine, variable modification: methionine oxidation, and at peptide N-termini and lysine side-chains with iTRAQ 8-plex. Data were analysed using Mascot (Matrix Science) to determine labelling efficiency. These data were excluded from the following results section due to low data density in comparison to the orbitrap Fusion dataset, and duplication therewith.

Data were processed from experiment 2 (section 2.6 3.5) using Proteome Discoverer V2.2 (Thermo Scientific) in combination with Mascot V2.6 (Matrix Science) search engine for protein identification. Peptide identification was performed against the reviewed entries from the Uniprot Human reference proteome (20328 entries) using a precursor mass tolerance of 10 ppm, a product mass tolerance of 0.6 Da, fixed methylthio cysteine modification and variable iTRAQ (8-plex) modification of lysine and the N-terminus. Peptide and protein identifies were accepted using 0.01 FDR for strict filtering and 0.05 % for relaxed filtering. Unique peptides were only accepted for quantification if all reporter channels were present, with an average S/N > 5. Sample normalisation was performed based on the spiked enolase standard. Resultant data were exported to Microsoft Excel.

2.6.3.7. Statistical analysis

The analysis of the data was carried out within Excel (Microsoft, version 10) and GraphPad (Prism 8), to plot relevant tables and graphs. Online Venn diagrams (Venny 2.1) (Oliveros, 2007) were used for data comparison. Data were analysed using descriptive statistics of box and whisker plot using GraphPad (Prism 8). Non-parametric statistical analysis was performed using GraphPad (Prism 8), using one sample Wilcoxon test, and two samples analysis using Mann-Whitney test.

2.7. Methods outline workflow

The major method in this chapter is summarised using outline workflow in Figure 2.6.

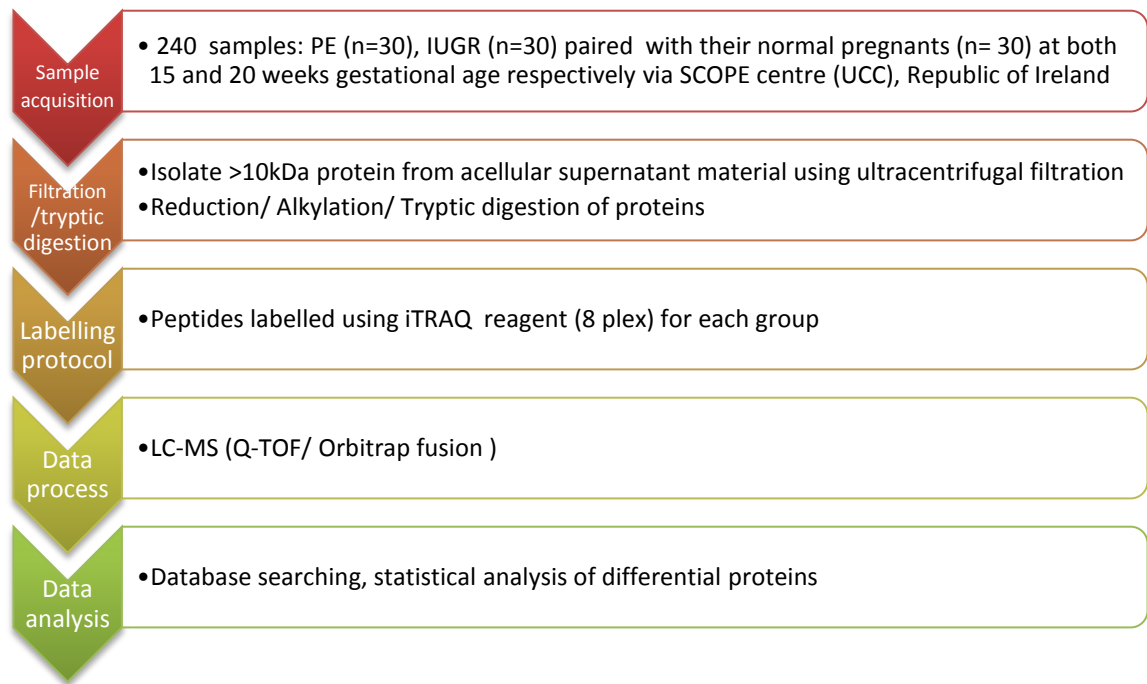


Figure 2. 6: Schematic chart flow illustrating typical proteomic method.

2.8. Results

2.8.1 Proteomic analysis for PE at 15 and 20 weeks gestational age

Proteome analysis of the iTRAQ-labelled urine tryptic peptides generated a dataset of protein identification and quantification data. 988 peptides in total were identified from 315 proteins as being differentially expressed. Proteins which are detected as differentially expressed are filtered using the following criteria: ≥ 2 peptides; Confidence Interval (CI) 99 %, FDR < 0.01. 173 / 826 (proteins / peptides) remained following this filtering with all 8 label channels being present. Peptides with significant changes for samples taken from women who went on to develop PE at a FDR < 0.01 and confidence interval 99 % are presented for samples collected at 15 and 20 weeks gestational age as Tables S 1 and S 2 respectively (see appendix 3). Identified proteins are also quantified,

with differential expression being determined using \log_2 fold change of abundance ratio (cases: control) at 15 and 20 weeks gestation for PE cases respectively compared with uncomplicated pregnancies, shown in Figure 2.7 and 2.8 respectively. All quantitation data underwent normalisation using a known amount of internal standard, adjusting the \log_2 fold change ratio for internal standard to zero (Figure 2.7, 2.8, and 2.9).

At PE at 15 weeks gestational age, 173 proteins were identified as showing significant changes in expression ($P < 0.01$) (Figure 2.7). These data are showed a median at (-0.576) with (3.5-(-4.08)) max. and min. values of dysregulated proteins using one sample Wilcoxon test (see Figure 2.7). For PE at 20 week of gestation, outcome showed highly statistically significant differences quantitatively ($P < 0.01$) at 95 % confidence interval of 173 proteins using \log_2 fold change of abundance ratio (case/ control) using one sample Wilcoxon test. The data have a median ratio of (0.089), max. and min. values at (2.91-(-1.639)) (Figure 2.8). A Mann-Whitney test comparing \log_2 fold changes ratio (case/ control) observed between 15 and 20 week gestation pregnancies for PE showed a significant difference ($P < 0.01$) at 95 % confidence interval for 173 proteins (Figure 2.9).

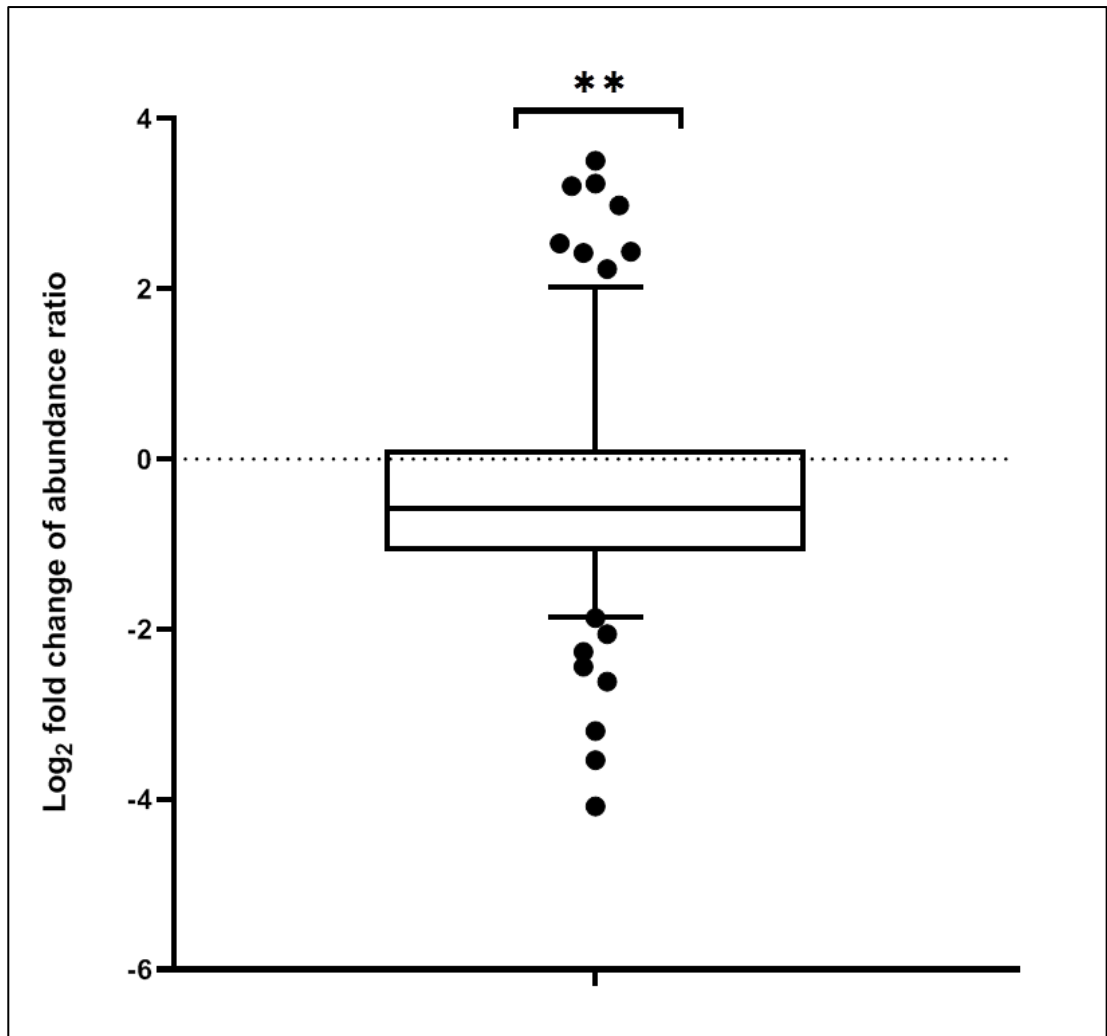


Figure 2. 7: Differentially expressed protein changes quantitatively for PE at 15 weeks' gestation

Plot displays log₂ fold change of abundance ratio of (case: control) for PE at 15 week gestational age. Data were analysed showed proteins differentially expressed changes quantitatively vs internal standard of enolase at zero value (dotted line). 173 Proteins are analysed are showed highly significant changes (P < 0.01) at confidence interval 95 % using one sample Wilcoxon test. Data were expressed as interquartile range of upper and lower quartiles (box), median (line inside box), lines at top and bottom represent the highest and lowest values (whiskers), confidence intervals (95 %), ** represents highly significant value (P < 0.01), black balls represent outliers points that showed highest and lowest ratio compared with internal standard.

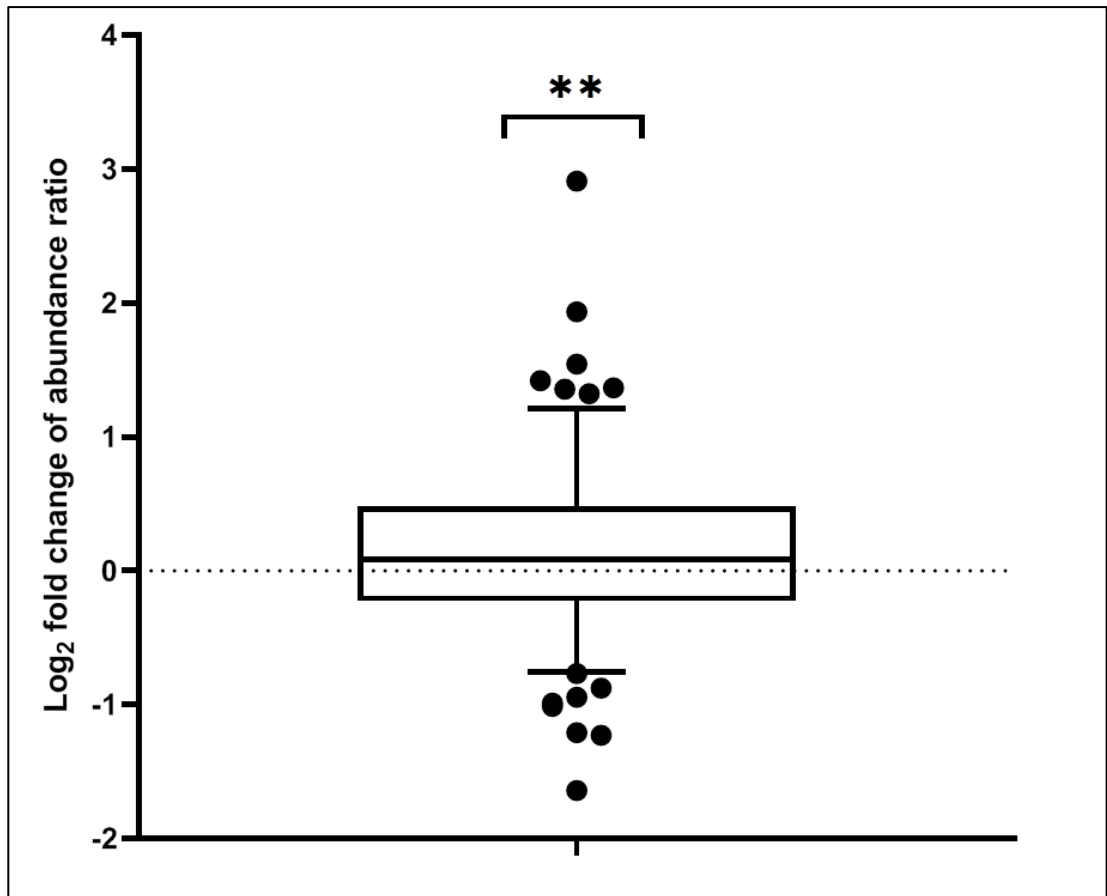


Figure 2. 8: Differentially expressed protein changes for PE cases at 20 weeks' gestation

Plot displays \log_2 fold change of abundance ratio of (case: control) for PE at 20 weeks' gestation. Differentially expressed are expressed as fold change in comparison to enolase tryptic digest internal standard (dotted line). 173 observed proteins showed highly significant changes ($P < 0.01$) at confidence interval 95 % using one sample Wilcoxon test. Data were expressed as interquartile range of upper and lower quartiles (box), median (line inside box), lines at top and bottom represent the highest and lowest values (whiskers), confidence intervals (95 %), ** represents highly significant value ($P < 0.01$), black balls represent outlier points with highest and lowest ratios compared with internal standard.

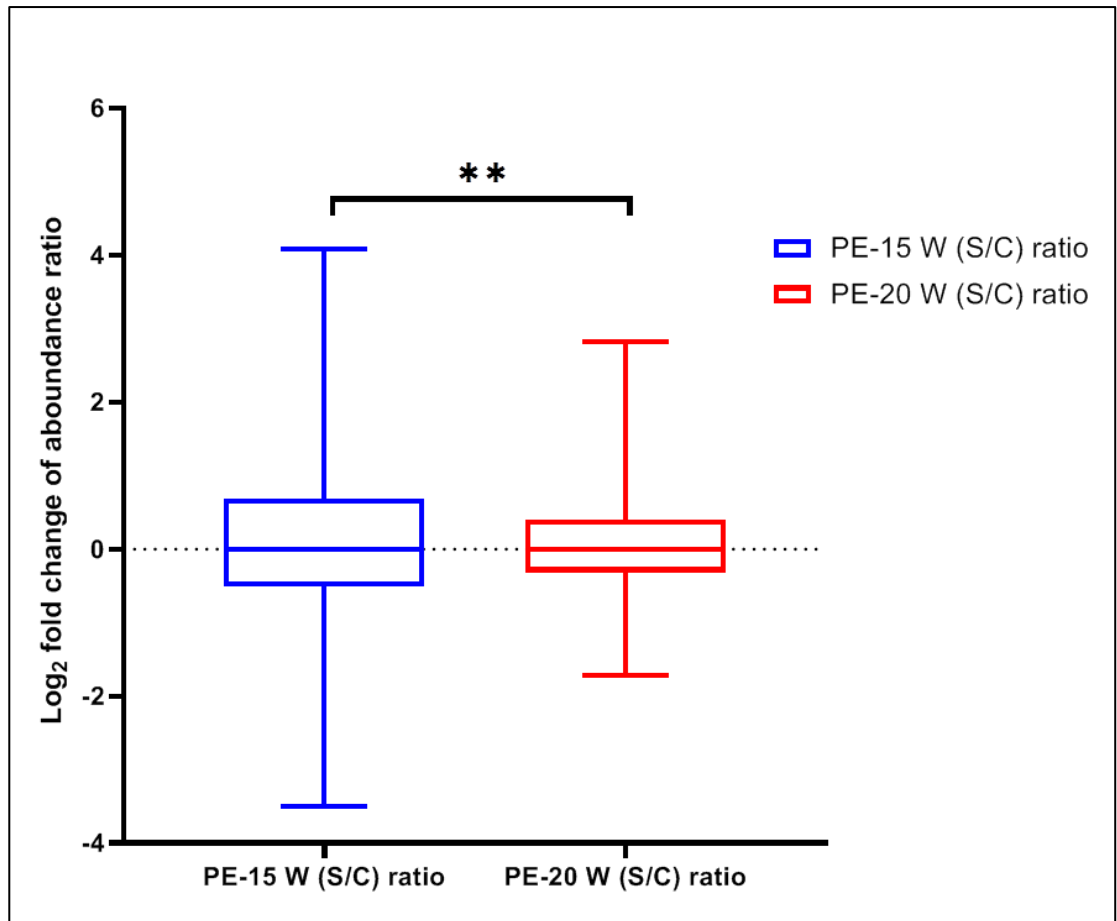


Figure 2. 9: Differentially expressed protein changes for PE cases at 15 and 20 weeks of pregnancy

Plot shows comparison analysis of \log_2 fold change of abundance ratio of (case: control) for PE at both 15 week (blue box) and 20 week (red box) gestational age respectively. Data are analysed show proteins expression differences quantitatively between 15 and 20 weeks of gestational age of PE. Significant differences between groups are observed ($P < 0.01$), confidence interval 95 % for the 173 proteins analysed using a non-parametric Mann-Whitney test of unpaired two samples groups. Data are expressed as interquartile range of upper and lower quartiles (box), median (line inside box), lines at top and bottom represent max and min values (whiskers). ** represents significant changes ($P < 0.01$), dotted line represents internal standard value of enolase.

The following section describes outlier/dysregulated proteins (highest / lowest) identified as significant in previous analysis at 15 and 20 weeks' gestation respectively

2.8.1.1. Highly dysregulated proteins for PE at 15 and 20 weeks

Differentially expressed proteins were identified as outliers (Figure 2.7 and 2.8) at 15 and 20 weeks gestation respectively for PE in comparison to uncomplicated pregnancy. 16/42 (proteins/ peptides) were identified as outliers in previous analyses (Figures 2.7) as highly dysregulated proteins at 15 weeks' gestation for PE compared with uncomplicated pregnancy. 50 % (n = 8) of these proteins are upregulated and downregulated respectively compared with uncomplicated pregnancy (Table 2.3). At 20 weeks' gestation, 15 / 96 (proteins/ peptides) were dysregulated in pregnancies complicated by PE vs. uncomplicated pregnancies. Of these, 7 proteins were upregulated and 8 proteins were downregulated compared with uncomplicated pregnancy Table 2.4. Three of the candidate dysregulated proteins showed identical profiles at 15 and 20 weeks gestation. These proteins are fibrinogen beta chain, haptoglobin and carboxypeptidase, and are highlighted in Table 2.3 and 2.4, and the comparison in Figure 2.10.

Table 2.3: Dysregulated proteins dataset of PE at 15 week gestational age compared with control

No.	Protein FDR Confidence: Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of PE at 15 W	Abundances of reporter ion of control at 15W	Abundance ratio (Case/Control) of PE at 15 W	Log ₂ fold change ratio	Up/Down
1	High	P01023	Alpha-2-macroglobulin OS=Homo sapiens OX=9606 GN=A2M PE=1 SV=3	2	163.2	234.4	20.7	11.33	3.502	Up
2	High	P02675	Fibrinogen beta chain OS=Homo sapiens OX=9606 GN=FGB PE=1 SV=2	2	55.9	159.4	16.9	9.415	3.235	Up
3	High	Q6GTX8	Leukocyte-associated immunoglobulin-like receptor 1 OS=Homo sapiens OX=9606 GN=LAIR1 PE=1 SV=1	2	31.4	143.4	15.6	9.223	3.205	Up
4	High	Q8IYS5	Osteoclast-associated immunoglobulin-like receptor OS=Homo sapiens OX=9606 GN=OSCAR PE=1 SV=3	2	30.5	42.4	5.4	7.89	2.980	Up
5	High	P00738	Haptoglobin OS=Homo sapiens OX=9606 GN=HP PE=1 SV=1	4	45.2	219.5	38	5.78	2.531	Up
6	High	P02788	Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF PE=1 SV=6	3	78.1	273.3	50.6	5.404	2.434	Up
7	High	P06733	Alpha-enolase OS=Homo sapiens OX=9606 GN=ENO1 PE=1 SV=2	4	47.1	98.4	18.4	5.351	2.420	Up

8	High	P13796	Plastin-2 OS=Homo sapiens OX=9606 GN=LCP1 PE=1 SV=6	4	70.2	177.4	37.8	4.692	2.230	Up
9	High	Q06830	Peroxiredoxin-1 OS=Homo sapiens OX=9606 GN=PRDX1 PE=1 SV=1	2	22.1	19	69.2	0.274	-1.868	Down
10	High	P10253	Lysosomal alpha-glucosidase OS=Homo sapiens OX=9606 GN=GAA PE=1 SV=4	4	105.3	14.5	60.5	0.24	-2.059	Down
11	High	Q9Y646	Carboxypeptidase Q OS=Homo sapiens OX=9606 GN=CPQ PE=1 SV=1	2	51.9	16.4	79	0.208	-2.265	Down
12	High	Q9P121	Neurotrimin OS=Homo sapiens OX=9606 GN=NTM PE=1 SV=1	2	37.9	13.5	72.9	0.184	-2.442	Down
13	High	P07195	L-lactate dehydrogenase B chain OS=Homo sapiens OX=9606 GN=LDHB PE=1 SV=2	2	36.6	13	79.8	0.163	-2.617	Down
14	High	Q8WZ75	Roundabout homolog 4 OS=Homo sapiens OX=9606 GN=ROBO4 PE=1 SV=1	2	107.4	7.8	71.6	0.109	-3.198	Down
15	High	Q9HCU0	Endosialin OS=Homo sapiens OX=9606 GN=CD248 PE=1 SV=1	2	80.8	3.3	38	0.086	-3.540	Down
16	High	Q07075	Glutamyl aminopeptidase OS=Homo sapiens OX=9606 GN=ENPEP PE=1 SV=3	3	109.2	1.7	28.2	0.059	-4.083	Down

Data identified at confidence interval 99 %, FDR (False Discovery Rate) < 0.01, Peptides no.≥ 2; Outlier of data that quantified at confidence interval 95

%. P < 0.05, highlighted data is showed identical in protein profile at both 15 and 20 weeks of PE.

Table 2.4: Dysregulated proteins dataset of PE at 20 week gestational age compared with control

No.	Protein FDR Confidence: Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of PE at 20 W	Abundances of reporter ion of control at 20W	Abundance ratio (case/control) of PE at 20 W	Log ₂ fold change ratio	Up/Down
1	High	P02675	Fibrinogen beta chain OS=Homo sapiens OX=9606 GN=FGB PE=1 SV=2	2	55.9	167.3	22.2	7.526	2.912	Up
2	High	P00738	Haptoglobin OS=Homo sapiens OX=9606 GN=HP PE=1 SV=1	4	45.2	183.4	48	3.823	1.935	Up
3	High	P02787	Serotransferrin OS=Homo sapiens OX=9606 GN=TF PE=1 SV=3	12	77	179.2	61.4	2.919	1.545	Up
4	High	P01861	Immunoglobulin heavy constant gamma 4 OS=Homo sapiens OX=9606 GN=IGHG4 PE=1 SV=1	3	35.9	201.1	75.1	2.677	1.421	Up

5	High	P01877	Immunoglobulin heavy constant alpha 2 OS=Homo sapiens OX=9606 GN=IGHA2 PE=1 SV=4	3	36.6	203	78.6	2.582	1.368	Up
6	High	P00450	Ceruloplasmin OS=Homo sapiens OX=9606 GN=CP PE=1 SV=1	13	122.1	227.8	88.8	2.564	1.358	Up
7	High	P01859	Immunoglobulin heavy constant gamma 2 OS=Homo sapiens OX=9606 GN=IGHG2 PE=1 SV=2	4	35.9	179.4	71.7	2.502	1.323	Up
8	High	P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens OX=9606 GN=KRT1 PE=1 SV=6	17	66	56	95.3	0.588	-0.766	Down
9	High	P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens OX=9606 GN=KRT4 PE=1 SV=4	12	57.3	59.2	108.7	0.545	-0.876	Down
10	High	Q9Y646	Carboxypeptidase Q OS=Homo sapiens OX=9606 GN=CPQ PE=1 SV=1	2	51.9	84.6	162.8	0.52	-0.943	Down
11	High	P07911	Uromodulin OS=Homo sapiens OX=9606 GN=UMOD PE=1 SV=1	7	69.7	94.1	186.4	0.505	-0.986	Down
12	High	P05062	Fructose-bisphosphate aldolase B OS=Homo sapiens OX=9606 GN=ALDOB PE=1 SV=2	2	39.4	57.7	116.4	0.496	-1.012	Down
13	High	P09211	Glutathione S-transferase P	2	23.3	54	124.9	0.433	-1.208	Down

			OS=Homo sapiens OX=9606 GN=GSTP1 PE=1 SV=2							
14	High	P07355	Annexin A2 OS=Homo sapiens OX=9606 GN=ANXA2 PE=1 SV=2	8	38.6	78.5	184	0.427	-1.228	Down
15	High	P63104	14-3-3 protein zeta/delta OS=Homo sapiens OX=9606 GN=YWHAZ PE=1 SV=1	5	27.7	37.6	116.9	0.321	-1.639	Down

Data identified at confidence interval 99 %, FDR (False Discovery Rate) < 0.01, Peptides no. ≥ 2 ; Outlier of data that quantified at confidence interval 95 %. P < 0.05, highlighted data is showed identical in protein profile at both 15 and 20 weeks of PE.

.

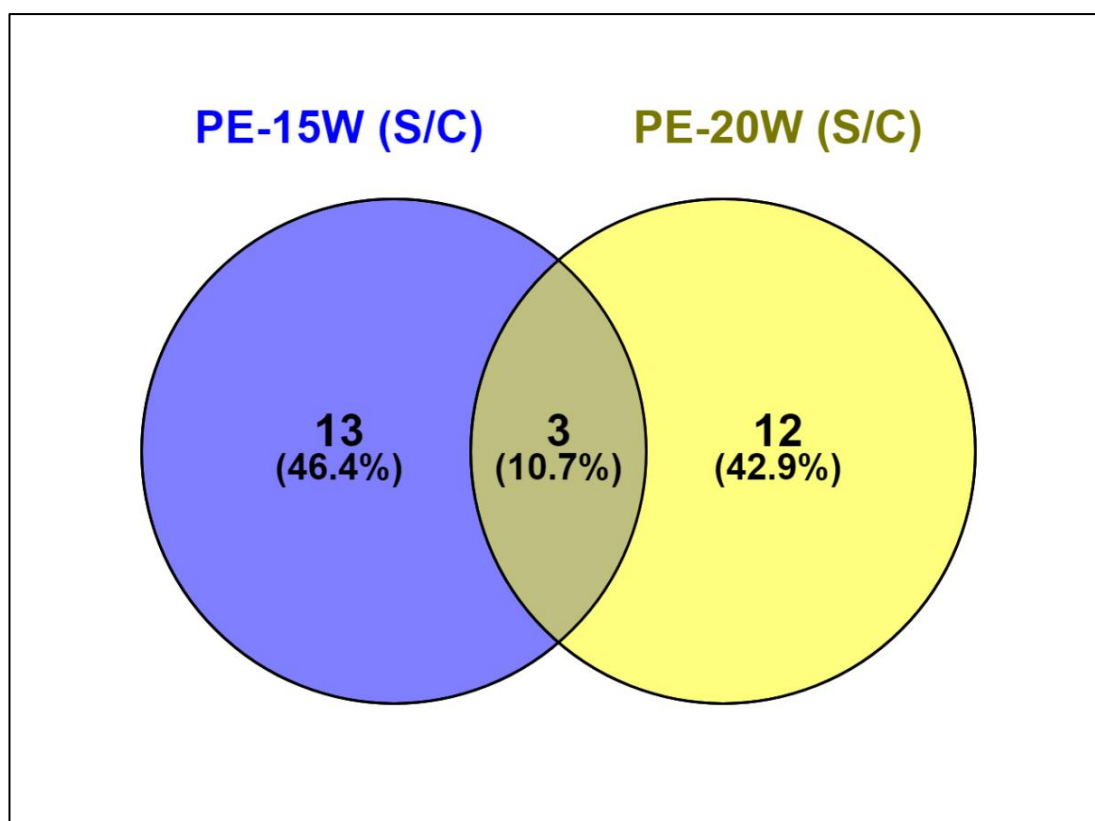


Figure 2. 10: Dysregulated proteins comparison between 15 and 20 weeks gestational age of PE

Venn diagram plot showing dysregulated proteins comparison between 15 and 20 week gestational age of PE. Outlier protein accession numbers were compared for each gestation. Data are presented as proteins numbers and % of differential protein profile at 15 and 20 weeks of PE.

2.8.2. Proteomic analysis for IUGR at 15 and 20 weeks gestational age

The full dataset of proteins identified as showing significant changes with a FDR < 0.01 and confidence interval 99 % related to IUGR at both 15 and 20 weeks gestational age respectively are shown in Table S 3 and S 4 (see Appendix 3). Total of 988 peptides from 315 proteins were identified. Proteins detected as differentially expressed were filtered using the following criteria: Peptide number ≥ 2 ; Confidence Interval (CI) 99 %, FDR <

0.01. 173/826 (proteins / peptides) following data filtering showed significant changes (FDR < 0.01 at confidence interval 99 %) in all 8 label channels at both 15 and 20 weeks gestational age of IUGR; these changes are shown in Table S 3 and S 4 (see Appendix 3).

Differentially expressed candidate proteins were also quantified using \log_2 fold change in abundance ratio (cases: control) at 15 and 20 weeks gestation for pregnancies subsequently complicated by IUGR respectively against uncomplicated controls. The comparisons at 15 and 20 weeks' gestation are shown in Figure 2.12 and 2.13 respectively. Quantitatively, all data are subjected to normalisation using an internal standard (yeast enolase), whose \log_2 fold change ratio was adjusted to zero (Figures 2.11, 2.12, and 2.13).

At 15 week gestational age, no statistically significant protein alterations quantitatively observed using one sample Wilcoxon test (Figure 2.11). For 20 week of gestation of IUGR, results showed highly statistically significant differences quantitatively ($P < 0.01$) at 95 % confidence interval using \log_2 fold change of abundance ratio (case/ control) using one sample Wilcoxon test. A median \log_2 fold change in ratio of (-0.353) was observed, with max. and min. values (1.549-(-2.44)) (Figure 2.12). Comparison ratios of \log_2 fold change (case/ control) between 15 and 20 weeks gestation pregnancies complicated by IUGR was showed a significant differences ($P < 0.01$) using a Mann-Whitney test at 95 % confidence interval is presented in Figure 2.13. .

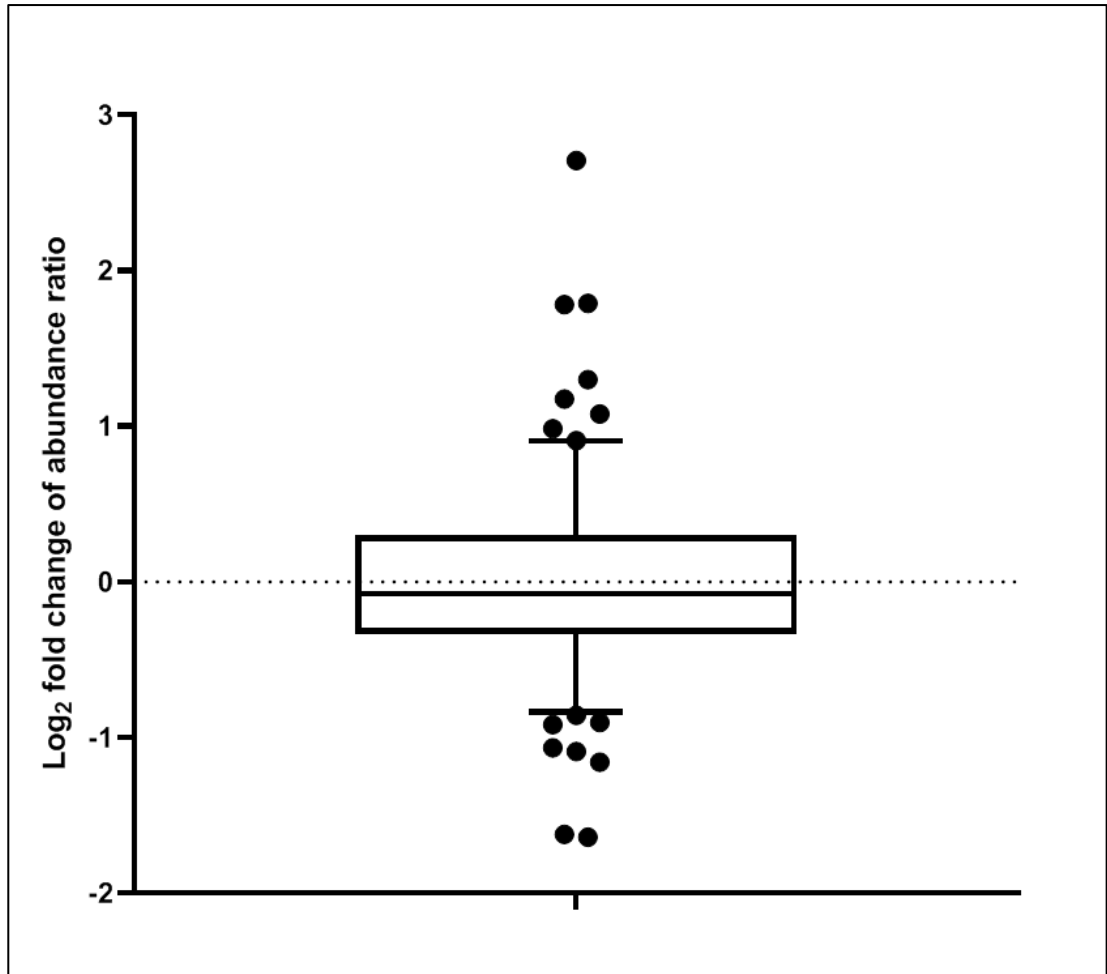


Figure 2. 11: Differentially expressed protein changes quantitatively for IUGR at 15 week of gestation

Plot displays \log_2 fold change of abundance ratio of (case: control) for IUGR at 15 week gestational age. Data were analysed showed differentially expressed proteins changes quantitatively vs internal standard of enolase at zero value (dotted line). No statistically significant change is observed for 173 proteins using one sample Wilcoxon test. Data were expressed as interquartile range of upper and lower quartiles (box), median (line inside box), lines at top and bottom represent the highest and lowest values (whiskers), confidence intervals (95 %), black balls represent outliers points that were showed highest and lowest ratio compared with internal standard.at 95 % confidence interval.

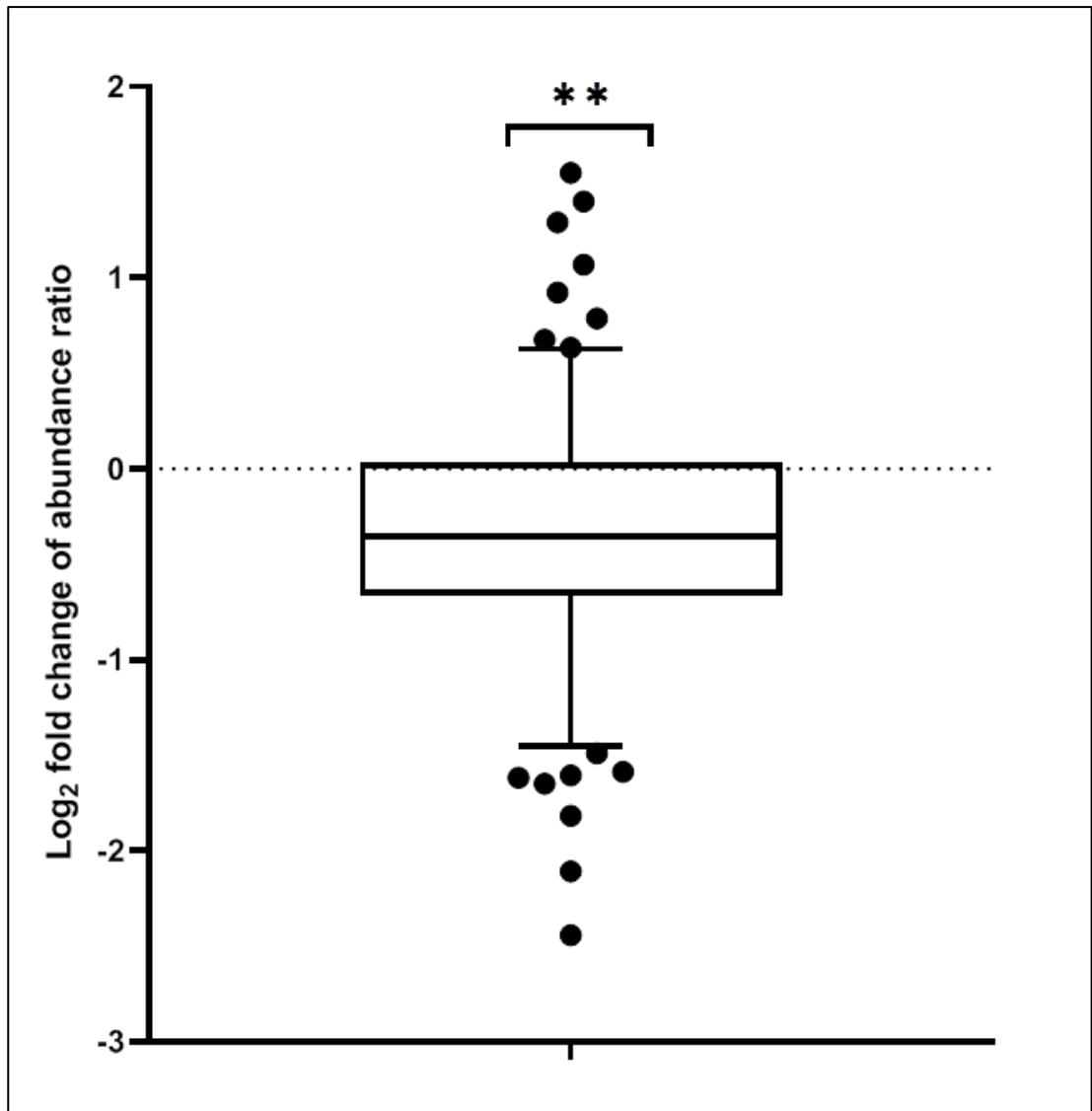


Figure 2. 12: Differentially expressed protein changes quantitatively for IUGR at 20 week of gestational age

Plot displays \log_2 fold change of abundance ratio of (case: control) for IUGR at 20 week' gestation. Data were analysed showed proteins differentially expressed changes quantitatively vs enolase internal standard at zero value (dotted line). 173 proteins are analysed, showed highly significant changes ($P < 0.01$) at confidence interval 95 % using one sample Wilcoxon test. Data are expressed as interquartile range of upper and lower quartiles (box), median (line inside box), lines at top and bottom represent the highest and lowest values (whiskers), confidence intervals (95 %), ** represents significant observed differences ($P < 0.01$), black balls represent outliers points that showed highest and lowest ratio compared with internal standard value.

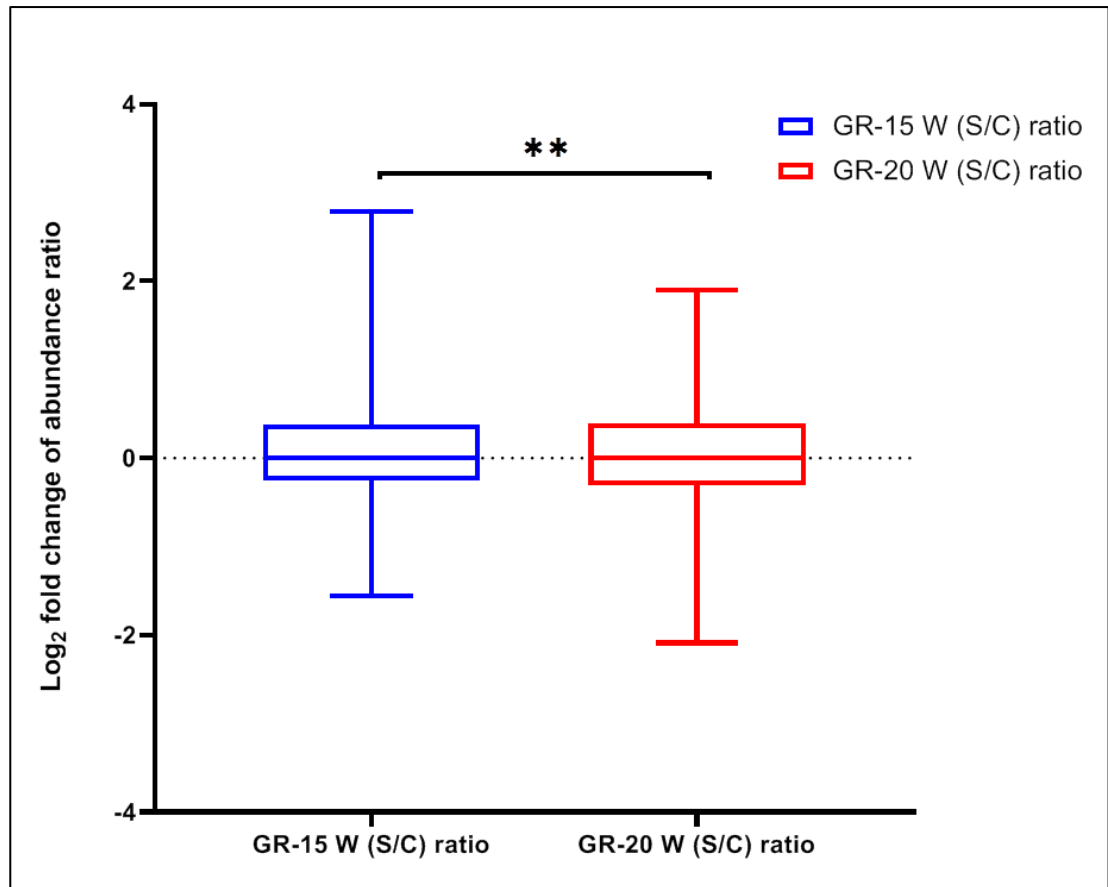


Figure 2. 13: Differentially expressed protein changes quantitatively for IUGR cases at 15 and 20 weeks' gestation

Plot shows comparison analysis of \log_2 fold change of abundance ratio of (case: control) for IUGR at 15 week (blue box) and 20 week (red box) gestational age respectively. Proteins showing significant changes in expression between 15 and 20 weeks ($P < 0.01$) are observed at confidence interval 95 %. 173 proteins at each gestational week were analysed using non-parametric analysis (Mann-Whitney test of unpaired two samples groups). Data are expressed as interquartile range of upper and lower quartiles (box), median (line inside box), lines at top and bottom are represented max and min values (whiskers). ** is represented extremely significant changes ($P < 0.01$), dotted line represents internal standard value of enolase.

The following section describes outliers/ dysregulated protein (highest / lowest) quantitatively of IUGR that determined at above section at 15 and 20 weeks of pregnancies respectively.

2.8.2.1 Highly dysregulated proteins for IUGR at 15 and 20 weeks

Differentially expressed proteins were identified as outlier observations in previous quantitative analysis (Figure 2.11 and 2.12), which showed highly dysregulated (up / down) proteins compared with uncomplicated pregnancies at 15 and 20 weeks of IUGR respectively. At 15 weeks gestation, outlier data were identified 16/ 74 (proteins/peptides) as being highly dysregulated for IUGR compared with uncomplicated pregnancies. Of which 50 % (n = 8) proteins are upregulated and downregulated respectively are showed in Table 2.5. For 20 weeks gestation, 16 / 60 (proteins/peptides) were showed dysregulation for IUGR-complicated pregnancies, of which 8 proteins are upregulated and 8 proteins are downregulated compared with uncomplicated pregnancies are showed in Table 2.6. Seven dysregulated proteins showing outlier behaviour had similar profiles at both 15 and 20 weeks gestation in IUGR-complicated pregnancies. These are zinc-alpha-2-glycoprotein, leucine-rich alpha-2-glycoprotein, alpha-1-acid glycoprotein 1, alpha-1-acid glycoprotein 2, histone H4, CD44 antigen and lysosome-associated membrane glycoprotein 2; these proteins are highlighted in Tables 2.5 and 2.6, and illustrated in Figure 2.14.

Table 2.5: Dysregulated proteins dataset of IUGR at 15 week gestational age compared with control

No.	Protein FDR Confidence: Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of IUGR at 15 W	Abundances of reporter ion of control at 15 W	Abundance ratio (case/control) of IUGR at 15W	Log ₂ fold change ratio	UP/ Down
1	High	P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=AZGP1 PE=1 SV=2	13	34.2	271.3	41.6	6.528	2.707	UP
2	High	P02750	Leucine-rich alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=LRG1 PE=1 SV=2	5	38.2	195.9	56.7	3.458	1.790	UP
3	High	P02763	Alpha-1-acid glycoprotein 1 OS=Homo sapiens OX=9606 GN=ORM1 PE=1 SV=1	3	23.5	236.2	68.6	3.442	1.783	UP
4	High	P19652	Alpha-1-acid glycoprotein 2 OS=Homo sapiens OX=9606 GN=ORM2 PE=1 SV=2	2	23.6	193.5	78.5	2.464	1.301	UP
5	High	P02765	Alpha-2-HS-glycoprotein OS=Homo sapiens OX=9606 GN=AHSG PE=1 SV=1	3	39.3	142.1	62.9	2.26	1.176	UP
6	High	P04217	Alpha-1B-glycoprotein OS=Homo sapiens OX=9606 GN=A1BG PE=1 SV=4	3	54.2	146.8	69.5	2.112	1.079	UP
7	High	P01009	Alpha-1-antitrypsin OS=Homo sapiens OX=9606 GN=SERPINA1 PE=1 SV=3	14	46.7	181.5	91.8	1.978	0.984	UP
8	High	P62805	Histone H4 OS=Homo sapiens OX=9606 GN=HIST1H4A PE=1 SV=2	3	11.4	149	79.4	1.876	0.908	UP

9	High	P07911	Uromodulin OS=Homo sapiens OX=9606 GN=UMOD PE=1 SV=1	7	69.7	82.1	148.7	0.552	-0.857	Down
10	High	P16070	CD44 antigen OS=Homo sapiens OX=9606 GN=CD44 PE=1 SV=3	3	81.5	78.6	146.8	0.535	-0.902	Down
11	High	P01877	Immunoglobulin heavy constant alpha 2 OS=Homo sapiens OX=9606 GN=IGHA2 PE=1 SV=4	3	36.6	82.3	155.3	0.53	-0.916	Down
12	High	P08637	Low affinity immunoglobulin gamma Fc region receptor III-A OS=Homo sapiens OX=9606 GN=FCGR3A PE=1 SV=2	2	29.1	110.9	232.1	0.478	-1.065	Down
13	High	P13473	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens OX=9606 GN=LAMP2 PE=1 SV=2	2	44.9	98.8	210.1	0.47	-1.089	Down
14	High	P31944	Caspase-14 OS=Homo sapiens OX=9606 GN=CASP14 PE=1 SV=2	2	27.7	85	189.9	0.448	-1.158	Down
15	High	P13796	Plastin-2 OS=Homo sapiens OX=9606 GN=LCP1 PE=1 SV=6	4	70.2	52.4	161.2	0.325	-1.621	Down
16	High	P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens OX=9606 GN=KRT2 PE=1 SV=2	5	65.4	89.2	277.4	0.321	-1.639	Down

Data identified at confidence interval 99 %, FDR (False Discovery Rate) < 0.01, Peptides no.≥ 2; Outlier of data that quantified at confidence interval 95

%. P < 0.05, highlighted data is showed identical in protein profile at both 15 and 20 weeks of IUGR.

Table 2.6: Dysregulated proteins dataset of IUGR at 20 week gestational age compared with control.

No.	Protein FDR Confidence: Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of IUGR at 20 W	Abundances of reporter ion of control at 20 W	Abundance ratio (case/control) of IUGR at 20W	Log ₂ fold change ratio	Up/Down
1	High	P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=AZGP1 PE=1 SV=2	13	34.2	171.4	58.6	2.927	1.549	Up
2	High	P02675	Fibrinogen beta chain OS=Homo sapiens OX=9606 GN=FGB PE=1 SV=2	2	55.9	133.2	50.5	2.638	1.399	Up
3	High	P02750	Leucine-rich alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=LRG1 PE=1 SV=2	5	38.2	173.2	70.8	2.445	1.290	Up
4	High	P02763	Alpha-1-acid glycoprotein 1 OS=Homo sapiens OX=9606 GN=ORM1 PE=1 SV=1	3	23.5	131.4	62.7	2.097	1.068	Up
5	High	P19652	Alpha-1-acid glycoprotein 2 OS=Homo sapiens OX=9606 GN=ORM2 PE=1 SV=2	2	23.6	127.6	67.3	1.895	0.922	Up
6	High	P16070	CD44 antigen OS=Homo sapiens OX=9606 GN=CD44 PE=1 SV=3	3	81.5	104.1	60.3	1.727	0.788	Up

7	High	Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1 OS=Homo sapiens OX=9606 GN=LYVE1 PE=1 SV=2	2	35.2	110.5	69.2	1.598	0.676	Up
8	High	Q01459	Di-N-acetylchitobiase OS=Homo sapiens OX=9606 GN=CTBS PE=1 SV=1	2	43.7	149.5	96.3	1.552	0.634	Up
9	High	P34059	N-acetylgalactosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GALNS PE=1 SV=1	2	58	62.4	175.3	0.356	-1.490	Down
10	High	P60174	Triosephosphate isomerase OS=Homo sapiens OX=9606 GN=TPI1 PE=1 SV=3	2	30.8	39.3	118	0.333	-1.586	Down
11	High	P62736	Actin, aortic smooth muscle OS=Homo sapiens OX=9606 GN=ACTA2 PE=1 SV=1	5	42	65.4	198.7	0.329	-1.604	Down
12	High	P11021	Endoplasmic reticulum chaperone BiP OS=Homo sapiens OX=9606 GN=HSPA5 PE=1 SV=2	5	72.3	59.4	182.3	0.326	-1.617	Down
13	High	P62805	Histone H4 OS=Homo sapiens OX=9606 GN=HIST1H4A PE=1 SV=2	3	11.4	42.8	134.2	0.319	-1.648	Down
14	High	P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens OX=9606 GN=KRT16 PE=1 SV=4	6	51.2	44.8	157.9	0.284	-1.816	Down

15	High	P13473	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens OX=9606 GN=LAMP2 PE=1 SV=2	2	44.9	50.5	218	0.232	-2.108	Down
16	High	P02788	Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF PE=1 SV=6	3	78.1	28.7	156.2	0.184	-2.442	Down

Data identified at confidence interval 99 %, FDR (False Discovery Rate) < 0.01, Peptides no.≥ 2; Outlier of data that quantified at confidence interval 95

%. P < 0.05, highlighted data is showed identical in protein profile at both 15 and 20 weeks of IUGR.

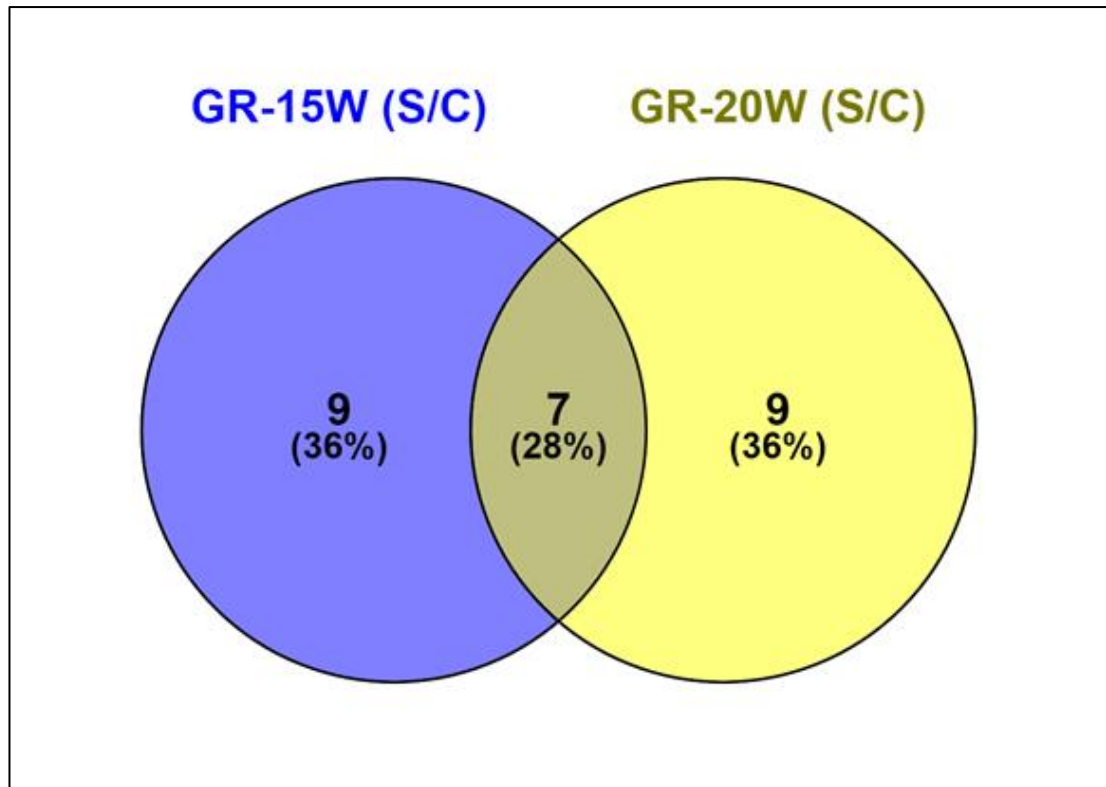


Figure 2. 14: Dysregulated proteins comparison between 15 and 20 weeks gestational age of IUGR

Venn diagram plot shows dysregulated proteins comparison associates with differences / matches of protein profile between 15 and 20 week gestational age of IUGR. Outlier data are associate with accession number of protein profile correspond with \log_2 fold change of abundance ratio of (case/ control) was used for this analysis. Data are presented as proteins numbers and % of protein profile at both 15 and 20 weeks of IUGR.

2.9. Discussion

The term “*biomarker*” is defined by the medical subject heading term: “*measurable and quantifiable biological parameters, (e.g., specific enzyme concentration, specific hormone concentration, specific gene phenotype distribution in a population, presence of biological substances) which serve as indices for health- and physiology-related assessments, such as disease risk, psychiatric disorders, environmental exposure and its effects, disease diagnosis, metabolic processes, substance abuse, pregnancy, etc.*” (Bergman and Bergquist, 2014). Desirable characteristics for a biomarker include that it should be specifically and quantitatively correlated with the state or disorder in question; it can be studied with particular and sensitive approaches, which provide reproducible outcomes (Bergman and Bergquist, 2014). In addition, the biomarker should be chemically stable during examination and storage, and should not be susceptible to alteration from outside sources (Bergman and Bergquist, 2014). Ideally, a biomarker is also available by non-invasive estimation, and in an individual the levels of biomarker should be somewhat constant (Rossi *et al.*, 2006).

Various biomarkers have been demonstrated during pregnancy complicated with preeclampsia. For instance, vascular endothelial growth factor (VEGF), placental growth factor (PlGF), and their antagonists such as fms-like tyrosine kinase 1 (sFlt-1) and endoglin (sEng) are common biomarkers which are significantly associated with preeclampsia (Grill *et al.*, 2009, Carty, 2012, Kar, 2014, Unal *et al.*, 2007). These biomarkers (VEGF and PlGF) have been shown significantly decreased in serum from the end of first trimester gestational age throughout pregnancy in women at high risk of condition. Buhimschi and colleagues used immunoassay method to study proteomic biomarkers in preeclampsia

(Buhimschi *et al.*, 2005). In this study, 68 women were enrolled prospectively categorised: non-pregnant reproductive age (NP-CTR n = 14), uncomplicated pregnant control (P-CTR n = 16), pregnant hypertensive and proteinuric women who did not meet criteria for severe preeclampsia (pHTN n = 21), and women with severe preeclampsia (sPE n = 17). The urinary output of antiangiogenic factor sFlt-1 was found to be increased, while output of placental growth factor PlGF decreased at the time of clinical manifestation in severe preeclampsia. Authors suggested these results provide a fast non-invasive screening of hypertensive women based on a sFlt / PlGF ratio. They concluded that the sFlt / PlGF ratio could be used as marker for severity of preeclampsia; this measure shows improved specificity and selectivity compared to random urinary protein measurements (Buhimschi *et al.*, 2005).

In the present study, outcome identified 988 peptides from 315 proteins totally in all 8-plex labelling samples at 99 confidence interval, FDR < 0.01, of which 173 proteins have a sufficient reporter ion signal for quantification related to PE and IUGR at each week respectively (see supplementary Tables S1-S4 in appendix 3). Proteins with < 2 peptides are excluded from identification / quantification analyses, due to the low confidence associated with these “one-hit wonders”; others have suggested that these data should be excluded from ratiometric determinations (Evans *et al.*, 2012).

Identified proteins showing differential expression, associated with log₂ fold change of abundance ratio (cases: control) are reported for 15 and 20 weeks of PE and IUGR respectively compared with uncomplicated pregnancies, (Figure 2.7 - 2.9 for PE, Figure 2.11 - 2.13 for IUGR). All data were normalised using enolase as an internal standard.

According to PE, dysregulated proteins (n=173) that shown highly significant differences ($P < 0.01$) in protein profile at 95 % confidence interval using one sample Wilcoxon test Figure 2.7 and 2.8 at 15 and 20 weeks of PE respectively. Additionally, quantitative data in Figure 2.9 are showed highly significant differences ($P < 0.01$) between 15 and 20 weeks of PE related to \log_2 fold change abundance ratio (case / control) at 95 % confidence interval using Mann-Whitney test. Similar study of urinary proteomic analysis using iTRAQ labelling worked by (Chen *et al.*, 2011), they identified 113 proteins. These proteins were expressed differentially between complicated hypertension (PE and hypertension pregnancies) and normal pregnancy. Of which 31/113 proteins revealed statistically significant difference between these study groups.

Outlier data in these analyses are represented as highly dysregulated proteins (up / down) compared with control for PE (Table 2.3 and 2.4). 16 proteins were identified as outliers suggesting highly dysregulated proteins compared with uncomplicated pregnancies using one sample Wilcoxon test (Figure 2.7), of which 50 % (n = 8) proteins are upregulated and downregulated respectively for PE during 15 week of pregnancy (Table 2.3). For 20 week of PE, 7 proteins were upregulated, and 8 downregulated proteins were identified (Table 2.4). However, in present study, the putative proteomic biomarkers (n=28) showed a highly dysregulated protein profile compared with the corresponding proteins in a control group of uncomplicated pregnancies for PE at both 15 and 20 weeks; those identified in previous studies are illustrated in Table 2.7. 3 Proteins identified with similar profiles at both 15 and 20 weeks of PE (fibrinogen beta chain, haptoglobin and carboxypeptidase) are highlighted in Table 2.3 and 2.4 and shown in

Figure 2.10. Similar identifications were observed in previous studies are highlighted in Table 2.7 below.

Table 2.7: Proteomic biomarkers for PE

Protein name	Sample source	References
Alpha-2-macroglobulin	Serum / plasma, urine	(Auer <i>et al.</i> , 2010), (Blumenstein <i>et al.</i> , 2009), (Chen <i>et al.</i> , 2011), (Carty <i>et al.</i> , 2011), (D'Silva <i>et al.</i> , 2018), (Little <i>et al.</i> , 2010)
Fibrinogen beta chain	Serum, urine	(Auer <i>et al.</i> , 2010), (Chen <i>et al.</i> , 2011)
Leukocyte-associated immunoglobulin-like receptor 1	Blood / bone marrow	(Van der Vuurst de Vries <i>et al.</i> , 1999)
Osteoclast-associated immunoglobulin-like receptor	Uterine cells	(Weiner <i>et al.</i> , 2010)
Haptoglobin	Serum, urine, amniotic fluid	(Auer <i>et al.</i> , 2010), (Chen <i>et al.</i> , 2011), (D'Silva <i>et al.</i> , 2018), (Tsangaris <i>et al.</i> , 2006)
Lactotransferrin	Serum	(Auer <i>et al.</i> , 2010)
Alpha-enolase	Urine	(Chen <i>et al.</i> , 2011)
Plastin-2	Serum	(Nagalla <i>et al.</i> , 2010)
Peroxiredoxin-1	Urine	(Chen <i>et al.</i> , 2011)
Lysosomal alpha-glucosidase	Serum / urine	(Jackson <i>et al.</i> , 1996)
Carboxypeptidase Q	urine	(Carty <i>et al.</i> , 2011)
Neurotrimin	Uterine cells	(Weiner <i>et al.</i> , 2010)
L-lactate dehydrogenase B chain	Urine	(Carty <i>et al.</i> , 2011)
Roundabout homolog 4	Urine, serum	(Chen <i>et al.</i> , 2011), (Nagalla <i>et al.</i> , 2010), (Myers <i>et al.</i> , 2013)
Endosialin	Blood	(Enquobahrie <i>et al.</i> , 2011)
Glutamyl aminopeptidase	Serum	(Mizutani <i>et al.</i> , 1987)
Serotransferrin	Urine, amniotic fluid	(Chen <i>et al.</i> , 2011), (Tsangaris <i>et al.</i> , 2006)
Immunoglobulin heavy constant gamma 4	Blood, cells/ DNA	(Ellison and Hood, 1982), (Hofker <i>et al.</i> , 1989)
Immunoglobulin heavy constant alpha 2	Cells / DNA	(Hofker <i>et al.</i> , 1989)
Ceruloplasmin	Serum, urine	(Auer <i>et al.</i> , 2010), (Chen <i>et al.</i> , 2011)

Immunoglobulin heavy constant gamma 2	Cells /DNA	(Hofker <i>et al.</i> , 1989)
Keratin, type II cytoskeletal 1	Serum, urine	(Auer <i>et al.</i> , 2010), (Chen <i>et al.</i> , 2011)
Keratin, type II cytoskeletal 4	Serum, urine	(Auer <i>et al.</i> , 2010), (Chen <i>et al.</i> , 2011), (Carty <i>et al.</i> , 2011)
Uromodulin	Urine	(Carty <i>et al.</i> , 2011)
Fructose-bisphosphate aldolase B	Blood	(Cox, 1994)
Glutathione S-transferase P	Serum	(Nagalla <i>et al.</i> , 2010)
Annexin A2	Placental cells	(Xin <i>et al.</i> , 2012)
14-3-3 protein zeta / delta	Serum	(Nagalla <i>et al.</i> , 2010)

Among proteomic biomarkers that are shown as highly dysregulated in their profile associated with complicated pregnancies such as PE compared with uncomplicated pregnancies (Table 2.7), for instance, alpha-2-macroglobulin, haptoglobin and roundabout were identified. According to alpha-2-macroglobulin, log₂ fold change abundances ratio (case: control) has shown upregulated at value (3.5) at 15 week of PE compared with uncomplicated pregnancies (control). Alpha-2-macroglobulin has antiplasmin activity and possibly the increasing in this levels protein in preeclampsia explain its important role in the intravascular coagulation of this condition (Horne *et al.*, 1970). Similar findings were reported in serum-based studies of proteomic biomarkers of PE (Auer *et al.*, 2010, Blumenstein *et al.*, 2009, D'Silva *et al.*, 2018). Studies of proteomic biomarkers for PE also agreed with findings of other groups using urine sample (Chen *et al.*, 2011, Carty *et al.*, 2011).

Another biomarker shown as high dysregulated in protein expression in this study is haptoglobin, log₂ fold change abundances ratio (case: control) showed upregulation (2.53-fold) in PE cases compared with uncomplicated pregnancies (control) at 15 weeks'

gestation. Haptoglobin is an alpha-2 sialoglycoprotein that is found in plasma, and is considered to have antioxidative properties (Tseng *et al.*, 2004). The major physiological role of haptoglobin is to capture, and combine with the free plasma hemoglobin that results from hemolysis (Langlois and Delanghe, 1996), thereby preventing the release of heme iron which can accumulate in the kidney and cause oxidative damage (Tolosano *et al.*, 2002) The finding is reported previously by (Sammour *et al.*, 2010). Previous studies have reported similar result using different types of samples e.g. urine and / or serum to study proteomic biomarkers in PE (Auer *et al.*, 2010, Blumenstein *et al.*, 2009, Chen *et al.*, 2011, Carty *et al.*, 2011, D'Silva *et al.*, 2018, Little *et al.*, 2010) (see Table 2.7).

Roundabout (ROBO4) showed downregulation (-3.19-fold) in protein expression for PE subjects at 15 weeks gestational age compared with non-complicated pregnancies (Table 2.3). Roundabout (ROBO4), is also known as axon guidance receptor, homolog 4 (*Drosophila*); the protein is reported to be a receptor for SLIT proteins, at least for SLIT2, and seems to be involved in angiogenesis and vascular patterning (Huminiacki *et al.*, 2002). This protein has previously been found in developing placental and embryonic tissues (Park *et al.*, 2003), and has been shown to inhibit endothelial cell migration (Park *et al.*, 2003). As such downregulation in specific protein expression could be an indication of poor vascular development in the growing fetoplacental unit, and hence a direct readout for the impaired fetal development observed in later pregnancy. Similar finding has been previously identified in urine (Chen *et al.*, 2011) and in serum / blood studies of preeclampsia (Nagalla *et al.*, 2010, Myers *et al.*, 2013) (see Table 2.7).

Similar analysis was performed using urine sample of IUGR condition during 15 and 20 weeks gestational age. According to IUGR at 15 weeks' gestation, no statistically

significant quantitative differences were detected for the 173 proteins observed, 16 differentially expressed proteins were identified as outliers. Eight proteins are (up/down) respectively compared with uncomplicated pregnancies were identified at confidence interval 95 % using one sample Wilcoxon test (see Figure 2.11). For 20 week of gestation of IUGR, results showed highly statistically significant differences quantitatively ($P < 0.01$) at 95 % confidence interval of \log_2 fold change of abundance ratio (case/control) using one sample Wilcoxon test within median (-0.353) within max. and min. values at (1.549-(-2.44)) of 173 proteins (Figure 2.12). Additionally, quantitative data were observed in Figure 2.13 had showed highly significant differences ($P < 0.01$) between 15 and 20 weeks of IUGR at 95 % confidence interval associates with \log_2 fold change of abundance ratio (case/control) using Mann-Whitney test. Similar study worked by (Auer *et al.*, 2010) using iTRAQ labelling coupled with MS analysis to identify proteomic biomarkers in IUGR using serum sample. Outcome revealed 166 proteins totally were identified, of which 31/128 identified proteins/ peptides showing expression changes in IUGR (14 up and 17 down-regulated) (Auer *et al.*, 2010).

Outlier data analysis identified a total of 74 peptides from 16 proteins as candidate dysregulated proteins. 50 % ($n = 8$) of these proteins are each upregulated and downregulated respectively at 15 weeks' gestation for IUGR subjects compared with uncomplicated pregnancies, shown in Table 2.5. At 20 weeks' gestation for IUGR cases, 16 / 60 (protein/peptides) were observed as highly dysregulated, again, 8 proteins are upregulated and 8 proteins are downregulated compared with uncomplicated pregnancies, shown in Table 2.6. Previous studies corroborate these findings (Table 2.8) in various different sample types such as urine (Chen *et al.*, 2011)(Carty *et al.*, 2011),

serum (Auer *et al.*, 2010), plasma (Blumenstein *et al.*, 2009), amniotic fluid (Tsangaris *et al.*, 2006) and placental cells (Mandò *et al.*, 2016, Xin *et al.*, 2012), to identify different proteomic biomarkers.

Table 2. 8: Proteomic biomarkers for IUGR

Protein name	Sample source	References
Zinc-alpha-2-glycoprotein	Serum, urine	(Auer <i>et al.</i> , 2010); (Blumenstein <i>et al.</i> , 2009); (Chen <i>et al.</i> , 2011); (Carty <i>et al.</i> , 2011)
Leucine-rich alpha-2-glycoprotein	Amniotic fluid, serum,	(Tsangaris <i>et al.</i> , 2006), (Kay <i>et al.</i> , 2009)
Alpha-1-acid glycoprotein 1	Serum	(Auer <i>et al.</i> , 2010)
Alpha-1-acid glycoprotein 2	Amniotic fluid	(Tsangaris <i>et al.</i> , 2006)
Alpha-2-HS-glycoprotein	Serum	(Auer <i>et al.</i> , 2010)
Alpha-1B-glycoprotein	Serum	(D'Silva <i>et al.</i> , 2018)
Alpha-1-antitrypsin	Urine	(Chen <i>et al.</i> , 2011)
Histone H4	Urine	(Chen <i>et al.</i> , 2011)
Uromodulin	Urine	(Chen <i>et al.</i> , 2011)
CD44 antigen	Placental cells	(Mandò <i>et al.</i> , 2016); (Banerji <i>et al.</i> , 1999)
Immunoglobulin heavy constant alpha 2	Cells	(Hofker <i>et al.</i> , 1989)
Low affinity immunoglobulin gamma Fc region receptor III-A	Cells	(Hofker <i>et al.</i> , 1989)
Lysosome-associated membrane glycoprotein 2	Cells	(Gough and Fambrough, 1997)
Caspase-14	Cytotrophoblastic cells	(White <i>et al.</i> , 2007)
Plastin-2	Serum	(Nagalla <i>et al.</i> , 2010)
Keratin, type II cytoskeletal 2 epidermal	Urine	(Chen <i>et al.</i> , 2011)
Fibrinogen beta chain	Urine	(Chen <i>et al.</i> , 2011)
Alpha-1-acid glycoprotein 1	Serum	(Auer <i>et al.</i> , 2010), (Tsangaris <i>et al.</i> , 2006)
N-acetylgalactosamine-6-sulfatase	Placental cells	(Nakashima <i>et al.</i> , 1994)
Triosephosphate isomerase	Plasma	(Little <i>et al.</i> , 2010)
Actin, aortic smooth muscle	Serum	(Nagalla <i>et al.</i> , 2010)
Endoplasmic reticulum chaperone BiP	Placental cells	(Burton <i>et al.</i> , 2009)

Keratin, type I cytoskeletal 16	Serum, urine	(Auer <i>et al.</i> , 2010), (Chen <i>et al.</i> , 2011)
Lactotransferrin	Milk	(Weiner <i>et al.</i> , 2010)

However, 7 dysregulated proteins detected here as outlier observations have shown identical profiles at both 15 and 20 weeks of IUGR: zinc-alpha-2-glycoprotein, leucine-rich alpha-2-glycoprotein, alpha-1-acid glycoprotein 1, alpha-1-acid glycoprotein 2, histone H4, CD44 antigen and lysosome-associated membrane glycoprotein 2. These proteins are highlighted in Table 2.5 and 2.6, and shown in Figure 2.14. Similar findings were observed in various previous studies, highlighted in Table 2.8 above

Among proteomic biomarkers which have shown dysregulation in their profile associated with pregnancies complicated by IUGR shown in Table 2.8, for example, Zinc-alpha-2-glycoprotein, log₂ fold change value has upregulated at (2.707) in Table 2.5. Zinc-alpha-2-glycoprotein, has various important functions in the human body, including fertilization (Ohkubo *et al.*, 1990) and lipid mobilization (Morse *et al.*, 2017). Zinc-alpha-2-glycoprotein stimulates lipid degradation in adipocytes and causes the extensive fat losses associated with some advanced cancers (Hirai *et al.*, 1998, Zhang *et al.*, 2018); and could bind with polyunsaturated fatty acids. Therefore, the upregulation in that protein's expression explains the impaired fetal growth observed in later pregnancy. Similar outcomes were identified by previous studies using serum or urine samples to study proteomic biomarkers in complicated pregnancies such as IUGR and/or PE (Auer *et al.*, 2010, Blumenstein *et al.*, 2009, Chen *et al.*, 2011, Carty *et al.*, 2011).

Leucine-rich alpha-2-glycoprotein-1 (Lrg1), another biomarker in IUGR has showed (1.79) upregulated (Table 2.5). Leucine-rich alpha-2-glycoprotein-1 (Lrg1) is an emerging

biomarker for angiogenesis (Lio *et al.*, 2018). Angiogenesis a process of new blood vessel formation that participates to the development and progression of several diseases, such as cancers, proliferative diabetic retinopathy and chronic wound healing (Carmeliet, 2005). Vascular endothelial growth factors (VEGFs) and their receptors play a crucial role in angiogenesis (Carmeliet *et al.*, 1996). This protein has recently been found up-regulated in ocular tissues in both human patients with proliferative diabetic retinopathy and rodent models of pathological angiogenesis (Zhang *et al.*, 2018). Similar finding was detected using serum sample to identify protein leucine-rich alpha-2-glycoprotein as a putative biomarker of recombinant human growth hormone (rhGH) abuse which worked by (Kay *et al.*, 2009). Another study has also identified similar protein using amniotic fluid to predict abnormal fetuses biomarkers has been worked by (Tsangaris *et al.*, 2006)

The presence of urothelial markers as differentially expressed proteins is an indication that the amount of epithelial shedding from different individual during micturition (urination) is playing a role in the biological variation observed; therefore, this influence needs to be excluded before further exploitation of these putative biomarkers is possible.

2.10. Conclusions

All in all, the passage of fetoplacental markers of hypoxia and vascular dysfunction into urine via the maternal blood in a manner enabling quantitative detection of their variation is a strong indicator that this strategy has merit for determination of potential markers for malfunction of the fetoplacental unit. This study identified 28 proteomic biomarkers have shown highly dysregulated (up / down) for PE, and 24 biomarkers for IUGR compared with uncomplicated pregnancies as shown in Table 2.7 and 2.8 respectively. The limitations of this study are that we have not performed independent

quantitative validation of these markers in a separate cohort of women who have the same conditions, nor have putative markers found in pooled samples been investigated in (preferably independent) individuals. Confirmation of these markers on an independent set of individual patient samples, using targeted mass spectrometric assay and/or biochemical assay methods such as ELISA is vital before these markers can be taken forward for clinical translation.

Chapter 3

3. Fourier transform infrared analysis of urine – optimisation method using animal model

3.1 Introduction Infrared Spectroscopy

Infrared (IR) radiation is a component of the electromagnetic spectrum. The IR spectral region covers a wide range of frequencies, typically segmented into three ranges, the near (NIR), mid (MID) and far (FIR) infrared ranges. This region of the electromagnetic spectrum interacts strongly with biological macromolecules. This property is exploited in infrared spectroscopy, where a beam of IR radiation typically interacts with a sample of interest and the incident spectra are detected by comparison with a background reference 'sample' (Griffiths and De Haseth, 2007). Energy within the IR range is absorbed by covalent bonds within molecules, inducing an excited quantum state and consequent bond vibration. The strength of vibrational bonds between atoms differs such that the excitation energy can be used as a distinguishing feature, i.e. each bond type absorbs at a different energy, which is read from an infrared spectra as a function of wavenumber ($1/\lambda$; being related to the energy of the bond via Equations 3.1 and 3.2). Different spectra are observed for molecular species, with the maximum absorbance for a species varying with the bonds/atoms involved and the strength of any intermolecular interactions. Hence the IR spectrum is a fingerprint of the chemical structure of a given sample (Griffiths and De Haseth, 2007).

$$c=f\lambda \qquad \text{Equation 3.1}$$

Where c is the speed of light in a vacuum, f is the frequency of vibration (in Hz) and λ is the wavelength of the vibration (in cm).

$$E=hf \qquad \text{Equation 3.2}$$

Where E is the energy of the bond Joules (J) and h is Plank's constant (6.626×10^{-34} J·s)

In a given molecule, the energy is partitioned into $3N$ (N being the number of atoms in the molecule) degrees of freedom; these describe three translational and three rotational motions in the molecule. In order to absorb IR radiation, the vibrational displacement of the molecular atoms must cause a change in dipole moment; therefore, not all molecular vibrations absorb IR radiation. For example, a carbon dioxide molecule allows four vibrational modes (Colthup, 2012). The CO_2 molecule is a linear molecule in shape, with consequently four vibrational modes, the symmetric stretching (st), the anti-symmetric or asymmetric (as) and two bending vibrational modes (Colthup, 2012). During symmetric stretching, both oxygen atoms of the molecule are stretched/ compressed at the same time, causing no change in net dipole (Figure 3.1 A). The symmetric stretch vibration of the molecule therefore does not absorb IR radiation, while both the anti-symmetric and bending vibrations of the CO_2 molecule show strong absorbance maxima (Figure 3.1).

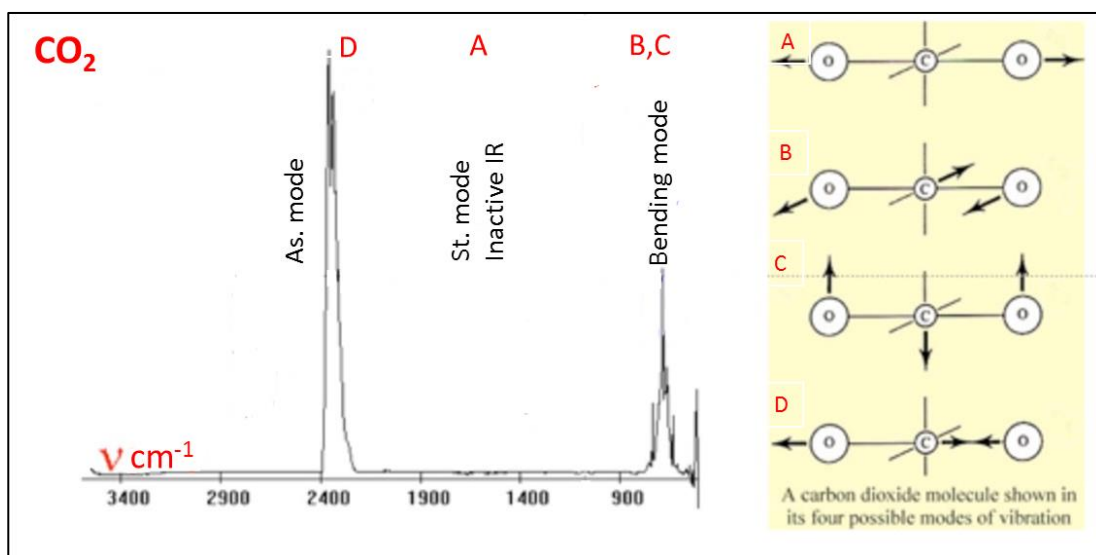


Figure 3. 1: IR spectra of CO_2 molecule and its vibrational modes

Main panel: IR spectrum of CO_2 showing absorbance maxima associated with varying modes of vibration (inset). Carbon dioxide has four theoretically possible vibrational modes: A: symmetric stretch (st), which lacks change in dipole and consequently spectral output, D: asymmetric stretch

(as), predicted at 2640 cm^{-1} and observed at 2345 cm^{-1} , whereas both B, C: are bending modes, predicted at 546 cm^{-1} vibrational mode (Colthup, 2012).

3.2 FTIR spectroscopy and Michelson interferometer

Fourier transform infrared (FTIR) spectroscopy utilises an interferometer based on the Michelson interferometer (Figure. 3.2). This device features a beam splitter, fixed mirror and a movable mirror, with adjustments in the path of the beam caused by the moveable mirror generating interference which is then recorded by the detector. This movement enables FTIR simultaneous measurement of all infrared frequencies (Griffiths and De Haseth, 2007). This speeds up the rate of analysis considerably in comparison to monochromatic approaches.

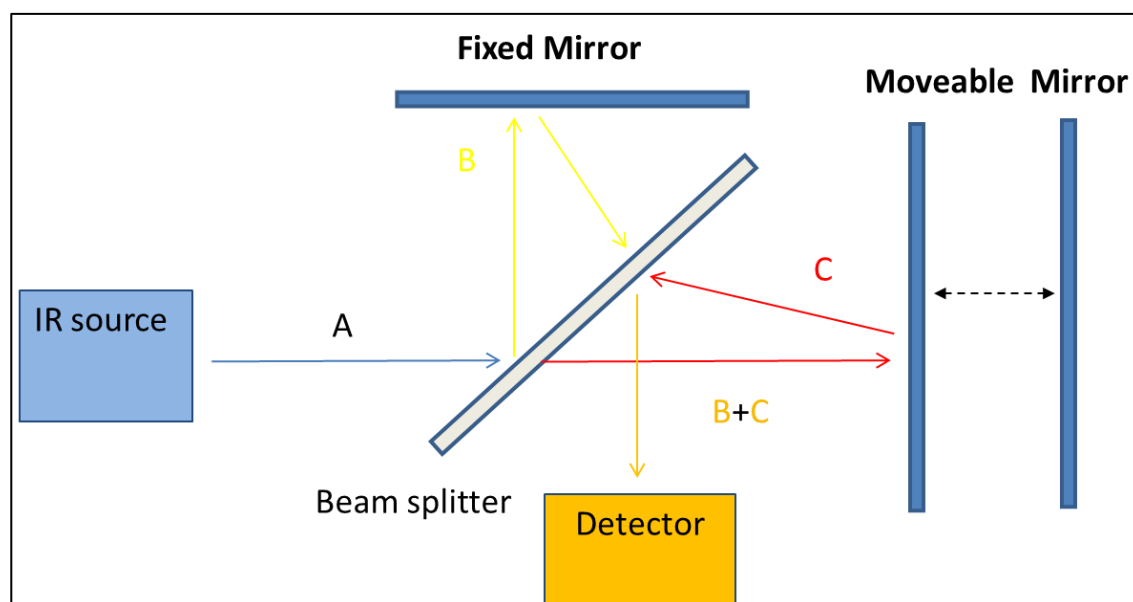


Figure 3. 2: Schematic diagram of a Michelson interferometer

An initial source beam (A) is split towards both fixed and moveable mirror by the beam splitter. Reflected beams are recombined using a one-sided mirror and the collimated beams detected.

3.3 Spectroscopy and Beer-Lambert law relationship

In spectroscopy wavenumber ($\tilde{\nu}$) is used to describe frequency divided by the speed of light in vacuum (Fuwa and Valle, 1963). Spectrophotometric analysis is typically used to quantitatively analyse a light absorbing analyte in solution according to the Beer-Lambert law. This relationship principally states that absorbance is directly proportional to the concentration of analyte in the sample, (Fuwa and Valle, 1963) described by Equation .3.3

$$A = \epsilon lc \quad \text{Equation (3.3)}$$

A refers to an absorbance, ϵ is the molar absorptivity at that wavelength or known as (molar extinction coefficient), **l** is the distance the light travels through the sample (the path length), and **c** is the concentration of the analyte in solution.

The light passed through a sample of interest has a fixed pathlength, **l**. The detector measures the incident intensity of the light **I**, and compares it to that of initial source light I_0 . The ratio of $[I / I_0]$ is a measure of the fraction of light that passes through a given sample and is called the transmittance (T) (Figure 3.3) (Fuwa and Valle, 1963). Absorbance is related to transmittance at Equation. (3.4)

$$A = -\log [I / I_0] \quad \text{Equation (3.4)}$$

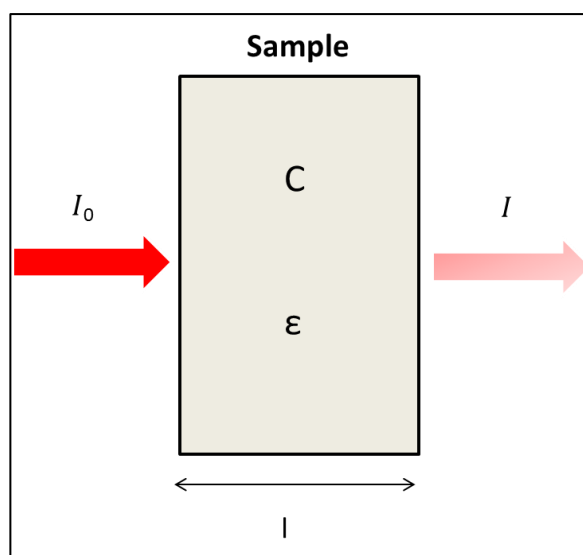


Figure 3. 3: Incident light vs. transmitted

Intensity of the transmitted light reduced due to sample absorption proportional to the specific absorbance coefficient ϵ and concentration of a given substance C . For a constant given pathlength l , the absorbance of a compound will have the mathematical relationship as above (Fuwa and Valle, 1963).

3.4 FTIR spectroscopy to study biological molecules

Biological samples typically contain biomolecules including proteins, lipids, carbohydrates and nucleic acids. These major macromolecule classes play critical roles in biological functions, such as, proliferation, differentiation and cell death (Movasaghi *et al.*, 2008). The vibrational frequencies of the principal functional groups of these macromolecules are well defined. As mentioned before, each biomolecule will give its own ‘fingerprint’ spectrum, based on its covalent bond makeup, and therefore each kind of biological sample will produce a signature spectrum (Movasaghi *et al.*, 2008). A representative spectrum of a generalised biological sample can be viewed below in Figure 3.4.

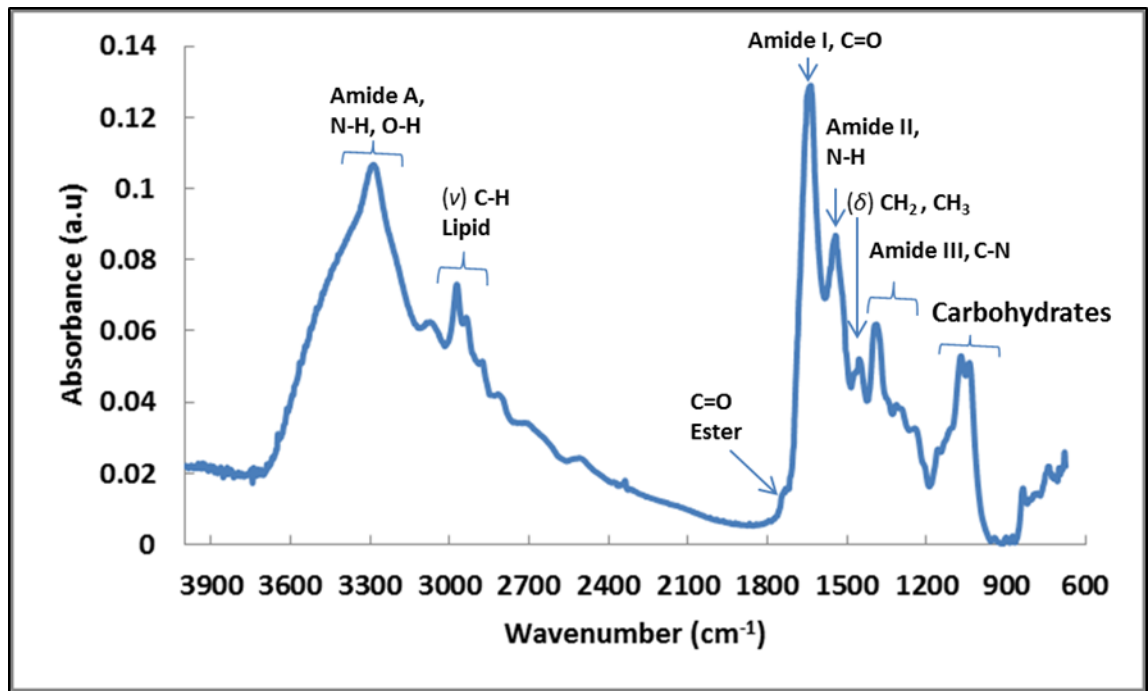


Figure 3. 4: Typical FTIR absorption spectrum of a biological sample

Generalised FTIR spectrum of a biological sample including molecular assignment of the main absorption bands in the mid FTIR spectrum (reproduced from Movasaghi *et al.*, 2008).

3.4.1 Protein FTIR

During protein synthesis, amino acid monomers within the growing polypeptide chain are linked via a condensation reaction termed peptide bond formation. Amino acid monomers are differentiated on the basis of their sidechain or R group composition (Figure 3.5) (Vasudevan *et al.*, 2013). Proteins can therefore be considered generally as polymers comprised of amino acids (bearing varying sidechains) linked by amide bonds.

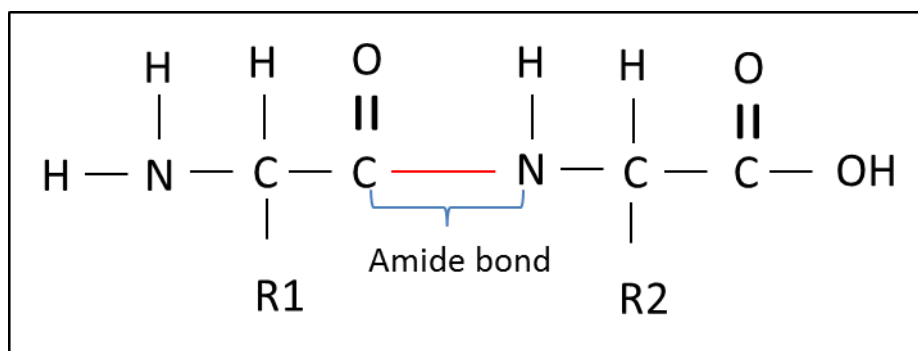


Figure 3. 5: Typical structure of amide bond via amino acids linking

Peptide bond is the name for an amide bond (red line) in a polypeptide. Protein characterises as a backbone of poly peptide chain via links of terminal base (amine) of one amino acid with acid termini of other amino acid, the condensation of these linkage will generate nascent protein molecules.

A high intensity broad peak for both amide I and II at mid FTIR wavenumber have previously been observed (reviewed by Romeo *et al.*, 2008). Amide I absorption band corresponds to the polypeptide backbone of protein, which provides the most information on protein geometric structure (Romeo *et al.*, 2008). Proteins have a number of levels of structure; primary structure, and 3 dimensional structures, both secondary and tertiary. Amide I corresponds to linear polypeptide (primary structure); secondary structure describes folding of proteins into two main patterns, α -helices and β -sheets (Roach *et al.*, 2005, Gauglitz and Vo-Dinh, 2014). FTIR micro-spectroscopy is able to provide important information about secondary structure of proteins. The stretching vibrations of the C=O bond within the amide I band and bending vibrations of the N-H bond with C-N stretching within amide II correspond to protein secondary structure components. Amide I is typically observed as a broad absorbance from $1700\text{-}1600\text{ cm}^{-1}$, with an intense maximum at 1650 cm^{-1} , and derives primarily from C=O stretching.

Amide II is observed as an intense absorbance at $\sim 1530\text{ cm}^{-1}$, and is associated with N-H and C-N bond bending vibrations (Movasaghi *et al.*, 2008). Amide III is observed around $1300\text{-}1180\text{ cm}^{-1}$ and is associated with C-N stretching and N-H bending. The amide A band (about 3500 cm^{-1}) and amide B (about 3100 cm^{-1}) originate from a Fermi resonance between the first overtone of amide II and the N-H stretching vibration (Movasaghi *et al.*, 2008). Proteins are thus a good example of the complex overlapping that can occur within FTIR spectra when attempting to investigate biomolecules (Gauglitz and Vo-Dinh, 2014).

3.4.2 Lipid FTIR

Lipids are another major component of biological samples, and can be grossly described as containing a long hydrophobic hydrocarbon chain, covalently attached to an acyl group. The hydrocarbon component of lipids is primarily comprised of C-C and C-H bonds, absorption maxima for which appear at range $3000\text{-}2830\text{ cm}^{-1}$ in IR spectra. All four major classes of biological macromolecule absorb in the region $3000\text{-}2830\text{ cm}^{-1}$. This is at least partly attributable to the presence of C-C and C-H bonds, veritably the “building blocks” of organic molecules. The overlapping contributions from the different macromolecule classes can be distinguished readily; lipid molecule acyl chain vibrations are of more significant intensity than the other three major macromolecular structures, due to the increased presence of C-C and C-H groups in lipids. IR vibrations in the region of $3000\text{-}2830\text{ cm}^{-1}$ are therefore primarily attributable to lipid molecular structures when

examining biological samples within mid IR spectra (Figure 3.4) (Mannock *et al.*, 2006, Dukor, 2002, Yonar *et al.*, 2018).

3.4.3 Carbohydrate and nucleic acid FTIR

Carbohydrates (CHO) are a diverse group of macromolecules, ranging from monomers (monosaccharides, e.g. glucose) to complex branched polysaccharides which may have a molecular weight of several kilodaltons. The primary structural components which characterise carbohydrates are alcohol and aldehyde functional groups. These groups are associated with wide IR absorbance bands at $\sim 3400\text{-}3380\text{ cm}^{-1}$ due to the OH stretching at $2930\text{-}2900\text{ cm}^{-1}$ corresponding to CH_2 groups undergoing C-H stretching. Additional peaks appear at $1240\text{-}950\text{ cm}^{-1}$, arising from C-O, C-C stretching and C-OH bending modes vibration (Figure 3.4).

Nucleic acids (DNA and RNA) represent a major class of biomolecules which include a sugar-phosphate backbone, which also contributes to the C-O absorbance observed in biological samples, observed around $1100\text{-}1005\text{ cm}^{-1}$ (Brandenburg and Seydel, 2002, Kačuráková and Mathlouthi, 1996). The contribution of nucleic acids is however of (relatively) limited relevance to our analyses due to the absence of cells in the samples.

3.5 FTIR spectroscopy studies of biofluids

Fourier Transform Infrared (FTIR) spectroscopy can be used to study the composition of biologically-derived materials, such as biofluids and human tissue samples. Biofluids such

as urine, whole blood, plasma and serum and saliva are a valuable biological sample type for clinical diagnostics due to their (variably) non-invasive collection, relative ease of handling and high biological information content. Infrared spectroscopy can provide a rapid, robust and high throughput tool for biological sample analysis.

Biofluid samples have previously been subjected to IR spectroscopy by various others (Harrigan *et al.*, 2004, Hoşafçı *et al.*, 2007). A study published by Hosafci *et al.* subjected blood plasma, whole blood, and urine to FTIR analysis. 100 urine samples and 400 blood samples were subjected to ATR-FTIR spectroscopy and clinical reference assays. Partial least-squares (PLS) regression calibration models were built for predicted sample components, namely albumin, total protein, glucose, urea, cholesterol, and triglycerides in whole blood and blood plasma; and urea, creatinine, inorganic phosphate and uric acid in urine. Additional calibration models were established using FTIR spectra of artificial urine components containing, ammonia, sulphate, hydroxybutyric acid, and acetoacetate (Hoşafçı *et al.*, 2007). Authors observed good correlation with their reference standards.

(Harrigan *et al.*, 2004) used FTIR spectroscopy to investigate urinary metabonomic components in an investigation of drug toxicological effects. Pre-dosing controls were compared with samples collected from rats subjected to treatment with a potential inflammatory agent, bacterial lipopolysaccharide (LPS) and the antacid ranitidine over a time course. Hepatotoxicity of the ranitidine was observed using this non-invasive FTIR-based metabonomics method (Harrigan *et al.*, 2004).

Centrifugal filtration devices, based retention of (predominantly) protein fractions using specific molecular weight cut-off filters are commonly used in bioanalysis workflows to

prepare biofluid samples prior to analysis. Concentration and relative purification of protein represents an important improvement for the application of vibrational spectroscopic technique (Bonnier *et al.*, 2014). Bonnier and co-workers (2014) were used commercially available centrifugal filters to segregate protein of blood serum according to molecular weight fractionation prior to FTIR analysis. Centrifugal concentration of the sample was shown to result in an improvement in both spectral intensity and quality. They suggested these devices are suited for application of vibrational spectroscopy to disease diagnostics purposes (Bonnier *et al.*, 2014).

In proteomic studies based on molecular weight fractions of samples of interest, other components have similar chemical moieties to peptides and proteins, such as urea and creatinine; interference from these components should be considered in order to avoid the spectral profile overlapping during analysis of the biomolecules of interest. Therefore, there is a need for appropriate method of sample collection and preparation prior to analysis in order to verify spectral profile of spectroscopic analysis.

The aim of this chapter is to establish a protocol for use with urine samples, using an animal model for sampling and method optimisation experiments. The use of locally-available animal-derived materials allowed us to develop methods using a freely-available waste material, with no significant ethical or governance concerns involved in their collection or storage. We aimed to develop a reproducible method, investigating the impact of sample volume upon analytical sensitivity, using whole urine sample of rat. Identification of the lower limit for routine operation using FTIR analysis; moreover, we sought to establish an appropriate methodology to identify FTIR spectral profiles for urine samples based on protein fractions with specific molecular weights. One of the major

factors in this study was the optimisation with a focus on the sample volume used for each repeat. For animal urine this would not be a problem but when moving onto human samples collected from a clinical trial it was essential to get the maximum data with reliable reproducibility. These are the factors highlighted in this optimisation chapter.

3.6 Chapter aims

1. To investigate threshold volumes for analysis of whole rat urine using FTIR spectroscopy.
2. To determine the influence of sample preparation procedures on FTIR spectra (whole rat urine, >10 kDa ultrafiltration urine sample without buffer exchange and >10 kDa after buffer exchange).
3. To establish a Standard Operating Protocol (SOP) based on use of 10 kDa nominal molecular weight cut-off membranes to segregate urine samples, compliant with methods used for proteomic analysis (Chapter 2).

3.7 Materials and methods

3.7.1 Urine sample collection

All animals used in this study were kept in accordance with UK Home Office legislation “Animals Scientific Procedures Act 1986 United Kingdom”, with *ad libitum* access to food and water, at (19-23) °C and 40-70 % relative humidity, maintained on a 12 hr light-dark cycle. Sprague Dawley rats were used for the study, being of the age 4-7 months at the time of collection. Urine collection was carried out by trained personnel according to

routine husbandry practices, under the supervision of the named Animal Care and Welfare Office, with approval by the local ethics committee. All samples were collected over 14 hrs from animals kept in metabolism cages. 10 samples of 20 mL volume each were collected from each animal, representing one sample over this time period. 10 samples from non-pregnant and 5 samples from pregnant rats at 17 day's gestation were collected.

3.7.2 Sample preparation

20 mL from each urine sample was centrifuged to pellet urinary tract cells (i.e. uroepithelial cells) (Eppendorf, Germany) at 3220 g for 10 min, 4 °C. The supernatant fraction of this step was aliquoted into 100 µL portions for storage at -80 °C. Thawed samples were mixed by vortexing before analysis to ensure homogeneous distribution of solutes throughout the sample. Aliquots were subjected to triplicate FTIR analysis each using different volumes of 1, 2, 5, and 10 µL respectively.

3.7.3 Ultrafiltration method

An ultrafiltration step was carried out on acellular supernatant material (20 mL) of rat urine to segregate samples on the basis of molecular size using ultrafiltration tubes polyethersulfone (PES) MWCO 10000, 20 mL, (Sartorius AG,). Ultrafiltration centrifuge tubes were washed 3 X with HPLC grade water to remove storage components which can contribute to spectral interference before applying a urine sample. Buffer exchange was performed on the >10 kDa retentate fraction, via the addition and spin-through of three cycles of 0.5 M triethylammonium bicarbonate (TEAB) (Sigma-Aldrich, UK) at 3220 g, 4 °C.

Concentrated retentates (~ 1mL) were collected, aliquoted and stored at -80 °C. 5 µL of sample at different conditions was spotted and dried for FTIR analysis.

3.7.4 FTIR instrument set-up

Urine samples were analysed using an FTIR spectrometer (ThermoFisher Scientific IS50). This spectrometer is fitted with a single bounce germanium ATR accessory. Prepared urine samples as per each (section 3.7.2 and 3.7.3) respectively above were spotted upon germanium crystal piece and air dried prior analyses to reduce the water contribution in the data collection. Spectra were collected at room temperature using Omnic software (version 2.0). 5-point adjacent averaging smoothing was used to improve (signal / noise) ratio with linear baseline offset spectral correction. Each spectrum was collected with 128 scans being averaged within Omnic, at a resolution of 4 cm⁻¹. Three separate spectra were recorded from each sample, with all samples being spotted in triplicate. Critical to the workflow was that a blank ATR background was collected between each sample. Peak height and area were measured and corrected with a linear baseline set between the upper and lower limits of the spectral region examined. Ultrafiltration-treated urine samples of pregnant (n = 4) and non-pregnant (n = 4) rat urine were subjected to FTIR analysis. 5 µL of each sample of rat urine was spotted on the ATR germanium crystal and air dried prior analyses to reduce the water contribution to FTIR spectra. Spectra were collected at room temperature using Omnic.

3.7.5 Statistical analysis

Data analysis was performed using Excel (version 10, Microsoft). Data were expressed as a mean and inter-sample variation expressed as standard error. Principal Component Analysis was used to interrogate differences between samples.

3.7.6 Principal Component Analysis

Principal Component Analysis (PCA) is an unsupervised multivariate algorithm. PCA is typically used to identify differences in multivariate data, such as spectral variation recorded from different biological samples. In order to perform PCA, a dataset must be transformed to ensure all points are distributed using a common set of axes. In the case of our FTIR data, this required mean centering of the original dataset matrix. The transformed matrix has different spectral absorption mean intensities distributed around zero. The centred mean matrix (X) which has the same original matrix dimensions $I \times J$ (I the number of samples and J the number of wavenumbers variable in the given dataset matrix; loadings (P) (square matrix that derivative from variables in the original data set which has the same number of rows and columns). This transformation generates new variables within the new modelled matrices, termed principal components (PCs). PCA is a dimensionality reduction method, which allows large numbers of variables (such as spectral wavenumbers and resultant intensities) to be transformed into a few principal components (that are calculated from the covariance of the analysed dataset). PCA and other multivariate methods enable us to understand and identify the origin of differences in datasets in an unbiased and unsupervised manner. A number of principal components can be generated to explain the origin of variance between samples within a dataset. PC1

is the component which describes the largest proportion of this variance, with higher ordinals describing lower variance components. Figure 3.6 summarises the major experimental approach in this chapter.

3.7.7 Summary workflow for methods

The following workflow illustrates the major experimental approach in this chapter is summarised in Figure 3.6 below:

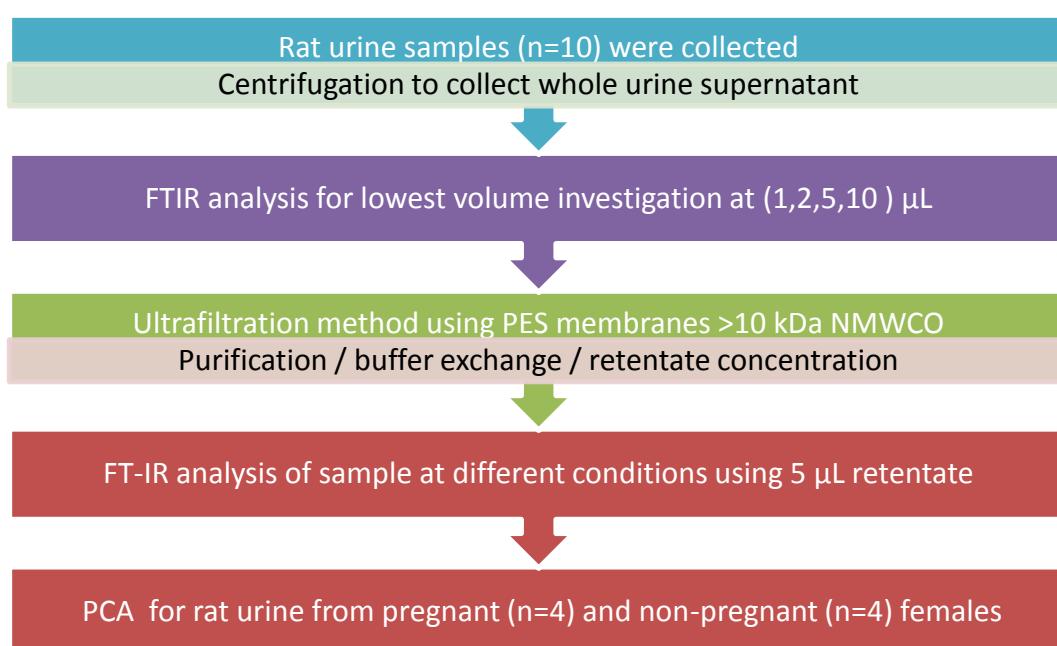


Figure 3. 6: Summary workflow protocol optimisation and application

3.8 Results

FTIR spectra were collected at mid-IR region using whole urine samples (n=10). Samples were subjected to analysis in triplicate analysis. Spectra were limited the range 1800-1500 cm^{-1} , corresponding to amide I and amide II, as shown in Figure 3.7. This region was selected due to its importance in protein spectroscopy, and all sensitivity/volume

determination experiments are based on observations of this region (section 3.8.1). Qualitative analysis of FTIR spectra was carried out to investigate the impact of ultrafiltration with different conditions for selection of an optimised protocol (Figure 3.9). The FTIR spectral profile of pregnant and non-pregnant rat urine after ultrafiltration process was then investigated, using PCA to observe emergent differences (Figure 3.10, Section 3.8.2).

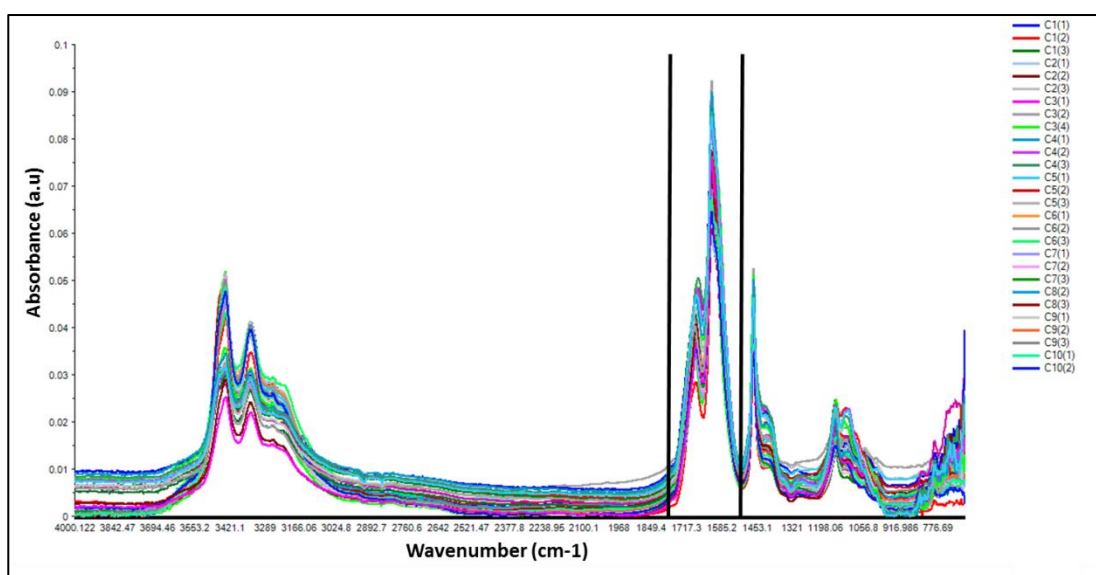


Figure 3. 7: FTIR spectra at particular range 1800-1500 cm⁻¹ of mid region using whole rat urine

FTIR spectra over mid-IR region 4000-650 cm⁻¹ for whole rat urine. 10 samples were each analysed in triplicate. Highlighted area of interest (delimited by black lines, 1800-500 cm⁻¹), corresponds to amide I and amide II region.

3.8.1. Volume sensitivity investigation of FTIR analysis using whole rat urine

Identification of spectra is typically used for qualitative analysis of substances, and involves the identification of absorbance at specific wavenumbers in a spectrum. Figure 3.7 illustrates typical FTIR spectra for of whole rat urine at mid-IR region. Subsequent analyses focussed on the region $1800-1500\text{cm}^{-1}$, – cropping spectra over the desired wavenumbers Figure 3.8. This region represents wavenumbers of amide I, II peaks in biological samples. Spectra were collected with varying volumes of whole rat urine, from 1, 2, 5, and 10 μL respectively. Figure 3.8 presents the observed variation in the range $1800-1500\text{ cm}^{-1}$ region in amides I, II band height and peak area for these different volumes. This region was analysed to provide clarity of peak maxima signal / noise and to see if peak area was a better method to use, as an integrated area is likely to present a lower sample to sample (repeated measures) variance.

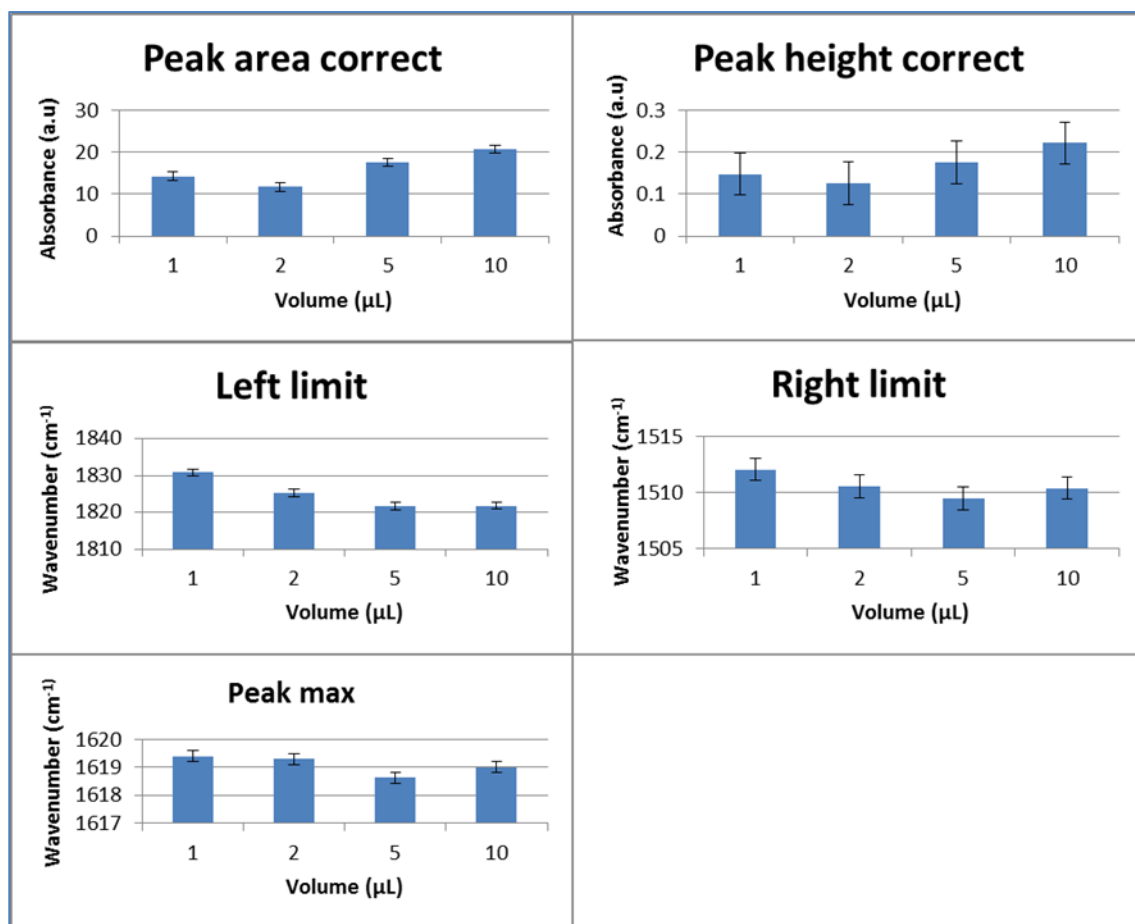


Figure 3. 8: Bar graphs of FTIR analysis at (1850-1500) cm⁻¹ range

A series of volumes of whole rat urine (1, 2, 5 and 10 µL) were measured, with spectral parameters being examined at amide I and II wavenumbers. Peak area is bordered by the spectrum when the X-axis limits of the area under curve and the baseline. Peak height was measured from the maxima at full absorbance, taking into account a linear baseline fitted to the extents of the wavenumber range intercepting the band; limits of peak (left and right) with peak max in each volume were identified. All data are represented as mean and error bar.

The methods optimised above demonstrated that sample volume of 5 µL robustly gave a good degree of reproducibility and high signal to noise. For this reason, all following processes used this volume as part of further sample preparation.

3.8.2. Influence of ultrafiltration preparation procedure on FTIR spectra using rat urine sample

FTIR data were collected at mid-IR range 4000-600 cm^{-1} using 5 μL in triplicate analysis of the same sample. Here, we wished to investigate the impact of sample preparation steps on spectral variability. Table 3.1 provides a list of band assignments describing the major macromolecules in urine. The results of this experiment show differences in spectral observations before using centrifugal filtration devices Figure 3.9 (A); and after using >10 kDa NMWO membrane filters for centrifugal ultrafiltration without buffer exchange Figure 3.9 (B), and after a buffer exchange step during ultrafiltration Figure 3.9 (C). Figure 3.9 (A and B) show spectra for the components in the urine with similar chemical moieties to peptides and proteins, therefore the spectral profile overlap with biomolecules of interest could occur. Following ultracentrifugation, spectra show observed improvement in characteristics related to protein identification. This can be attributed to removal of interference from small molecules and salts by buffer exchange Figure 3.9 (C).

Table 3. 1: Major FTIR bands of macromolecules described in urine

Bands cm^{-1}	Functional groups	Description
3450-3070	N-H or H_2O	Broad peak of N-H bending vibration and water
2990-2840	CH_2 group	Asym. (2990-2950) cm^{-1} ; sym. (2880-2850) cm^{-1} , stretching vibration
2920-2850	CH_3 group	Stretching vibration of asym.(2985-2950) cm^{-1} ; sym. (2885-2865) cm^{-1}
1800-1750	C=O ester	Stretching vibration of saturated (1750-1730) cm^{-1} , unsaturated (1730-1715) cm^{-1}
1650	Amide I	Stretching vibration of C=O hydrogen bonding , which range (1600-1700) cm^{-1}
1541	Amide II	Coupling of (C-N) stretching and (C-N-H bending) range (1585-1500) cm^{-1}
1463	C-N; N-C-N	Wagging vibration; asym. stretching

		vibration
1155	NH ₂	Rocking vibration
1055	NH ₂	Asym. bending vibration
1000	C-N; N-C-N	Stretching vibration (sym.; asym)
788	C-O; NH ₂	Bending vibration; sym. stretching vibration

Sym: Symmetric; Asym: Asymmetric.

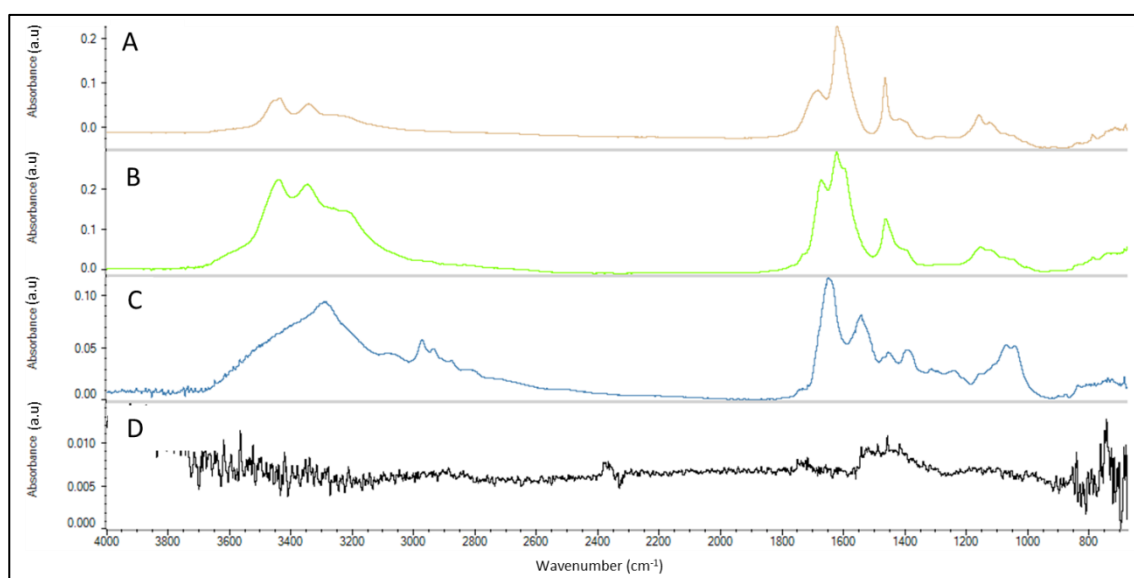


Figure 3. 9: Typical FTIR spectra comparison of sample preparation in different conditions

FTIR spectra were taken at (4000-600) cm⁻¹ region. Spectrum (A): whole urine sample; spectrum (B): > 10 kDa NMWCO filtration alone (without buffer exchange); spectrum (C): urine sample following buffer exchange into TEAB; spectrum (D): triethyl ammonium bicarbonate (0.5 M) alone.

Clear differences in spectral profile were observed related to amide I at 1700-1600 cm⁻¹ (corresponding to protein bands) under varying sample preparation conditions (Figure 3.9 (A), (B), (C) and (D)). The observed spectral profile using ultrafiltration device coupled with buffer exchange condition (Figure 3.9 C) gave the most pronounced improvement in

spectral profile compared with using ultrafiltration device alone (Figure 3.9 B) or when using whole rat urine sample (Figure 3.9 A) respectively. Similar findings were observed for amide A and OH bands at $3650\text{-}3500\text{ cm}^{-1}$ and also C-H stretching band at $3000\text{-}2830\text{ cm}^{-1}$ in IR (Fig 3.9 A-C). TEAB buffer analysed in the absence of urine sample shows no peaks of significant intensity within this region Figure 3.9 (D).

The results above indicated that ultrafiltration protocol improved the observed spectral profile for FTIR measurements of $5\text{ }\mu\text{L}$ urine. This supported further analyses of spectra with PCA to investigate spectral profile of FTIR for pregnant and non-pregnant rat urine. The following analysis will focus on the mid-IR region, specifically the protein fingerprint region ($1800\text{-}1000\text{ cm}^{-1}$), and the lipid region ($3100\text{-}2700\text{ cm}^{-1}$).

3.8.3. Mean differences of spectral profiles for pregnant and non-pregnant rat urine at mid-FTIR

A difference in averaged spectral profile between pregnant ($n=4$) and non-pregnant ($n=4$) is presented in Figure 3.10. Data were collected over the mid-IR range ($4000\text{-}600\text{ cm}^{-1}$), with absorbance intensity being plotted vs wavenumber. Mean spectral profiles for pregnant rat urine showed increased absorbance intensity compared with those for non-pregnant rats (Figure 3.10). These observed differences include peaks within the range $3500\text{-}3100\text{ cm}^{-1}$ related to amide A, N-H and O-H stretching vibrations (Figure 3.10). Different intensities in spectral profile relating to protein fingerprint at range $1800\text{-}1000\text{ cm}^{-1}$ is presented in Figure 3.10. Amide I and amide II show sample-dependent differences within the range $1700\text{-}1450\text{ cm}^{-1}$ and at amide III ($1235\text{-}1224\text{ cm}^{-1}$). Peaks

within the C-H stretching bond (lipid region), over 3100-2700 cm^{-1} are also observed in Figure 3.10. A small peak at 1730 cm^{-1} , attributed to (C=O) stretching bond of ester; and bending vibration of (CH_2) (1465 cm^{-1}) were also observed, as were “stretch band” signals at range 1100-1050 cm^{-1} related to both (C–O) and (C–O–C) of carbohydrates (Figure 3.10).

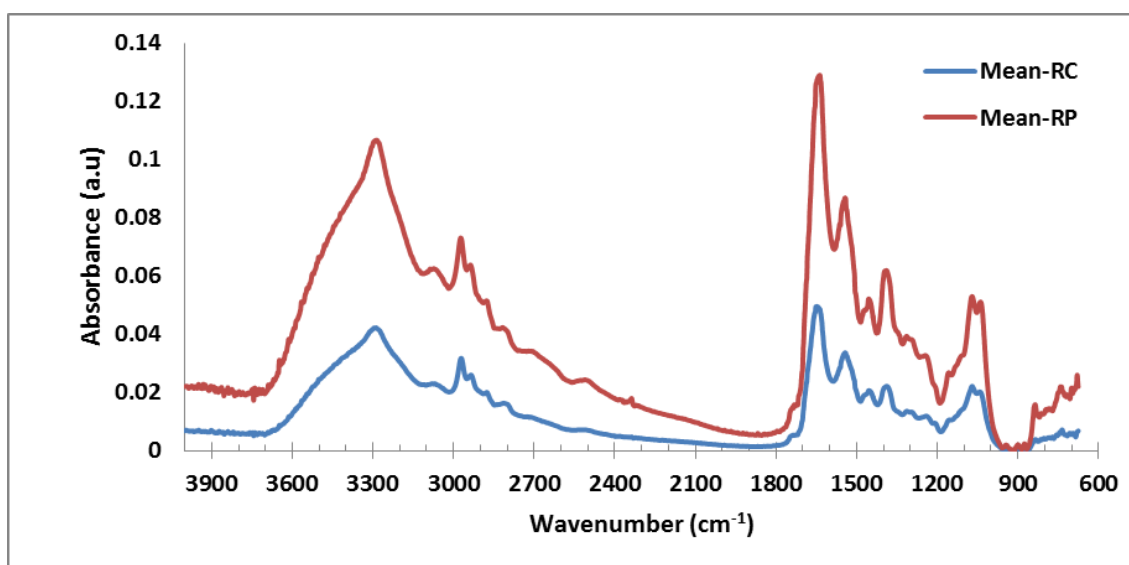


Figure 3. 10: Mean differences of FTIR spectra intensity for rat urine from pregnant and non-pregnant females

Mean of FTIR spectra for each group of urine samples taken from pregnant (n=4) and non-pregnant (n=4) female rats were analysed. FTIR spectra were taken at (4000-600) cm^{-1} region. Wavenumber (cm^{-1}) vs absorbance (a.u.) is used for this analysis. Red=mean of pregnant rat (n=4), blue= mean of non-pregnant rat (n=4).

3.8.4. Spectral profile of protein fingerprint region of urine from pregnant and non-pregnant rats

Variation in observed spectral profiles for rat urine from pregnant and non –pregnant females is shown in Figure 3.11, a PCA plot of the protein fingerprint region. For these

data, 77 % of variance was accounted for by PC1 vs. 14 % in PC2 (Figure 3.11 A) for pregnant and non-pregnant rat urine. Data were analysed at a 95% confidence interval over the protein fingerprint region at (1800-1000 cm^{-1}), with triplicate analysis being performed for each sample in each group. Scores plot (Figure 3.11 A) presents the observation that some data points remain as outliers using a 95% confidence interval. PC1 and PC2 loading spectra were examined, and observed absorbance variation corresponds to amide I at range 1700-1500 cm^{-1} , amide II at range 1500-1350 cm^{-1} , and amide III at range \sim 1335-1220 cm^{-1} (Figure 3.11 B).

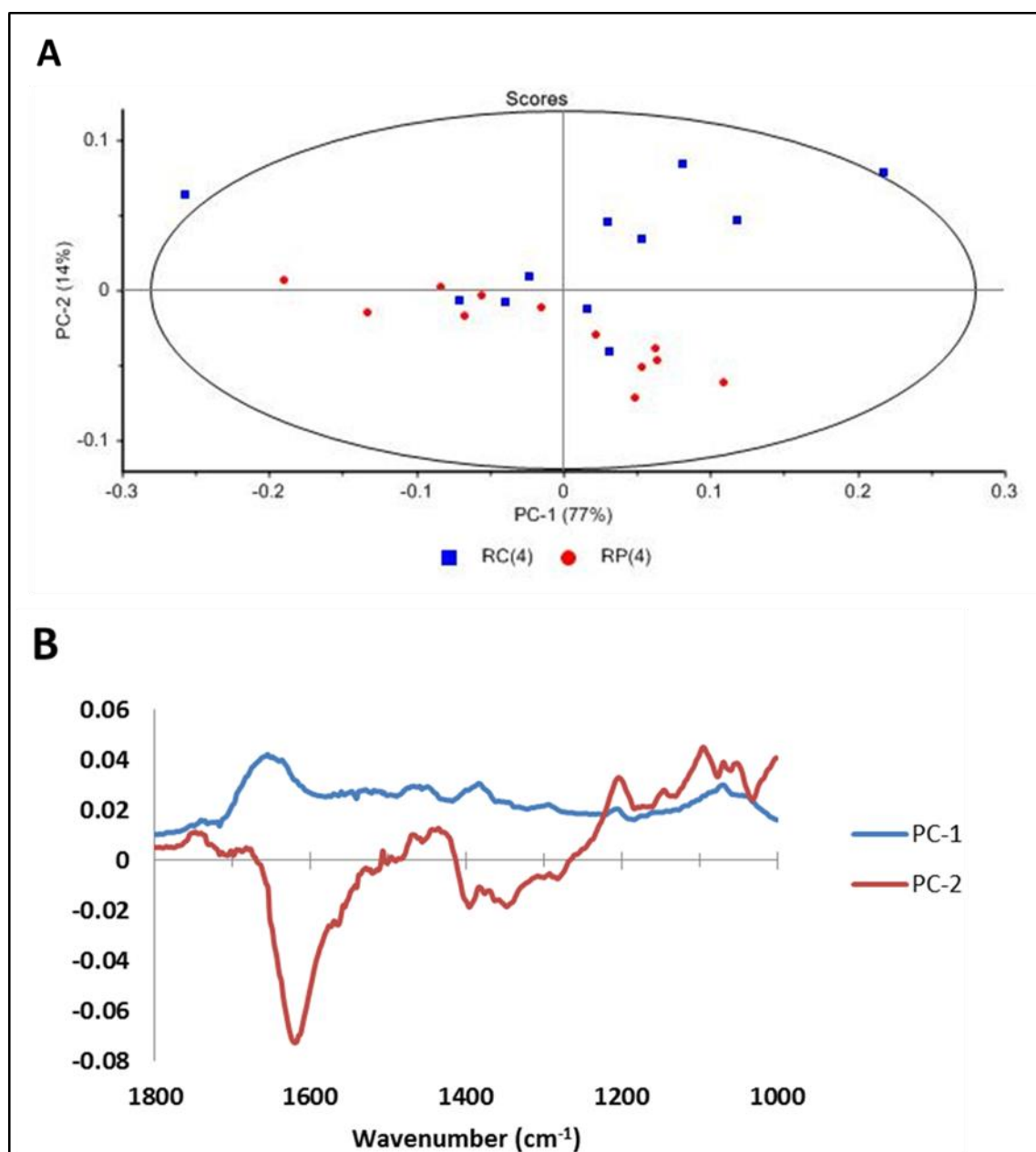


Figure 3. 11: PCA plot for fingerprint protein region of pregnant and non-pregnant rat urine

PCA plot of PC1 and PC2 score distribution around a normalised scale (**A**) for urine from pregnant and non-pregnant rats. Protein fingerprint region (1800-1000 cm⁻¹) was examined in triplicate, with a confidence interval of 95 %. Observed variation of absorbance intensities between PC1 and PC2 loading spectra (**B**) were. **A**: Scores of IR spectral waves, RC = Non pregnant rat, RP = Pregnant rat, **B**: Loading score of IR vibrational bands at 1800-1000 cm⁻¹ region; PC1 (blue); PC2 (red).

3.8.5. Spectral profile of lipid region for pregnant and non-pregnant rat urine

PCA reveals that emergent differences in the spectral profile of urine samples from pregnant and non-pregnant rats are related to the lipid region ($3100\text{--}2700\text{ cm}^{-1}$), as shown in Figure 3.12. Within this region, 91 % of inter sample variance was accounted for in PC1, and 7 % in PC2 (Figure 3.12 A) within these triplicate analyses. Examination of the loading plot showed differentially-observed C-H stretching bond peaks, associated with lipids ($2980\text{--}2950\text{ cm}^{-1}$), and additionally (CH), $2880\text{--}2865\text{ cm}^{-1}$ and $2860\text{--}2840\text{ cm}^{-1}$ for (asymmetric and symmetric) vibrational bonds respectively of CH_3 and CH_2 (Figure 3.12 B).

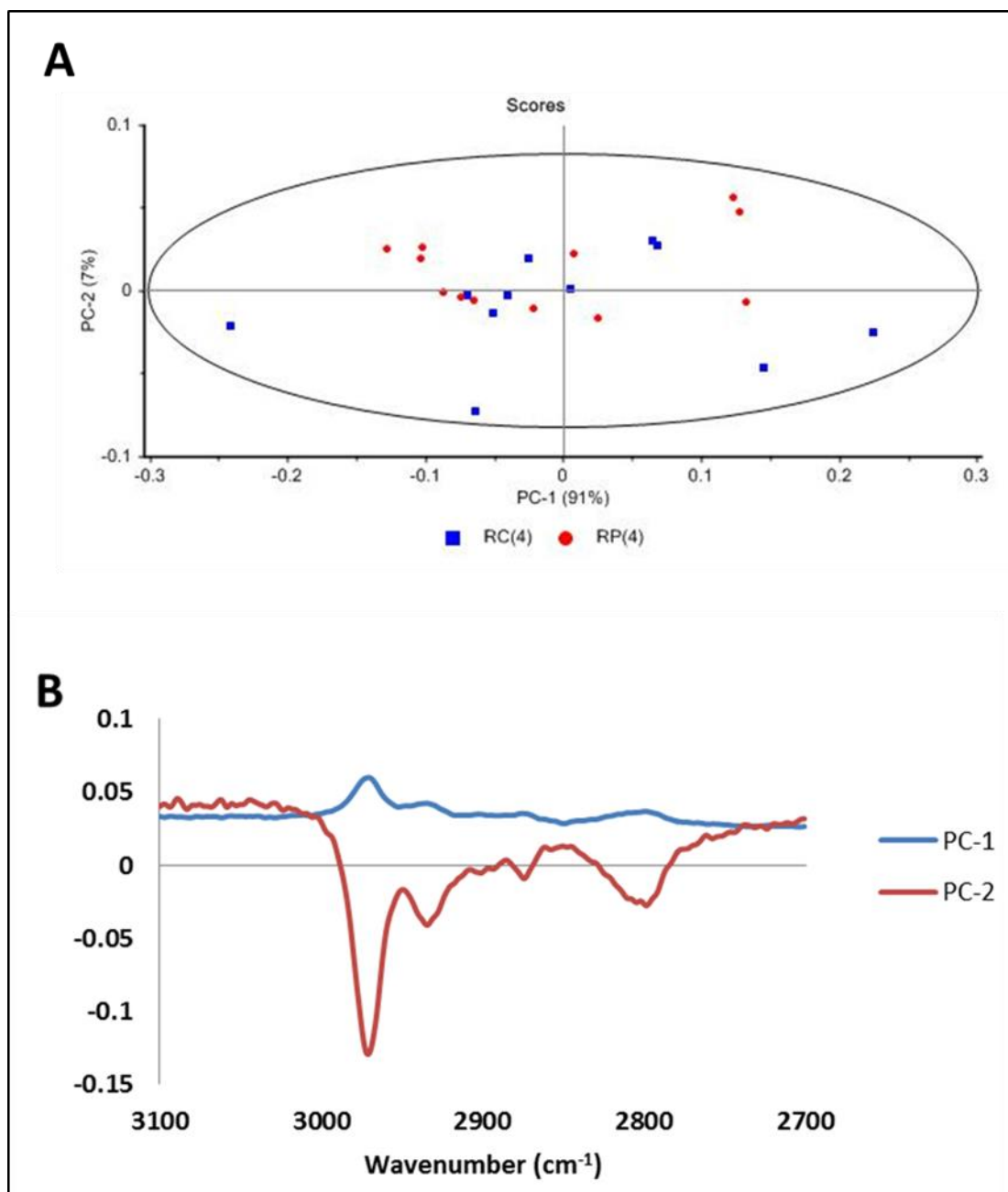


Figure 3. 12: PCA plot for lipid region of pregnant and non-pregnant rat urine

PCA plot displays scores distribution of PC1 and PC2 around normalised scale (**A**) for pregnant and non-pregnant rat urine. Lipid region data (3100-2700 cm⁻¹) were investigated at a confidence interval of 95 %. Different absorbance intensities between PC1 and PC2 loading spectra (**B**) were observed. **A**: Scores of IR spectral waves, RC = Non pregnant rat, RP = Pregnant rat, **B**: Loading score of IR vibrational bands at 3100-2700 cm⁻¹ region; PC1 (blue); PC2 (red).

3.9 Discussion

This study was performed to derive an optimised protocol for FTIR analysis of (rat) urine as an animal model. Due to the extensive requirements for ethical approval and sample handling and storage associated with sampling of human biofluids, the use of these materials for method optimisation is prohibitively difficult. Human samples are therefore expensive and available in limited amounts, therefore rat urine was used as an alternative to human samples for method optimisation experiments, as the latter is available locally and with little-to no impact on the animals involved as this is their waste. Protein typically represents a crucial component in bio-fluid samples (Alberts *et al.*, 2002, Moore, 2014), with a high intensity broad peak for amide I and amide II being observed at ~ 1800 - 1500 cm^{-1} (Romeo *et al.*, 2006). Spectra of interest were cropped at amide I and amide II (Figures 3.7) to study parameters such as peak area, peak height, and peak max relation to different volume of sample. Data presented in Figure 3.8 show the variation at range 1800 - 1500 cm^{-1} region in amides I and amide II band relation to both height and peak area at different volumes. The results indicated that volumes as low as $5\text{ }\mu\text{L}$ generated reproducible data with high signal to noise (Figure 3.8).

Use of ultrafiltration protocol improved spectral identification/ classification (Figure 3.9). This type of sample pre-fractionation was previously used by Bonnier, *et al.*, giving good discrimination of blood sample (Bonnier *et al.*, 2014). For instance, spectral differences are clearly identified in the range 1700 - 1650 cm^{-1} with a distinct peak at $\sim 1650\text{ cm}^{-1}$ (amide I). A band of amide II observed at 1650 - 1500 cm^{-1} , amide III peak at ~ 1530 - 1220 cm^{-1} was also observed (Lacombe *et al.*, 2015). Similar findings corresponding to amide A and OH bands at 3650 - 3500 cm^{-1} and also C-H stretching band at 3000 - 2830 cm^{-1} were

reported for blood samples (Lacombe *et al.*, 2015). No distinct peaks were observed in the mid-IR region for TEAB buffer (Figure 3.9 (D)), thus encouraging the use of this buffer as part of a unified sample preparation method in line with our proteomics methods.

Prior to urine application, PES centrifugal filters were washed well with HPLC grade water to remove any contaminants such as traces of storage (e.g. glycerine), which could give rise to spectral interference (Bonnier *et al.*, 2014). This data showed good indication that the spectral bands observed arise from urine samples rather than overlapping vibrations pertaining to buffer components or to PES storage materials (Bonnier *et al.*, 2014). However, no attempt has been made to quantify the spectral characteristics corresponding to the constituent components in the urine; the optimisation experiments demonstrated development of Standard Operating Protocol (SOP) for further analysis of clinical samples.

Very limited discrimination was seen between urine from pregnant and non- pregnant rats within the protein fingerprint (Figure 3.11 A) and lipid (Figure 3.12 A) regions, although the differences in mean spectral intensities were observed between urine from pregnant and non- pregnant rats (Figure 3.10). This study was not without limitations; specifically, the repeated use of a single bounce germanium ATR accessory during FTIR spectra collection could cause technical variation. The limited cohort size for this study was a further limitation.

3.10 Conclusions

This work presents proof of concept, using a rat urine model for optimisation experiment design to measure volume sensitivity and built up a standard operating protocol for the

further study of human clinical urine samples. The 5 μ L loading volume showed appropriate sensitivity for the mid FTIR region. Moreover, ultrafiltration protocol displayed robust and simple preparation method for protein and lipid identification by FTIR within a urine sample.

Chapter 4

4. Study of spectral profile for pregnancy complications using Fourier transform infrared spectroscopy

4.1. Application of FTIR in biological studies

FTIR micro-spectroscopy (FTIRMS) instruments are typically used to perform spectroscopic analysis, generating spectra of biological samples at high speed and sensitivity. Biological samples are often of limited volume and heterogeneous in nature, thus micro-spectroscopic methods were developed to enable microscale FTIR measurements to be obtained. Numerous studies have successfully employed FTIR to investigate molecular biomarkers (Cakmak et al., 2006, Manoharan et al., 1993, Lin et al., 2007, Baker et al., 2014, Lacombe et al., 2015). As mentioned in chapter 3 (section 3.4) a generalised biological sample is comprised of several classes of biomolecules; such as proteins, lipids, carbohydrates and nucleic acids. The principal assignment of spectroscopic vibrations to biomolecules of clinical relevance has been previously described (Lacombe et al., 2015). Table 4.1 demonstrates a list of band assignments of the major macromolecules in blood plasma reproduced from (Lacombe et al., 2015). FTIR spectroscopy of biological molecules was discussed previously in detail in chapter 3 (section 3.4).

Table 4. 1: Assignment of main FTIR absorption bands of plasma

Bands (cm⁻¹)	Major assignment for plasma content
3300	ν (N-H) of proteins (amide A band)
3055–3090	ν (=CH) of lipids and proteins
2950–2960	ν as (CH ₃) of lipids and proteins
2920–2930	ν as (CH ₂) of lipids and proteins
2865–2880	ν s (CH ₃) of lipids and proteins
2840–2860	ν s (CH ₂) of lipids and proteins
1730–1760	ν (C=O) of fatty acids
1660	ν (C=O) of proteins (amide I band)
1550	δ (N-H) of proteins (amide II band)
1400	ν (COO ⁻) of amino acids
1240	ν as (P=O) of nucleic acids
1170–1120	ν (C-O) and ν (C-O-C) of carbohydrates

v, stretching vibrations; δ , bending vibrations; s, symmetric; as, asymmetric.

A previous study by (Mukherjee et al., 2014) used FTIR and NMR analysis of serum to investigate metabolic profiles in patient samples collected from women with early and late onset PE. 80 subjects were enrolled in this study, classified into 4 sub-groups according to gestational age at onset of symptoms: early onset PE (gestational age at onset <34 weeks, n=20), late onset PE (onset >34 weeks, n=20), and 20 subjects with uncomplicated pregnancy outcome at the same gestational age of these cases. 13 peaks showed significant variation in early onset PE ($P < 0.001$). These peaks had wavenumbers associated with carbohydrate, protein and lipid region. Fisher's linear discriminant analysis (FLDA) identified five vibrational bands of highly significant wave-numbers (maxima at 1078, 1088, 1122, 1169 and 1171 cm^{-1} respectively); these spectral numbers gave strong predictive capacity of outcome ($\geq 80\%$ overall accuracy compared with control). Hierarchical clustering was able to separate early from late onset PE, although PCA demonstrated that the main separation of the wavenumbers resulted from gestational stage. The authors suggested that further validation of putative markers using orthogonal techniques such as liquid chromatography–mass spectrometry (LC–MS) in a larger population was required prior to clinical adoption (Mukherjee *et al.*, 2014).

Despite many studies exploring metabolic profiles in preeclampsia and IUGR, no studies to date have identified robust biomarkers to predict PE and IUGR disorders at early gestation. The main predictor of outcome remains previous pregnancy outcome, which is not relevant for nulliparous women (who are in their first pregnancy). Here, we lack any predictive diagnostic tools. Chapter 3 described the derivation of a (SOP) for the analysis

of urine samples using an animal model. This Chapter describes the use of FTIR with our SOP to explore the spectral profiles of PE and IUGR vs. uncomplicated pregnancies.

4.2. Chapter aims

1. To investigate spectral profiles of urine samples collected in a case/control study of PE related at 15 and 20 weeks gestational age respectively using FTIR.
2. To investigate spectral profile of urine samples collected in a case/control study of IUGR cases vs. normal birthweight controls at 15 and 20 weeks of pregnancy respectively using FTIR.

4.3. Materials and methods

4.3.1. Urine sample collection

240 urine samples (450 μ L each.) were obtained from subjects of PE and IUGR respectively. 60 samples for each pregnancy condition (PE and IUGR respectively); and 120 from healthy pregnant women were collected. All subjects were prospectively recruited and their samples collected at 15 and 20 weeks' gestation pregnancy via the SCOPE (Screening for Pregnancy Outcomes) study arm at University College Cork (UCC), Republic of Ireland (see section 2.3.5), with approval by the local ethics committee (see appendices for local approval letter).

4.3.2. Sample preparation

4.3.2.1. Centrifugal filtrate and protein ultrafiltration protocol

450 μL from each sample was centrifuged to pellet urinary tract cells (i.e. uroepithelial cells) using centrifugation at 3220 g for 10 min, 4° C (Eppendorf, Germany). Non-pelleted matter was decanted into clean labelled eppendorf tubes. Clarified samples were subjected to ultrafiltration (10 kDa MWCO filters, Polyethersulfone (PES), 2 mL) (Sartorius AG). Prior to their use, membranes were cleaned by spinning 3 X with 2 mL HPLC grade water to remove storage buffer contaminants from PES membranes and resultant interference spectra. Urine samples were then added to membranes and ultrafiltration carried out. The retained urine fraction following ultrafiltration was further subjected to buffer exchange, spinning (3 X) with 0.5M triethylammonium bicarbonate (TEAB) (Sigma-Aldrich, UK) at 3220 g, 4° C. Concentrated retentates (\sim 150 μL) were collected, aliquotted and stored at -80° C until FTIR analysis.

4.3.3. FTIR experiment setup

The FTIR optimisation described in Chapter 3 was carried out using ATR-FTIR, due to both time constraints and availability of new equipment part way through the project, this chapter employs micro-FTIR using reflection rather than ATR. Samples of urine were treated exactly as reported in Chapter 3, with volumes of 5 μL being found to give reliability for reflection-based spectroscopy analysis.

Ultrafiltration-treated urine samples were subjected to FTIR analysis. 5 μL of each sample was spotted on an aluminium micro-spot slide (ThermoFisher Scientific, Winsford,

Cheshire UK) (see Fig 4.1 B). Samples were then analysed using the FTIR (ThermoFisher Scientific iN10). This FTIR instrument, housed at Loughborough University, was fitted with an imaging infrared microscope with a 10 X objective and computer controlled x-y-z stage (Figure. 4.1 A). The instrument was equipped with liquid nitrogen cooled Mercury Cadmium Telluride (MCT) detector. Infrared spectral images were collected in reflection mode with 128 scans being averaged at 8 cm^{-1} resolution, and a physical $400 \times 400\ \mu\text{m}$ aperture size being used to cover the spatial field of view. Instrument background analysis and subtraction was performed between each sample for calibration using the same scan number (128) used for the biological samples. These background calibration spectra were used to minimise the influence of ambient CO_2 on observed spectra. Prepared urine samples as in (section 4.3.2.1) above were spotted upon 12-position aluminium micro spot slides (Figure 4.1 B) and air dried prior FTIR analyses to reduce the water contribution in the data collection. Spectra were collected saved in map files. Composite files were split into their individual constituent spectra individual files of each sample using Omnic Picta software (version 2.0, ThermoFisher Scientific).

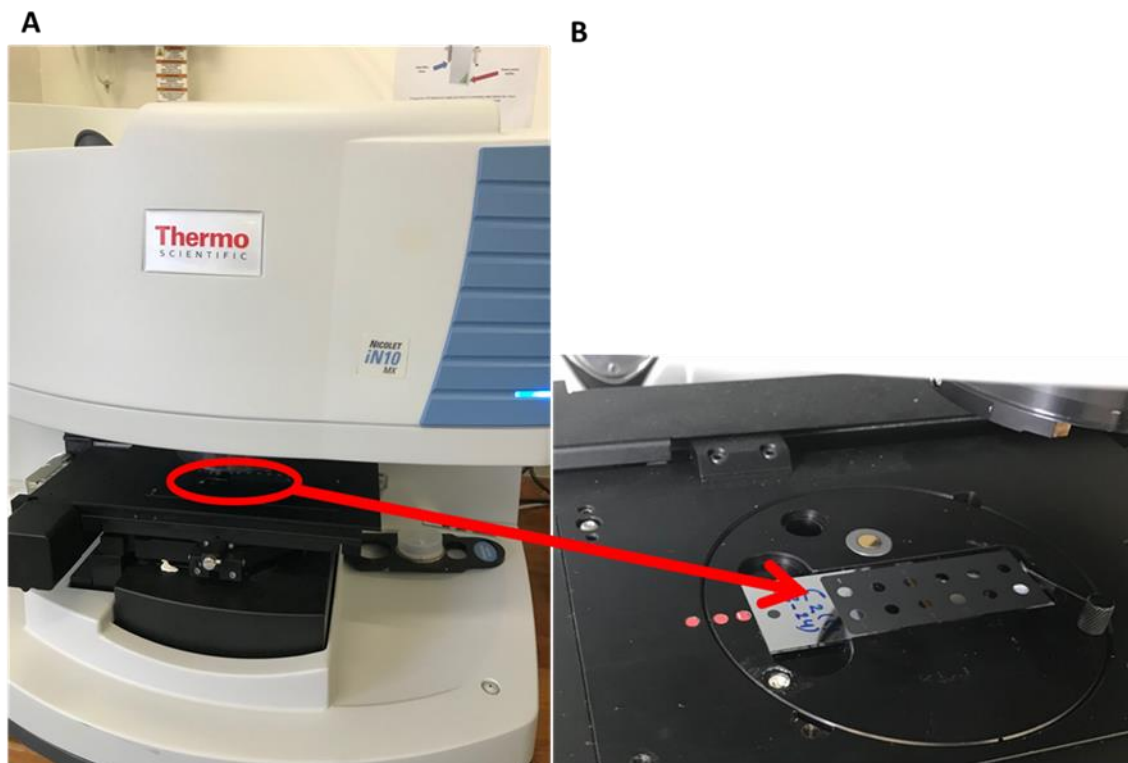


Figure 4. 1: FTIR instrument and its sampling platform.

(A): IR spectra recorded using Nicolet iN10 MX FTIR (ThermoFisher Scientific iN10). (B): 12-position aluminium micro spot slides used as sampling platform for urine samples.

4.3.4. Spectral transformation

Spectra are subjected to baseline subtraction using offset correction to remove baseline shifts caused by scattering. To decrease the signal-to-noise impact on spectra, 5-point Savitzky-Golay smoothing was used. Recorded spectra are normalised using vector unit normalisation (Brereton, 2009). Transformation of spectral data was carried out using Unscrambler[®]X software (version 10.4. Camo, Norway).

4.3.5. Data analysis

Principal Component Analysis (PCA), unsupervised multivariate algorithm typically used to identify differences in multivariate data, more detail in chapter 3 (section 3.7.6). Data analysis and plotting were performed using Excel (version 10, Microsoft). The global experimental approach in this chapter is summarised in Figure 4.2.

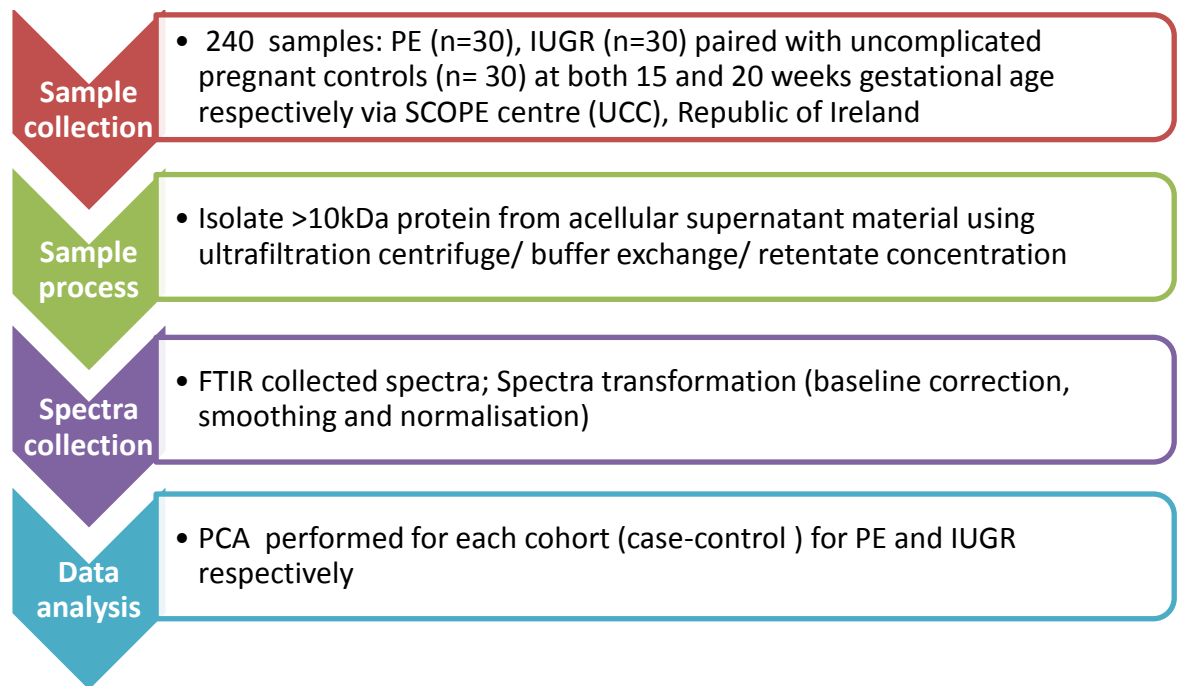


Figure 4. 2: Summary workflow of methods

4.4. Results

4.4.1. FTIR spectral profile of PE and control at 15 week gestational age

FTIR spectroscopy was performed for urine samples obtained from women who went on to develop PE (n=30) and those with uncomplicated pregnancies (control) (n=30) at 15

weeks' gestation. Figure S1 (see appendix 4) shows general FTIR spectra in PE and control at 15 weeks gestation. There is very little visible differences between these spectra over FTIR mid-range of 4000-600 cm^{-1} . Identification of emergent spectral differences therefore requires mean spectra difference. Differences in averaged spectral profiles between PE (n=30) and non-complicated (n=30) samples is shown in Figure 4.3. A mean spectral profile for PE shows lower intensity absorbance compared with the corresponding spectral profile of uncomplicated pregnancies at 15 weeks gestation (Figure 4.3). Peaks ranging from 3500-3100 cm^{-1} related to amide A, N-H and O-H stretching vibration delineate cases vs. controls (Figure 4.3). Other peaks related to protein fingerprint (1800-1000 cm^{-1}) also show variation in intensities between PE and control (Figure 4.3). Amide I and amide II show case/control-related differences (1700-1450 cm^{-1}), as do amide III at band range 1235-1224 cm^{-1} . Peaks related to C-H stretching bonds (lipid region, 3100-2700 cm^{-1}) are also observed Figure 4.3. A small peak at 1730 cm^{-1} associated with (C=O) stretching bond of ester; bending vibration of (CH_2) at peak around 1465 cm^{-1} , and stretch band peaks (1100-1050 cm^{-1}), related to both (C-O) and (C-O-C) of carbohydrates are also observed in Figure 4.3.

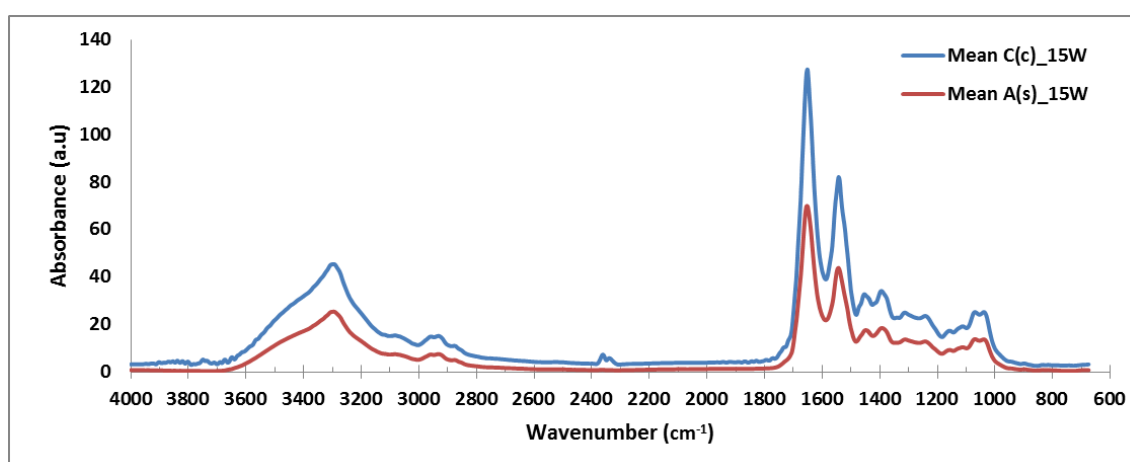


Figure 4. 3: Averaged spectral profile (means) for PE and control at 15 week gestational age at mid FTIR range. (n = 30) samples

Averaged (mean) FTIR spectral profiles for PE (n=30) and control (n=30) at 15 weeks' gestation. FTIR spectra were taken at mid-IR (4000-600) cm^{-1} region. Wavenumber is plotted vs. absorbance intensity. Red A (s) =mean of PE (n=30) at 15 week of pregnancy, blue C (c) = mean of control (n=30) at 15 week of pregnancy.

The following section describes the spectral profiles of PE and control at 15 weeks of pregnancy using PCA description related to protein fingerprint (1800-1000 cm^{-1}) and lipid (3100-2700 cm^{-1}) regions respectively.

4.4.2. Spectral profile study of protein fingerprint region for PE and control at 15 weeks gestational age

Data presented in Figure 4.4A show PCA plot for PE and control FTIR data at 15 weeks gestation over the protein fingerprint region (1800-1000 cm^{-1}). 99 % of inter-sample variation is accounted for by PC1 and 1 % in PC2. The score plot that in Figure 4.4B shows somewhat some clusters related to PE and control respectively at protein fingerprint region of 1800-1000 cm^{-1} . Small outlier group of control subjects are outlier when a 95 % confidence interval is applied. Loading plot (Figure 4.4B) shows different intensity peaks related to amide I at range of 1700-1600 cm^{-1} and amide II at range of \sim 1665-1530 cm^{-1} . Data associated with both PC1 and PC2 of PE and control at 15 week gestational age are presented in Figure 4.4B.

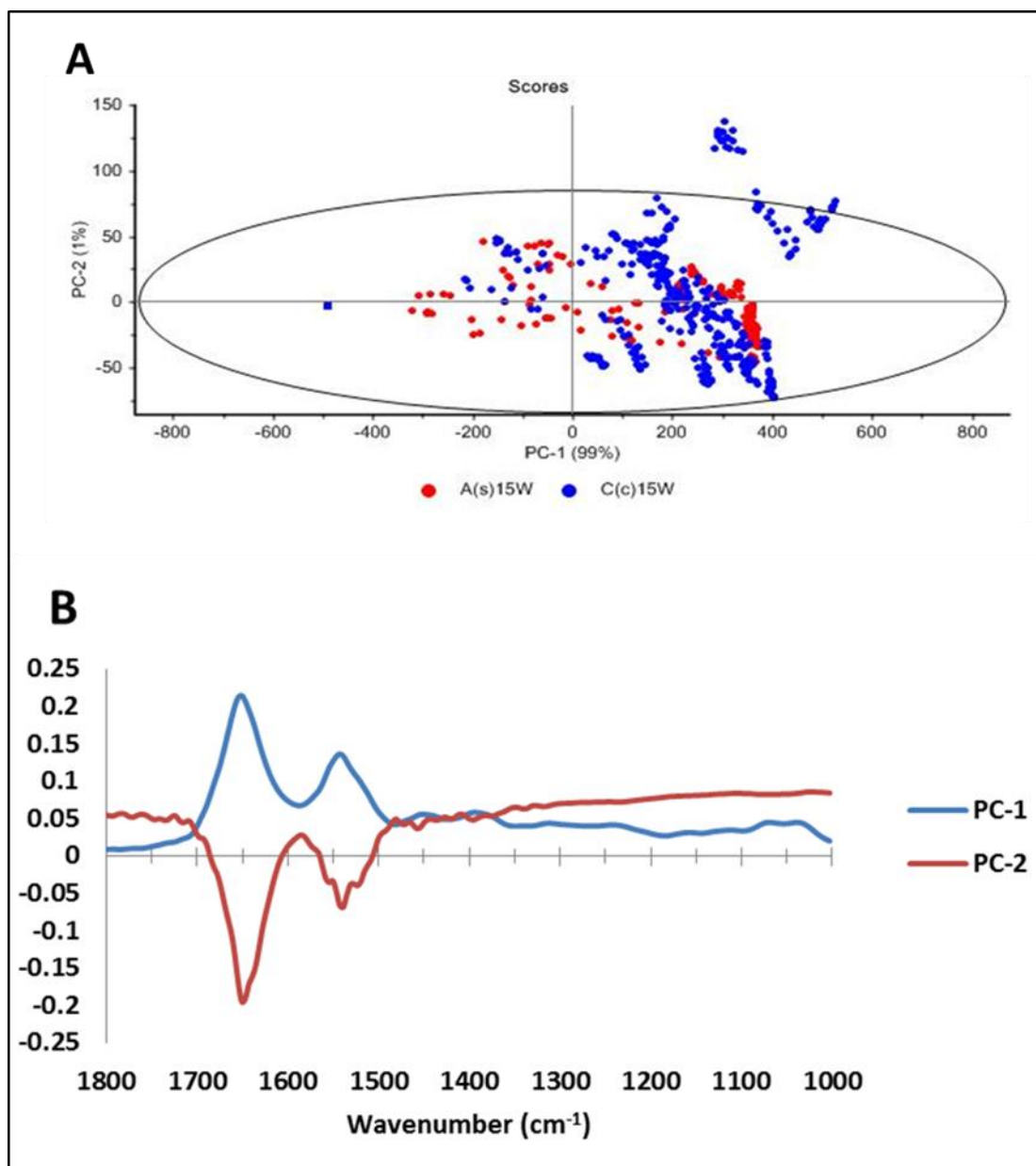


Figure 4. 4: PCA plot of PE and control during 15 week of gestation at protein region of 1800-1000 cm⁻¹

PCA plot shows scores distribution of PC1 and PC2 around normalised scale (**A**) for PE and control. Data were analysed for protein fingerprint area at 1800-1000 cm⁻¹ using confidence interval at 95 %. Different absorbance intensities associates with PC1 and PC2 loading spectra (**B**) corresponds to amide I at range 1700-1500 cm⁻¹ , amide II at range 1500-1350 cm⁻¹ , and amide III at range ~1335-1220 cm⁻¹ related to PE and control are observed. **A**: Scores of IR spectral waves, A (s) = PE at 15 week, C (c) = non-complicated pregnancies (control) at 15 week, **B**: Loading score of IR vibrational bands at 1800-1000 cm⁻¹ region; PC1 (blue); PC2 (red).

4.4.3. Spectral profile study of lipid region for PE and control at 15 weeks gestational age

PCA plots for PE and control FTIR data at 15 weeks gestational age, lipid region (3100-2700 cm^{-1}) (Figure 4.5A) 99 % of population variance is accounted for by PC1 and 1 % by PC2. Some clusters within the data are observed (Figure 4.5A); however there is overlap between some cases and some controls. Figure 4.5A demonstrates the observation of an outlier group of control subject samples at a 95 % confidence interval. Loading plot (Figure 4.5B) demonstrated peaks of C-H stretching at asym.(2985-2950) cm^{-1} ; sym. (2885-2865) cm^{-1} for CH_3 group, and at range (2990-2950) cm^{-1} and (2850-2780) cm^{-1} (asymmetric and symmetric) vibrational bonds respectively for CH_2 group are responsible for the significant PC2 variation associated with this outlier separation. This region was associated with both PC1 and PC2; intensities related to PE and control at 15 week gestational age for lipid region of 3100-2700 cm^{-1} are observed in Figure 4.5B.

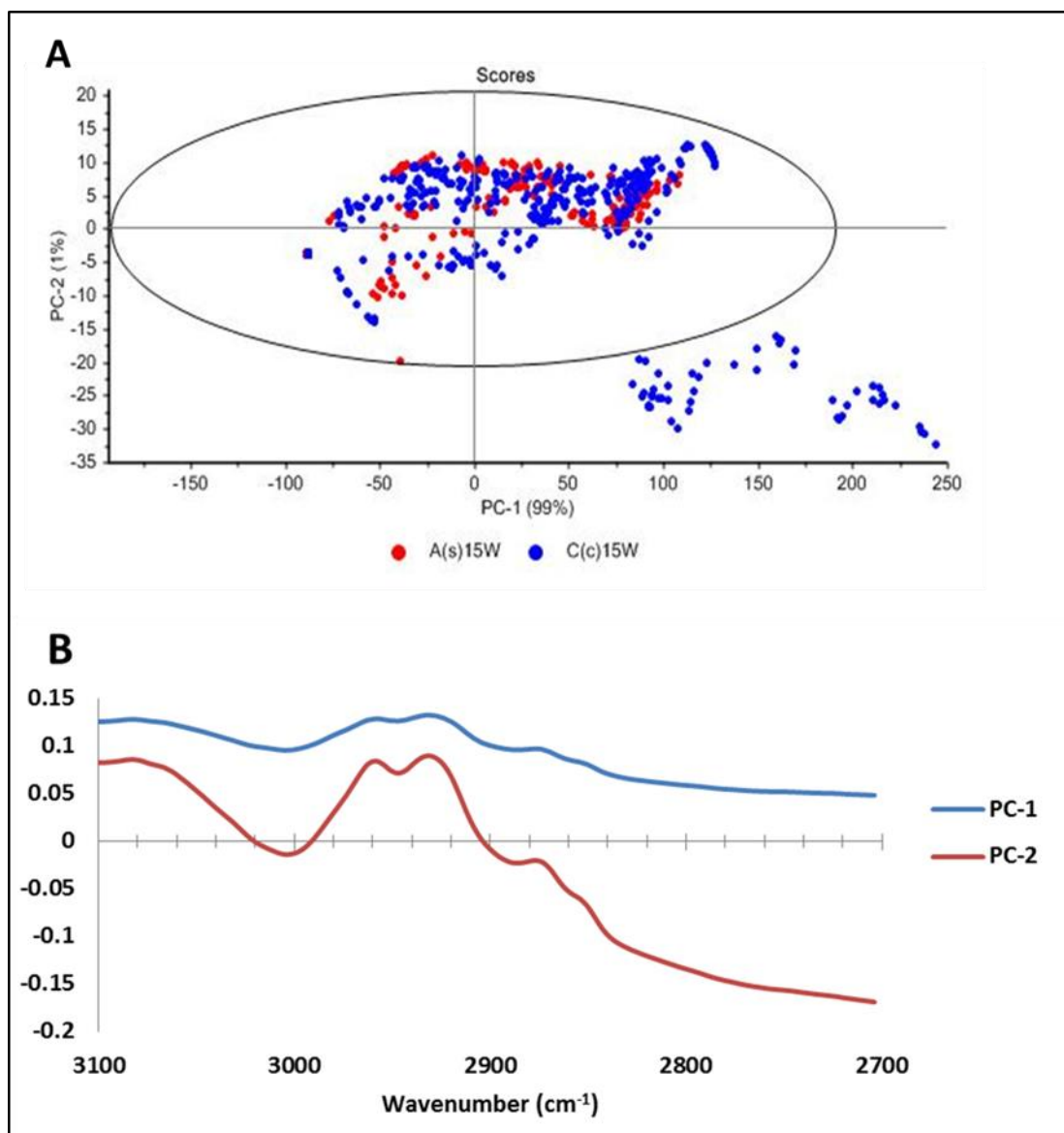


Figure 4. 5 : PCA and loadings plot of lipid region of (3100-2700 cm⁻¹) for 15 week gestation samples in PE vs. control subjects

PCA plot displays score distribution of PC1 and PC2 around normalised scale (**A**) for PE and control. Data were analysed using lipid region at 3100-2700 cm⁻¹ using 95 % confidence interval. Different absorbance intensities between PC1 and PC2 loading spectra (**B**) were observed. **A**: Scores of IR spectral waves, A (s) = PE at 15 week, C (c) = non-complicated pregnancies (control) at 15 week, **B**: Loading score of IR vibrational bands at 3100-2700 cm⁻¹ region; PC1 (blue); PC2 (red).

4.4.4. Spectral profile study for PE and control subjects at 20 weeks' gestation

Mid-IR region FTIR spectral profiles ($4000\text{-}600\text{cm}^{-1}$) were collected for samples provided at 20 weeks' gestation from women with PE and non-complicated pregnancies (controls). Mean spectral profiles for these groups are plotted in Figure 4.6. Higher intensity FTIR spectral profiles are observed for PE subjects compared with controls (Figure 4.6). The primary observed differences are in peaks in specific ranges; amide A, N-H and O-H stretching vibration ($3500\text{-}3100\text{ cm}^{-1}$), amide I and amide II ($1700\text{-}1450\text{ cm}^{-1}$), amide III ($1235\text{-}1224\text{ cm}^{-1}$), C-H stretching bonds of lipids ($2960\text{-}2950\text{ cm}^{-1}$, =CH), $2880\text{-}2865\text{ cm}^{-1}$ and $2860\text{-}2840\text{ cm}^{-1}$ for (asymmetric and symmetric) vibrational bonds respectively (CH_3 and CH_2) all show increased absorbance in cases as compared to controls at this gestation (Figure 4.6). A smaller peak at 1730 cm^{-1} ((C=O) stretching bond of ester), alongside bending vibration of (CH_2 , $\sim 1465\text{ cm}^{-1}$), and carbohydrate-related (C-O) and (C-O-C) stretch band peaks ($1170\text{-}1000\text{ cm}^{-1}$) (Figure 4.6).

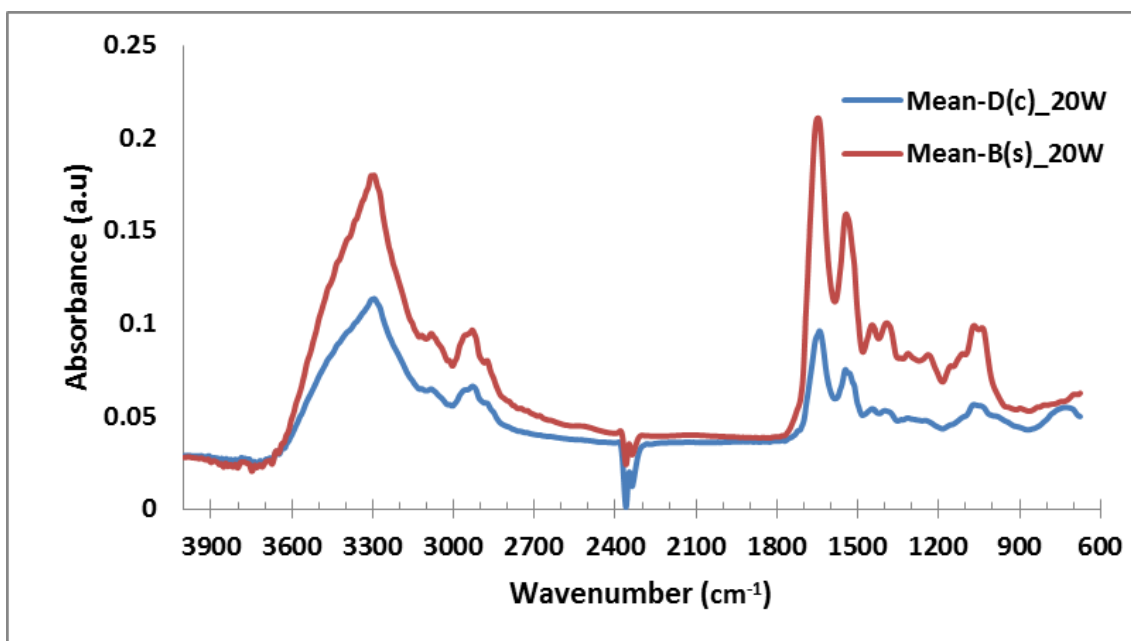


Figure 4. 6: Averaged mid FTIR spectral profiles for PE and control samples at 20 weeks' gestation

FTIR spectral profile averages were calculated for PE (n=30) and control (n=30) samples. Red B (s) =mean of PE (n=30) at 20 week of pregnancy, blue D (c) = mean of control (n=30) at 20 week of pregnancy.

The following section describes the spectral profiles of PE and control at 20 weeks of pregnancy using PCA description related to protein fingerprint of 1800-1000 cm^{-1} and lipid at 3100-2700 cm^{-1} regions respectively.

4.4.5. Spectral profile study of protein fingerprint region for PE and control at 20 weeks gestational age

PCA plots of protein fingerprint region of 1800-1000 cm^{-1} are presented in Figure 4.7A for PE and control samples at 20 weeks of pregnancy at 95 % of observed variation is related to PC1 vs 4 % for PC2. The score plot that presented in Figure 4.7A shows some degree of

separation related to PE and control over this protein fingerprint region. Some data fall outside the 95 % confidence interval (outside black line). The associated loading plot (Figure 4.7B) demonstrates differences in the observed intensity of peaks related to amide I ($1700\text{-}1600\text{ cm}^{-1}$), a band at $\sim 1650\text{ cm}^{-1}$, and amide II ($\sim 1685\text{-}1530\text{ cm}^{-1}$).

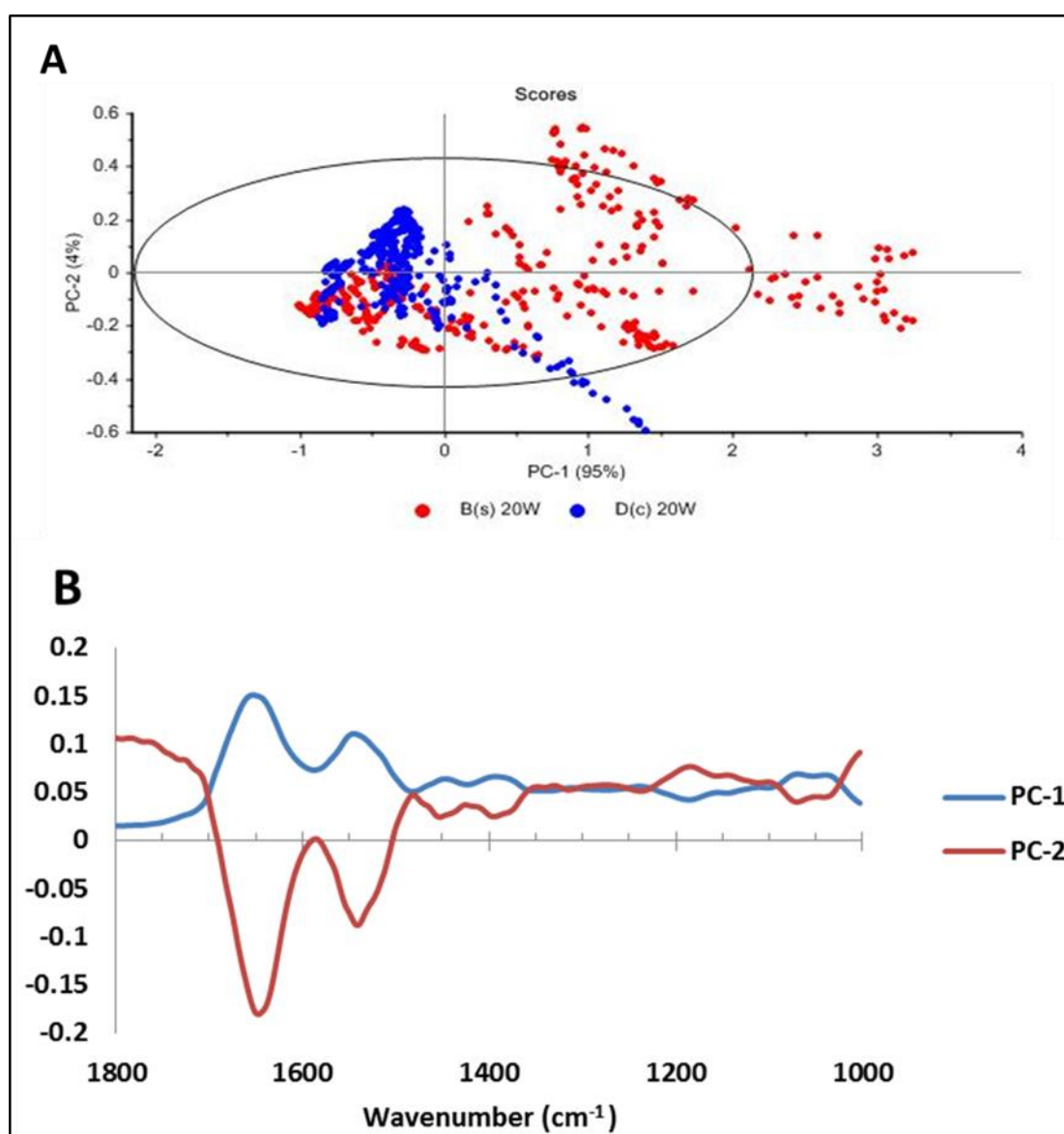


Figure 4. 7: PCA plot protein region ($1800\text{-}1000\text{ cm}^{-1}$) for PE and control samples at 20 weeks' gestation

PCA plot shows scores distribution of PC1 and PC2 around normalised scale (A) for PE and control. Data are presented for protein fingerprint area ($1800\text{-}1000\text{ cm}^{-1}$) with a 95% confidence interval. Outlier data lie out with the 95% confidence interval. **A:** Scores of IR spectral waves, B (s) = PE at

20 week, D (c) = non-complicated pregnancies (control) at 20 week, **B**: Loading score of IR vibrational bands at 1800-1000 cm^{-1} region; PC1 (blue); PC2 (red).

4.4.6. Spectral profile study of lipid region for PE and control at 20 weeks gestational age

Data presented in Figure 4.8A as a score plot for PE and control urine samples' lipid region (3100-2700 cm^{-1}) at 20 weeks' gestation . The scores plot in Figure 4.8A shows 97 % of populations are related to PC1 and 3 % in PC2. Some clusters differentiating PE and controls are observed in this region (Figure 4.8A). A number of outliers are observed which are outside the PCA 95% confidence interval. The associated loading plot (Figure 4.8B) shows a number of differential peaks: asymmetric and symmetric CH_3 C-H stretching (2985-2950 and 2885-2865 cm^{-1} respectively); and also asymmetric and symmetric vibrations CH_2 group for CH_3 group (2990-2950 and 2850-2780 cm^{-1} respectively). Peaks with significant contributions to both PC1 and PC2 were observed at these wavenumbers showing different intensities related to PE and control at 20 week gestation (Figure 4.8B).

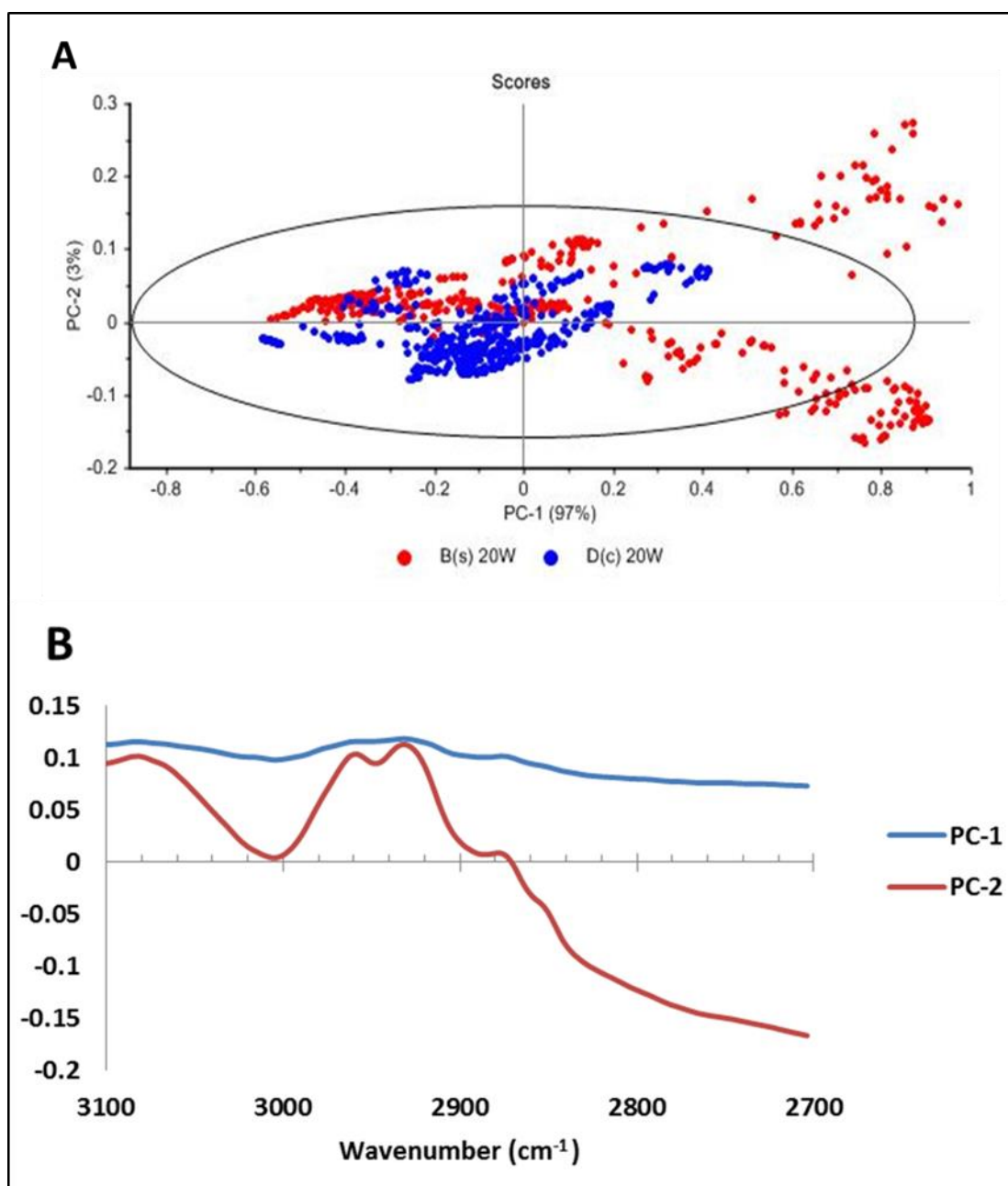


Figure 4. 8: PCA plot of PE and control during 20 week of gestation at lipid region of 3100-2700 cm^{-1}

PCA plot showing score distribution of PC1 and PC2 around normalised scale (A) for PE and control at 20 weeks of pregnancy. Data were analysed using lipid region at 3100-2700 cm^{-1} using confidence interval 95 %. Outlier data related to PE at 20 week gestational age of PCA scores using 95 % confidence interval are observed. Different absorbance intensities between PC1 and PC2 loading spectra (B) were observed related to C-H stretching vibrational bond of asym. (2985-2950 cm^{-1}) cm^{-1} ; sym. (2885-2865) cm^{-1} for CH_3 group, and at range (2990-2950) cm^{-1} and (2850-2780) cm^{-1} (asymmetric and symmetric) vibrational bonds respectively for CH_2 group. **A:** Scores of

IR spectral waves, B (s) = PE at 20 week, D (c) = non-complicated pregnancies (control) at 20 week, B: Loading score of IR vibrational bands at 3100-2700 cm^{-1} region; PC1 (blue); PC2 (red).

4.4.7. FTIR spectral profile of IUGR and control at 15 weeks' gestation

Spectral profiling by FTIR was carried out on urine samples collected from subjects with IUGR and those with uncomplicated pregnancies (controls) at 15 weeks' gestation over the mid-IR region (4000-600 cm^{-1}). Figure S2 (see appendix 5) shows overlaid FTIR spectra in IUGR and control at 15 weeks gestation. There are few visible differences; indicating mean spectral difference analysis is merited to observe emergent differences between cases and controls. Figure 4.9 shows differences between mean spectral profiles for IUGR (n=30) and uncomplicated pregnancies (n=30). Data are presented as absorbance intensity vs wavenumber. Mean spectral profile for IUGR shows overall higher intensity absorbance compared with that of uncomplicated pregnancy (Figure 4.9). The principal differences are peaks related to amide A, N-H and O-H stretching vibrations (3500-3100 cm^{-1} , Figure 4.9). Different intensities in spectral profile between IUGR and control related to protein fingerprint are also observed (1800-1000 cm^{-1}), as are amide I and amide II (1700-1450 cm^{-1}); amide III (1235-1224 cm^{-1}), C-H stretching (lipid region, 3100-2700 cm^{-1}). Smaller peaks at 1730 cm^{-1} (C=O stretching bond of ester); bending vibration of (CH_2) at 1465 cm^{-1} , and carbohydrate-related signals (1100-1050 cm^{-1} , (C-O) and (C-O-C)) were also observed as differential signals (Figure 4.9).

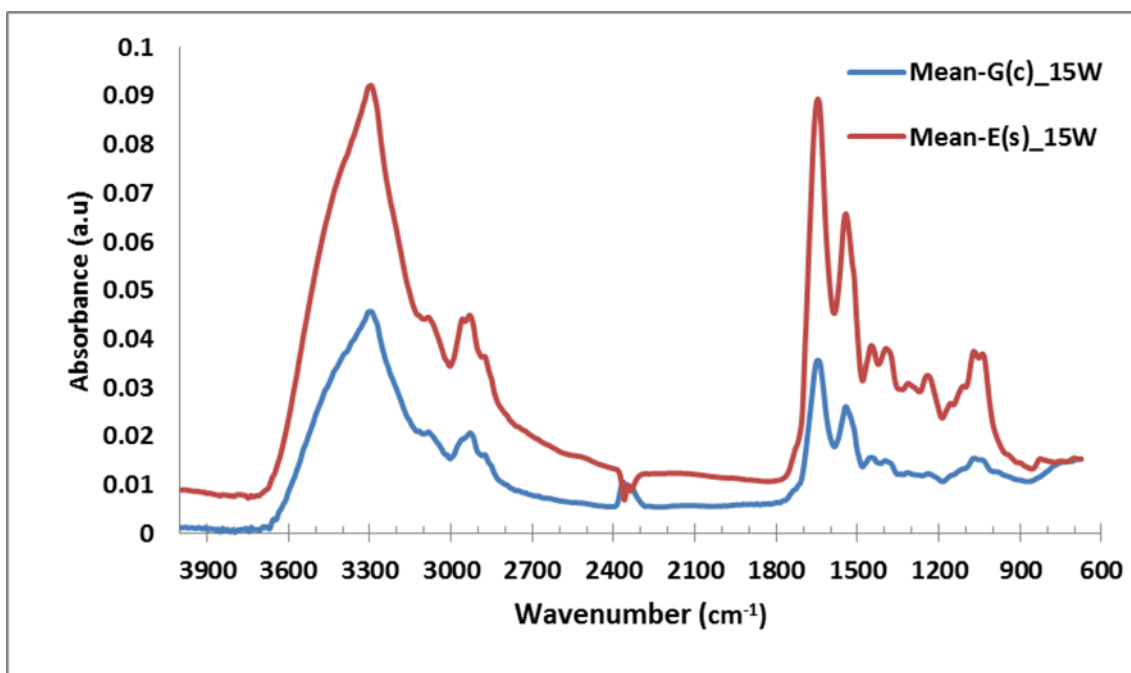


Figure 4. 9: Mean spectral profiles for mid-IR analysis of IUGR and control urine samples at 15 weeks' gestation

FTIR spectral profile means for IUGR (n=30) and control (n=30) samples were calculated. Wavenumber is plotted vs. absorbance intensity. Red E (s) =mean of IUGR samples (n=30) at 15 week of pregnancy, blue G (c) = mean of control samples (n=30).

The following sections describe the spectral profiles of IUGR and control at 15 weeks of pregnancy using PCA of protein fingerprint (1800-1000 cm^{-1}) and lipid (3100-2700 cm^{-1}) regions.

4.4.8. Spectral profile investigation of protein fingerprint region for IUGR and control urine samples at 15 weeks' gestation

Figure 4.10A shows a score plot for IUGR and control urine sample FTIR at 15 weeks' gestational age at protein fingerprint region (1800-1000 cm^{-1}). The score plot in Figure

4.10A shows 92 % of signal variance is related to PC1 and 6 % to PC2. Some outcome-related clusters are observed within this region (Figure 4.10A). A group of observations as outliers beyond the 95 % confidence interval were observed for the IUGR subjects. Loading plot (Figure 4.10B) attributes the differences in observed intensity to peaks related to amide I at (1700-1600 cm^{-1}), amide II ($\sim 1665\text{-}1530 \text{ cm}^{-1}$), with lower-scale differences related to amide III (1229-1220 cm^{-1}). These changes were across both PC1 and PC2 (Figure 4.10B).

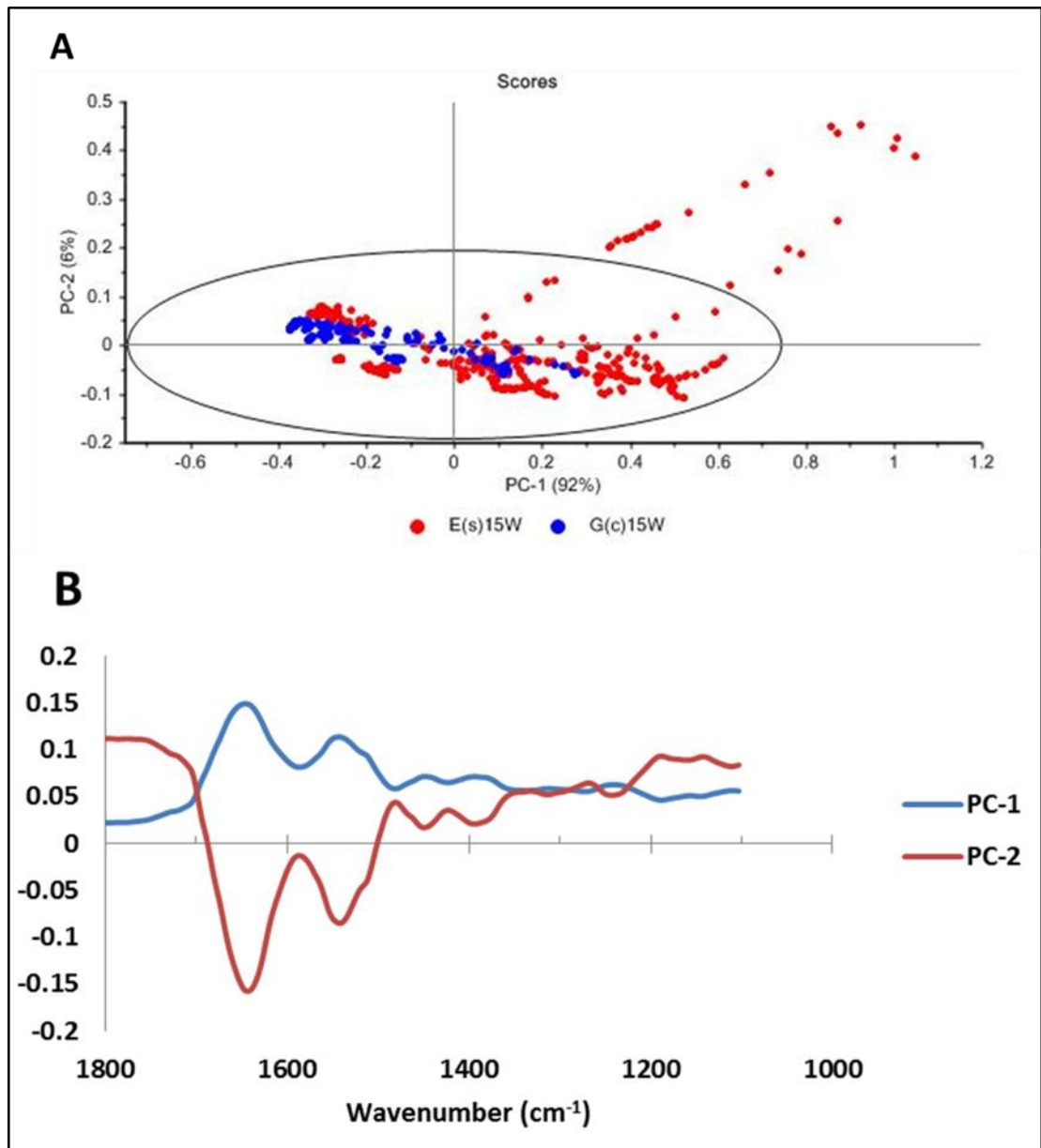


Figure 4. 10: PCA plot of IUGR and control urine samples collected at 15 weeks' gestation for protein region 1800-1000 cm⁻¹

PCA plot shows score distribution of PC1 and PC2 around normalised scale (A) for IUGR and control. Data were analysed using protein fingerprint area at 1800-1000 cm⁻¹ at confidence interval at 95 %. Outlier data related to IUGR at 15 week gestational age of PCA scores using 95 % confidence interval are observed. Different absorbance intensities associate with PC1 and PC2 loading spectra (B). A: Scores of IR spectral waves, E (s) = IUGR at 15 week, G (c) = uncomplicated pregnancies (control) at 15 week, B: Loading score of IR vibrational bands at 1800-1000 cm⁻¹ region; PC1 (blue); PC2 (red).

4.4.9. Spectral profile investigation of lipid region for IUGR and control urine samples at 15 weeks' gestation

Figure 4.11A shows score plot for IUGR and control urine samples at 15 weeks of pregnancy over the lipid region ($3100\text{-}2700\text{ cm}^{-1}$). 99 % of observed variation is related to PC1 and 1 % to PC2. Some differential clustering of cases and controls is observed (Figure 4.11A); with a subgroup of outlier observations pertaining to IUGR cases lie out with the PCA 95 % confidence interval. The loading plot associated with this PCA (Figure 4.11B) shows that differences are attributable to C-H stretching peaks (CH_2 and CH_3 (symmetric and asymmetric) $\sim 2950\text{-}2850\text{ cm}^{-1}$). Intensity differences are observed in both PC1 and PC2 (Figure 4.11B).

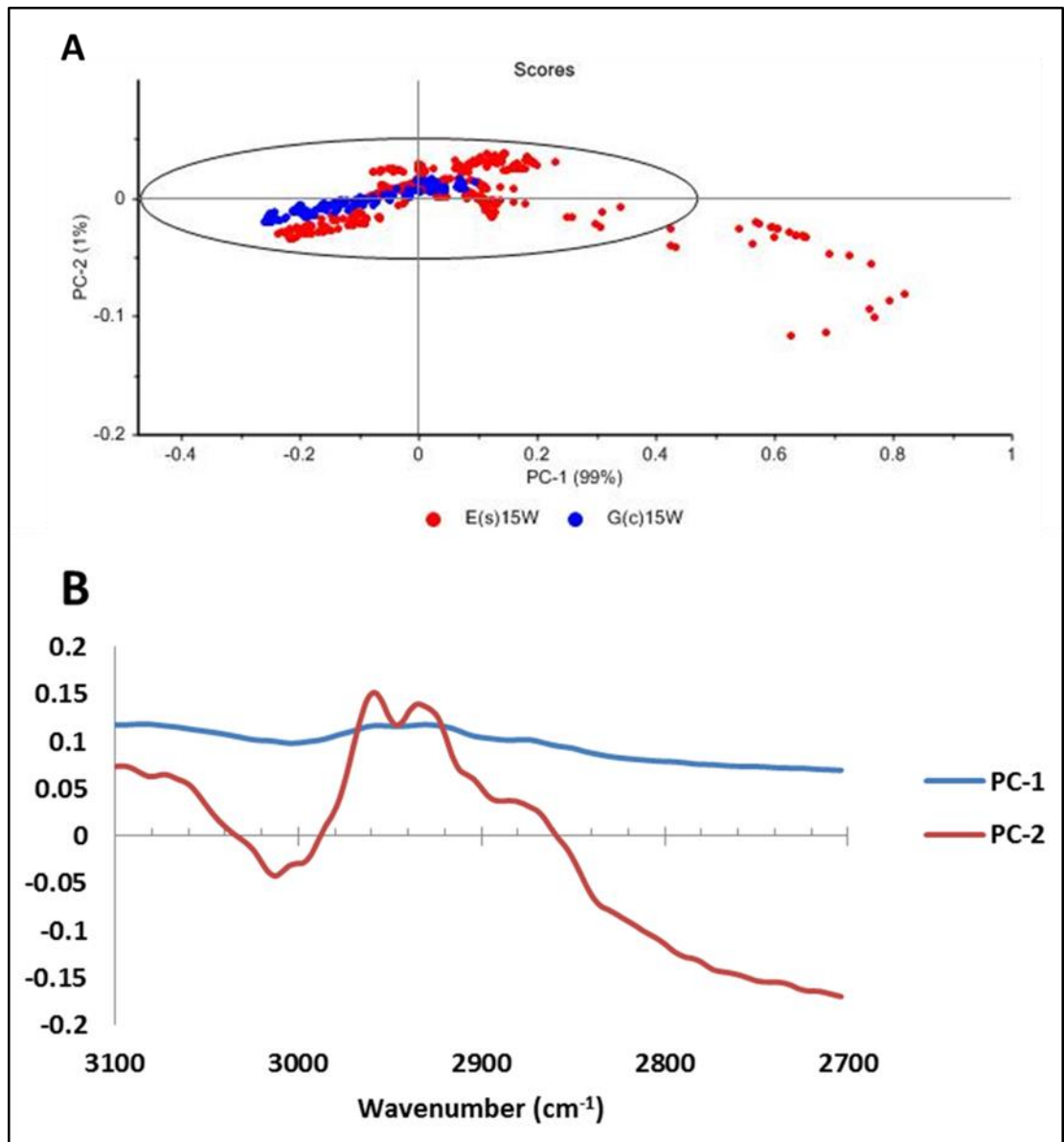


Figure 4. 11: PCA plot of IUGR and control urine samples collected at 15 weeks' gestation over lipid region of 3100-2700 cm⁻¹

PCA plot shows score distribution of PC1 and PC2 around normalised scale (A) for IUGR and control urine at 15 weeks of pregnancy. Lipid region at 3100-2700 cm⁻¹ plotted, at a 95% confidence interval PC1 and PC2 loading spectra (B). **A:** Scores of IR spectral waves, E (s) = IUGR at 15 week, G (c) = non-complicated pregnancies (control) at 15 week, **B:** Loading plot of IR vibrational bands at 3100-2700 cm⁻¹ region; PC1 (blue); PC2 (red).

4.4.10. Spectral profiling of urine samples collected from IUGR and control subjects at 20 weeks' gestation

FTIR analysis was performed for urine samples obtained from subjects with IUGR and uncomplicated pregnancies (controls) at 20 weeks' gestation. Limited difference was observed in mean spectral profile between IUGR (n=30) and control (n=30) samples over the mid-IR range (4000-600 cm^{-1} , Figure 4.12). Observationally, minor differences are present related to amide A, N-H and O-H stretching vibration (3500-3100 cm^{-1}) and C-H stretching (2960–2950 cm^{-1}) (Figure 4.12). More evident are protein finger print region amide I and amide II variations (1700-1000 cm^{-1}). Increased variation of intensity is observed for (C=O) stretching bond of ester (small peak at 1730 cm^{-1}); bending vibration of CH_2 (1465 cm^{-1}), and carbohydrate-related (C–O) and (C–O–C) stretch (1170-1000 cm^{-1}).

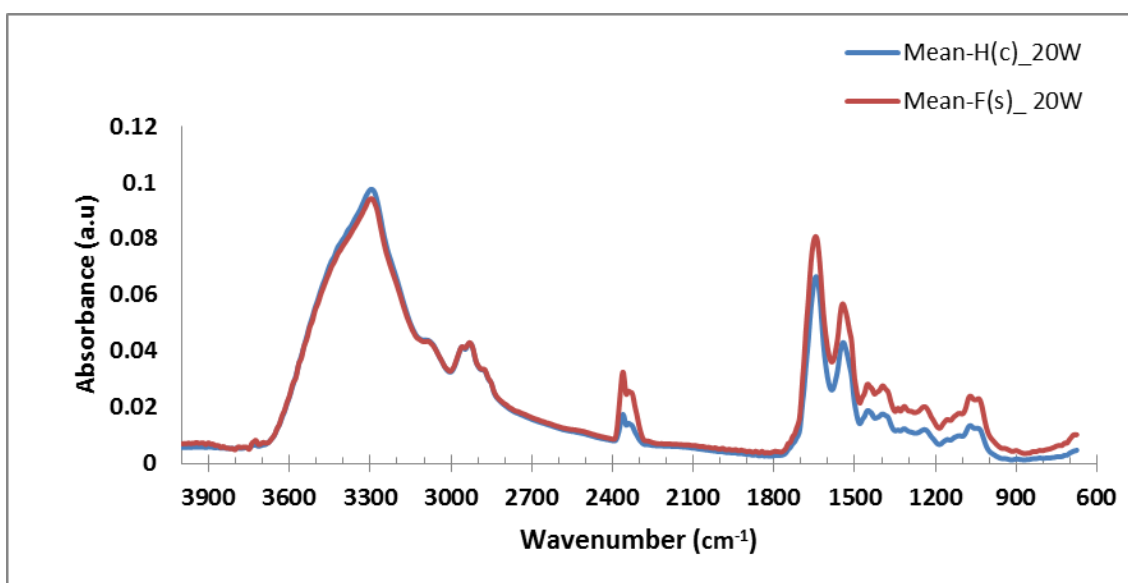


Figure 4. 12: Spectral profile comparison of mean FTIR of IUGR and control samples at 20 weeks' gestation

Mean FTIR spectral profiles plotted for IUGR (n=30) and control (n=30) urine samples collected at week 20 of pregnancies. Red F (s) =mean of IUGR (n=30) at 20 week of pregnancy, blue H (c) = mean of control (n=30) at 20 week of pregnancy.

The following section describes the spectral profiles of IUGR and control at 20 weeks of pregnancy using PCA description divided according to protein fingerprint (1800-1000 cm^{-1}) and lipid (3100-2700 cm^{-1}) regions respectively.

4.4.11. Spectral profile investigation of protein fingerprint region for urine samples taken from IUGR cases and controls at 20 weeks' gestation

Figure 4.13A is a PCA score plot for IUGR and control samples at 20 weeks of pregnancy over the protein fingerprint region of 1800-1000 cm^{-1} . 91% of observed variance was accounted for by PC1, and 8% in PC2. Figure 4.13A shows some overlapping clusters related to IUGR and control observations. A number of case observations are outlier out with the 95% confidence interval. Associated loading plot for this analysis (Figure 4.13B) shows a number of variant peaks related to amide I (1700-1600 cm^{-1}), and amide II (~1585-1480 cm^{-1}), associated with PC1. PC2 shows a small shift related to amide I (1670-1575 cm^{-1}), and amide II (1560-1475 cm^{-1}).

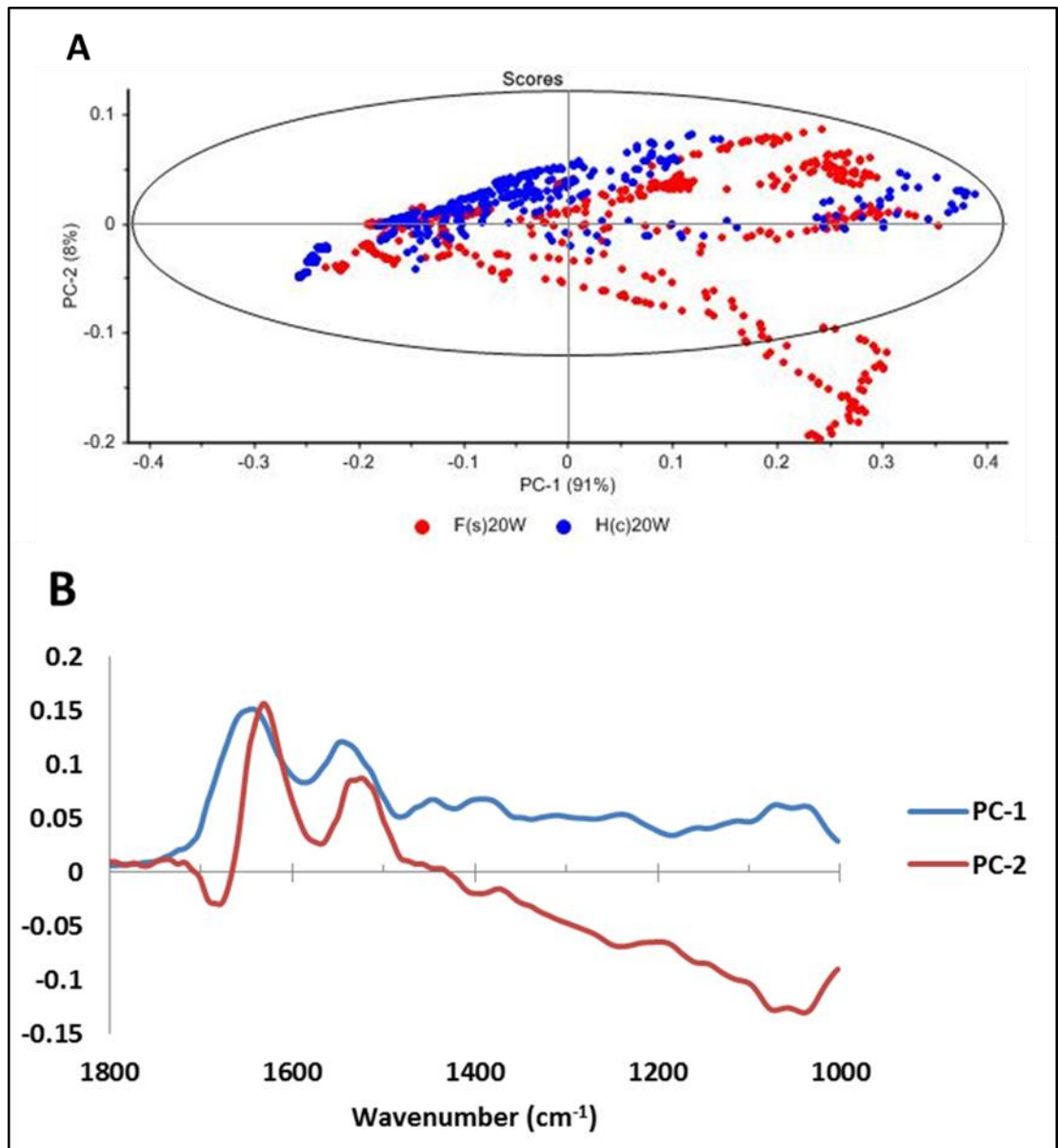


Figure 4. 13: PCA plot of IUGR and control urine samples collected at 20 weeks' gestation over protein region, 1800-1000 cm⁻¹

PCA plot shows score distribution of PC1 and PC2 around normalised scale (A) for IUGR and control. Data were analysed using protein fingerprint area at 1800-1000 cm⁻¹ at a 95% confidence interval A: Scores of IR spectral waves, F (s) = IUGR at 20 week, H (c) = non-complicated pregnancies (control) at 20 week, B: Loading score of IR vibrational bands at 1800-1000 cm⁻¹ region; PC1 (blue); PC2 (red).

4.4.12. Spectral profile investigation of lipid region for IUGR and control at 20 weeks' gestation

Data presented in Figure 4.14A show score plots for IUGR and control urine samples at 20 week of pregnancy at lipid region of 3100-2700 cm^{-1} . Score plot (Figure 4.14A) shows 93 % of variation is related to PC1 and 6 % to PC2. Data show no clear pattern, with both cases and control observations being spread around scale at a 95% confidence interval. A small number of both case and control outlier observations sit out with the PCA 95% confidence interval. Loading plot (Figure 4.14 B) showed differences could be attributed to broad peaks of C-H stretching (2950-2850 cm^{-1}), corresponding to CH_2 and CH_3 (symmetric and asymmetric).

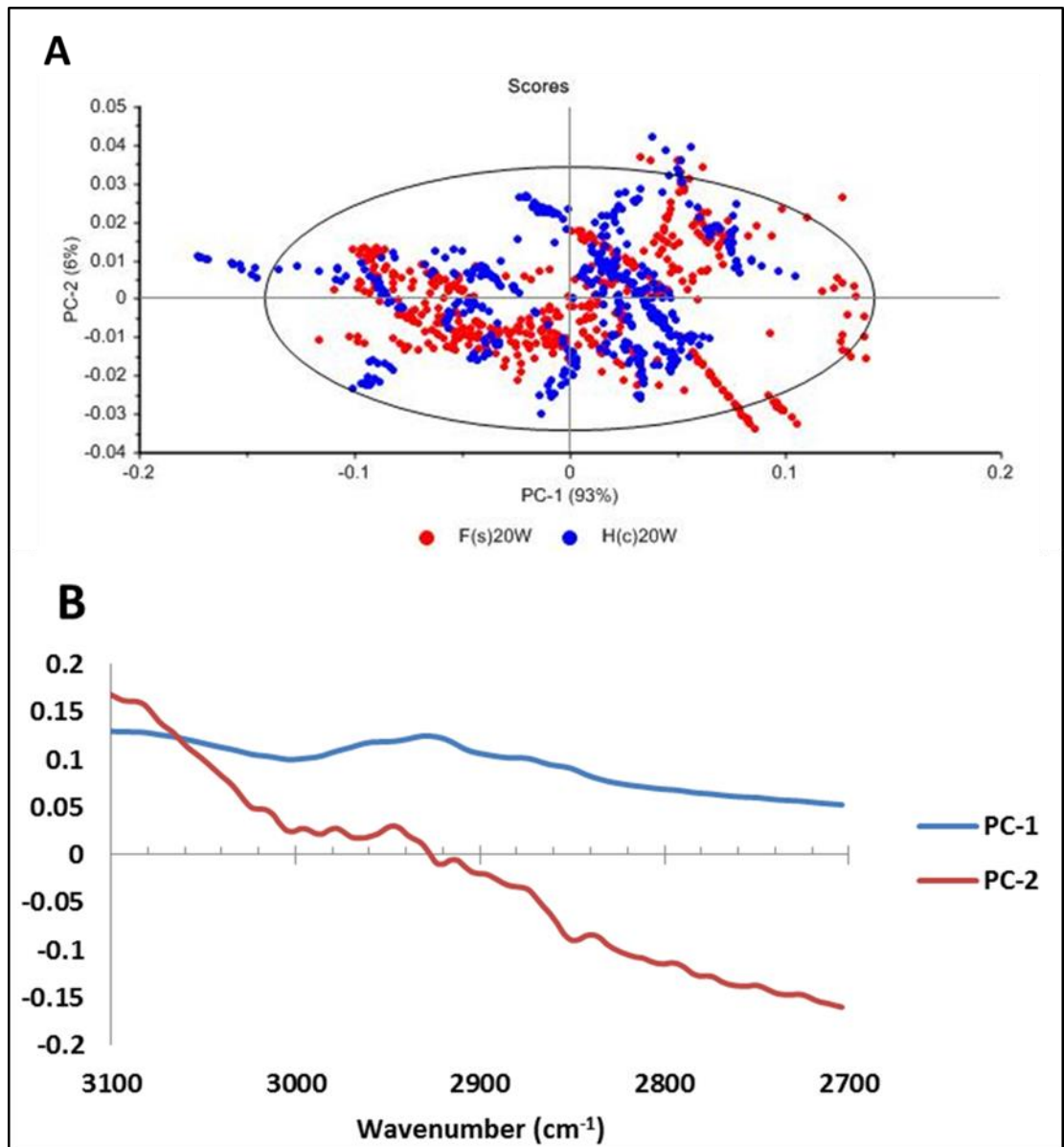


Figure 4. 14: PCA plot of IUGR and control subjects for 20 week's gestation using lipid region of 3100-2700 cm⁻¹

PCA plot shows score distribution of PC1 and PC2 around normalised scale (**A**) for IUGR and control at 20 weeks of pregnancy. Data were analysed using lipid region at 3100-2700 cm⁻¹ using confidence interval 95 %. **A:** Scores of IR spectral waves, F (s) = IUGR at 20 week, H (c) = non-complicated pregnancies (control) at 20 week, **B:** Loading score of IR vibrational bands at 3100-2700 cm⁻¹ region; PC1 (blue); PC2 (red).

4.5. Discussion

This study was performed to identify whether changes in IR spectral profile could be observed for PE and IUGR disorders when compared to uncomplicated control pregnancies. The aim of this Chapter was to report the investigation of mid-IR spectral differences between urine samples from women clinical presenting with these disorders during 15 and 20 weeks of gestational age using FTIR. No clinically useful screening test currently exists to predict the progress of PE prior to onset of clinically-observable symptoms (Conde-Agudelo *et al.*, 2004). According to the changes observed hypothesised to be within the protein (and lipid) profiles of these diseases; not only might these profiles change in regard to the composition of these species, but also due to their structural changes, i.e. quaternary and secondary structures of proteins.

Differences in urine profiles may be indicative of implantation abnormality in PE (Roberts, 1998, Powell *et al.*, 2018) or IUGR (Conde-Agudelo *et al.*, 2013, Powell *et al.*, 2018). These differences could manifest as either up or down regulation of profile expression (Takacs *et al.*, 2003, Powers *et al.*, 2005). In order to be of prognostic value, biomarkers need to show predictable and robust changes before and during clinical manifestation of PE (Hubel, 1999, Krauss *et al.*, 1997, Taylor *et al.*, 1991).

It has been suggested by others that PE is a disturbance of antioxidant inadequacy appearing when the normal antioxidant equilibrium is upset (Stark, 1993). Oxidative damage induced by free radicals or Reactive Oxygen Species (ROS) could cause protein modification and lipid peroxidation (Sinclair *et al.*, 1990), and consequently alterations in membrane properties and cell damage (Spickett *et al.*, 1998). When free radicals attack polyunsaturated fatty acids or cholesterol in membrane and lipoproteins, they generate

high reactive species of lipid peroxidation and lipid hydroperoxides (Sinclair *et al.*, 1990). In normal physiological function, lipid hydroperoxides are regulated by enzymes and redox-sensitive genes (Marnett, 1992, Sen and Packer, 1996). Uncontrolled lipid peroxidation could be causing cellular dysfunction and damage which associates with abnormal placentation such as in PE and IUGR (Hubel, 1999, Hubel *et al.*, 1990, Walsh, 1998, Maisonneuve *et al.*, 2015). Therefore, oxidative stress results in such damage, which normally prevented by wide-ranging and manifold antioxidant substances. Antioxidant also known as defence system that having substances with different molecular weight (Sinclair *et al.*, 1990).

Albumin represents major circulating antioxidant components in plasma, allowing exposing to continuous oxidative stress (Soriani *et al.*, 1994). Albumin has direct protective effects vs mortality risk due to the opposite relationship between serum albumin level and mortality risk has suggested (Goldwasser and Feldman, 1997, Phillips *et al.*, 1989, Halliwell, 1990). Alterations in the structure of albumin perhaps result in impairments of its biological functions (Terawaki *et al.*, 2004). Modifications of albumin functions may occur in many pathological conditions associated with early incidence of vascular complications conditions such in case of PE and insulin dependent diabetes mellitus patients and IUGR.

On the basis of the above consideration, we can explain the major alteration in the band intensities of protein finger print region of 1800-1000 cm^{-1} , amide I C=O stretching at range $\sim 1750\text{-}1560\text{ cm}^{-1}$, amide II N-H bending at range around $1550\text{-}1430\text{ cm}^{-1}$, and amide III (complex vibrations bonds) at range about $(1335\text{-}1220)\text{ cm}^{-1}$. The scores of the PCA plot showed a degree of data separation between case and control related to PE at 20

weeks of pregnancy can be seen in Figure 4.7A. Similarly, some clusters have been observed relating to patients complicated with IUGR during 15 and 20 weeks respectively compared to normal pregnancies (Figure 4.10A and 4.13A). Alterations in secondary structure of proteins could be associated with a lack in an antioxidant system. These findings agree with a study presented by Raouf and co-workers (2011) who used FTIR analysis of serum to study protein structure of PE at different gestational ages (Raouf *et al.*, 2011).

It has been suggested that there is a positive correlation between early pregnancy dyslipidemia and PE (Enquobahrie *et al.*, 2004). Women with a history of PE show a significant difference in lipid profile and high susceptibility to lipoprotein oxidation compared with uncomplicated pregnancies (Gratacós *et al.*, 2003, Mittal *et al.*, 2017). Therefore, disturbance of lipoprotein metabolism could be a characteristic of major causes of hypertension and proteinuria in PE (Gratacós *et al.*, 2003). The correlation between women complicated with IUGR and lipid peroxidation results in oxidative stress has been reported by a number of studies (Hracsko *et al.*, 2008, Lutnicki *et al.*, 1998, Karowicz-Bilińska, 2006, Dikbas *et al.*, 2017). For example, study reported by Fraile and colleagues (2003) showed that the ratio between amide II/I was higher in pure albumin than with albumin interacting with lipids (Fraile *et al.*, 2003). They suggested that structure of pure albumin has destabilised by lipids, which is now a well-understood and accepted mechanism of protein stabilisation (Fraile *et al.*, 2003)

The current research in this chapter has shown variation in a lipid region of range 3100-2700 cm^{-1} in PE at 20 week gestational age (Figure 4.8) and IUGR at 15 week of pregnancy (Figure 4.11) respectively. These bands correspond to C-H stretching bonds of

lipids components at range 2960–2950 cm^{-1} (as); CH_3 , 2880–2865 cm^{-1} , and at 2860–2840 cm^{-1} for (as and s), and bending vibration of CH_2 at peak around 1465 cm^{-1} of lipid acyl group, and ester (C=O) stretching at $\sim 1750\text{-}1700$ cm^{-1} . Results in this chapter obtained from urine samples in complicated pregnancies agreed with a previous report suggesting that strongly correlation exists between maternal complications and abnormal lipid metabolism of pre-eclamptic patients (Aziz and Mahboob, 2007).

Other research agreed with our findings (Lorentzen *et al.*, 1994). Their results revealed that the concentration of Free Fatty Acid (FFA), triglycerides (TG), and lipolytic activity were significantly increased in pre-eclamptic women at 16-18 weeks of gestational age compared to normal pregnancies (Lorentzen *et al.*, 1994). Elsewhere, studies by Kamath, *et al.*, stated that the possible increases of lipid peroxidation of malondialdehyde (MDA) and protein oxidation in both maternal and fetal erythrocytes for intrauterine growth retardation (Kamath *et al.*, 2006). Therefore, lipid profile abnormalities could be used as discriminatory biomarkers for maternal pathogenic conditions such in the case of PE and IUGR (Lorentzen *et al.*, 1994, Alvino *et al.*, 2008).

4.6. Conclusions

The results presented in this chapter highlight key spectral differences between controls and PE and IUGR clinical urine samples collected at various time points post conception. PCA is not completely discriminatory but scores plots for these samples highlight areas of difference, particularly in the protein bond regions. This is supported by other evidence from literature and aligns with other spectral differences observed in the lipid range and with the proteomics data in Chapter 2, where specific proteins were observed to show

altered expression patterns in PE and IUGR. Further study will be required in the future for integrative validation in a larger cohort.

Chapter 5

5. General discussion, conclusions and future work

5.1 Aim and direction of work

The overarching aim of the work presented in this thesis was to investigate the potential for urine-based biomarkers for both PE and IUGR studies at early stages of pregnancy. Here the results are discussed in context with the current state-of-the-art in proteomic biomarker analysis. Using two analytical techniques to independently assess a range of clinical samples, liquid chromatography mass spectrometry (LC-MS) and FTIR were used to investigate possible biomarkers, with the purpose of leading towards improved diagnostics and earlier management of PE and IUGR.

5.2 FTIR optimisation for urine-based samples

FTIR interrogation of urine samples was carried out in order to optimise methodology; it was important to consider the subsequent clinical use of this FTIR methodology with focus on the volumes of samples obtained (Lovergne *et al.*, 2015). Human samples are costly, have specific storage and handling requirements and are available in limited amounts, therefore rat urine was used as alternative to human samples for optimisation experiment work (Chapter 3); this enabled testing of urine volumes needed to produce reproducible spectra from as low volumes as possible.

Previous research has indicated differences in both protein bands (main amide bands have focussed on the region (1700-1550 cm^{-1}) and lipid regions ($\sim 3100\text{-}2700 \text{ cm}^{-1}$) (Raouf *et al.*, 2011). Since the thesis hypothesis was focussed on protein/ peptide biomarkers being found in urine from patients with PE and IUGR, the amide region was used for optimisation of signal clarity. A bands of amide I, II at range (1800-1500 cm^{-1}) were assessed by peak area, height peak, and wavenumber of peak maximum to understand

both the noise levels in terms of absorbance intensity and band reproducibility of band position, (Figure. 3.7 and 3.8). Spectra obtained from urine samples clearly showed amide bands with good sample repeat consistency. Peak height showed some degree of variance (Figure 3.8) having a mean of 0.175 with standard error of 0.053, which was most reliable with volumes of 5 μL . As this is a measure of in essence a single point within the spectra the level of noise can be expected. For this reason, the area (integrating under the amide profile) was used; this takes into account any variance across the whole region of interest and so ~ 60 data points were assessed rather than just 1 at the maximum. This is a common route used for the comparison of such spectra (Lovergne *et al.*, 2015). In addition to this, optimisation of data processing showed the reduced signal / noise, but importantly maintained consistency of data, when spectral smoothing was applied. There are always concerns that use of smoothing algorithms introduces spectral artefacts (Mukherjee *et al.*, 2014), and for this reason a low level 5 point adjacent averaging smoothing function was applied (Baker *et al.*, 2010).

Using amide I as the main band upon which methods were optimised, this work demonstrated a very good signal to noise ratio across the whole mid-IR range from samples with volumes as low as 5 μL . Figure (3.8) showed variability of observed signal according to volumes of 1, 2, 5 and 10 μL . However, results indicate a higher sensitivity related to most parameters at 5 and 10 μL vs 1 and 2 μL Figure (3.8). Previous research by (Hoşafçı *et al.*, 2007) used 5 μL of sample to investigate FTIR spectral vibration of urine and blood samples, other studies also used this volume of biofluid typically blood sample for FTIR analysis (Bonnier *et al.*, 2014, Lovergne *et al.*, 2015). Bonnier and co-workers (2014) revealed good improvement in FTIR spectra identification after protein fractions

using centrifugal filtrate at different molecular weight cut-off using blood sample at 5 μL , especially after centrifugal filtrate purification prior sampling. This was a very positive outcome since the total volumes expected from human samples (after processing; protein fractionation, etc.) and protein fractionation to establish the lowest reliable volumes which could be used to assess clinical samples, and which would be the best method to possibly identify spectral differences. A simple ultrafiltration protocol, based on molecular weight fractionation was used, enabling spectral profile determination of samples with improved observation of amide band (Bonnier *et al.*, 2014)

The outcome of optimisation of FTIR sampling demonstrated lower limits of reliable FTIR sampling. The second optimisation experiment was carried out to determine the influence of sample preparation procedure on FTIR spectra at different conditions including (whole rat urine, >10 kDa ultrafiltration urine sample without buffer exchange and >10 kDa after sample buffer exchange). (Bonnier *et al.*, 2014) has used ultrafiltration fractions bases on molecular weight cut-off to improve spectral profile outcome of blood sample. However, centrifugal filtrate bases on molecular weight cut-off coupled with buffer exchange method (Hart *et al.*, 2015, Çorbacı and Uçar, 2018) enabling to introduce a simple and robust methodology by removing unwanted molecules (Hall, 2015) that has a similar functional group and could be overlapping with the sample of interest. Ultrafiltration experiments displayed spectra peaks visibly of retentate of > 10 kDa compare with spectra of whole urine sample (Figure 3.9). Protein bands are clearly observed after retentate buffer exchange (see Figure.3.9 C) at spectra range 1700-1650 cm^{-1} with distinct peak at $\sim 1650 \text{ cm}^{-1}$ (amide I). A band of amide II observed at 1650-1500 cm^{-1} , amide III peak at $\sim 1330\text{-}1220 \text{ cm}^{-1}$ was also observed (Lacombe *et al.*, 2015).

Similarly, the spectra correspond to amide A and OH bands at $3650\text{-}3500\text{ cm}^{-1}$ and also C-H stretching band at $3000\text{-}2830\text{ cm}^{-1}$ in IR (Lacombe *et al.*, 2015). Although no attempt has been made to quantify the spectral characteristics corresponds to the constituent components in the urine, the study rather demonstrates that using such centrifugal fractionation procedure there is a passivity for ultimately developing medical diagnostics using this methodology.

In Chapter 4 the SOP that was established in Chapter 3 for urine sample analysis was applied to clinical samples. This strategy was used to investigate spectral profile of clinical samples in different pregnancy time points using FTIR technique. The first aim was to investigate spectral profile of PE related to case and control during both 15 and 20 weeks of gestational age respectively. The second aim was to investigate spectral profile of IUGR related to case and control during both 15 and 20 weeks of pregnancy respectively.

5.3 FTIR spectral profile for PE and IUGR

Chapter 4 presents the results of FTIR analysis, showing differences in intensity of mean spectra between PE ($n=30$) and controls ($n=30$) at 15 and 20 weeks gestational age (Figure 4.3 and 4.6) respectively using mid FTIR range of $4000\text{-}600\text{ cm}^{-1}$. Same observations related to IUGR ($n=30$) compared with controls ($n=30$) at 15 and 20 weeks of pregnancy respectively were detected (Figure 4.9 and 4.12). Significant differences in protein finger print region of $1800\text{-}1000\text{ cm}^{-1}$, amide I C=O stretching at range $\sim 1750\text{-}1560\text{ cm}^{-1}$, amide II N-H bending at range around $1550\text{-}1430\text{ cm}^{-1}$, and amide III (complex vibrations bonds) at range about $(1335\text{-}1220)\text{ cm}^{-1}$. The scores of the PCA plot showed a degree of data

separation between case and control related to PE at 20 weeks of pregnancy can be seen in Figure 4.7A. Similarly, some clusters have been observed also related to patients complicated with IUGR during 15 and 20 weeks respectively compared to normal pregnancies (Figure 4.10A and 4.13A). The alterations in secondary structure of proteins could be associated with a lack in an antioxidant system. These findings agree with a study presented by Raouf and co-workers (2011) who used FTIR analysis of serum to study protein structure of PE at different gestational ages (Raouf *et al.*, 2011).

A positive correlation between early pregnancy dyslipidemia and PE has been suggested (Enquobahrie *et al.*, 2004). The current study in this chapter has shown variation in a range of mid-IR spectral bands mainly corresponds to lipid region at range of 3100-2700 cm^{-1} in PE at 20 week gestational age (Figure 4.8) and in IUGR at 15 week of pregnancy (Figure 4.11) respectively. These bands correspond to C-H stretching bonds for lipids components at range 2960–2950 cm^{-1} (as); CH_3 , 2880–2865 cm^{-1} , and at 2860–2840 cm^{-1} for (as and s), and bending vibration of CH_2 at peak around 1465 cm^{-1} of lipid acyl group, and to ester (C=O) stretching at ~ 1750 – 1700 cm^{-1} . Results in this chapter for complicated pregnancies urine agreed with a previous report suggesting that strongly correlation exists between maternal complicates conditions and abnormal lipid metabolism of pre-eclamptic patients (Aziz and Mahboob, 2007). These spectral changes could be due to the potential impact of protein oxidant damage and lipid peroxidation caused by oxidative stress (Kamath *et al.*, 2006, Maisonneuve *et al.*, 2015, Dickinson *et al.*, 2017). Similar study revealed that women with a history of PE show a significant difference in lipid profile and high susceptibility to lipoprotein oxidation compared with uncomplicated pregnancies (Gratacós *et al.*, 2003, Mittal *et al.*, 2017). Another study which agreed with

our finding was carried out by (Enquobahrie et al., 2004), and found a high correlation between early pregnancy dyslipidemia and PE. Work by Raouf and co-workers (2011) investigating FTIR spectral profiles in PE at different time points of pregnancies using blood sample shows broad agreement with the present study. Correlation was observed in Rauf's work between lipoprotein oxidation and maternal implantation abnormality. Results in chapter 2 also indicated that urothelial markers were observed as differentially expressed proteins, suggesting that the amount of epithelial shedding from different individual during micturition (Kumar *et al.*, 2013) may play a role in the biological variation observed (see Tables S1-S4 in the appendix 3).

5.4 MS process methodology

LC-MS is a powerful method for proteomic analysis, especially when combined with iTRAQ labelling technique (Long *et al.*, 2018). This method enables concurrent identification and quantification of putative proteomic biomarkers; in the case of this study, with the aim of early pregnancy prediction of risk of late pregnancy complications such as PE or/and in utero growth restriction (IUGR). The ultrafiltration protocol that was used in present study has been used elsewhere (Tirumalai *et al.*, 2003). iTRAQ labelling coupled with LC-MS analysis were performed, also used by others in proteomic analyses biomarker study in PE using urine sample (Carty *et al.*, 2011). Data-dependent iTRAQ analysis with LC-MS improves throughput by analysing [differently labelled] pooled samples simultaneously (Carty *et al.*, 2011). This reduces the potential impact of inter-run technical variation. The application of the iTRAQ methodology to urine samples is

dependent on removal of low molecular weight solutes such as sugar, urea, amino acid in organic salts which are freely filtered by the kidney into urine (Hall, 2015); these components could interfere with labelling and subsequent analysis. Data-dependent analysis using the fusion mass spectrometer was performed at present study using a total sample loading of ~400 ng. The MS acquisition program followed the multi-notch MS 3 method developed previously by (McAlister *et al.*, 2014). Unique peptides were only accepted for quantification when all reporter channels were presents with an average S/N > 5. In effect, as there is a low level of contribution of isotopomeric impurities to adjacent peaks (i.e. 114 has a low level of contribution to 115 and so on), this means that in practice no peaks with significant reporter ion intensity will be missed. Peptide and protein identifies were accepted using 0.01 FDR for strict filtering at a 99 % confidence interval. Inter-sample normalisation was performed based on the spiked enolase standard (Cutillas, 2010).

5.4.1 Proteomic biomarkers of PE

PE improves immediately after delivery, but presently no specific intervention, save delivery of often premature infants, is available for PE. Prophylactic treatment with low dose aspirin has, however, been demonstrated to be of clinical use if administered to high-risk patients during early stages of pregnancy (6-16 weeks). (Bujold *et al.*, 2014) This strategy reduces overall risk of PE for these individuals. For primiparous women however the risk of PE is unknown, thus such treatment cannot be prospectively given. Urine is an

ideal biofluid sample for clinical investigation of PE and IUGR due to its ease of collection.

The findings obtained in this thesis can be summarised as the following:

The results outlined in Chapter 2 focused on the use of LC-MS for the proteomics analysis of urine, using iTRAQ methodology. This strategy was implemented for the investigation of peptides having a molecular weight fraction >10 kDa. Samples were analysed at different gestational age time points at 15 and 20 weeks for PE and IUGR. Results highlighted 988 unique peptides from a total of 315 proteins in each 8-plex labelling sample at 99 % confidence interval with 0.01 false discovery rate cut-off, of which 173 /826 (protein / peptides) proteins have sufficient reporter ion signal for quantification related to PE and IUGR at each week respectively (see Tables S1-S4 in the appendix 3). Proteins with < 2 peptides are excluded from identification / quantification analysis, due to reduced confidence of protein/peptide identification; these data are routinely excluded from ratiometric determinations made by others (Evans *et al.*, 2012).

Identified proteins were subjected to relative quantification, with differentially-expressed proteins being associated with \log_2 fold change of abundance ratio (cases: control) during 15 and 20 weeks of PE and IUGR respectively compared with uncomplicated pregnancies (Figures 2.7 - 2.9 for PE; Figures 2.11 - 2.13 for IUGR). Quantitatively, all data were normalised within a known amount of internal standard of enolase which \log_2 fold change ratio equal to zero (supplementary Tables S1 - S4 in Appendix 3, Figure (2.7 -2.9) for PE and in Figure (2.11 - 2.13) for IUGR respectively).

In PE, dysregulated proteins (n=173) that shown highly significant differences ($P < 0.01$) in protein profile at 95 % confidence interval using one sample Wilcoxon test (Figure 2.7 and 2.8 at 15 and 20 weeks' gestation respectively). Results revealed about 26.7 % (n=46)

proteins were upregulated, and 73.3 % (n= 126) proteins were downregulated at 15 weeks' gestation in PE cases compared with uncomplicated pregnancies (Table S 1 in appendix 3). At 20 weeks' gestation in pregnancies complicated by PE, about 60.5 % (n=104) proteins were upregulated and 39.5 % (n = 68) proteins were downregulated compared with uncomplicated pregnancies (Table S 2). Quantitative data in Figure 2.9 showed highly significant differences ($P < 0.01$) between 15 and 20 weeks of PE related to \log_2 fold change abundance ratio (case / control) at 95 % confidence interval using Mann-Whitney test. In a similar study of urinary proteomic analysis using iTRAQ labelling work by (Chen *et al.*, 2011) 113 proteins were identified. These proteins were expressed differentially between complicated hypertension (PE and hypertension pregnancies) and normal pregnancy. 31/113 of these proteins revealed statistically significant differences between these study groups.

Outlier data in these analysis are represented highly dysregulated proteins (up / down) compared with control for PE are showed in Table 2.3 and 2.4. 16 proteins were identified at outlier are showed highly dysregulated proteins compared with uncomplicated pregnancies using one sample Wilcoxon test are showed at Figure 2.7, of which 50 % (n=8) proteins are upregulated and downregulated respectively for PE during 15 week of pregnancy (Table 2.3). For 20 week of PE, 7 proteins were upregulated, and 8 proteins are downregulated were identified in Table 2.4. However, in present study, the putative proteomic biomarkers which showed highly dysregulated profiles compared with the control group of uncomplicated pregnancies have found in previous studies, which are listed in Table 2.7 (section 2.9). Three proteins were identified as having very similar profile changes at both 15 and 20 weeks in pregnancies subsequently complicated by PE

(fibrinogen beta chain, haptoglobin and carboxypeptidase), highlighted in Table 2.3 and 2.4 and shown in Figure 2.10. Similar proteins identified in previous studies are highlighted in Table 2.7.

Among putative biomarkers that showing highly dysregulated expression profiles associates with complicated pregnancies such as PE compared with uncomplicated pregnancies (Table 2.7), for instant, alpha-2-macroglobulin, haptoglobin and roundabout were identified. Alpha-2-macroglobulin showed a \log_2 fold change abundances ratio change of +3.5 (case: control) at 15 weeks gestation in PE subjects compared with uncomplicated pregnancies (control). Alpha-2-macroglobulin has antiplasmin activity; the increased level of protein observed in preeclampsia may explain the intravascular coagulation observed in this condition (Horne *et al.*, 1970). Similar findings have been previously observed in serum samples (Auer *et al.*, 2010, Blumenstein *et al.*, 2009, D'Silva *et al.*, 2018). Elsewhere, studies of proteomic biomarkers in PE also agreed with this finding using urine samples (Chen *et al.*, 2011, Carty *et al.*, 2011).

Haptoglobin is another biomarker shown as highly dysregulated in protein expression, with a \log_2 fold change abundances ratio of +2.53 (case: control) for PE cases compared with uncomplicated pregnancies (control) at 15 weeks' gestation. Many other previous studies have reported similar result using different types of samples e.g. urine and / or serum to study proteomic biomarkers in PE (Auer *et al.*, 2010, Blumenstein *et al.*, 2009, Chen *et al.*, 2011, Carty *et al.*, 2011, D'Silva *et al.*, 2018, Little *et al.*, 2010) (see Table 2.7). Roundabout (ROBO4) showed downregulation (-3.19) in protein expression of PE at 15 weeks gestational age compared with normotensive pregnancies (Table 2.3). This protein has previously been found in developing placenta and embryonic tissues (Park *et al.*,

2003), and has been shown to inhibit endothelial cell migration (Park *et al.*, 2003). Downregulation in the protein expression could be an indication of impaired fetal development observed in later pregnancy. Similar findings have been previously identified in urine (Chen *et al.*, 2011) or in blood and serum (Nagalla *et al.*, 2010, Myers *et al.*, 2013) (see Table 2.7).

5.4.2 Proteomic biomarkers of IUGR

In parallel with the urine sample analysis in PE, samples were evaluated with respect to patients suffering from IUGR. This dataset revealed several interesting potential biomarkers. At the 15 week gestation time point, no statistically significant quantitative differences were observed in the 173 identified proteins, however 16 outlier observations were tentatively assigned as differentially expressed proteins (n=173) were identified as 43 % (n=74) upregulated, and 98 proteins (57 %) downregulated in Table S3 (Appendix 3). These dysregulated proteins showed a median expression fold change value of (-0.077), and max.-to-min. values of (2.7-(-1.639)) of 95 % interval using a one sample Wilcoxon test (Figure 2.11). Of the 16 outlier proteins, eight were categorised as highly dysregulated (up/ down) compared with uncomplicated pregnancies (Figure 2.11 and Table 2.5).

At 20 weeks' gestation in participants who went on to develop IUGR, highly statistically significant differences were observed ($P < 0.01$) using one sample Wilcoxon test. Median fold-change ratio at (-0.353) and max. and min. values at (1.549-(-2.44)) were determined for the 173 proteins identified (Figure 2.12). 27.9 % (n=48) were upregulated, while

around 72.1 % (n=124) were downregulated compared with uncomplicated pregnancies (Table S4, Appendix 3). Additionally, quantitative data (Figure 2.13) showed highly significant differences ($P < 0.01$) between 15 and 20 weeks of IUGR using a Mann-Whitney test. Similar work using iTRAQ labelling coupled with MS analysis identified proteomic biomarkers in IUGR using serum (Auer *et al.*, 2010). There, 166 proteins were identified, of which 31/128 identified proteins/ peptides showed expression changes in IUGR (14 up and 17 down-regulated) (Auer *et al.*, 2010).

A total of 74 peptides were revealed as outlier observations from 16 proteins. 50 % (n = 8) of these proteins are upregulated and eight downregulated respectively in samples taken at 15 weeks' gestation for of IUGR compared with uncomplicated pregnancies; these proteins are shown in Table 2.5. At 20 weeks' gestation, 16 / 60 (protein/peptides) were observed as being highly dysregulated in IUGR compared to uncomplicated pregnancy; 8 proteins are upregulated and 8 are downregulated (Table 2.6). Similar identified proteins were observed in previous studies, summarised Table 2.8, using different sample types such as urine (Carty *et al.*, 2011, Chen *et al.*, 2011), serum (Auer *et al.*, 2010), plasma (Blumenstein *et al.*, 2009), amniotic fluid (Tsangaris *et al.*, 2006) and placental cells (Mandò *et al.*, 2016, Xin *et al.*, 2012).

Seven of the dysregulated proteins identified as outlier observations showed highly similar profiles at both 15 and 20 weeks' gestation. These were zinc-alpha-2-glycoprotein, leucine-rich alpha-2-glycoprotein, alpha-1-acid glycoprotein 1, alpha-1-acid glycoprotein 2, histone H4, CD44 antigen and lysosome-associated membrane glycoprotein 2, highlighted in Table 2.5 and 2.6, and shown in Figure 2.14. Similar findings have been made by various previous studies, again which are highlighted in Table 2.8.

Among proteomic biomarkers which have shown dysregulation in their profile associated with pregnancies complicated by IUGR which are shown in Table 2.8., for example, Zinc-alpha-2-glycoprotein, \log_2 fold change value has upregulated at (2.707) in Table 2.5. Zinc-alpha-2-glycoprotein, a protein play various important functions in the human body, including fertilization (Ohkubo *et al.*, 1990) and lipid mobilization (Morse *et al.*, 2017). Zinc-alpha-2-glycoprotein stimulates lipid degradation in adipocytes and causes the extensive fat losses associated with some advanced cancers (Hirai *et al.*, 1998, Zhang *et al.*, 2018); this protein could bind with polyunsaturated fatty acids. Upregulation of zinc-alpha-2-glycoprotein may therefore be a key regulator of fetal growth in later pregnancy. The protein has been identified by others using serum or urine samples to study proteomic biomarkers in complicated pregnancies such as IUGR and PE (Auer *et al.*, 2010, Blumenstein *et al.*, 2009)(Chen *et al.*, 2011),(Carty *et al.*, 2011).

Leucine-rich alpha-2-glycoprotein-1 (Lrg1), another putative IUGR biomarker was upregulated (+1.79 \log_2 fold-changes, Table 2.5). Leucine-rich alpha-2-glycoprotein-1 (Lrg1) is an emerging biomarker for angiogenesis (Lio *et al.*, 2018). Angiogenesis is the process of new blood vessel formation, and is involved in the development and progression of several diseases, including cancers, proliferative diabetic retinopathy and chronic wound healing (Carmeliet, 2005). Vascular endothelial growth factors (VEGFs) and their receptors play a crucial role in angiogenesis (Carmeliet *et al.*, 1996). Similar finding was detected using serum sample to identify protein leucine-rich alpha-2-glycoprotein as a putative biomarker of recombinant human growth hormone abuse (Kay *et al.*, 2009). Another study has also identified similar protein using amniotic fluid to predict abnormal fetal biomarkers has been previously performed (Tsangaris *et al.*, 2006).

The presence of urothelial markers as differentially expressed proteins is an indication that the amount of epithelial shedding from different individuals during micturition (urination) is playing a role in the observed biological variation; therefore, this influence needs to be excluded before further exploitation of these putative biomarkers is possible.

5.5 Conclusions

This work has presented a proof of concept investigation using a rat urine model for optimisation experiment design to measure volume sensitivity and built up a standard operating protocol for the further study of clinical samples. We conclude that spectra using 5 μ L volume showed appropriate sensitivity for the mid FTIR region. Furthermore, the ultrafiltration protocol displayed robust and simple preparation method for protein identification at urine sample. However, FTIR instrumental limitation is that it is somewhat slow and presents long sampling times due to drying of samples (required for reproducibility). This makes it is poorly-suited to measure larger sample sizes. Therefore, a higher throughput and more rapid spectroscopic instrument would be needed, such as FT-IMR to achieve this objective.

LC-MS coupling with iTRAQ technique introducing high throughput analysis enabling to identified and quantified proteomic biomarkers for both PE and IUGR simultaneously. This study identified 28 proteomic biomarkers that have showed highly dysregulated proteins (up /down) for PE shown in Table 2.7, and 24 possible biomarkers for IUGR compared with uncomplicated pregnancies respectively, shown in Table 2.8. The fetoplacental markers of hypoxia and vascular dysfunction into urine via the maternal

blood in a manner enabling quantitative detection of their variation is a strong indicator that this strategy has merit for determination of potential markers for malfunction of the fetoplacental unit. The limitations of this study are, however, clear – we have yet to interrogate these markers using quantitative methods in an independent cohort of women who have the same conditions, or to investigate putative markers found in individual samples in (preferably independent) individuals. Independent confirmation of these markers using targeted mass spectrometric assay and/or biochemical assay methods such as ELISA is essential before these markers can be taken forward for clinical translation.

5.6 Future work

Mortality and morbidity associated with both maternal hypertension and children born following pregnancies complicated by pre-eclampsia and its consequences are a major burden (Villar et al., 2001, Van Lerberghe et al., 2005). The World Health Organization estimates that around 15 % of maternal deaths are directly associated with preeclampsia and eclampsia in low- and middle-income countries (Van Lerberghe *et al.*, 2005), and over recent decades the condition remains the major causes of maternal death in the UK (Lewis, 2012, Hibbard and Milner, 1994). The Confidential Enquiry into Stillbirths and Deaths in Infancy (CESDI) 5th annual report (Sidebotham, 1998) revealed 1 in 6 stillbirths and around 1 in 6 neonatal deaths occurred due to gestational problems associated with maternal hypertension; gestational hypertension was found to be responsible for ~20% of special care costs at neonatal units in the UK (Sidebotham, 1998). The morbidity burden

of prematurity does not cease at the neonatal stage; children born early as a result of PE have increased rates of hypertension, heart disorders, and disability in adulthood (Meyer and Zhang, 2007, Barker et al., 1989). Significant effort has been invested in fetal growth restriction studies; however no unifying hypothesis has emerged to explain its occurrence (Lin and Santolaya-Forgas, 1998). In addition to adverse health consequences, PE is also has economic burden due to the medical services needed to take care of pregnant and postnatal women and their babies, who are often born preterm (Obstetricians and Gynecologists, 2013, Medicine, 2007). Stevens *et al.*, (2017) documented the short-term medical costs associated with PE. State hospital discharge data, birth document data, economic insurance claims data, and nationally representative Healthcare Cost and Utilization Project (HCUP) data were interrogated to allocate nationally representative estimates of the additional cost of care for women with PE and their infants compared to women experiencing uncomplicated birth. This cost was estimated at \$2.18 billion USD in 2012, including \$1.03 billion in costs of caring for mothers, and \$1.15 billion for newborns (Stevens *et al.*, 2017). Elsewhere, study reported by (Fox *et al.*, 2017), estimated the cost of PE in a socialised medical system using secondary data from the SCOPE study. The Irish cohort in Cork was recruited between November 2008 and February 2011. Within the cohort of 1774 women, 68 developed PE (3.8%); these cases were compared with 171 randomly-selected controls. Pregnant women affected by PE had higher service use levels (Fox *et al.*, 2017). The average service cost for women complicated by PE was €5243 per case vs. ~€2452 per case for women without complications. The national cost of PE is thus estimated as between €6.5 and €9.1 million annually using prevalence rate 5-7% (Fox *et al.*, 2017). Ongoing effort to establish early pregnancy markers for this condition (and to therefore reduce the impact of PE on women and the healthcare sector)

is thus a valid direction of future investigation. The following future plans are summarised below:

- ❖ The development of targeted assays for pregnancy under particular complications can be assisted by the use of detailed urinary protein analysis. The influence of biological variability necessitates further study of a larger, independent cohort using more targeted methods, such as immunoassay or mass spectrometric multiple reaction monitoring.
- ❖ A urinary metabolomics study for PE and IUGR respectively during (15 and 20) weeks gestational age at <10 kDa molecular weight cut off using MS and FTIR techniques.
- ❖ Genetic study such as DNA and RNA at (15 and 20) weeks of gestational age for each PE & IUGR disorders respectively using cells derived from urine samples.
- ❖ Raman spectroscopic analysis for PE and IUGR disorders at (15 and 20) weeks of gestational using urine sample at molecular weight higher and lower than 10 kDa cut off respectively.
- ❖ NMR spectroscopic study for PE and IUGR during both (15 and 20) weeks of gestational age for <10 kDa molecular weight cut off using urine sample.

References:

- ABIAN, J., OOSTERKAMP, A. & GELPI, E. 1999. Comparison of conventional, narrow-bore and capillary liquid chromatography/mass spectrometry for electrospray ionization mass spectrometry: practical considerations. *Journal of Mass Spectrometry*, 34, 244-254.
- AGGARWAL, K., CHOE, L. H. & LEE, K. H. 2005. Quantitative analysis of protein expression using amine-specific isobaric tags in *Escherichia coli* cells expressing rhsA elements. *Proteomics*, 5, 2297-2308.
- ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. & WALTER, P. 2002. *Molecular biology of the Cell* ((New York: Garland Science Taylor & Francis Group). Inc., ZDA.
- ALVAREZ, M. M., KHOURY, J. T., SCHAAFF, T. G., SHAFIGULLIN, M. N., VEZMAR, I. & WHETTEN, R. L. 1997. Optical absorption spectra of nanocrystal gold molecules. *The Journal of Physical Chemistry B*, 101, 3706-3712.
- ALVINO, G., COZZI, V., RADAELLI, T., ORTEGA, H., HERRERA, E. and CETIN, I. 2008. Maternal and fetal fatty acid profile in normal and intrauterine growth restriction pregnancies with and without preeclampsia. *Pediatric Research*, 64, 615.
- AMBIHAPATHY, K., YALCIN, T., LEUNG, H. W. & HARRISON, A. G. 1997. Pathways to immonium ions in the fragmentation of protonated peptides. *Journal of Mass Spectrometry*, 32, 209-215.
- AUER, J., CAMOIN, L., GUILLONNEAU, F., RIGOURD, V., CHELBI, S. T., LEDUC, M., LAPARRE, J., MIGNOT, T. M. & VAIMAN, D. 2010. Serum profile in preeclampsia and intra-uterine growth restriction revealed by iTRAQ technology. *J Proteomics*, 73, 1004-17.
- AZIZ, R. and MAHBOOB, T. 2007. Pre-eclampsia and lipid profile. *Pakistan Journal of Medical Sciences*, 23, 751.
- BAKER, M. J., TREVISAN, J., BASSAN, P., BHARGAVA, R., BUTLER, H. J., DORLING, K. M., FIELDEN, P. R., FOGARTY, S. W., FULLWOOD, N. J. and HEYS, K. A. 2014. Using Fourier transform IR spectroscopy to analyze biological materials. *Nature Protocols*, 9, 1771-1791.
- BALCERZAK, M. 2003. An overview of analytical applications of time of flight-mass spectrometric (TOF-MS) analyzers and an inductively coupled plasma-TOF-MS technique. *Analytical Sciences*, 19, 979-989.

- BANERJI, S., NI, J., WANG, S.-X., CLASPER, S., SU, J., TAMMI, R., JONES, M. & JACKSON, D. G. 1999. LYVE-1, a new homologue of the CD44 glycoprotein, is a lymph-specific receptor for hyaluronan. *The Journal of Cell Biology*, 144, 789-801.
- BARH, D., BLUM, K. & MADIGAN, M. A. 2011. *OMICS: biomedical perspectives and applications*, CRC Press.
- BARKER, D. J., OSMOND, C., WINTER, P., MARGETTS, B. and SIMMONDS, S. J. 1989. Weight in infancy and death from ischaemic heart disease. *The Lancet*, 334, 577-580.
- BATEMAN, R. H. & HOYES, J. B. 2000. Methods and apparatus for tandem mass spectrometry. Google Patents.
- BATTAGLIA, F. C. and LUBCHENCO, L. O. 1967. A practical classification of newborn infants by weight and gestational age. *The Journal of Pediatrics*, 71, 159-163.
- BERGMAN, N. & BERGQUIST, J. 2014. Recent developments in proteomic methods and disease biomarkers. *Analyst*, 139, 3836-3851.
- BLACKSTOCK, W. P. & WEIR, M. P. 1999. Proteomics: quantitative and physical mapping of cellular proteins. *Trends in Biotechnology*, 17, 121-127.
- BLANKLEY, R. T., FISHER, C., WESTWOOD, M., NORTH, R., BAKER, P. N., WALKER, M. J., WILLIAMSON, A., WHETTON, A. D., LIN, W. C., MCCOWAN, L., ROBERTS, C. T., COOPER, G. J. S., UNWIN, R. D. & MYERS, J. E. 2013. A Label-free Selected Reaction Monitoring Workflow Identifies a Subset of Pregnancy Specific Glycoproteins as Potential Predictive Markers of Early-onset Pre-eclampsia. *Molecular & Cellular Proteomics*, 12, 3148-3159.
- BLUMENSTEIN, M., MCMASTER, M. T., BLACK, M. A., WU, S., PRAKASH, R., COONEY, J., MCCOWAN, L. M., COOPER, G. J. & NORTH, R. A. 2009. A proteomic approach identifies early pregnancy biomarkers for preeclampsia: novel linkages between a predisposition to preeclampsia and cardiovascular disease. *Proteomics*, 9, 2929-2945.
- BONNIER, F., BAKER, M. J. & BYRNE, H. J. 2014. Vibrational spectroscopic analysis of body fluids: avoiding molecular contamination using centrifugal filtration. *Analytical Methods*, 6, 5155-5160.
- BRANDENBURG, K. & SEYDEL, U. 2002. Vibrational spectroscopy of carbohydrates and glycoconjugates. *Handbook of Vibrational Spectroscopy*.
- BUHIMSCHI, C. S., NORWITZ, E. R., FUNAI, E., RICHMAN, S., GULLER, S., LOCKWOOD, C. J. & BUHIMSCHI, I. A. 2005. Urinary angiogenic factors cluster hypertensive disorders and identify women with severe preeclampsia. *Am J Obstet Gynecol*, 192, 734-41.
- BUHIMSCHI, I. A., ZHAO, G. M., FUNAI, E. F., HARRIS, N., SASSON, I. E., BERNSTEIN, I. M., SAADE, G. R. & BUHIMSCHI, C. S. 2008. Proteomic profiling of urine identifies

- specific fragments of SERPINA1 and albumin as biomarkers of preeclampsia. *American Journal of Obstetrics and Gynecology*, 199, 551. e1-551. e16.
- BUJOLD, E., ROBERGE, S. & NICOLAIDES, K. H. 2014. Low-dose aspirin for prevention of adverse outcomes related to abnormal placentation. *Prenat Diagn*, 34, 642-8.
- BURTON, G., YUNG, H.-W., CINDROVA-DAVIES, T. & CHARNOCK-JONES, D. 2009. Placental endoplasmic reticulum stress and oxidative stress in the pathophysiology of unexplained intrauterine growth restriction and early onset preeclampsia. *Placenta*, 30, 43-48.
- CABRAL, E., SOARES, H., GUIMARÃES, H., VITORINO, R., FERREIRA, R. & HENRIQUES-COELHO, T. 2017. Prediction of cardiovascular risk in preterm neonates through urinary proteomics: An exploratory study. *Porto Biomedical Journal*, 2, 287-292.
- CAKMAK, G., TOGAN, I. and SEVERCAN, F. 2006. 17 β -Estradiol induced compositional, structural and functional changes in rainbow trout liver, revealed by FT-IR spectroscopy: a comparative study with nonylphenol. *Aquatic Toxicology*, 77, 53-63.
- CAMM, E. J., HANSELL, J. A., KANE, A. D., HERRERA, E. A., LEWIS, C., WONG, S., MORRELL, N. W. and GIUSSANI, D. A. 2010. Partial contributions of developmental hypoxia and undernutrition to prenatal alterations in somatic growth and cardiovascular structure and function. *American Journal of Obstetrics and Gynecology*, 203, 495. e24-495.
- CARMELIET, P. 2005. Angiogenesis in life, disease and medicine. *Nature*, 438, 932.
- CARMELIET, P., FERREIRA, V., BREIER, G., POLLEFEYT, S., KIECKENS, L., GERTSENSTEIN, M., FAHRIG, M., VANDENHOECK, A., HARPAL, K. & EBERHARDT, C. 1996. Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature*, 380, 435.
- CARTY, D. M. 2012. Pre-eclampsia: early prediction and long-term consequences. University of Glasgow.
- CARTY, D. M., SIWY, J., BRENNAND, J. E., ZURBIG, P., MULLEN, W., FRANKE, J., MCCULLOCH, J. W., ROBERTS, C. T., NORTH, R. A., CHAPPELL, L. C., MISCHAK, H., POSTON, L., DOMINICZAK, A. F. & DELLES, C. 2011. Urinary proteomics for prediction of preeclampsia. *Hypertension*, 57, 561-9.
- CHAHROUR, O., COBICE, D. & MALONE, J. 2015. Stable isotope labelling methods in mass spectrometry-based quantitative proteomics. *Journal of Pharmaceutical and Biomedical Analysis*.
- CHEN, G., ZHANG, Y., JIN, X., ZHANG, L., ZHOU, Y., NIU, J., CHEN, J. & GU, Y. 2011. Urinary proteomics analysis for renal injury in hypertensive disorders of pregnancy with iTRAQ labeling and LC-MS/MS. *Proteomics Clin Appl*, 5, 300-10.
- CLAYDON, A. J. & BEYNON, R. J. 2011. Protein turnover methods in single-celled organisms: dynamic SILAC. *Yeast Systems Biology*. Springer.

- COLTHUP, N. 2012. Introduction to infrared and Raman spectroscopy, Elsevier.
- CONDE-AGUDELO, A., PAPAGEORGHIU, A., KENNEDY, S. and VILLAR, J. 2013. Novel biomarkers for predicting intrauterine growth restriction: a systematic review and meta-analysis. *BJOG: An International Journal of Obstetrics and Gynaecology*, 120, 681-694.
- CONDE-AGUDELO, A., VILLAR, J. & LINDHEIMER, M. 2004. World Health Organization systematic review of screening tests for preeclampsia. *Obstet Gynecol*, 104, 1367-91.
- COTTRELL, J. S. & LONDON, U. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20, 3551-3567.
- ÇORBACI, C. & UÇAR, F. B. 2018. Purification, characterization and in vivo biocontrol efficiency of killer toxins from *Debaryomyces hansenii* strains. *International Journal of Biological Macromolecules*, 119, 1077-1082.
- COX, T. M. 1994. Aldolase B and fructose intolerance. *The FASEB journal*, 8, 62-71.
- CUTILLAS, P. R. 2010. Analysis of peptides in biological fluids by LC-MS/MS. *LC-MS/MS in Proteomics*. Springer.
- DASS, C. & BRODBELT, J. S. 2001. Principles and practice of biological mass spectrometry. *Applied Spectroscopy*, 55, 296.
- DAVIS, E. F., LEWANDOWSKI, A. J. & LEESON, P. 2012. Cardiac Dysfunction and Preeclampsia Can Imaging Give Clues to Mechanism? *Circulation: Cardiovascular Imaging*, 5, 691-692.
- DAWSON, P. H. 2013. Quadrupole mass spectrometry and its applications, Elsevier.
- DE HOFFMANN, E. 1996. Tandem mass spectrometry: a primer. *Journal of Mass Spectrometry*, 31, 129-137.
- DICKINSON, E., ARNOLD, J. R. and FISHER, J. 2017. Determination of glucose exchange rates and permeability of erythrocyte membrane in preeclampsia and subsequent oxidative stress-related protein damage using dynamic-19 F-NMR. *Journal of Biomolecular NMR*, 67, 145-156.
- DIKBAS, L., YAPCA, O. E., DIKBAS, N. and GUNDOGDU, C. 2017. Paraoxonase-2 and paraoxonase-3: comparison of mRNA expressions in the placentae of unexplained intrauterine growth restricted and noncomplicated pregnancies. *The Journal of Maternal-Fetal and Neonatal Medicine*, 30, 1200-1206.
- DOHERTY, M. K., WHITEHEAD, C., MCCORMACK, H., GASKELL, S. J. & BEYNON, R. J. 2005. Proteome dynamics in complex organisms: using stable isotopes to monitor individual protein turnover rates. *Proteomics*, 5, 522-533.

- DOUGLAS, D. J., FRANK, A. J. & MAO, D. 2005. Linear ion traps in mass spectrometry. *Mass Spectrometry Reviews*, 24, 1-29.
- D'SILVA, A. M., HYETT, J. A. & COORSEN, J. R. 2018. Proteomic analysis of first trimester maternal serum to identify candidate biomarkers potentially predictive of spontaneous preterm birth. *Journal of Proteomics*, 178, 31-42.
- DUKOR, R. K. 2002. Vibrational spectroscopy in the detection of cancer. *Handbook of Vibrational Spectroscopy*.
- DULEY, L. The global impact of pre-eclampsia and eclampsia. *Seminars in perinatology*, 2009. Elsevier, 130-137.
- ENQUOBAHRIE, D. A., WILLIAMS, M. A., BUTLER, C. L., FREDERICK, I. O., MILLER, R. S. and LUTHY, D. A. 2004. Maternal plasma lipid concentrations in early pregnancy and risk of preeclampsia. *American Journal of Hypertension*, 17, 574-581.
- ELLISON, J. & HOOD, L. 1982. Linkage and sequence homology of two human immunoglobulin gamma heavy chain constant region genes. *Proceedings of the National Academy of Sciences*, 79, 1984-1988.
- ENQUOBAHRIE, D. A., QIU, C., MUHIE, S. Y. & WILLIAMS, M. A. 2011. Maternal peripheral blood gene expression in early pregnancy and preeclampsia. *International Journal of Molecular Epidemiology and Genetics*, 2, 78.
- EVANS, C., NOIREL, J., OW, S. Y., SALIM, M., PEREIRA-MEDRANO, A. G., COUTO, N., PANDHAL, J., SMITH, D., PHAM, T. K. & KARUNAKARAN, E. 2012. An insight into iTRAQ: where do we stand now? *Analytical and Bioanalytical Chemistry*, 404, 1011-1027.
- FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246, 64-71.
- FEULNER, G., GRAY, J., KIRSCHMAN, J., LEHNER, A., SADOSKY, A., VLAZNY, D., ZHANG, J., ZHAO, S. & HILL, C. 1990. Structure of the rhsA locus from *Escherichia coli* K-12 and comparison of rhsA with other members of the rhs multigene family. *Journal of Bacteriology*, 172, 446-456.
- FLISER, D., NOVAK, J., THONGBOONKERD, V., ARGILES, A., JANKOWSKI, V., GIROLAMI, M. A., JANKOWSKI, J. & MISCHAK, H. 2007. Advances in urinary proteome analysis and biomarker discovery. *J Am Soc Nephrol*, 18, 1057-71.
- FOX, A., MCHUGH, S., BROWNE, J., KENNY, L. C., FITZGERALD, A., KHASHAN, A. S., DEMPSEY, E., FAHY, C., O'NEILL, C. & KEARNEY, P. M. 2017. Estimating the cost of preeclampsia in the healthcare system: cross-sectional study using data from SCOPE study (Screening for Pregnancy End Points). *Hypertension*, 70, 1243-1249.
- FRAILE, M., BLANCO-MELGAR, M., MUÑOZ, R. M. N., LÓPEZ-RODRÍGUEZ, G., GALLEGONICASIO, J. and CARMONA, P. 2003. Structure and interactions of albumin-lipid

- systems as studied by infrared spectroscopy. *Journal of Molecular Structure*, 651, 231-236.
- FULLER, H. & MORRIS, G. 2012. Quantitative proteomics using iTRAQ labeling and mass spectrometry. *Integrative Proteomics, InTech, Croatia*, 347-362.
- FUWA, K. & VALLE, B. 1963. The Physical Basis of Analytical Atomic Absorption Spectrometry. The Pertinence of the Beer-Lambert Law. *Analytical Chemistry*, 35, 942-946.
- GASKELL, S. J. 1997. Electrospray: principles and practice. *Journal of Mass Spectrometry*, 32, 677-688.
- GAUGLITZ, G. & VO-DINH, T. 2014. *Handbook of Spectroscopy, 4 Volume Set*, Wiley-VCH Verlag GmbH & Company.
- GHAZALPOUR, A., BENNETT, B., PETYUK, V. A., OROZCO, L., HAGOPIAN, R., MUNGRUE, I. N., FARBER, C. R., SINSHEIMER, J., KANG, H. M. & FURLOTTE, N. 2011. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet*, 7, e1001393.
- GILSTRAP, L. & RAMIN, S. 2002. Diagnosis and management of preeclampsia and eclampsia. *ACOG Pract Bull*, 33, 1-9.
- GOLDENBERG, R. L. & ROUSE, D. J. 1998. Prevention of premature birth. *N Engl J Med*, 339, 313-20.
- GOLDWASSER, P. and FELDMAN, J. 1997. Association of serum albumin and mortality risk. *Journal of Clinical Epidemiology*, 50, 693-703.
- GOUGH, N. R. & FAMBROUGH, D. M. 1997. Different steady state subcellular distributions of the three splice variants of lysosome-associated membrane protein LAMP-2 are determined largely by the COOH-terminal amino acid residue. *The Journal of Cell Biology*, 137, 1161-1169.
- GRATACÓS, E., CASALS, E., GÓMEZ, O., LLURBA, E., MERCADER, I., CARARACH, V. and CABERO, L. 2003. Increased susceptibility to low density lipoprotein oxidation in women with a history of pre-eclampsia. *BJOG: An International Journal of Obstetrics and Gynaecology*, 110, 400-404.
- GRIFFITHS, P. R. & DE HASETH, J. A. 2007. *Fourier transform infrared spectrometry*, John Wiley & Sons.
- HÄGGLUND, P., BUNKENBORG, J., MAEDA, K., FINNIE, C. & SVENSSON, B. 2014. Identification of Thioredoxin Target Disulfides Using Isotope-Coded Affinity Tags. *Plant Proteomics*. Springer.
- HALLIWELL, B. 1990. How to characterize a biological antioxidant. *Free Radical Research Communications*, 9, 1-32.

- HAN, X., ASLANIAN, A. & YATES, J. R. 2008. Mass spectrometry for proteomics. *Current Opinion in Chemical Biology*, 12, 483-490.
- HAN, Y., LU, C., ZHANG, K., TIAN, S., FAN, E., CHEN, L., HE, X. & ZHANG, Y. 2015. Quantitative characterization of histone post-translational modifications using a stable isotope dimethyl-labeling strategy. *Analytical Methods*, 7, 3779-3785.
- HARRIGAN, G. G., LAPLANTE, R. H., COSMA, G. N., COCKERELL, G., GOODACRE, R., MADDOX, J. F., LUYENDYK, J. P., GANEY, P. E. & ROTH, R. A. 2004. Application of high-throughput Fourier-transform infrared spectroscopy in toxicology studies: contribution to a study on the development of an animal model for idiosyncratic toxicity. *Toxicology Letters*, 146, 197-205.
- HART, S. R., KENNY, L. C., MYERS, J. E. & BAKER, P. N. 2015. Electron transfer dissociation of native peptides facilitates enhanced identification of urinary peptides. *International Journal of Mass Spectrometry*, 391, 41-46.
- HART, S., KENNY, L., MYERS, J. & BAKER, P. 2015. [82-OR]: Novel proteomic method to identify urinary proteins which may predict pre-eclampsia: Proof of concept. *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*, 5, 44-45.
- HALL, J. E. 2015. *Guyton and Hall textbook of medical physiology e-Book*, Elsevier Health Sciences.
- HART-SMITH, G., LAMMENS, M., DU PREZ, F. E., GUILHAUS, M. & BARNER-KOWOLLIK, C. 2009. ATRP poly (acrylate) star formation: A comparative study between MALDI and ESI mass spectrometry. *Polymer*, 50, 1986-2000.
- HIRAI, K., HUSSEY, H. J., BARBER, M. D., PRICE, S. A. & TISDALE, M. J. 1998. Biological evaluation of a lipid-mobilizing factor isolated from the urine of cancer patients. *Cancer Research*, 58, 2359-2365.
- HOFKER, M. H., WALTER, M. A. & COX, D. W. 1989. Complete physical map of the human immunoglobulin heavy chain constant region gene complex. *Proceedings of The National Academy of Sciences*, 86, 5567-5571.
- HOŞAFCI, G., KLEIN, O., OREMEK, G. & MÄNTELE, W. 2007. Clinical chemistry without reagents? An infrared spectroscopic technique for determination of clinically relevant constituents of body fluids. *Analytical and Bioanalytical Chemistry*, 387, 1815.
- HIBBARD, B. and MILNER, D. 1994. Reports on confidential enquiries into maternal deaths: an audit of previous recommendations. *Health Trends*, 26, 26-28.
- HIRAI, K., HUSSEY, H. J., BARBER, M. D., PRICE, S. A. & TISDALE, M. J. 1998. Biological evaluation of a lipid-mobilizing factor isolated from the urine of cancer patients. *Cancer Research*, 58, 2359-2365.
- HORNE, C., HOWIE, P. & GOUDIE, R. 1970. Serum alpha2-macroglobulin, transferrin, albumin, and IgG levels in preeclampsia. *Journal of Clinical Pathology*, 23, 514-516.

- HOŞAFÇI, G., KLEIN, O., OREMEK, G. & MÄNTELE, W. 2007. Clinical chemistry without reagents? An infrared spectroscopic technique for determination of clinically relevant constituents of body fluids. *Analytical and Bioanalytical Chemistry*, 387, 1815.
- HRACSKO, Z., ORVOS, H., NOVAK, Z., PAL, A. and VARGA, I. S. 2008. Evaluation of oxidative stress markers in neonates with intra-uterine growth retardation. *Redox Report*, 13, 11-16.
- HU, Q., NOLL, R. J., LI, H., MAKAROV, A., HARDMAN, M. & GRAHAM COOKS, R. 2005. The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry*, 40, 430-443.
- HUBEL, C. A. 1999. Oxidative stress in the pathogenesis of preeclampsia. *Proceedings of the Society for Experimental Biology and Medicine*, 222, 222-235.
- HUBEL, C., ROBERTS, J., TAYLOR, R., MUSCI, T., ROGERS, G. and MCLAUGHLIN, M. 1990. Lipid peroxidation in pregnancy: New perspectives on preeclampsia. *International Journal of Gynecology and Obstetrics*, 32, 188-188.
- HUMINIECKI, L., GORN, M., SUCHTING, S., POULSOM, R. & BICKNELL, R. 2002. Magic roundabout is a new member of the roundabout receptor family that is endothelial specific and expressed at sites of active angiogenesis. *Genomics*, 79, 547-552.
- IRGENS, H. U., ROBERTS, J. M., REISÆTER, L., IRGENS, L. M. & LIE, R. T. 2001. Long term mortality of mothers and fathers after pre-eclampsia: population based cohort study. *Pre-eclampsia and cardiovascular disease later in life: who is at risk? Bmj*, 323, 1213-1217.
- JACKSON, D. W., SCISCIONE, A., HARTLEY, T. L., HAYNES, A. L., CARDER, E. A., BLAKEMORE, K. J., IDRISA, A. & GLEW, R. H. 1996. Lysosomal enzymuria in preeclampsia. *American Journal of Kidney Diseases*, 27, 826-833.
- KAY, R., BARTON, C., VELLOSO, C., BROWN, P., BARTLETT, C., BLAZEVIČH, A., GODFREY, R., GOLDSPINK, G., REES, R. & BALL, G. 2009. High-throughput ultra-high-performance liquid chromatography/tandem mass spectrometry quantitation of insulin-like growth factor-I and leucine-rich α -2-glycoprotein in serum as biomarkers of recombinant human growth hormone administration. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, 23, 3173-3182.
- KAČURÁKOVÁ, M. & MATHLOUTHI, M. 1996. FTIR and laser-Raman spectra of oligosaccharides in water: characterization of the glycosidic bond. *Carbohydrate Research*, 284, 145-157.
- KAMATH, U., RAO, G., KAMATH, S. U. and RAI, L. 2006. Maternal and fetal indicators of oxidative stress during intrauterine growth retardation (IUGR). *Indian Journal of Clinical Biochemistry*, 21, 111.

- KUMAR, A., KAPOOR, S. & GUPTA, R. 2013. Comparison of urinary protein: creatinine index and dipsticks for detection of microproteinuria in diabetes mellitus patients. *Journal of Clinical and Diagnostic Research: JCDR*, 7, 622.
- KARAS, M., BACHMANN, D., BAHR, U. E. & HILLENKAMP, F. 1987. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes*, 78, 53-68.
- KAROWICZ-BILIŃSKA, A. 2006. Lipid peroxidation in women with gestational hypertension complicated by asymmetric intrauterine growth retardation. *Ginekologia Polska*, 77, 435-440.
- KOLCH, W., NEUSÜß, C., PELZING, M. & MISCHAK, H. 2005. Capillary electrophoresis–mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. *Mass Spectrometry Reviews*, 24, 959-977.
- KOLIALEXI, A., MAVRELI, D., TOUNTA, G., MAVROU, A. & PAPANTONIOU, N. 2015. Urine proteomic studies in preeclampsia. *PROTEOMICS-Clinical Applications*.
- KORTE, A. R., YANDEAU-NELSON, M. D., NIKOLAU, B. J. & LEE, Y. J. 2015. Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer. *Analytical and Bioanalytical Chemistry*, 407, 2301-2309.
- KRAUSS, T., KUHN, W., LAKOMA, C. and AUGUSTIN, H. G. 1997. Circulating endothelial cell adhesion molecules as diagnostic markers for the early identification of pregnant women at risk for development of preeclampsia. *American Journal of Obstetrics and Gynecology*, 177, 443-449.
- LACOMBE, C., UNTEREINER, V., GOBINET, C., ZATER, M., SOCKALINGUM, G. D. and GARNOTEL, R. 2015. Rapid screening of classic galactosemia patients: a proof-of-concept study using high-throughput FTIR analysis of plasma. *Analyst*, 140, 2280-2286.
- LANGLEY-EVANS, S. C. 2009. Nutritional programming of disease: unravelling the mechanism. *Journal of Anatomy*, 215, 36-51.
- LANGLOIS, M. R. & DELANGHE, J. R. 1996. Biological and clinical significance of haptoglobin polymorphism in humans. *Clinical Chemistry*, 42, 1589-1600.
- LAU, H.-T., SUH, H. W., GOLKOWSKI, M. & ONG, S.-E. 2014. Comparing SILAC-and stable isotope dimethyl-labeling approaches for quantitative proteomics. *Journal of Proteome Research*, 13, 4164-4174.
- LAU, K. W., HART, S. R., LYNCH, J. A., WONG, S. C., HUBBARD, S. J. & GASKELL, S. J. 2009. Observations on the detection of b-and y-type ions in the collisionally activated decomposition spectra of protonated peptides. *Rapid Communications in Mass Spectrometry*, 23, 1508-1514.
- LAW, K. P., HAN, T. L., TONG, C. & BAKER, P. N. 2015. Mass spectrometry-based proteomics for pre-eclampsia and preterm birth. *Int J Mol Sci*, 16, 10952-85.

- LAWRIE, L., CURRAN, S., MCLEOD, H., FOTHERGILL, J. & MURRAY, G. 2001. Application of laser capture microdissection and proteomics in colon cancer. *Molecular Pathology*, 54, 253.
- LEE, S. M., PARK, J. S., NORWITZ, E. R., KIM, S. M., KIM, B. J., PARK, C.-W., JUN, J. K. & SYN, H. C. 2011. Characterization of discriminatory urinary proteomic biomarkers for severe preeclampsia using SELDI-TOF mass spectrometry. *Journal of Perinatal Medicine*, 39, 391-396.
- LEHNERT, S., JESSE, S., RIST, W., STEINACKER, P., SOININEN, H., HERUKKA, S.-K., TUMANI, H., LENTER, M., OECKL, P. & FERGER, B. 2012. iTRAQ and multiple reaction monitoring as proteomic tools for biomarker search in cerebrospinal fluid of patients with Parkinson's disease dementia. *Experimental Neurology*, 234, 499-505.
- LEWIS, G. Saving Mothers' Lives: the continuing benefits for maternal health from the United Kingdom (UK) Confidential Enquires into Maternal Deaths. *Seminars in perinatology*, 2012. Elsevier, 19-26.
- LIN, C.-C. and SANTOLAYA-FORGAS, J. 1998. Current concepts of fetal growth restriction: part I. Causes, classification, and pathophysiology. *Obstetrics and Gynecology*, 92, 1044-1055.
- LIN, S.-Y., LI, M.-J. and CHENG, W.-T. 2007. FT-IR and Raman vibrational microspectroscopies used for spectral biodiagnosis of human tissues. *Journal of Spectroscopy*, 21, 1-30.
- LIO, D. C. S., LIU, C., WIRAJA, C., QIU, B., FHU, C. W., WANG, X. & XU, C. 2018. Molecular beacon-gold nanosensors for Leucine-rich alpha-2-glycoprotein-1 (Lrg1) detection in pathological angiogenesis. *ACS Sensors*.
- LITTLE, K. M., SMALLEY, D. M., HARTHUN, N. L. & LEY, K. The plasma microparticle proteome. *Seminars in thrombosis and hemostasis*, 2010. © Thieme Medical Publishers, 845-856.
- LIU, B., AHMAD, W. & ARONSON JR, N. N. 1999. Structure of the human gene for lysosomal di-N-acetylchitobiase. *Glycobiology*, 9, 589-593.
- LORENTZEN, B., ENDRESEN, M. J., CLAUSEN, T. and HENRIKSEN, T. 1994. Fasting serum free fatty acids and triglycerides are increased before 20 weeks of gestation in women who later develop preeclampsia. *Hypertension in Pregnancy*, 13, 103-109.
- LONG, R., GAO, Y., SUN, H., ZHANG, T., LI, X., LI, M., SUN, Y., KANG, J., WANG, Z. & DING, W. 2018. Quantitative proteomic analysis using iTRAQ to identify salt-responsive proteins during the germination stage of two *Medicago* species. *Scientific Reports*, 8, 9553.
- LOVERGNE, L., CLEMENS, G., UNTEREINER, V., LUKASZWESKI, R. A., SOCKALINGUM, G. D. & BAKER, M. J. 2015. Investigating optimum sample preparation for infrared spectroscopic serum diagnostics. *Analytical Methods*, 7, 7140-7149.

- LUTNICKI, K., ZYCH, I., ROLA, R., PASZKOWSKI, T. and WRÓBEL, J. 1998. Levels of lipid peroxidation products in blood serum and placental tissue of women with pregnancy induced hypertension. *Ginekologia Polska*, 69, 895-901.
- MAISONNEUVE, E., DELVIN, E., OUELLET, A., MORIN, L., DUBÉ, J., BOUCOIRAN, I., MOUTQUIN, J.-M., FOURON, J.-C., KLAM, S. and LEVY, E. 2015. Oxidative conditions prevail in severe IUGR with vascular disease and Doppler anomalies. *The Journal of Maternal-Fetal and Neonatal Medicine*, 28, 1471-1475.
- MANDÒ, C., RAZINI, P., NOVIELLI, C., ANELLI, G. M., BELICCHI, M., ERRATICO, S., BANFI, S., MEREGALLI, M., TAVELLI, A. & BACCARIN, M. 2016. Impaired angiogenic potential of human placental mesenchymal stromal cells in intrauterine growth restriction. *Stem Cells Translational Medicine*, 5, 451-463.
- MANNOCK, D. A., LEWIS, R. N. & MCELHANEY, R. N. 2006. Comparative calorimetric and spectroscopic studies of the effects of lanosterol and cholesterol on the thermotropic phase behavior and organization of dipalmitoylphosphatidylcholine bilayer membranes. *Biophysical Journal*, 91, 3327-3340.
- MANOHARAN, R., BARAGA, J. J., RAVA, R. P., DASARI, R. R., FITZMAURICE, M. and FELD, M. S. 1993. Biochemical analysis and mapping of atherosclerotic human artery using FT-IR microspectroscopy. *Atherosclerosis*, 103, 181-193.
- MARNETT, L. J. 1992. Aspirin and the potential role of prostaglandins in colon cancer. *Cancer Research*, 52, 5575-5589.
- MARUSIC, J., PRUSAC, I. K., TOMAS, S. Z., KARARA, J. R. & ROJE, D. 2013. Expression of inflammatory cytokines in placentas from pregnancies complicated with preeclampsia and HELLP syndrome. *The Journal of Maternal-Fetal & Neonatal Medicine*, 26, 680-685.
- MCALISTER, G. C., NUSINOW, D. P., JEDRYCHOWSKI, M. P., WÜHR, M., HUTTLIN, E. L., ERICKSON, B. K., RAD, R., HAAS, W. & GYGI, S. P. 2014. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Analytical Chemistry*, 86, 7150-7158.
- MCLUCKEY, S. A. 1992. Principles of collisional activation in analytical mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 3, 599-614.
- MEDICINE, I. O. 2007. Preterm birth: Causes, consequences, and prevention. National Academies Press (US), National Academy of Sciences Washington (DC).
- MEDZIHRADESKY, K. F. 2005. Peptide sequence analysis. *Methods in Enzymology*, 402, 209-244.
- MEYER, K. and ZHANG, L. 2007. Fetal programming of cardiac function and disease. *Reproductive Sciences*, 14, 209-216.
- MIAO, J., CHEN, F., DUAN, S., GAO, X., LIU, G., CHEN, Y., DIXON, W., XIAO, H. & CAO, Y. 2015. iTRAQ-Based Quantitative Proteomic Analysis of the Antimicrobial Mechanism of Peptide F1 against *Escherichia coli*. *J Agric Food Chem*, 63, 7190-7.

- MITTAL, M., KULKARNI, C., PANCHONIA, A. and MITTAL, R. 2017. Evaluation of serum lipid profile in cases of pre-eclampsia and eclampsia. *International Journal of Reproduction, Contraception, Obstetrics and Gynecology*, 3, 732-734.
- MIZUTANI, S., YAMADA, R., KURAUCHI, O., ITO, Y., NARITA, O. & TOMODA, Y. 1987. Serum aminopeptidase A (AAP) in normal pregnancy and pregnancy complicated by pre-eclampsia. *Archives of Gynecology*, 240, 27-31.
- MOORE, D. S. 2014. *Handbook of Spectroscopy*, John Wiley & Sons.
- MORSE, K. W., ASTBURY, N. M., WALCZYSZYN, A., HASHIM, S. A. & GELIEBTER, A. 2017. Changes in zinc- α 2-glycoprotein (ZAG) plasma concentrations pre and post Roux-En-Y gastric bypass surgery (RYGB) or a very low calorie (VLCD) diet in clinically severe obese patients: Preliminary Study. *Integrative Obesity and Diabetes*, 3.
- MOVASAGHI, Z., REHMAN, S. & UR REHMAN, D. I. 2008. Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews*, 43, 134-179.
- MUKHERJEE, R., RAY, C. D., RAY, S., DASGUPTA, S. and CHAUDHURY, K. 2014. Altered metabolic profile in early and late onset preeclampsia: an FTIR spectroscopic study. *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*, 4, 70-80.
- MURAOKA, S., KUME, H., WATANABE, S., ADACHI, J., KUWANO, M., SATO, M., KAWASAKI, N., KODERA, Y., ISHITOBI, M. & INAJI, H. 2012. Strategy for SRM-based verification of biomarker candidates discovered by iTRAQ method in limited breast cancer tissue samples. *Journal of Proteome Research*, 11, 4201-4210.
- MYERS, J. E., TUYTTEN, R., THOMAS, G., LAROY, W., KAS, K., VANPOUCKE, G., ROBERTS, C. T., KENNY, L. C., SIMPSON, N. A., BAKER, P. N. & NORTH, R. A. 2013. Integrated proteomics pipeline yields novel biomarkers for predicting preeclampsia. *Hypertension*, 61, 1281-8.
- NAGALLA, S. R., RASANEN, J. & GRAVETT, M. 2010. Maternal serum biomarkers for detection of pre-eclampsia. Google Patents.
- NAKASHIMA, Y., TOMATSU, S., HORI, T., FUKUDA, S., SUKEGAWA, K., KONDO, N., SUZUKI, Y., SHIMOZAWA, N. & ORII, T. 1994. Mucopolysaccharidosis IV A: molecular cloning of the human N-acetylgalactosamine-6-sulfatase gene (GALNS) and analysis of the 5'-flanking region. *Genomics*, 20, 99-104.
- NGUYEN, H. P. & SCHUG, K. A. 2008. The advantages of ESI-MS detection in conjunction with HILIC mode separations: Fundamentals and applications. *Journal of Separation Science*, 31, 1465-1480.
- OH, K. J., PARK, J. S., NORWITZ, E. R., KIM, S. M., KIM, B. J., PARK, C.-W., JUN, J. K. & SYN, H. C. 2012. Proteomic biomarkers in second trimester amniotic fluid that identify women who are destined to develop preeclampsia. *Reproductive Sciences*, 19, 694-703.

- OBSTETRICIANS, A. C. O. & GYNECOLOGISTS 2013. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' task force on hypertension in pregnancy. *Obstetrics and Gynecology*, 122, 1122.
- OHKUBO, I., NIWA, M., TAKASHIMA, A., NISHIKIMI, N., GASA, S. & SASAKI, M. 1990. Human seminal plasma Zn- α 2-glycoprotein: Its purification and properties as compared with human plasma Zn- β 2-glycoprotein. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1034, 152-156.
- OLIVEROS, J. C. 2007. VENNY. An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>.
- ONG, S.-E. & MANN, M. 2006. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nature Protocols*, 1, 2650-2660.
- PAN, H.-T., GUO, M.-X., XIONG, Y.-M., REN, J., ZHANG, J.-Y., GAO, Q., KE, Z.-H., XU, G.-F., TAN, Y.-J. & SHENG, J.-Z. 2015. Differential proteomic analysis of umbilical artery tissue from preeclampsia patients, using iTRAQ isobaric tags and 2D nano LC-MS/MS. *Journal of Proteomics*, 112, 262-273.
- PARIKH, S. M. & KARUMANCHI, S. A. 2008. Putting pressure on pre-eclampsia. *Nature Medicine*, 14, 810-812.
- PARK, K. W., MORRISON, C. M., SORENSEN, L. K., JONES, C. A., RAO, Y., CHIEN, C.-B., WU, J. Y., URNESS, L. D. & LI, D. Y. 2003. Robo4 is a vascular-specific receptor that inhibits endothelial migration. *Developmental Biology*, 261, 251-267.
- PHANSTIEL, D., UNWIN, R., MCALISTER, G. C. & COON, J. J. 2009. Peptide quantification using 8-plex isobaric tags and electron transfer dissociation tandem mass spectrometry. *Analytical Chemistry*, 81, 1693-1698.
- PHILLIPS, A., SHAPER, A. G. and WHINCUP, P. 1989. Association between serum albumin and mortality from cardiovascular disease, cancer, and other causes. *The Lancet*, 334, 1434-1436.
- PLEBANI, M. 2005. Proteomics: the next revolution in laboratory medicine? *Clinica Chimica Acta*, 357, 113-122.
- POWELL, K. L., CARROZZI, A., STEPHENS, A. S., TASEVSKI, V., MORRIS, J. M., ASHTON, A. W. and DONA, A. C. 2018. Utility of metabolic profiling of serum in the diagnosis of pregnancy complications. *Placenta*.
- POWERS, R., ROBERTS, J., COOPER, K., GALLAHER, M., FRANK, M., HARGER, G. and NESS, R. 2005. Maternal serum soluble fms-like tyrosine kinase 1 concentrations are not increased in early pregnancy and decrease more slowly postpartum in women who develop preeclampsia. *American Journal of Obstetrics and Gynecology*, 193, 185-191.
- PRACTICE, A. C. O. O. 2002. ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. American College of Obstetricians and Gynecologists. *International journal of gynaecology and*

obstetrics: The Official Organ of the International Federation of Gynaecology and Obstetrics, 77, 67.

- RAMÍREZ-BOO, M., NÚÑEZ, E., JORGE, I., NAVARRO, P., FERNANDES, L. T., SEGALÉS, J., GARRIDO, J. J., VÁZQUEZ, J. & MORENO, Á. 2011. Quantitative proteomics by 2-DE, 16O/18O labelling and linear ion trap mass spectrometry analysis of lymph nodes from piglets inoculated by porcine circovirus type 2. *Proteomics*, 11, 3452-3469.
- RAOUF, G. A., AL-MALKI, A.-R. L., MANSOURI, N. and MAHMOUDI, R. M. 2011. Preliminary study in diagnosis and early prediction of preeclampsia by using FTIR spectroscopy technique. *Life Science Journal*, 8, 453-464.
- RAUNIYAR, N. & YATES III, J. R. 2014. Isobaric labeling-based relative quantification in shotgun proteomics. *Journal of Proteome Research*, 13, 5293-5309.
- ROACH, P., FARRAR, D. & PERRY, C. C. 2005. Interpretation of protein adsorption: surface-induced conformational changes. *Journal of the American Chemical Society*, 127, 8168-8173.
- ROBERTS, J. M. Endothelial dysfunction in preeclampsia. *Seminars in reproductive endocrinology*, 1998. Copyright© 1998 by Thieme Medical Publishers, Inc., 5-15.
- ROMEO, M. J., DUKOR, R. K. & DIEM, M. 2008. Introduction to spectral imaging, and applications to diagnosis of lymph nodes. *Handbook of Vibrational Spectroscopy*.
- ROSSI, R., DALLE-DONNE, I., MILZANI, A. & GIUSTARINI, D. 2006. Oxidized forms of glutathione in peripheral blood as biomarkers of oxidative stress. *Clinical Chemistry*, 52, 1406-1414.
- SAMMOUR, R. N., NAKHOUL, F. M., LEVY, A. P., MILLER-LOTAN, R., NAKHOUL, N., AWAD, H. R., GONEN, R. & OHEL, G. 2010. Haptoglobin phenotype in women with preeclampsia. *Endocrine*, 38, 303-308.
- SANKARALINGAM, S., ARENAS, I. A., LALU, M. M. & DAVIDGE, S. T. 2006. Preeclampsia: current understanding of the molecular basis of vascular dysfunction. *Expert Reviews in Molecular Medicine*, 8, 1-20.
- SCHAUB, S., WILKINS, J., WEILER, T., SANGSTER, K., RUSH, D. & NICKERSON, P. 2004. Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney International*, 65, 323-332.
- SEN, C. K. and PACKER, L. 1996. Antioxidant and redox regulation of gene transcription. *The FASEB journal*, 10, 709-720.
- SIDEBOTHAM, M. 1998. CESDI: the fifth annual report. *British Journal of Midwifery*, 6, 692-694.
- SINCLAIR, A., BARNETT, A. and LUNEC, J. 1990. Free radicals and antioxidant systems in health and disease. *British Journal of Hospital Medicine*, 43, 334-344.

- SORIANI, M., PIETRAFORTE, D. and MINETTI, M. 1994. Antioxidant potential of anaerobic human plasma: role of serum albumin and thiols as scavengers of carbon radicals. *Archives of Biochemistry and Biophysics*, 312, 180-188.
- SPICKETT, C. M., REGLINSKI, J., SMITH, W. E., WILSON, R., WALKER, J. J. and MCKILLOP, J. 1998. Erythrocyte glutathione balance and membrane stability during preeclampsia. *Free Radical Biology and Medicine*, 24, 1049-1055.
- STARK, J. 1993. Pre-eclampsia and cytokine induced oxidative stress. *BJOG: An International Journal of Obstetrics and Gynaecology*, 100, 105-109.
- STEVENS, W., SHIH, T., INCERTI, D., TON, T. G., LEE, H. C., PENEVA, D., MACONES, G. A., SIBAI, B. M. & JENA, A. B. 2017. Short-term costs of preeclampsia to the United States health care system. *American Journal of Obstetrics and Gynecology*, 217, 237-248. e16.
- STIMSON, E., TRUONG, O., RICHTER, W., WATERFIELD, M. & BURLINGAME, A. 1997. Enhancement of charge remote fragmentation in protonated peptides by high-energy CID MALDI-TOF-MS using "cold" matrices. *International journal of Mass Spectrometry and Ion Processes*, 169, 231-240.
- TAKACS, P., GREEN, K. L., NIKAEIO, A. and KAUMA, S. W. 2003. Increased vascular endothelial cell production of interleukin-6 in severe preeclampsia. *American Journal of Obstetrics and Gynecology*, 188, 740-744.
- TSANGARIS, G. T., KOLIALEXI, A., KARAMESSINIS, P. M., ANAGNOSTOPOULOS, A. K., ANTSAKLIS, A., FOUNTOULAKIS, M. & MAVROU, A. 2006. The normal human amniotic fluid supernatant proteome. *In Vivo*, 20, 479-490.
- TAYLOR, R. N., CROMBLEHOLME, W. R., FRIEDMAN, S. A., JONES, L. A., CASAL, D. C. and ROBERTS, J. M. 1991. High plasma cellular fibronectin levels correlate with biochemical and clinical features of preeclampsia but cannot be attributed to hypertension alone. *American Journal of Obstetrics and Gynecology*, 165, 895-901.
- TERAWAKI, H., YOSHIMURA, K., HASEGAWA, T., MATSUYAMA, Y., NEGAWA, T., YAMADA, K., MATSUSHIMA, M., NAKAYAMA, M., HOSOYA, T. and ERA, S. 2004. Oxidative stress is enhanced in correlation with renal dysfunction: examination with the redox state of albumin. *Kidney International*, 66, 1988-1993.
- THEODORESCU, D., WITTKKE, S., ROSS, M. M., WALDEN, M., CONAWAY, M., JUST, I., MISCHAK, H. & FRIERSON, H. F. 2006. Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. *The lancet Oncology*, 7, 230-240.
- THOMPSON, A., SCHÄFER, J., KUHN, K., KIENLE, S., SCHWARZ, J., SCHMIDT, G., NEUMANN, T. & HAMON, C. 2003. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry*, 75, 1895-1904.

- TISS, A., SMITH, C., CAMUZEUX, S., KABIR, M., GAYTHER, S., MENON, U., WATERFIELD, M., TIMMS, J., JACOBS, I. & CRAMER, R. 2007. Serum peptide profiling using MALDI mass spectrometry. *Proteomics*, 7, 77-89.
- TIRUMALAI, R. S., CHAN, K. C., PRIETO, D. A., ISSAQ, H. J., CONRADS, T. P. & VEENSTRA, T. D. 2003. Characterization of the low molecular weight human serum proteome. *Molecular & Cellular Proteomics*, 2, 1096-1103.
- TOLOSANO, E., FAGOONEE, S., HIRSCH, E., BERGER, F. G., BAUMANN, H., SILENGO, L. & ALTRUDA, F. 2002. Enhanced splenomegaly and severe liver inflammation in haptoglobin/hemopexin double-null mice after acute hemolysis. *Blood*, 100, 4201-4208.
- VAN BAAR, B. L. 2000. Characterisation of bacteria by matrix-assisted laser desorption/ionisation and electrospray mass spectrometry. *FEMS Microbiology Reviews*, 24, 193-219.
- VAN DER VUURST DE VRIES, A. R., CLEVERS, H., LOGTENBERG, T. & MEYAARD, L. 1999. Leukocyte-associated immunoglobulin-like receptor-1 (LAIR-1) is differentially expressed during human B cell differentiation and inhibits B cell receptor-mediated signaling. *European Journal of Immunology*, 29, 3160-3167.
- VAN LERBERGHE, W., MANUEL, A., MATTHEWS, Z. and CATHY, W. 2005. The World Health Report 2005-make every mother and child count, World Health Organization.
- VASUDEVAN, D. M., SREEKUMARI, S. & VAIDYANATHAN, K. 2013. Textbook of biochemistry for medical students, JP Medical Ltd.
- VAZ, F. M., PRAS-RAVES, M., BOOTSMA, A. H. & VAN KAMPEN, A. H. 2015. Principles and practice of lipidomics. *Journal of Inherited Metabolic Disease*, 38, 41-52.
- VILLAR, J., BA'AQEEL, H., PIAGGIO, G., LUMBIGANON, P., BELIZÁN, J. M., FARNOT, U., ALMAZROU, Y., CARROLI, G., PINOL, A. and DONNER, A. 2001. WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care. *The Lancet*, 357, 1551-1564.
- VILLAR, J., CARROLI, G., WOJDYLA, D., ABALOS, E., GIORDANO, D., BA'AQEEL, H., FARNOT, U., BERGSJØ, P., BAKKETEIG, L. and LUMBIGANON, P. 2006. Preeclampsia, gestational hypertension and intrauterine growth restriction, related or independent conditions? *American Journal of Obstetrics and Gynecology*, 194, 921-931.
- VON DADELSZEN, P., MAGEE, L. A. and ROBERTS, J. M. 2003. Subclassification of preeclampsia. *Hypertension in Pregnancy*, 22, 143-148.
- WALSH, S. W. Maternal-placental interactions of oxidative stress and antioxidants in preeclampsia. *Seminars in reproductive endocrinology*, 1998. Copyright© 1998 by Thieme Medical Publishers, Inc., 93-104.
- WALTHER, T. C. & MANN, M. 2010. Mass spectrometry-based proteomics in cell biology. *The Journal of Cell Biology*, 190, 491-500.

- WANG, F., SHI, Z., WANG, P., YOU, W. & LIANG, G. 2013. Comparative proteome profile of human placenta from normal and preeclamptic pregnancies. *PLoS One*, 8.
- WASHBURN, M. P., ULASZEK, R., DECIU, C., SCHIELTZ, D. M. & YATES, J. R. 2002. Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Analytical Chemistry*, 74, 1650-1657.
- WATSON, J. T. & SPARKMAN, O. D. 2007. *Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation*, John Wiley & Sons.
- WEN, Q., LIU, L. Y., YANG, T., ALEV, C., WU, S., STEVENSON, D. K., SHENG, G., BUTTE, A. J. & LING, X. B. 2013. Peptidomic identification of serum peptides diagnosing preeclampsia. *PLoS ONE* 8(6): e65571.
- WEINER, C. P., MASON, C. W., DONG, Y., BUHIMSCHI, I. A., SWAAN, P. W. & BUHIMSCHI, C. S. 2010. Human effector/initiator gene sets that regulate myometrial contractility during term and preterm labor. *American Journal of Obstetrics and Gynecology*, 202, 474. e1-474. e20.
- WHITE, L., DHARMARAJAN, A. & CHARLES, A. 2007. Caspase-14: a new player in cytotrophoblast differentiation. *Reproductive Biomedicine Online*, 14, 300-307.
- WILEY, W. & MCLAREN, I. H. 1955. Time-of-flight mass spectrometer with improved resolution. *Review of Scientific Instruments*, 26, 1150-1157.
- WILKINS, M. R. 1997. *Proteome research: new frontiers in functional genomics*, Springer Science & Business Media.
- WILLIAMSON, J. C., EDWARDS, A. V., VERANO-BRAGA, T., SCHWÄMMLE, V., KJELDSSEN, F., JENSEN, O. N. & LARSEN, M. R. 2016. High-performance hybrid Orbitrap mass spectrometers for quantitative proteome analysis: Observations and implications. *Proteomics*, 16, 907-914.
- WINGERT, N., AG NUNES, M., T BARDEN, A., GOMES, P., I MULLER, E., MM FLORES, E. & STEPPE, M. 2015. Ultra-performance LC-ESI/Q-TOF MS for the Rapid Analysis of Rivaroxaban: Method Validation Using Experimental Design for Robustness Evaluation. *Current Analytical Chemistry*, 11, 124-129.
- XIN, H., ZHANG, Y., WANG, H. & SUN, S. 2012. Alterations of profibrinolytic receptor annexin A2 in pre-eclampsia: a possible role in placental thrombin formation. *Thrombosis Research*, 129, 563-567.
- YATES, J. R. 2000. Mass spectrometry: from genomics to proteomics. *Trends in Genetics*, 16, 5-8.
- YONAR, D., OCEK, L., TIFTIKCIOGLU, B. I., ZORLU, Y. & SEVERCAN, F. 2018. Relapsing-Remitting Multiple Sclerosis diagnosis from cerebrospinal fluids via Fourier transform infrared spectroscopy coupled with multivariate analysis. *Scientific Reports*, 8, 1025.

- YOSHIMURA, T., SUZUKI, A. & IWASAKI, N. 2015. Mechanism of O and C isotope fractionation in magnesian calcite skeletons of Octocorallia corals and an implication on their calcification response to ocean acidification. *Biogeosciences Discussions*, 12, 389-412.
- ZHANG, Q., HUANG, R., TANG, Q., YU, Y., HUANG, Q., CHEN, Y., WANG, G. & WANG, X. 2018. leucine-rich alpha-2-glycoprotein-1 is up-regulated in colorectal cancer and is a tumor promoter. *OncoTargets and Therapy*, 11, 2745.

6 Appendices:

6.1 Appendix 1 : Material transfer agreement

 UCC <small>Coláiste na hOllscoile Corcaigh, Oir The College of Cork, Ireland</small>	Office of Technology Transfer
MATERIALS TRANSFER AGREEMENT	

This Agreement is made on _____ (hereinafter the
“Effective Date”) between

1	UNIVERSITY COLLEGE CORK - NATIONAL UNIVERSITY OF IRELAND, CORK, of Western Road, Cork, Ireland	“UCC”
2	UNIVERSITY OF KEELE, a university established by the University of Keele Act 1962 (10 & 11 Eliz. 2 Ch Xv) and the granting of a Royal Charter in 1962, of Keele, Staffordshire ST5 5BG, United Kingdom Each a ‘Party’ and together the ‘Parties’;	“RECIPIENT”

Recipient desires to obtain from UCC certain materials being banked urine samples, gestational age 15 and 20 weeks as more fully described in Annex 1 (“Materials”) and information including un-identifying clinical data including risk factors, socioeconomic, coupled with biomarker data directly relating to the Materials (“Information”) for the performance of urine proteomic biomarker discovery in pregnant women subsequently diagnosed with pre-eclampsia or intrauterine growth restriction (IUGR), and compare this to matched normotensive controls as more fully described at Annex 2 in the Work-Plan hereinafter (“the Purpose”).

1. TRANSFER AND USE OF MATERIALS AND INFORMATION

1.1 UCC agrees to make available to Recipient the Materials and the Information on a non-exclusive basis for the Purpose for a term of 24 months (unless terminated earlier in accordance with the provisions of this Agreement) and Recipient agrees to receive the Materials and the Information subject to the following terms and conditions.

1.2 Recipient agrees and undertakes that the Materials and the Information shall: a) be used for the Purpose only as detailed in the Annex 2; b) be stored securely and not removed from Recipient’s research premises; (c) be accessed only by Recipient and by Recipient’s authorised personnel; d) be used with all due skill and care in accordance with good laboratory practice and shall ensure compliance with applicable laws, regulations including those relating to Data Protection and administrative and biobanking best practices governing the storage, use and disposal of the Materials and the Information; e) not be altered, changed or mutated in any way; f) be transported, stored and used in compliance with all relevant laws, regulations and guidelines relating to the use of the Materials and Information; ; g) not be disclosed or transferred to any party outside of Recipient’s organisation except with UCC’s prior written consent and on such conditions as UCC may specify; (h) in relation to Materials, not be used on humans and if they are used on animals, no animal on which the Materials are used and no product which is derived from that animal shall be used for food, for commercial breeding purposes or for diagnostic or therapeutic purposes; and (i)

where the Materials and Information are to be used in conducting tests in vitro and on laboratory research animals, that research must be conducted by a person regularly engages in such tests and are qualified by training and/or experience to conduct such tests on the Materials.

1.3 Recipient shall keep UCC regularly informed in writing of the status of, and results obtained from, its use of the Materials and Information.

1.4 Recipient shall keep confidential the Materials and the Information, except for any of the Information which: a) (as evidenced by Recipient’s written records) was in Recipient’s possession or control prior to the date of disclosure; b) was in the public domain or enters into the public domain through no improper act on Recipient’s part or on the part of any Recipient’s employees; or c) is lawfully given to Recipient from sources independent of UCC.

1.5 Recipient may only disclose the Material and the Information to Recipient’s authorized personnel (“Recipient’s Personnel”) to the extent required in order to permit the fulfilment by Recipient of the Purpose provided that Recipient obtains from each of Recipient’s Personnel an undertaking to the effect that s/he will comply with the obligations set out in clause 1.4 hereof. Recipient shall procure that Recipient’s Personnel shall comply with the undertakings on their part required by this paragraph and agrees that the acts and omissions of Recipient’s Personnel in connection with the Materials and the Information will be treated as acts and omissions of Recipient. The Recipient shall under no circumstances disclose the Information and shall use their best efforts to ensure that it is kept securely with only limited access to authorized Personnel.

1.6 Recipient shall provide UCC with a copy of any reports or publications which describe any work undertaken by Recipient using the Materials, and UCC shall be entitled to use all such data, reports and publications and make them available to third parties following Recipient’s prior written consent and on such conditions as Recipient may specify.

1.7 The Recipient will not try to link the Information or Materials provided by the UCC to other SCOPE CONSORTIUM Information held by different recipients or by the same Recipient for different projects. The Recipient will not make any attempts to use SCOPE Information or Material to identify individual SCOPE participants.

2. COST & DELIVERY

The Materials and the Information are provided to Recipient at no cost but Recipient shall reimburse UCC for any costs incurred by UCC in providing the Materials and the Information to Recipient.

The cost of collating and the provision of the Materials and Information shall be no greater than € 4200 (inclusive of VAT) (four thousand, two hundred euro). Shipping costs, separate to the costs above, to be met by the Recipient. These costs shall be invoiced at the earliest when all the Materials and the Information are collected from UCC premises for delivery to those of the Recipient. The Recipient shall within 30 days of the date of the invoice reimburse UCC.

3. UCC shall not be liable whatsoever for the transport, delivery and Recipient handling, of the Materials.

OWNERSHIP OF MATERIALS AND INFORMATION

3.1 UCC retains title and all rights to the Materials and the Information. No licence of UCC's intellectual property in the Information is granted or implied by this Agreement. Recipient acknowledges that it does not obtain any right, title or interest to the Information.

3.2 Nothing contained in this Agreement shall prevent UCC from being able to distribute the Materials and the Information to other commercial or non-commercial entities.

4. USE

4.1 It is understood by the Parties hereto that the Materials provided hereunder is **solely for the stated Purpose** as detailed in Annex 2 and that the Recipient is not to perform any activities, work or development utilizing the Materials other than for the stated Purpose. Further, whilst efforts will be made to provide the Materials and Information as described in Annex 1, this provision will be subject to the availability and/or number of aliquots remaining, participant consent and study design and the provision of Materials and Information of the nature and quantities requested will be at the sole discretion of UCC.

4.2 If Recipient makes or observes an invention, improvement or discovery (whether or not patentable) ("Invention") through use of the Materials outside of the Purpose then Recipient shall promptly notify UCC of that Invention. UCC shall own, and Recipient shall deliver to UCC, all documents, samples, information reports, test results and the like prepared by or for Recipient and which relate to the Invention or any other matter potentially the subject of intellectual property rights. Recipient shall disclose to UCC the name(s) of the inventor(s) of any potentially patentable Invention to enable UCC to file for appropriate patent protection if it wishes to do so.

4.3 For Invention(s) created under 4.2: the Recipient hereby assigns to UCC by way of future assignment all copyrights and other intellectual property rights so arising in any Invention referred to under 4.2 (and waives or will procure the waiver of any equivalent moral rights) immediately on their coming into existence. To the extent that full legal title to any Invention so arising shall fail automatically to belong to UCC by virtue of the above provisions Recipient shall hold such right on trust for UCC absolutely and shall immediately at UCC's request execute or procure the execution of any documents that UCC requires to vest in it (or as it shall direct) the full legal title to such intellectual property right to enable it (or its nominee) to enjoy the benefit of such right.

4.4 Recipient shall notify UCC of any Invention that it may make and shall not disclose any details of the Invention to any third party without the prior written consent of UCC.

5. PUBLICATION

5.1 In order to protect UCC's proprietary and/or patent rights to the Materials and Information, Recipient shall provide UCC with an advance copy of any proposed publication. If in the opinion of UCC any such publication describes a patentable invention of UCC, UCC shall be entitled to request that Recipient delay publication until such time as UCC has filed a patent application (which period of delay shall not exceed ninety (90) days).

5.2 Recipient shall acknowledge UCC and the SCOPE Consortium as the source of the Materials and Information in any publication that

refers to the Materials and/or Information and/or shall give credit to UCC scientists, as scientifically appropriate, for any direct contribution they may have made to work undertaken by Recipient using the Materials and/or Information.

5.3 Each of the Parties shall be named as co-authors on all publications. All proposed publications shall be submitted by the Recipient to UCC within 30 days prior to the proposed publication. This preview period for approval shall apply to publications including abstracts for conference presentation and shall follow the process for review, as set out below in clause 5.1.

6. USE OF UCC'S NAME

Except as provided in clause 5, the Parties shall not use the name of the other Party, or variant thereof, in any advertising or publicity materials without prior written approval from the relevant Party.

7. WARRANTY DISCLAIMER

Recipient acknowledges that the Materials and Information are experimental in nature and that the Materials and the Information are provided without warranty of merchantability, fitness for a particular purpose or safety for use or any other warranty, express or implied, and without any representation or warranty that UCC owns all necessary property and other rights in the Materials and the Information and that the use or supply of the Materials or the Information will not infringe any patent, copyright, trade mark or other intellectual property right of any third party.

8. LIABILITY

8.1 Recipient hereby assumes all liability with respect to any risk arising from transporting, receiving, keeping and using the Materials and the Information. To the fullest extent permitted by law, UCC excludes all liability to Recipient for any loss, claim or demand made by Recipient or made against Recipient by a third party arising from transporting, receiving, keeping or using the Materials and the Information.

8.2 Subject to Article 8.4, UCC shall not be liable in contract, tort or otherwise for any indirect damages or losses, or to any loss of profits, loss of revenue, loss of data, loss of contracts or otherwise, whether direct or indirect, even if the Recipient bringing the claim has advised the other of the possibility of those losses, or if they were within the other Party's contemplation.

8.3 Recipient hereby agrees to defend, indemnify and hold UCC and its officers, employees and representatives harmless from any loss, claim, damage or liability of whatever kind or nature, which may arise from or in connection with this Agreement or the keeping or use of the Materials and the Information.

8.4. Nothing in this Agreement limits or excludes either Party's liability for: (i) death or personal injury cause by that Party's negligence; and (ii) any fraud or for any sort of liability that, by law, cannot be limited or excluded.

9. GOVERNING LAW

This Agreement shall be governed by the laws of Ireland. The courts of Ireland shall have exclusive jurisdiction to settle any dispute arising out of or in connection with this Agreement and the parties submit to the jurisdiction of the Irish courts for that purpose.

10. TERM

This Agreement may be extended with the written agreement of UCC following expiry of the Term. Any request by Recipient for an extension of the term must be received by UCC in writing at least 30 days before the expiry of the Term

11. TERMINATION

11.1 This Agreement may be terminated by either Party on thirty (30) days' prior written notice to the other Party but termination shall not relieve Recipient of its obligations under this Agreement.

11.2 UCC shall be entitled to terminate this Agreement immediately: a) in the event of the insolvency or bankruptcy of the Recipient; or b) upon written notice to the Party in material breach of a term of this Agreement.

11.3 Clauses 1.4, 1.5, 1.6, 1.7, 3, 4, 5, 6, 7, 8, 9, 11 and 14 shall survive the termination or expiry of this Agreement.

12. RETURN OF MATERIAL AND INFORMATION

12.1 On written request and direction of UCC, Recipient shall return all of the Materials and the Information to UCC: a) immediately on termination or expiry of this Agreement; b) immediately if Recipient has not made use of the Materials and Information within six months of the receipt of same or does not intend to use, or to continue to use, the Materials or the Information; or c) immediately, at any other time within 14 days of receipt of a written request from UCC at their sole discretion.

12.2 At the request of UCC, Recipient shall destroy or return the Materials (unused) and all data, records and other information comprising or containing or derived from the Information, or copies thereof, extracts or derivatives of any part of the Information and shall provide to UCC certification of that destruction, signed on its behalf by one of its officers.

13. ASSIGNMENT

This Agreement and any rights therein may not be assigned or sub-licensed by the Recipient without the prior written consent of UCC. UCC shall be entitled to assign or transfer this Agreement, or any of its rights and obligations herein, without the consent of Recipient.

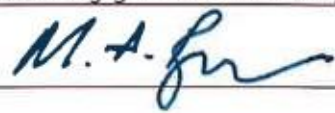
14. ENTIRE AGREEMENT

This Agreement sets out the entire understanding between the Parties and cannot be changed or amended except by written agreement between the Parties.

15. COUNTERPARTS

This Agreement may be executed in two or more counterparts, each of which shall be deemed to be an original but which together shall constitute one and the same instrument.

IN WITNESS OF WHICH, UCC AND RECIPIENT have executed this Agreement on the date first written above.

UCC	
Name	Dr Tim Roche
Title	Director - Office of Technology Transfer
Signature	
University of Keele UK	
Name	Dr Mark Bacon
Title:	Director - Engagement & Partnerships
Signature	

ANNEX 1 – THE MATERIALS

The materials are as per SCOPE RAF3.45 SCOPE Urine 1 aliquots (450ul) per woman for each time point 15 w & 20w from a total of 120 participants. The participants include women whose pregnancies were complicated by IUGR (n=30) and with PE (n=30) as compared to uncomplicated pregnancies, matched for ethnicity and BMI (n=30 in each case) (total 240 urine aliquots, if available).

Specimen Type (e.g., EDTA plasma)	Gestation of Specimen (15w, 20w or Time of disease/late control)	Number of Aliquots per woman*	Total of Aliquots in Expt
Urine	15w	1	120
Urine	20 w	1	120

Information:

a) Sociodemographic and physical data

The current simplest predictor of pre-eclampsia is a previous obstetric history of pre-eclampsia, however for nulliparous women, this parameter is irrelevant. For IUGR, risk factors include maternal malnutrition and/or metabolic disruption (such as gestational diabetes) and smoking, fetal abnormality, however these alone lack sensitivity for clinical risk prediction. We will utilise information on patient

information regarding socioeconomic and other risk factors as part of a multicomponent model alongside biomarker data. Available parameters include age, ethnicity, parity, BMI, Socio-economic status (measured using the Index of Multiple Deprivation¹⁰ and by classification of head of household's occupation using the NS-SEC system¹¹) and cigarette intake.

ANNEX 2 - THE WORK PLAN

The work plan as outlined in SCOPE RAF3.45:

Purpose:

To develop a predictive algorithm identifying those women at increased risk of IUGR and pre-eclampsia using protein-based urine screening tests.

b) *Proteomic biomarker discovery.*

Sample preparation and analysis: We have developed a Standard Operating Procedure (SOP) for sample preparation from urine samples by ultracentrifugation and retention of proteins via ultra-filtration [Hart *et al.*, Submitted, J Proteome Res]. Isotopic labelling using amine-reactive tags (iTRAQ or TMT) can then be used to tag proteins from between 4 and 10 subjects or pooled samples, combining analysis of isotopically-labelled proteins to enable multiplexed LC-MS/MS identification and quantification of a large number of individuals on a compressed timescale in a more effective quantitative manner. The resultant derivatized peptides will be subjected to high-performance liquid chromatography separations (ion exchange, hydrophobic ligand interaction and reversed-phase) to increase coverage depth and sensitivity of analysis by chemical/temporal separation of analytes with coincident mass-to-charge ratios. Eluting peptides will be subjected to collision-induced dissociation (CID) to enable both identification (on the basis of peptide backbone fragmentation and database matching) and quantification (based on reporter ion signal integration). Quality control samples (acquired by pooling urinary peptides from all subjects) will be run as part of every multiplexed experiment.

c) *Data analysis and construction of a predictive algorithm:*

We will utilise the SCOPE database; centralised, secure, Internet-based databases (FDA and HIPPA compliant) have been established by MedSciNet, and allow continuous data entry and monitoring of study progress. Completeness of clinical data and specimen collection is constantly monitored, with incomplete fields 'flagged' for attention. Storage of all specimen aliquots is tracked. To comply with biobank regulatory issues, patients are only identified by a unique study number. Identifying information about participants will be kept in a separate, secure local database.

Data analysis will be performed using Mascot for identification of proteolytic peptides, with the Scaffold Q+ algorithm being used for integration of large datasets and quantitative data comparison, using ANOVA to compare data across parallel experiments (vs. an identical pooled control sample which will be included in all analyses).

A rigorous statistical analysis will be required to help predict which combinations of candidate markers will have the best

sensitivity and specificity as predictive tools. We have substantial expertise in this area.

Summary of predicted outcome: We anticipate that this case control study will generate a predictive algorithm capable of discriminating between those pregnant women whose pregnancies are complicated by IUGR and PE, vs. those whose pregnancies continue to term and whose babies are delivered at normal birthweight.

6.2 Appendix.2: Ethical review

ETHICAL REVIEW PANEL

APPLICATION FORM (STAFF AND PGR STUDENT RESEARCH PROJECTS)

Section A - Applicant's details

A1	Project title	Urinary proteome analysis in Pre-eclamptic and Intrauterine Growth Restricted Pregnancies
A2	Name of researcher	Dr Sarah Hart/ Miss Wafaa Al-Jasim
	Research Institute or School	ISTM
A4	Correspondence address	Huxley Building, School of Life Sciences, Keele University, ST5 5GB
A5	Keele E-mail address	s.r.hart@keele.ac.uk
	Work telephone number	01782 733759
	Type of application	PGR Student Project
A7	Please give supervisor name and contact details if PGR application	Dr Sarah Hart
	Project start date	July 2016
	Project end date	September 2017
A10		

Section B - Project details

B1	<p>In lay terms provide a brief summary of the project including the background and rationale for the proposed research and the hypotheses or research question(s) (max 500 words).</p> <p>Project Team: This application concerns a PhD studentship (MOHESR, Wafaa Al Jasim) supervised by Dr Sarah Hart and Dr Paul Roach, Keele University.</p> <p>Pre-eclampsia (PE) is a pregnancy complication which affects ~5% of first pregnancies. In second and subsequent pregnancies the risks are somewhat quantifiable (largely based upon PE occurrence in first pregnancy); whilst some pre-disposing risk factors are known (maternal age, BMI, ethnicity), in first pregnancies the risk is largely unquantifiable. Intrauterine growth restriction (IUGR) is where babies are not growing as fast as they should in the womb; those babies who are less than 10% of the normal birth weight are termed "small for gestational age" (SGA). IUGR has many causes, including poor placental function, and can be further complicated by PE.</p> <p>Because of the prevalence of PE and IUGR, and a lack of predictive algorithms, screening antenatally for these conditions forms a large part of the current obstetric care program in the UK, at a significant cost to the NHS (typically during a first pregnancy, an average woman has 10 antenatal visits). Developing risk prediction tools is therefore of strong potential benefit to the national and global health economy.</p> <p>Urine is a biofluid which is routinely used in during pregnancy for the diagnosis of metabolic problems and pre-eclampsia (by detection of elevated protein levels). Use of this fluid in clinical diagnostics has high acceptability to patients (in contrast with blood testing); a urine-based test is therefore a highly desirable outcome.</p> <p>This study aims to characterise large biological molecules (proteins and peptides) within urine samples from women, comparing those with pre-eclampsia and intrauterine growth restriction with those with normal pregnancy outcomes. For this study we are acting in collaboration with the INFANT centre, University of Cork's SCOPE biobank of fully-consented patient samples and data; this is an ongoing large-scale study of pregnancy outcomes. Our working hypothesis is that protein differences will be present in the urinary output of women who go on to develop pre-eclampsia, predating any observable symptoms. Detection of differences</p>
----	---

will enable us to begin to develop new diagnostic tools which could predict risk of pre-eclampsia for first-time pregnant mothers.

SCOPE is a cohort of healthy first time mums, with the main outcomes of interest being late pregnancy complications such as Small for Gestational age (SGA), Pre-eclampsia (PE) and Pre-term Labour (PTL). Through the collaboration of women and scientists across the globe, SCOPE vision was to predict and prevent the major diseases of late pregnancy. With the valued participation of women around the world, the SCOPE study has established a unique pregnancy biobank for scientific discovery to develop tests that predict these conditions. The biobank is rich with clinical data and studies can be designed based on outcomes and according to the participants' wishes for inclusion in studies based on their consents. The SCOPE Patient Information Leaflet and Consent Form are appended to this application.

B2 In lay terms outline the type of procedure(s) and/or research methodology (eg observational, questionnaire, interviews, experimental) to be employed (**max 500 words**).

This experimental study will involve the measurement of biological macromolecules (proteins and peptides) in the urine of pregnant human donors with pre-eclampsia, small for gestational age fetuses or normal pregnancy outcome in a laboratory setting (Keele University, School of Life Sciences) using biological mass spectrometry. A total of 120 human donor samples will be analysed, all of which are supplied by the SCOPE biobank, specifically for research purposes. Samples will be stored in HTA-designated freezers dedicated for laboratory work, and any resulting waste disposed of in accordance with local guidelines.

Samples will be processed by spinning out any uroepithelial cells (rendering the samples acellular) prior to undertaking protein separation via ultracentrifugal filtration and analysis by spectroscopic methods (mass spectrometry and infrared spectroscopy).

No amplification or analysis of human DNA will take place.

Note: this study will not involve any recruitment or interaction with human subjects and will rely on previously collected, ethically approved samples from a pre-existing biobank, with access to anonymised, coded data derived from participant's medical records within a database being provided on request via the SCOPE co-ordinators. The clinical data together with analytical data will ultimately be used to build a preliminary risk model. The applicants have previous experience in the techniques relevant to this application.

Diagram Appended

SCOPE PIL and Consent form appended to application

Diagrams or flow charts that would aid clarification of the research should be attached if appropriate (these attachments will not be included in the word count).

Remember to attach questionnaire or interview questions.

B3 Describe the characteristics of the participant group, and the inclusion and exclusion criteria. Indicate the sample size, with an explanation of how this sample size was decided/calculated.

This study will utilise urine samples, collected from fully consented donors, and obtained via the SCOPE team; one aim of the study is supplying human samples for pregnancy research purposes, therefore broad approval was gained at the time of consent. The proposal for this research has been approved by the SCOPE board (as RAF 3.45), which includes permission for use of these samples.

Sample size calculations:

The use of multi-omics data as an endpoint, and the novelty of our investigation, gives unique challenges in determining sample size. Our previous experiences, and those of the SCOPE consortium, therefore inform our sample size and proteome coverage depth respectively.

Section C - Ethically sensitive, challenging or issues of risk

C1	<p>Will the research involve deceased persons, body parts or other human elements such as blood, hair or tissue samples (including saliva and waste products)?</p>	Yes <input checked="" type="checkbox"/>	No <input type="checkbox"/>
<p>If yes, please discuss this project with Dr Alan Harper, Human Tissue Officer on 01782 674472 / 734654 or e-mail a.g.s.harper@keele.ac.uk. Please cite the reference number given by Dr Harper for this research project below:-</p> <p>1706AGSH2016</p> <p>If yes, please give details with reference to the Human Tissue Act 2004.</p> <p>Urine samples are considered relevant materials under the HTA 2004. Appropriate logging of daughter fractions generated from these samples has been discussed in detail with the HTA officer, with guidance and a protocol for sample tracking and logging of sample processing and disposal being provided. The initial step in the preparation of all samples will be centrifugation to pellet and remove uroepithelial material shed from the urinary tract – at this point the samples will be rendered effectively acellular, significantly reducing the risks associated with their use. Protein extraction procedures will then be undertaken on the residual acellular material.</p>			
<p>Human Tissue Act can be accessed via https://www.hta.gov.uk/human-tissue-act-2004</p>			

C2	<p>Outline any potential risks to individuals, participants and members of the research team; the measures that will be taken to minimise risk; and the procedures that will be adopted in the event of an adverse event.</p>		
<p>Human samples will only be handled in a designated area within appropriately equipped laboratories at Keele University, observing or exceeding all biological safety rules in place at the time of experimentation. Standard operating protocols will be followed for all laboratory procedures.</p>			
<p>The University's Lone Working Policy can be accessed via http://www.keele.ac.uk/dohs/a2z/loneworking/</p>			

C3	<p>Will the research be undertaken overseas?</p>	Yes <input type="checkbox"/>	No <input checked="" type="checkbox"/>
<p>If yes, have you consulted the foreign and commonwealth office website for guidance/travel advice and is it safe to travel there?</p>			
		Yes <input type="checkbox"/>	No <input type="checkbox"/>
<p>Have you completed and submitted a risk assessment form?</p>			
		Yes <input type="checkbox"/>	No <input type="checkbox"/>
<p>Foreign and Commonwealth Office travel advice website: https://www.gov.uk/foreign-travel-advice</p> <p>Overseas Travel Policy and risk assessment form (covers both Staff and PGR students) is available from http://www.keele.ac.uk/finance/insurance/travelinsurance/travellingoverseas-policyriskassessment/</p>			

C4	<p>Will the research include children or vulnerable adults such as individuals with a learning disability or individuals with cognitive impairment?</p>	Yes <input type="checkbox"/>	No <input checked="" type="checkbox"/>
----	---	------------------------------	--

C5	<p>Will participants be deceived in any way about the study?</p>	Yes <input type="checkbox"/>	No <input checked="" type="checkbox"/>
----	--	------------------------------	--

If **yes**, describe the nature and extent of deception involved, including how and when this deception will be revealed and who will administer this feedback (debrief).

SECTION D - Recruitment & consent process

D1 Indicate how potential participants will be identified, approached and recruited and outline any relationship between the researcher and potential participant.

This study will utilise material collected by a biobanking study (SCOPE) and as such no patient recruitment is necessary.

Remember to attach copies of posters, advertisements, invitation letters/e-mails to be used as part of the recruitment process with version numbers included in the footer.

D2 Describe the process that will be used to seek and obtain informed consent.

All samples were obtained within the SCOPE study, with full informed consent to use the samples for future research being obtained from participants by the originating investigators at UCC. Ethical approval was obtained from the local ethics committee at Cork (ECM5(10)05/02/08); all women provided written informed consent, which is retained by the originating investigators.

Broad ethical approval for commercial and non-commercial research outwith the original aims of SCOPE was consented at the time of original recruitment and consenting (see Participant Consent Form, attached). Recruitment has now finished, but during recruitment participants were made aware that they could withdraw from the study at any point (Patient Information Leaflet p2, Participant Consent Form Q2).

If you answered yes to C4 (your participants are children or vulnerable adults) explain how you will ensure that individuals in these groups are competent to give consent to participate in this study.

N/A

Remember to attach your information sheet and consent form with versions numbers & date included in the footer
Templates available from <http://www.keele.ac.uk/researchsupport/researchethics/>

D3 Will consent be sought to use the data for other research?

Yes

No

Will consent be sought to contact the individual to participate in future research?

Yes

No

D4 Can participants withdraw from the research?

Yes

No

If **yes**, state up to what point participants are able to withdraw from the research

If **yes**, outline how participants will be informed of their right to withdraw, how they can do this and what will happen to their data if they withdraw.

If **no**, explain why they cannot withdraw (eg anonymous survey).

Samples are entirely anonymised, detaching them from any patient-identifiable information, with full informed consent having been taken during the sample collection phase of the project. Thus, at the point where samples are provided to Keele, re-contact is not possible. All samples are denoted by a unique identifier provided within the SCOPE database.

SECTION E - Confidentiality and anonymity

E1	<p>Outline the procedures that will be used to protect, as far as possible, the anonymity of participants and/or confidentiality of data during the conduct of the research and in the release of its findings.</p> <p>Our SCOPE collaborators (from whom we are purchasing the samples) provide no patient-identifiable information within database entries, thus anonymity is ensured.</p>
----	--

SECTION F - Storage, access to, management of, and disposal of data

F1	<p>Describe the research data that will be stored; where it will be stored and for how long; the measures that will be put in place to ensure the security of data; who will have access to the data; long term data management plans following completion of the project; and how/when data will be disposed of.</p> <p>Spectroscopic data (raw files) and identification (Mascot and Scaffold) data will be stored at Keele University on a password protected hard drive, with robust backup facilities via networked storage. No patient identifiable data will be available to the research team. Data will be accessible initially to Dr Hart and Ms Al-Jasim, with sharing of outcomes with collaborators at Cork, Manchester and Leicester, but will be deposited in a public data archive, for example the PRoteomics IDentifications (PRIDE) Archive hosted by EBI, coinciding with publication (in accordance with the guidelines of most scientific journals). Data will be retained for a minimum of 5 years following the conclusion of this project.</p>
----	---

SECTION G - Other ethical issues raised by the research

G1	<p>Are there any other ethical issues that may be raised by the research? Yes <input type="checkbox"/> No <input checked="" type="checkbox"/></p> <p>If yes, please give details.</p>
----	---

SECTION H - Other approvals required


H1	<p>Does the project require researcher(s) to have a Disclosure and Barring Service (DBS) check? Yes <input type="checkbox"/> No <input checked="" type="checkbox"/></p> <p>If yes, have you attached a confirmation of satisfactory DBS check memo? Yes <input type="checkbox"/> No <input type="checkbox"/></p> <p>Does the project require National Offender Management Service (NOMS) approval? Yes <input type="checkbox"/> No <input checked="" type="checkbox"/></p> <p>Does the project require NHS Research Development (R&D) Approval? Yes <input type="checkbox"/> No <input checked="" type="checkbox"/></p> <p>Does the project require approval from another organisation? Yes <input type="checkbox"/> No <input checked="" type="checkbox"/></p>
----	--


SECTION I - Checklist


I1	<p>Please list the documents attached to this application</p> <table border="1"><tr><td>DocumentVersion number Schematic workflow1.0</td><td>Date 27/6/16</td></tr></table>	DocumentVersion number Schematic workflow1.0	Date 27/6/16
DocumentVersion number Schematic workflow1.0	Date 27/6/16		

SECTION J - Declarations

J1	<p>Declaration by researcher</p>
----	---

	<p>I confirm that:-</p> <ul style="list-style-type: none"> • The form is accurate to the best of my knowledge • I will abide by the University's ethical requirements • I will inform the panel of any changes to the project • I am aware of my responsibility to be up to date and comply with the requirements of the law and any relevant professional guidelines
	<p>Researcher name (in capitals) Miss Wafaa Al Jasim</p>
	<p>Researcher signature </p>
	<p>Date 4/8/16</p>

J2	<p>Declaration by supervisor (PGR applications only)</p> <p>I confirm that:-</p> <ul style="list-style-type: none"> • The application has been appropriately peer reviewed • I have read the application and am happy for it to proceed for ethical review • The application is accurate to the best of my knowledge • The project will comply with the University's ethical requirements • The applicant will inform the panel of any changes to the project • I am aware of my responsibility to ensure that the applicant is familiar with and complies with the requirements of the law and any relevant professional guidelines
	<p>Supervisor name (in capitals) Dr Sarah Hart</p>
	<p>Supervisor signature </p>
	<p>Date 4/8/16</p>

J3	<p>Declaration by Faculty Research Office/Research Institute Director, Centre/Theme Head, or Head of School (Staff applications only)</p> <p>I confirm that:-</p> <ul style="list-style-type: none"> • The application has been appropriately peer reviewed • I have read the application and am happy for it to proceed for ethical review • The form is accurate to the best of my knowledge • The project will comply with the University's ethical requirements • The applicant will inform the panel of any changes to the project • I am aware of my responsibility to ensure that the applicant is familiar with and complies with the requirements of the law and any relevant professional guidelines
	<p>Name (in capitals) PROF NICHOLAS R FORSYTH</p>
	<p>Post INSTITUTE DIRECTOR - ISTM</p>
	<p>Signature </p>
	<p>Date 29/06/16</p>



No

Ref: ERP380

8th August 2016

Dr Sarah Hart/Miss Wafaa Al-Jasiim
ISTM
Keele University

Dear Sarah and Wafaa,

Re: Urinary proteome analysis in Pre-eclamptic and Intrauterine Growth restricted Pregnancies

Thank you for submitting your revised application for review. I am pleased to inform you that your application has been approved by the Ethics Review Panel. The following documents have been reviewed and approved by the panel as follows:

Document(s)	Version Number	Date
ERP Application Form	1	05-10-2016
Schematic Workflow	1	05-10-2016

If the fieldwork goes beyond the date stated in your application, **30th September 2017**, or there are any other amendments to your study you must submit an 'application to amend study' form to the ERP administrator at research.erps@keele.ac.uk stating **ERP3** in the subject line of the e-mail. This form is available via <http://www.keele.ac.uk/researchsupport/researchethics/>

If you have any queries, please do not hesitate to contact me via the ERP administrator on research.erps@keele.ac.uk stating **ERP3** in the subject line of the e-mail.

Yours sincerely

Dr Helen Price
Vice Chair – Ethical Review Panel

CC RI Manager

Directorate of Engagement & Partnerships
T: +44(0)1782 734467

Keele University, Staffordshire ST5 5BG, UK
www.keele.ac.uk +44 (0)1782 732000

6.3 Appendix 3: Dataset of all protein expression data for PE and IUGR with their controls

Protein expression dataset that shows significant changes compared with control using confidence interval 99 % at FDR< 0.01 for PE and IUGR respectively at both 15 and 20 week gestational age are presented in supplementary Tables (S2-S4) respectively as below:

Table S 1: Dataset of protein expression for PE and control at 15 weeks gestational age

No.	Protein FDR Confidence : Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of PE at 15 W	Abundances of reporter ion of control at 15 W	Abundance ratio (Case / Control) of PE at 15 W	Log ₂ fold change
1	High	P01023	Alpha-2-macroglobulin OS=Homo sapiens OX=9606 GN=A2M PE=1 SV=3	2	163.2	234.4	20.7	11.33	3.502
2	High	P02675	Fibrinogen beta chain OS=Homo sapiens OX=9606 GN=FGB PE=1 SV=2	2	55.9	159.4	16.9	9.415	3.235
3	High	Q6GTX8	Leukocyte-associated immunoglobulin-like receptor 1 OS=Homo sapiens OX=9606 GN=LAIR1 PE=1 SV=1	2	31.4	143.4	15.6	9.223	3.205
4	High	Q8IYS5	Osteoclast-associated immunoglobulin-like receptor OS=Homo sapiens OX=9606 GN=OSCAR PE=1 SV=3	2	30.5	42.4	5.4	7.89	2.980
5	High	P00738	Haptoglobin OS=Homo sapiens OX=9606 GN=HP PE=1 SV=1	4	45.2	219.5	38	5.78	2.531
6	High	P02788	Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF PE=1 SV=6	3	78.1	273.3	50.6	5.404	2.434
7	High	P06733	Alpha-enolase OS=Homo sapiens OX=9606 GN=ENO1 PE=1 SV=2	4	47.1	98.4	18.4	5.351	2.420

8	High	P13796	Plastin-2 OS=Homo sapiens OX=9606 GN=LCP1 PE=1 SV=6	4	70.2	177.4	37.8	4.692	2.230
9	High	Q02487	Desmocollin-2 OS=Homo sapiens OX=9606 GN=DSC2 PE=1 SV=1	3	99.9	38.7	10.2	3.795	1.924
10	High	P62805	Histone H4 OS=Homo sapiens OX=9606 GN=HIST1H4A PE=1 SV=2	3	11.4	156.6	45.8	3.416	1.772
11	High	P08185	Corticosteroid-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA6 PE=1 SV=1	3	45.1	113	36.1	3.132	1.647
12	High	P60174	Triosephosphate isomerase OS=Homo sapiens OX=9606 GN=TPI1 PE=1 SV=3	2	30.8	76	28.3	2.686	1.425
13	High	P18510	Interleukin-1 receptor antagonist protein OS=Homo sapiens OX=9606 GN=IL1RN PE=1 SV=1	2	20	37	15.4	2.4	1.263
14	High	P05164	Myeloperoxidase OS=Homo sapiens OX=9606 GN=MPO PE=1 SV=1	3	83.8	110.7	47.2	2.347	1.231
15	High	P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens OX=9606 GN=KRT2 PE=1 SV=2	5	65.4	57.7	24.8	2.329	1.220
16	High	P25815	Protein S100-P OS=Homo sapiens OX=9606 GN=S100P PE=1 SV=2	2	10.4	109.3	47.6	2.293	1.197
17	High	Q9ULI3	Protein HEG homolog 1 OS=Homo sapiens OX=9606 GN=HEG1 PE=1 SV=3	3	147.4	44.9	21.9	2.048	1.034
18	High	P00450	Ceruloplasmin OS=Homo sapiens OX=9606 GN=CP PE=1 SV=1	13	122.1	82	40.6	2.022	1.016
19	High	P02787	Serotransferrin OS=Homo sapiens OX=9606 GN=TF PE=1 SV=3	12	77	88.1	45.2	1.952	0.965
20	High	P19652	Alpha-1-acid glycoprotein 2 OS=Homo sapiens OX=9606 GN=ORM2 PE=1 SV=2	2	23.6	47.5	25.2	1.886	0.915
21	High	Q93088	Betaine--homocysteine S- methyltransferase 1 OS=Homo sapiens OX=9606 GN=BHMT PE=1 SV=2	3	45	57.5	32.8	1.754	0.811

22	High	P60709	Actin, cytoplasmic 1 OS=Homo sapiens OX=9606 GN=ACTB PE=1 SV=1	6	41.7	80.5	49.6	1.623	0.699
23	High	P01859	Immunoglobulin heavy constant gamma 2 OS=Homo sapiens OX=9606 GN=IGHG2 PE=1 SV=2	4	35.9	94.7	60.8	1.557	0.639
24	High	P02763	Alpha-1-acid glycoprotein 1 OS=Homo sapiens OX=9606 GN=ORM1 PE=1 SV=1	3	23.5	29.2	19.2	1.526	0.610
25	High	P04075	Fructose-bisphosphate aldolase A OS=Homo sapiens OX=9606 GN=ALDOA PE=1 SV=2	2	39.4	80.1	52.7	1.52	0.604
26	High	P29972	Aquaporin-1 OS=Homo sapiens OX=9606 GN=AQP1 PE=1 SV=3	2	28.5	40.4	26.9	1.501	0.586
27	High	P02768	Serum albumin OS=Homo sapiens OX=9606 GN=ALB PE=1 SV=2	19	69.3	91	61.1	1.491	0.576
28	High	P0DOY2	Immunoglobulin lambda constant 2 OS=Homo sapiens OX=9606 GN=IGLC2 PE=1 SV=1	3	11.3	43.2	31.7	1.366	0.450
29	High	P01876	Immunoglobulin heavy constant alpha 1 OS=Homo sapiens OX=9606 GN=IGHA1 PE=1 SV=2	5	37.6	66	49.2	1.342	0.424
30	High	Q14515	SPARC-like protein 1 OS=Homo sapiens OX=9606 GN=SPARCL1 PE=1 SV=2	4	75.2	35.9	27	1.33	0.411
31	High	P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens OX=9606 GN=HSPA8 PE=1 SV=1	9	70.9	71.5	54.8	1.304	0.383
32	High	P04217	Alpha-1B-glycoprotein OS=Homo sapiens OX=9606 GN=A1BG PE=1 SV=4	3	54.2	62.5	48.7	1.282	0.358
33	High	P01019	Angiotensinogen OS=Homo sapiens OX=9606 GN=AGT PE=1 SV=1	2	53.1	45.9	37.7	1.216	0.282
34	High	P02750	Leucine-rich alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=LRG1 PE=1 SV=2	5	38.2	31	26	1.193	0.255

35	High	P06703	Protein S100-A6 OS=Homo sapiens OX=9606 GN=S100A6 PE=1 SV=1	3	10.2	76.1	64.1	1.188	0.249
36	High	P01614	Immunoglobulin kappa variable 2D-40 OS=Homo sapiens OX=9606 GN=IGKV2D-40 PE=1 SV=2	2	13.3	52.8	44.5	1.188	0.249
37	High	P07476	Involucrin OS=Homo sapiens OX=9606 GN=IVL PE=1 SV=2	9	68.4	25.9	21.8	1.187	0.247
38	High	P01871	Immunoglobulin heavy constant mu OS=Homo sapiens OX=9606 GN=IGHM PE=1 SV=4	2	49.4	48.1	42	1.147	0.198
39	High	P01834	Immunoglobulin kappa constant OS=Homo sapiens OX=9606 GN=IGKC PE=1 SV=2	3	11.8	46.6	41.4	1.126	0.171
40	High	P01857	Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1 SV=1	7	36.1	60.4	54.2	1.115	0.157
41	High	P10599	Thioredoxin OS=Homo sapiens OX=9606 GN=TXN PE=1 SV=3	3	11.7	70.6	63.4	1.115	0.157
42	High	P80188	Neutrophil gelatinase-associated lipocalin OS=Homo sapiens OX=9606 GN=LCN2 PE=1 SV=2	3	22.6	61.5	55.7	1.104	0.143
43	High	P09211	Glutathione S-transferase P OS=Homo sapiens OX=9606 GN=GSTP1 PE=1 SV=2	2	23.3	33.4	30.4	1.097	0.134
44	High	P12273	Prolactin-inducible protein OS=Homo sapiens OX=9606 GN=PIP PE=1 SV=1	2	16.6	31.1	29.5	1.057	0.080
45	High	P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens OX=9606 GN=PTGDS PE=1 SV=1	3	21	47.2	45.7	1.033	0.047
46	High	Q08345	Epithelial discoidin domain-containing receptor 1 OS=Homo sapiens OX=9606 GN=DDR1 PE=1 SV=1	4	101.1	64.8	64.8	1.001	0.001
47	High	P00924	Enolase 1 OS=Saccharomyces	14	46.8	100	100	1	0.000

			cerevisiae (strain ATCC 204508 / S288c) GN=ENO1 PE=1 SV=3						
48	High	B9A064	Immunoglobulin lambda-like polypeptide 5 OS=Homo sapiens OX=9606 GN=IGLL5 PE=2 SV=2	3	23	47.7	47.8	0.999	-0.001
49	High	P31947	14-3-3 protein sigma OS=Homo sapiens OX=9606 GN=SFN PE=1 SV=1	5	27.8	58.4	59.6	0.981	-0.028
50	High	P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=AZGP1 PE=1 SV=2	13	34.2	22.8	23.7	0.963	-0.054
51	High	Q8NBJ4	Golgi membrane protein 1 OS=Homo sapiens OX=9606 GN=GOLM1 PE=1 SV=1	7	45.3	34.3	35.7	0.96	-0.059
52	High	Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1 OS=Homo sapiens OX=9606 GN=LYVE1 PE=1 SV=2	2	35.2	31.8	34.1	0.932	-0.102
53	High	P02774	Vitamin D-binding protein OS=Homo sapiens OX=9606 GN=GC PE=1 SV=1	5	52.9	55.2	59.4	0.93	-0.105
54	High	Q96NY8	Nectin-4 OS=Homo sapiens OX=9606 GN=NECTIN4 PE=1 SV=1	3	55.4	51.7	57.3	0.903	-0.147
55	High	P05060	Secretogranin-1 OS=Homo sapiens OX=9606 GN=CHGB PE=1 SV=2	2	78.2	28.1	31.2	0.902	-0.149
56	High	P01011	Alpha-1-antichymotrypsin OS=Homo sapiens OX=9606 GN=SERPINA3 PE=1 SV=2	6	47.6	34.2	38.6	0.884	-0.178
57	High	P01009	Alpha-1-antitrypsin OS=Homo sapiens OX=9606 GN=SERPINA1 PE=1 SV=3	14	46.7	35.1	39.8	0.881	-0.183
58	High	P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens OX=9606 GN=KRT9 PE=1 SV=3	5	62	44.4	51.6	0.86	-0.218
59	High	Q9BY67	Cell adhesion molecule 1 OS=Homo sapiens OX=9606 GN=CADM1 PE=1 SV=2	3	48.5	35.3	41.2	0.857	-0.223
60	High	P01591	Immunoglobulin J chain OS=Homo	2	18.1	38.4	45	0.854	-0.228

			sapiens OX=9606 GN=JCHAIN PE=1 SV=4						
61	High	P04083	Annexin A1 OS=Homo sapiens OX=9606 GN=ANXA1 PE=1 SV=2	6	38.7	15.4	18	0.852	-0.231
62	High	P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens OX=9606 GN=KRT10 PE=1 SV=6	14	58.8	57.4	67.6	0.85	-0.234
63	High	P01040	Cystatin-A OS=Homo sapiens OX=9606 GN=CSTA PE=1 SV=1	2	11	29.1	34.7	0.84	-0.252
64	High	P10451	Osteopontin OS=Homo sapiens OX=9606 GN=SPP1 PE=1 SV=1	11	35.4	16.4	19.9	0.828	-0.272
65	High	P06702	Protein S100-A9 OS=Homo sapiens OX=9606 GN=S100A9 PE=1 SV=1	7	13.2	44.8	55	0.814	-0.297
66	High	P01877	Immunoglobulin heavy constant alpha 2 OS=Homo sapiens OX=9606 GN=IGHA2 PE=1 SV=4	3	36.6	53.4	66.9	0.799	-0.324
67	High	P06396	Gelsolin OS=Homo sapiens OX=9606 GN=GSN PE=1 SV=1	6	85.6	28.6	35.9	0.797	-0.327
68	High	Q6UXB8	Peptidase inhibitor 16 OS=Homo sapiens OX=9606 GN=PI16 PE=1 SV=1	2	49.4	31.8	41.2	0.77	-0.377
69	High	P24855	Deoxyribonuclease-1 OS=Homo sapiens OX=9606 GN=DNASE1 PE=1 SV=1	3	31.4	25.2	32.9	0.768	-0.381
70	High	P05543	Thyroxine-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA7 PE=1 SV=2	6	46.3	35.3	46.3	0.761	-0.394
71	High	P11021	Endoplasmic reticulum chaperone BiP OS=Homo sapiens OX=9606 GN=HSPA5 PE=1 SV=2	5	72.3	39.6	52	0.761	-0.394
72	High	P01833	Polymeric immunoglobulin receptor OS=Homo sapiens OX=9606 GN=PIGR PE=1 SV=4	7	83.2	34.5	46.2	0.748	-0.419
73	High	P02765	Alpha-2-HS-glycoprotein OS=Homo	3	39.3	34.6	46.3	0.747	-0.421

			sapiens OX=9606 GN=AHSG PE=1 SV=1						
74	High	P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens OX=9606 GN=KRT1 PE=1 SV=6	17	66	59.2	79.4	0.746	-0.423
75	High	P50995	Annexin A11 OS=Homo sapiens OX=9606 GN=ANXA11 PE=1 SV=1	3	54.4	41.9	56.6	0.741	-0.432
76	High	Q9UBG3	Cornulin OS=Homo sapiens OX=9606 GN=CRNN PE=1 SV=1	4	53.5	23.9	32.4	0.736	-0.442
77	High	P07355	Annexin A2 OS=Homo sapiens OX=9606 GN=ANXA2 PE=1 SV=2	8	38.6	30.9	42.8	0.722	-0.470
78	High	P15144	Aminopeptidase N OS=Homo sapiens OX=9606 GN=ANPEP PE=1 SV=4	12	109.5	43.1	60.4	0.713	-0.488
79	High	P05109	Protein S100-A8 OS=Homo sapiens OX=9606 GN=S100A8 PE=1 SV=1	5	10.8	44.2	63.7	0.694	-0.527
80	High	P34059	N-acetylgalactosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GALNS PE=1 SV=1	2	58	22.3	32.4	0.686	-0.544
81	High	P02751	Fibronectin OS=Homo sapiens OX=9606 GN=FN1 PE=1 SV=4	11	262.5	41.1	60.2	0.683	-0.550
82	High	P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens OX=9606 GN=LRP2 PE=1 SV=3	6	521.6	42.6	62.4	0.683	-0.550
83	High	Q6UVK1	Chondroitin sulfate proteoglycan 4 OS=Homo sapiens OX=9606 GN=CSPG4 PE=1 SV=2	2	250.4	39.3	57.8	0.68	-0.556
84	High	P12830	Cadherin-1 OS=Homo sapiens OX=9606 GN=CDH1 PE=1 SV=3	8	97.4	25.1	37	0.679	-0.559
85	High	O60494	Cubilin OS=Homo sapiens OX=9606 GN=CUBN PE=1 SV=5	3	398.5	46.6	68.8	0.676	-0.565
86	High	O00187	Mannan-binding lectin serine protease 2 OS=Homo sapiens OX=9606 GN=MASP2 PE=1 SV=4	3	75.7	19.5	29	0.673	-0.571
87	High	P13647	Keratin, type II cytoskeletal 5 OS=Homo	11	62.3	36.4	54.3	0.671	-0.576

			sapiens OX=9606 GN=KRT5 PE=1 SV=3						
88	High	P04746	Pancreatic alpha-amylase OS=Homo sapiens OX=9606 GN=AMY2A PE=1 SV=2	7	57.7	35.2	52.5	0.671	-0.576
89	High	P50895	Basal cell adhesion molecule OS=Homo sapiens OX=9606 GN=BCAM PE=1 SV=2	3	67.4	16.4	25	0.655	-0.610
90	High	O00391	Sulfhydryl oxidase 1 OS=Homo sapiens OX=9606 GN=QSOX1 PE=1 SV=3	5	82.5	23.1	35.6	0.648	-0.626
91	High	P05090	Apolipoprotein D OS=Homo sapiens OX=9606 GN=APOD PE=1 SV=1	3	21.3	51.3	80.1	0.641	-0.642
92	High	P13473	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens OX=9606 GN=LAMP2 PE=1 SV=2	2	44.9	28.7	44.9	0.64	-0.644
93	High	P62873	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1 OS=Homo sapiens OX=9606 GN=GNB1 PE=1 SV=3	2	37.4	45.3	71.1	0.638	-0.648
94	High	P05154	Plasma serine protease inhibitor OS=Homo sapiens OX=9606 GN=SERPINA5 PE=1 SV=3	4	45.6	33.5	53.1	0.631	-0.664
95	High	P09603	Macrophage colony-stimulating factor 1 OS=Homo sapiens OX=9606 GN=CSF1 PE=1 SV=2	4	60.1	21.9	35.2	0.622	-0.685
96	High	P15586	N-acetylglucosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GNS PE=1 SV=3	3	62	39.9	64.4	0.619	-0.692
97	High	P78324	Tyrosine-protein phosphatase non-receptor type substrate 1 OS=Homo sapiens OX=9606 GN=SIRPA PE=1 SV=2	3	54.9	32.2	52.5	0.613	-0.706
98	High	P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens OX=9606 GN=KRT16 PE=1 SV=4	6	51.2	23.1	37.8	0.611	-0.711
99	High	Q68CJ9	Cyclic AMP-responsive element-binding	2	49	28.4	46.6	0.608	-0.718

			protein 3-like protein 3 OS=Homo sapiens OX=9606 GN=CREB3L3 PE=1 SV=2						
100	High	P02790	Hemopexin OS=Homo sapiens OX=9606 GN=HPX PE=1 SV=2	5	51.6	51.5	85.8	0.6	-0.737
101	High	A8K2U0	Alpha-2-macroglobulin-like protein 1 OS=Homo sapiens OX=9606 GN=A2ML1 PE=1 SV=3	6	161	32.6	54.6	0.597	-0.744
102	High	P06870	Kallikrein-1 OS=Homo sapiens OX=9606 GN=KLK1 PE=1 SV=2	3	28.9	37.5	62.9	0.597	-0.744
103	High	P62736	Actin, aortic smooth muscle OS=Homo sapiens OX=9606 GN=ACTA2 PE=1 SV=1	5	42	23.1	38.8	0.595	-0.749
104	High	P26992	Ciliary neurotrophic factor receptor subunit alpha OS=Homo sapiens OX=9606 GN=CNTFR PE=1 SV=2	2	40.6	35.4	60.4	0.587	-0.769
105	High	P07339	Cathepsin D OS=Homo sapiens OX=9606 GN=CTSD PE=1 SV=1	6	44.5	41.7	71.9	0.58	-0.786
106	High	Q01459	Di-N-acetylchitobiase OS=Homo sapiens OX=9606 GN=CTBS PE=1 SV=1	2	43.7	23.2	40.3	0.576	-0.796
107	High	P08571	Monocyte differentiation antigen CD14 OS=Homo sapiens OX=9606 GN=CD14 PE=1 SV=2	4	40.1	25.4	44.2	0.574	-0.801
108	High	Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens OX=9606 GN=ITIH4 PE=1 SV=4	4	103.3	22.7	39.6	0.573	-0.803
109	High	P0C0L5	Complement C4-B OS=Homo sapiens OX=9606 GN=C4B PE=1 SV=2	3	192.6	31.4	54.8	0.573	-0.803
110	High	P07911	Uromodulin OS=Homo sapiens OX=9606 GN=UMOD PE=1 SV=1	7	69.7	31.5	56.8	0.555	-0.849
111	High	Q9UQ72	Pregnancy-specific beta-1-glycoprotein 11 OS=Homo sapiens OX=9606 GN=PSG11 PE=2 SV=3	7	37.1	28.3	51.1	0.553	-0.855

112	High	P10909	Clusterin OS=Homo sapiens OX=9606 GN=CLU PE=1 SV=1	4	52.5	34.7	62.7	0.553	-0.855
113	High	P19835	Bile salt-activated lipase OS=Homo sapiens OX=9606 GN=CEL PE=1 SV=3	3	79.3	35.2	63.7	0.553	-0.855
114	High	P05155	Plasma protease C1 inhibitor OS=Homo sapiens OX=9606 GN=SERPING1 PE=1 SV=2	6	55.1	30.4	55.5	0.547	-0.870
115	High	Q9UBC9	Small proline-rich protein 3 OS=Homo sapiens OX=9606 GN=SPRR3 PE=1 SV=2	5	18.1	27.5	50.3	0.546	-0.873
116	High	Q9NZP8	Complement C1r subcomponent-like protein OS=Homo sapiens OX=9606 GN=C1RL PE=1 SV=2	3	53.5	30.1	55.7	0.541	-0.886
117	High	P02760	Protein AMBP OS=Homo sapiens OX=9606 GN=AMBP PE=1 SV=1	5	39	28.1	52.5	0.535	-0.902
118	High	P12109	Collagen alpha-1(VI) chain OS=Homo sapiens OX=9606 GN=COL6A1 PE=1 SV=3	10	108.5	33.8	63.3	0.534	-0.905
119	High	Q13510	Acid ceramidase OS=Homo sapiens OX=9606 GN=ASA1 PE=1 SV=5	3	44.6	41.2	77.4	0.532	-0.911
120	High	Q08380	Galectin-3-binding protein OS=Homo sapiens OX=9606 GN=LGALS3BP PE=1 SV=1	10	65.3	30.5	57.9	0.526	-0.927
121	High	P01042	Kininogen-1 OS=Homo sapiens OX=9606 GN=KNG1 PE=1 SV=2	6	71.9	34.8	67.7	0.514	-0.960
122	High	O43451	Maltase-glucoamylase, intestinal OS=Homo sapiens OX=9606 GN=MGAM PE=1 SV=5	4	209.7	39.1	77.4	0.505	-0.986
123	High	Q8NFZ8	Cell adhesion molecule 4 OS=Homo sapiens OX=9606 GN=CADM4 PE=1 SV=1	2	42.8	36.8	72.9	0.504	-0.989
124	High	Q9BRK3	Matrix remodeling-associated protein 8 OS=Homo sapiens OX=9606	4	49.1	26.6	52.9	0.503	-0.991

			GN=MXRA8 PE=1 SV=1						
125	High	P09668	Pro-cathepsin H OS=Homo sapiens OX=9606 GN=CTSH PE=1 SV=4	2	37.4	25.9	51.5	0.502	-0.994
126	High	P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens OX=9606 GN=KRT6B PE=1 SV=5	20	60	24.7	51	0.484	-1.047
127	High	P11464	Pregnancy-specific beta-1-glycoprotein 1 OS=Homo sapiens OX=9606 GN=PSG1 PE=1 SV=1	4	47.2	29.3	60.8	0.482	-1.053
128	High	P39059	Collagen alpha-1(XV) chain OS=Homo sapiens OX=9606 GN=COL15A1 PE=1 SV=2	2	141.6	44.7	94.2	0.475	-1.074
129	High	Q16270	Insulin-like growth factor-binding protein 7 OS=Homo sapiens OX=9606 GN=IGFBP7 PE=1 SV=1	3	29.1	23	48.6	0.473	-1.080
130	High	P15289	Arylsulfatase A OS=Homo sapiens OX=9606 GN=ARSA PE=1 SV=3	6	53.6	27.8	59	0.471	-1.086
131	High	P19440	Glutathione hydrolase 1 proenzyme OS=Homo sapiens OX=9606 GN=GGT1 PE=1 SV=2	2	61.4	20.5	43.6	0.471	-1.086
132	High	P31944	Caspase-14 OS=Homo sapiens OX=9606 GN=CASP14 PE=1 SV=2	2	27.7	19.9	42.7	0.465	-1.105
133	High	P43121	Cell surface glycoprotein MUC18 OS=Homo sapiens OX=9606 GN=MCAM PE=1 SV=2	3	71.6	28.2	60.9	0.462	-1.114
134	High	P30086	Phosphatidylethanolamine-binding protein 1 OS=Homo sapiens OX=9606 GN=PEBP1 PE=1 SV=3	3	21	31.7	68.9	0.461	-1.117
135	High	P16070	CD44 antigen OS=Homo sapiens OX=9606 GN=CD44 PE=1 SV=3	3	81.5	43.9	95.9	0.458	-1.127
136	High	Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens OX=9606	3	40.2	20.1	43.9	0.458	-1.127

			GN=LMAN2 PE=1 SV=1						
137	High	Q8TDQ0	Hepatitis A virus cellular receptor 2 OS=Homo sapiens OX=9606 GN=HAVCR2 PE=1 SV=3	2	33.4	26.2	57.9	0.453	-1.142
138	High	P14384	Carboxypeptidase M OS=Homo sapiens OX=9606 GN=CPM PE=1 SV=2	2	50.5	28.9	65.3	0.443	-1.175
139	High	P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens OX=9606 GN=KRT4 PE=1 SV=4	12	57.3	22.7	52.9	0.43	-1.218
140	High	P29508	Serpin B3 OS=Homo sapiens OX=9606 GN=SERPINB3 PE=1 SV=2	6	44.5	18	42.6	0.421	-1.248
141	High	P15941	Mucin-1 OS=Homo sapiens OX=9606 GN=MUC1 PE=1 SV=3	3	122	31.1	74	0.42	-1.252
142	High	Q92820	Gamma-glutamyl hydrolase OS=Homo sapiens OX=9606 GN=GGH PE=1 SV=2	3	35.9	29.9	73.5	0.407	-1.297
143	High	P55290	Cadherin-13 OS=Homo sapiens OX=9606 GN=CDH13 PE=1 SV=1	3	78.2	25.5	63.2	0.404	-1.308
144	High	Q7KYR7	Butyrophilin subfamily 2 member A1 OS=Homo sapiens OX=9606 GN=BTN2A1 PE=1 SV=3	3	59.6	29.5	74	0.399	-1.326
145	High	Q9NQ84	G-protein coupled receptor family C group 5 member C OS=Homo sapiens OX=9606 GN=GPRC5C PE=1 SV=2	2	48.2	25.7	64.9	0.396	-1.336
146	High	Q16651	Prostasin OS=Homo sapiens OX=9606 GN=PRSS8 PE=1 SV=1	2	36.4	18.7	47.7	0.392	-1.351
147	High	P01008	Antithrombin-III OS=Homo sapiens OX=9606 GN=SERPINC1 PE=1 SV=1	4	52.6	22.7	58.1	0.391	-1.355
148	High	P63104	14-3-3 protein zeta/delta OS=Homo sapiens OX=9606 GN=YWHAZ PE=1 SV=1	5	27.7	39.8	102.8	0.388	-1.366
149	High	P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens OX=9606 GN=KRT13 PE=1 SV=4	18	49.6	19.4	50.3	0.386	-1.373

150	High	P01133	Pro-epidermal growth factor OS=Homo sapiens OX=9606 GN=EGF PE=1 SV=2	11	133.9	24.3	63.2	0.385	-1.377
151	High	P27482	Calmodulin-like protein 3 OS=Homo sapiens OX=9606 GN=CALML3 PE=1 SV=2	4	16.9	16.8	43.8	0.383	-1.385
152	High	P98160	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens OX=9606 GN=HSPG2 PE=1 SV=4	12	468.5	29.9	79.5	0.376	-1.411
153	High	P62979	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens OX=9606 GN=RPS27A PE=1 SV=2	3	18	24.3	65	0.373	-1.423
154	High	Q99988	Growth/differentiation factor 15 OS=Homo sapiens OX=9606 GN=GDF15 PE=1 SV=3	3	34.1	26.8	72.5	0.37	-1.434
155	High	P10153	Non-secretory ribonuclease OS=Homo sapiens OX=9606 GN=RNASE2 PE=1 SV=2	2	18.3	28.6	77.3	0.37	-1.434
156	High	P08637	Low affinity immunoglobulin gamma Fc region receptor III-A OS=Homo sapiens OX=9606 GN=FCGR3A PE=1 SV=2	2	29.1	12.4	34	0.364	-1.458
157	High	P54802	Alpha-N-acetylglucosaminidase OS=Homo sapiens OX=9606 GN=NAGLU PE=1 SV=2	5	82.2	22.1	63.5	0.348	-1.523
158	High	P05937	Calbindin OS=Homo sapiens OX=9606 GN=CALB1 PE=1 SV=2	3	30	25.3	72.8	0.347	-1.527
159	High	O75882	Attractin OS=Homo sapiens OX=9606 GN=ATRN PE=1 SV=2	4	158.4	20.7	61.2	0.338	-1.565
160	High	O94919	Endonuclease domain-containing 1 protein OS=Homo sapiens OX=9606 GN=ENDOD1 PE=1 SV=2	3	55	19.1	57.6	0.332	-1.591
161	High	P02538	Keratin, type II cytoskeletal 6A	19	60	14.8	49.7	0.297	-1.751

			OS=Homo sapiens OX=9606 GN=KRT6A PE=1 SV=3						
162	High	P01861	Immunoglobulin heavy constant gamma 4 OS=Homo sapiens OX=9606 GN=IGHG4 PE=1 SV=1	3	35.9	25	85.1	0.293	-1.771
163	High	P05062	Fructose-bisphosphate aldolase B OS=Homo sapiens OX=9606 GN=ALDOB PE=1 SV=2	2	39.4	25.7	89.4	0.287	-1.801
164	High	Q03154	Aminoacylase-1 OS=Homo sapiens OX=9606 GN=ACY1 PE=1 SV=1	2	45.9	21.1	76.3	0.277	-1.852
165	High	Q6EMK4	Vasorin OS=Homo sapiens OX=9606 GN=VASN PE=1 SV=1	5	71.7	19.5	70.6	0.276	-1.857
166	High	Q06830	Peroxiredoxin-1 OS=Homo sapiens OX=9606 GN=PRDX1 PE=1 SV=1	2	22.1	19	69.2	0.274	-1.868
167	High	P10253	Lysosomal alpha-glucosidase OS=Homo sapiens OX=9606 GN=GAA PE=1 SV=4	4	105.3	14.5	60.5	0.24	-2.059
168	High	Q9Y646	Carboxypeptidase Q OS=Homo sapiens OX=9606 GN=CPQ PE=1 SV=1	2	51.9	16.4	79	0.208	-2.265
169	High	Q9P121	Neurotrimin OS=Homo sapiens OX=9606 GN=NTM PE=1 SV=1	2	37.9	13.5	72.9	0.184	-2.442
170	High	P07195	L-lactate dehydrogenase B chain OS=Homo sapiens OX=9606 GN=LDHB PE=1 SV=2	2	36.6	13	79.8	0.163	-2.617
171	High	Q8WZ75	Roundabout homolog 4 OS=Homo sapiens OX=9606 GN=ROBO4 PE=1 SV=1	2	107.4	7.8	71.6	0.109	-3.198
172	High	Q9HCU0	Endosialin OS=Homo sapiens OX=9606 GN=CD248 PE=1 SV=1	2	80.8	3.3	38	0.086	-3.540
173	High	Q07075	Glutamyl aminopeptidase OS=Homo sapiens OX=9606 GN=ENPEP PE=1 SV=3	3	109.2	1.7	28.2	0.059	-4.083

Data identified at confidence interval (CI): 99 %, FDR < 0.01; Peptides no. ≥ 2.

Table S 2: Dataset of protein expression for PE and control at 20 weeks gestation

No.	Protein FDR Confidence: Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of PE at 20 W	Abundances of reporter ion of control at 20 W	Abundance ratio (Case / Control) of PE at 20 W	Log ₂ fold change
1	High	P02675	Fibrinogen beta chain OS=Homo sapiens OX=9606 GN=FGB PE=1 SV=2	2	55.9	167.3	22.2	7.526	2.912
2	High	P00738	Haptoglobin OS=Homo sapiens OX=9606 GN=HP PE=1 SV=1	4	45.2	183.4	48	3.823	1.935
3	High	P02787	Serotransferrin OS=Homo sapiens OX=9606 GN=TF PE=1 SV=3	12	77	179.2	61.4	2.919	1.545
4	High	P01861	Immunoglobulin heavy constant gamma 4 OS=Homo sapiens OX=9606 GN=IGHG4 PE=1 SV=1	3	35.9	201.1	75.1	2.677	1.421
5	High	P01877	Immunoglobulin heavy constant alpha 2 OS=Homo sapiens OX=9606 GN=IGHA2 PE=1 SV=4	3	36.6	203	78.6	2.582	1.368
6	High	P00450	Ceruloplasmin OS=Homo sapiens OX=9606 GN=CP PE=1 SV=1	13	122.1	227.8	88.8	2.564	1.358
7	High	P01859	Immunoglobulin heavy constant gamma 2 OS=Homo sapiens OX=9606 GN=IGHG2 PE=1 SV=2	4	35.9	179.4	71.7	2.502	1.323
8	High	P13796	Plastin-2 OS=Homo sapiens OX=9606 GN=LCP1 PE=1 SV=6	4	70.2	105.6	45.6	2.316	1.212
9	High	P04217	Alpha-1B-glycoprotein OS=Homo sapiens OX=9606 GN=A1BG PE=1 SV=4	3	54.2	223	96.3	2.316	1.212
10	High	P01023	Alpha-2-macroglobulin OS=Homo sapiens OX=9606 GN=A2M PE=1 SV=3	2	163.2	128.1	55.4	2.312	1.209

11	High	P01009	Alpha-1-antitrypsin OS=Homo sapiens OX=9606 GN=SERPINA1 PE=1 SV=3	14	46.7	213	95	2.242	1.165
12	High	P11021	Endoplasmic reticulum chaperone BiP OS=Homo sapiens OX=9606 GN=HSPA5 PE=1 SV=2	5	72.3	115.5	54.5	2.119	1.083
13	High	Q6GTX8	Leukocyte-associated immunoglobulin- like receptor 1 OS=Homo sapiens OX=9606 GN=LAIR1 PE=1 SV=1	2	31.4	158	77.9	2.029	1.021
14	High	P02768	Serum albumin OS=Homo sapiens OX=9606 GN=ALB PE=1 SV=2	19	69.3	154.6	76.4	2.025	1.018
15	High	P02750	Leucine-rich alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=LRG1 PE=1 SV=2	5	38.2	165	81.5	2.024	1.017
16	High	Q93088	Betaine--homocysteine S- methyltransferase 1 OS=Homo sapiens OX=9606 GN=BHMT PE=1 SV=2	3	45	159.1	79.5	2.002	1.001
17	High	P18510	Interleukin-1 receptor antagonist protein OS=Homo sapiens OX=9606 GN=IL1RN PE=1 SV=1	2	20	171.5	85.8	1.998	0.999
18	High	P01614	Immunoglobulin kappa variable 2D-40 OS=Homo sapiens OX=9606 GN=IGKV2D-40 PE=1 SV=2	2	13.3	154.7	80.1	1.932	0.950
19	High	P04746	Pancreatic alpha-amylase OS=Homo sapiens OX=9606 GN=AMY2A PE=1 SV=2	7	57.7	175.7	93.5	1.88	0.911
20	High	P07476	Involucrin OS=Homo sapiens OX=9606 GN=IVL PE=1 SV=2	9	68.4	203.7	110	1.851	0.888
21	High	P0DOY2	Immunoglobulin lambda constant 2 OS=Homo sapiens OX=9606 GN=IGLC2 PE=1 SV=1	3	11.3	167	90.7	1.841	0.880
22	High	P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens OX=9606 GN=PTGDS	3	21	190.8	105	1.817	0.862

			PE=1 SV=1						
23	High	P01876	Immunoglobulin heavy constant alpha 1 OS=Homo sapiens OX=9606 GN=IGHA1 PE=1 SV=2	5	37.6	160.3	90.3	1.774	0.827
24	High	P04083	Annexin A1 OS=Homo sapiens OX=9606 GN=ANXA1 PE=1 SV=2	6	38.7	259.2	146.8	1.765	0.820
25	High	P08185	Corticosteroid-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA6 PE=1 SV=1	3	45.1	171.9	100.2	1.715	0.778
26	High	Q02487	Desmocollin-2 OS=Homo sapiens OX=9606 GN=DSC2 PE=1 SV=1	3	99.9	188.4	114	1.653	0.725
27	High	P01871	Immunoglobulin heavy constant mu OS=Homo sapiens OX=9606 GN=IGHM PE=1 SV=4	2	49.4	206.5	126.9	1.628	0.703
28	High	P05060	Secretogranin-1 OS=Homo sapiens OX=9606 GN=CHGB PE=1 SV=2	2	78.2	168.4	103.9	1.62	0.696
29	High	P01011	Alpha-1-antichymotrypsin OS=Homo sapiens OX=9606 GN=SERPINA3 PE=1 SV=2	6	47.6	178	110.1	1.617	0.693
30	High	P25815	Protein S100-P OS=Homo sapiens OX=9606 GN=S100P PE=1 SV=2	2	10.4	110.6	68.5	1.613	0.690
31	High	P01019	Angiotensinogen OS=Homo sapiens OX=9606 GN=AGT PE=1 SV=1	2	53.1	145.2	95.3	1.523	0.607
32	High	P01591	Immunoglobulin J chain OS=Homo sapiens OX=9606 GN=JCHAIN PE=1 SV=4	2	18.1	157.7	104.4	1.51	0.595
33	High	P06733	Alpha-enolase OS=Homo sapiens OX=9606 GN=ENO1 PE=1 SV=2	4	47.1	141.1	93.8	1.505	0.590
34	High	P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens OX=9606 GN=HSPA8 PE=1 SV=1	9	70.9	138.1	92.1	1.5	0.585
35	High	P29508	Serpin B3 OS=Homo sapiens OX=9606	6	44.5	153.3	103.3	1.483	0.569

			GN=SERPINB3 PE=1 SV=2						
36	High	Q9HCU0	Endosialin OS=Homo sapiens OX=9606 GN=CD248 PE=1 SV=1	2	80.8	221.1	150.3	1.471	0.557
37	High	B9A064	Immunoglobulin lambda-like polypeptide 5 OS=Homo sapiens OX=9606 GN=IGLL5 PE=2 SV=2	3	23	152.2	105.2	1.446	0.532
38	High	P15144	Aminopeptidase N OS=Homo sapiens OX=9606 GN=ANPEP PE=1 SV=4	12	109.5	150.6	104.9	1.436	0.522
39	High	P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=AZGP1 PE=1 SV=2	13	34.2	123.9	86.8	1.428	0.514
40	High	P50995	Annexin A11 OS=Homo sapiens OX=9606 GN=ANXA11 PE=1 SV=1	3	54.4	112	78.7	1.423	0.509
41	High	Q96NY8	Nectin-4 OS=Homo sapiens OX=9606 GN=NECTIN4 PE=1 SV=1	3	55.4	139.2	98	1.421	0.507
42	High	Q13510	Acid ceramidase OS=Homo sapiens OX=9606 GN=ASA11 PE=1 SV=5	3	44.6	152.9	108.2	1.414	0.500
43	High	P04075	Fructose-bisphosphate aldolase A OS=Homo sapiens OX=9606 GN=ALDOA PE=1 SV=2	2	39.4	107.1	76.1	1.406	0.492
44	High	P10599	Thioredoxin OS=Homo sapiens OX=9606 GN=TXN PE=1 SV=3	3	11.7	138.9	99.7	1.393	0.478
45	High	P05090	Apolipoprotein D OS=Homo sapiens OX=9606 GN=APOD PE=1 SV=1	3	21.3	143.8	103.3	1.392	0.477
46	High	P01834	Immunoglobulin kappa constant OS=Homo sapiens OX=9606 GN=IGKC PE=1 SV=2	3	11.8	147.7	107	1.381	0.466
47	High	P06703	Protein S100-A6 OS=Homo sapiens OX=9606 GN=S100A6 PE=1 SV=1	3	10.2	119.2	86.4	1.379	0.464
48	High	O00187	Mannan-binding lectin serine protease 2 OS=Homo sapiens OX=9606 GN=MASP2 PE=1 SV=4	3	75.7	183.8	133.3	1.378	0.463
49	High	P01833	Polymeric immunoglobulin receptor	7	83.2	146.3	109.9	1.331	0.413

			OS=Homo sapiens OX=9606 GN=PIGR PE=1 SV=4						
50	High	P05937	Calbindin OS=Homo sapiens OX=9606 GN=CALB1 PE=1 SV=2	3	30	116.2	87.9	1.323	0.404
51	High	P01857	Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1 SV=1	7	36.1	124.1	94.2	1.317	0.397
52	High	P05543	Thyroxine-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA7 PE=1 SV=2	6	46.3	159	121.5	1.309	0.388
53	High	Q07075	Glutamyl aminopeptidase OS=Homo sapiens OX=9606 GN=ENPEP PE=1 SV=3	3	109.2	195.5	150.5	1.299	0.377
54	High	P07339	Cathepsin D OS=Homo sapiens OX=9606 GN=CTSD PE=1 SV=1	6	44.5	126.9	100.4	1.264	0.338
55	High	P11464	Pregnancy-specific beta-1-glycoprotein 1 OS=Homo sapiens OX=9606 GN=PSG1 PE=1 SV=1	4	47.2	162.3	130	1.249	0.321
56	High	Q7KYR7	Butyrophilin subfamily 2 member A1 OS=Homo sapiens OX=9606 GN=BTN2A1 PE=1 SV=3	3	59.6	125.3	100.4	1.248	0.320
57	High	Q6EMK4	Vasorin OS=Homo sapiens OX=9606 GN=VASN PE=1 SV=1	5	71.7	125.6	101.2	1.241	0.312
58	High	P05155	Plasma protease C1 inhibitor OS=Homo sapiens OX=9606 GN=SERPING1 PE=1 SV=2	6	55.1	143.2	115.5	1.24	0.310
59	High	P10153	Non-secretory ribonuclease OS=Homo sapiens OX=9606 GN=RNASE2 PE=1 SV=2	2	18.3	113.4	91.4	1.24	0.310
60	High	Q8TDQ0	Hepatitis A virus cellular receptor 2 OS=Homo sapiens OX=9606 GN=HAVCR2 PE=1 SV=3	2	33.4	128.7	104.2	1.235	0.305
61	High	Q9UQ72	Pregnancy-specific beta-1-glycoprotein	7	37.1	162.2	133	1.219	0.286

			11 OS=Homo sapiens OX=9606 GN=PSG11 PE=2 SV=3						
62	High	Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1 OS=Homo sapiens OX=9606 GN=LYVE1 PE=1 SV=2	2	35.2	173.5	142.5	1.218	0.285
63	High	P19652	Alpha-1-acid glycoprotein 2 OS=Homo sapiens OX=9606 GN=ORM2 PE=1 SV=2	2	23.6	142.9	117.4	1.217	0.283
64	High	Q08345	Epithelial discoidin domain-containing receptor 1 OS=Homo sapiens OX=9606 GN=DDR1 PE=1 SV=1	4	101.1	115.5	95.1	1.215	0.281
65	High	P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens OX=9606 GN=KRT9 PE=1 SV=3	5	62	116.5	96.7	1.205	0.269
66	High	P02760	Protein AMBP OS=Homo sapiens OX=9606 GN=AMBP PE=1 SV=1	5	39	146.2	121.7	1.201	0.264
67	High	O43451	Maltase-glucoamylase, intestinal OS=Homo sapiens OX=9606 GN=MGAM PE=1 SV=5	4	209.7	110.3	91.9	1.199	0.262
68	High	Q9UBC9	Small proline-rich protein 3 OS=Homo sapiens OX=9606 GN=SPRR3 PE=1 SV=2	5	18.1	129.4	108	1.198	0.261
69	High	P98164	Low-density lipoprotein receptor- related protein 2 OS=Homo sapiens OX=9606 GN=LRP2 PE=1 SV=3	6	521.6	131.4	110.2	1.193	0.255
70	High	Q8NBJ4	Golgi membrane protein 1 OS=Homo sapiens OX=9606 GN=GOLM1 PE=1 SV=1	7	45.3	130.9	110.5	1.184	0.244
71	High	P02774	Vitamin D-binding protein OS=Homo sapiens OX=9606 GN=GC PE=1 SV=1	5	52.9	120.9	102.6	1.178	0.236
72	High	P60709	Actin, cytoplasmic 1 OS=Homo sapiens OX=9606 GN=ACTB PE=1 SV=1	6	41.7	119.1	101.8	1.17	0.227
73	High	P55290	Cadherin-13 OS=Homo sapiens OX=9606 GN=CDH13 PE=1 SV=1	3	78.2	147.3	126.3	1.166	0.222
74	High	Q9NZP8	Complement C1r subcomponent-like	3	53.5	115.9	99.7	1.162	0.217

			protein OS=Homo sapiens OX=9606 GN=C1RL PE=1 SV=2						
75	High	P02763	Alpha-1-acid glycoprotein 1 OS=Homo sapiens OX=9606 GN=ORM1 PE=1 SV=1	3	23.5	135.4	117.3	1.154	0.207
76	High	Q99988	Growth/differentiation factor 15 OS=Homo sapiens OX=9606 GN=GDF15 PE=1 SV=3	3	34.1	98.5	85.5	1.153	0.205
77	High	A8K2U0	Alpha-2-macroglobulin-like protein 1 OS=Homo sapiens OX=9606 GN=A2ML1 PE=1 SV=3	6	161	132.8	116	1.145	0.195
78	High	P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens OX=9606 GN=KRT6A PE=1 SV=3	19	60	97.4	85.4	1.139	0.188
79	High	P10909	Clusterin OS=Homo sapiens OX=9606 GN=CLU PE=1 SV=1	4	52.5	118.7	104.7	1.134	0.181
80	High	Q14515	SPARC-like protein 1 OS=Homo sapiens OX=9606 GN=SPARCL1 PE=1 SV=2	4	75.2	134.4	118.8	1.132	0.179
81	High	Q03154	Aminoacylase-1 OS=Homo sapiens OX=9606 GN=ACY1 PE=1 SV=1	2	45.9	120.8	106.8	1.131	0.178
82	High	Q6UXB8	Peptidase inhibitor 16 OS=Homo sapiens OX=9606 GN=PI16 PE=1 SV=1	2	49.4	143.5	127.4	1.126	0.171
83	High	Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens OX=9606 GN=LMAN2 PE=1 SV=1	3	40.2	144.8	128.7	1.124	0.169
84	High	P24855	Deoxyribonuclease-1 OS=Homo sapiens OX=9606 GN=DNASE1 PE=1 SV=1	3	31.4	116.4	105.2	1.107	0.147
85	High	P01040	Cystatin-A OS=Homo sapiens OX=9606 GN=CSTA PE=1 SV=1	2	11	130	118.3	1.099	0.136
86	High	P02788	Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF PE=1 SV=6	3	78.1	66.6	62.4	1.067	0.094
87	High	P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens OX=9606 GN=KRT16	6	51.2	120	112.8	1.064	0.089

			PE=1 SV=4						
88	High	P19440	Glutathione hydrolase 1 proenzyme OS=Homo sapiens OX=9606 GN=GGT1 PE=1 SV=2	2	61.4	130.6	122.7	1.064	0.089
89	High	P09603	Macrophage colony-stimulating factor 1 OS=Homo sapiens OX=9606 GN=CSF1 PE=1 SV=2	4	60.1	127.2	119.9	1.061	0.085
90	High	P12109	Collagen alpha-1(VI) chain OS=Homo sapiens OX=9606 GN=COL6A1 PE=1 SV=3	10	108.5	124.2	117.2	1.06	0.084
91	High	P08571	Monocyte differentiation antigen CD14 OS=Homo sapiens OX=9606 GN=CD14 PE=1 SV=2	4	40.1	145.7	137.6	1.059	0.083
92	High	P09668	Pro-cathepsin H OS=Homo sapiens OX=9606 GN=CTSH PE=1 SV=4	2	37.4	132.1	124.7	1.059	0.083
93	High	Q9ULI3	Protein HEG homolog 1 OS=Homo sapiens OX=9606 GN=HEG1 PE=1 SV=3	3	147.4	142.6	135.5	1.053	0.075
94	High	P54802	Alpha-N-acetylglucosaminidase OS=Homo sapiens OX=9606 GN=NAGLU PE=1 SV=2	5	82.2	144.4	137.6	1.049	0.069
95	High	Q6UVK1	Chondroitin sulfate proteoglycan 4 OS=Homo sapiens OX=9606 GN=CSPG4 PE=1 SV=2	2	250.4	130.5	124.9	1.045	0.064
96	High	O94919	Endonuclease domain-containing 1 protein OS=Homo sapiens OX=9606 GN=ENDOD1 PE=1 SV=2	3	55	125.1	120	1.043	0.061
97	High	P14384	Carboxypeptidase M OS=Homo sapiens OX=9606 GN=CPM PE=1 SV=2	2	50.5	102	97.9	1.042	0.059
98	High	Q16651	Prostasin OS=Homo sapiens OX=9606 GN=PRSS8 PE=1 SV=1	2	36.4	116.9	112.7	1.037	0.052
99	High	P60174	Triosephosphate isomerase OS=Homo sapiens OX=9606 GN=TPI1 PE=1 SV=3	2	30.8	121.7	119	1.023	0.033

100	High	Q9BY67	Cell adhesion molecule 1 OS=Homo sapiens OX=9606 GN=CADM1 PE=1 SV=2	3	48.5	130.1	128.5	1.013	0.019
101	High	Q9UBG3	Cornulin OS=Homo sapiens OX=9606 GN=CRNN PE=1 SV=1	4	53.5	120.5	119.7	1.007	0.010
102	High	P01008	Antithrombin-III OS=Homo sapiens OX=9606 GN=SERPINC1 PE=1 SV=1	4	52.6	114.9	114.1	1.007	0.010
103	High	P34059	N-acetylgalactosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GALNS PE=1 SV=1	2	58	118.4	117.7	1.006	0.009
104	High	P06396	Gelsolin OS=Homo sapiens OX=9606 GN=GSN PE=1 SV=1	6	85.6	153.9	153.5	1.003	0.004
105	High	P00924	Enolase 1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ENO1 PE=1 SV=3	14	46.8	100	100	1	0.000
106	High	P39059	Collagen alpha-1(XV) chain OS=Homo sapiens OX=9606 GN=COL15A1 PE=1 SV=2	2	141.6	112.8	114.2	0.988	-0.017
107	High	P01042	Kininogen-1 OS=Homo sapiens OX=9606 GN=KNG1 PE=1 SV=2	6	71.9	108.9	111.2	0.98	-0.029
108	High	P05164	Myeloperoxidase OS=Homo sapiens OX=9606 GN=MPO PE=1 SV=1	3	83.8	94.6	96.7	0.978	-0.032
109	High	Q8IYS5	Osteoclast-associated immunoglobulin-like receptor OS=Homo sapiens OX=9606 GN=OSCAR PE=1 SV=3	2	30.5	110.3	114.5	0.963	-0.054
110	High	P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens OX=9606 GN=KRT6B PE=1 SV=5	20	60	100	104.1	0.96	-0.059
111	High	O00391	Sulfhydryl oxidase 1 OS=Homo sapiens OX=9606 GN=QSOX1 PE=1 SV=3	5	82.5	119.1	124.9	0.954	-0.068
112	High	Q68CJ9	Cyclic AMP-responsive element-binding protein 3-like protein 3 OS=Homo	2	49	114.2	120.4	0.949	-0.076

			sapiens OX=9606 GN=CREB3L3 PE=1 SV=2						
113	High	P02751	Fibronectin OS=Homo sapiens OX=9606 GN=FN1 PE=1 SV=4	11	262.5	112.3	119.5	0.94	-0.089
114	High	Q06830	Peroxiredoxin-1 OS=Homo sapiens OX=9606 GN=PRDX1 PE=1 SV=1	2	22.1	117.6	125.4	0.938	-0.092
115	High	P62873	Guanine nucleotide-binding protein G(l)/G(s)/G(t) subunit beta-1 OS=Homo sapiens OX=9606 GN=GNB1 PE=1 SV=3	2	37.4	109.2	116.7	0.936	-0.095
116	High	P12830	Cadherin-1 OS=Homo sapiens OX=9606 GN=CDH1 PE=1 SV=3	8	97.4	135.5	144.9	0.935	-0.097
117	High	P02765	Alpha-2-HS-glycoprotein OS=Homo sapiens OX=9606 GN=AHSG PE=1 SV=1	3	39.3	161.4	173	0.933	-0.100
118	High	P43121	Cell surface glycoprotein MUC18 OS=Homo sapiens OX=9606 GN=MCAM PE=1 SV=2	3	71.6	109.5	119.1	0.919	-0.122
119	High	P15289	Arylsulfatase A OS=Homo sapiens OX=9606 GN=ARSA PE=1 SV=3	6	53.6	118.2	128.9	0.917	-0.125
120	High	P10451	Osteopontin OS=Homo sapiens OX=9606 GN=SPP1 PE=1 SV=1	11	35.4	150.8	165	0.914	-0.130
121	High	P98160	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens OX=9606 GN=HSPG2 PE=1 SV=4	12	468.5	114.2	126	0.907	-0.141
122	High	P19835	Bile salt-activated lipase OS=Homo sapiens OX=9606 GN=CEL PE=1 SV=3	3	79.3	97.7	108.8	0.899	-0.154
123	High	P15586	N-acetylglucosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GNS PE=1 SV=3	3	62	123.6	137.4	0.899	-0.154
124	High	Q9P121	Neurotrimin OS=Homo sapiens OX=9606 GN=NTM PE=1 SV=1	2	37.9	94.6	106.3	0.89	-0.168
125	High	P78324	Tyrosine-protein phosphatase non-	3	54.9	112.5	126.5	0.889	-0.170

			receptor type substrate 1 OS=Homo sapiens OX=9606 GN=SIRPA PE=1 SV=2						
126	High	P06870	Kallikrein-1 OS=Homo sapiens OX=9606 GN=KLK1 PE=1 SV=2	3	28.9	89.7	101.3	0.886	-0.175
127	High	P15941	Mucin-1 OS=Homo sapiens OX=9606 GN=MUC1 PE=1 SV=3	3	122	122.7	139.1	0.882	-0.181
128	High	P80188	Neutrophil gelatinase-associated lipocalin OS=Homo sapiens OX=9606 GN=LCN2 PE=1 SV=2	3	22.6	95.7	110.1	0.869	-0.203
129	High	P06702	Protein S100-A9 OS=Homo sapiens OX=9606 GN=S100A9 PE=1 SV=1	7	13.2	109.6	126.3	0.868	-0.204
130	High	Q8NFZ8	Cell adhesion molecule 4 OS=Homo sapiens OX=9606 GN=CADM4 PE=1 SV=1	2	42.8	89.1	104.2	0.855	-0.226
131	High	P05154	Plasma serine protease inhibitor OS=Homo sapiens OX=9606 GN=SERPINA5 PE=1 SV=3	4	45.6	125.5	146.9	0.854	-0.228
132	High	P05109	Protein S100-A8 OS=Homo sapiens OX=9606 GN=S100A8 PE=1 SV=1	5	10.8	96.2	113.6	0.846	-0.241
133	High	P16070	CD44 antigen OS=Homo sapiens OX=9606 GN=CD44 PE=1 SV=3	3	81.5	123.4	146.9	0.84	-0.252
134	High	P31944	Caspase-14 OS=Homo sapiens OX=9606 GN=CASP14 PE=1 SV=2	2	27.7	111.2	132.6	0.838	-0.255
135	High	O60494	Cubilin OS=Homo sapiens OX=9606 GN=CUBN PE=1 SV=5	3	398.5	105.4	126.1	0.836	-0.258
136	High	Q08380	Galectin-3-binding protein OS=Homo sapiens OX=9606 GN=LGALS3BP PE=1 SV=1	10	65.3	112.8	135.4	0.833	-0.264
137	High	Q16270	Insulin-like growth factor-binding protein 7 OS=Homo sapiens OX=9606 GN=IGFBP7 PE=1 SV=1	3	29.1	109.4	131.8	0.83	-0.269
138	High	P31947	14-3-3 protein sigma OS=Homo sapiens	5	27.8	113.9	138.5	0.823	-0.281

			OX=9606 GN=SFN PE=1 SV=1						
139	High	Q8WZ75	Roundabout homolog 4 OS=Homo sapiens OX=9606 GN=ROBO4 PE=1 SV=1	2	107.4	110.5	136.1	0.812	-0.300
140	High	P01133	Pro-epidermal growth factor OS=Homo sapiens OX=9606 GN=EGF PE=1 SV=2	11	133.9	98.9	122.4	0.808	-0.308
141	High	P10253	Lysosomal alpha-glucosidase OS=Homo sapiens OX=9606 GN=GAA PE=1 SV=4	4	105.3	99.8	124.2	0.804	-0.315
142	High	Q9BRK3	Matrix remodeling-associated protein 8 OS=Homo sapiens OX=9606 GN=MXRA8 PE=1 SV=1	4	49.1	97.5	122.1	0.798	-0.326
143	High	P26992	Ciliary neurotrophic factor receptor subunit alpha OS=Homo sapiens OX=9606 GN=CNTFR PE=1 SV=2	2	40.6	93.7	117.4	0.798	-0.326
144	High	P50895	Basal cell adhesion molecule OS=Homo sapiens OX=9606 GN=BCAM PE=1 SV=2	3	67.4	115.1	145.7	0.79	-0.340
145	High	P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens OX=9606 GN=KRT13 PE=1 SV=4	18	49.6	85.6	108.6	0.789	-0.342
146	High	O75882	Attractin OS=Homo sapiens OX=9606 GN=ATRN PE=1 SV=2	4	158.4	106.9	136	0.786	-0.347
147	High	P07195	L-lactate dehydrogenase B chain OS=Homo sapiens OX=9606 GN=LDHB PE=1 SV=2	2	36.6	110.7	141.1	0.784	-0.351
148	High	P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens OX=9606 GN=KRT5 PE=1 SV=3	11	62.3	87.7	112.8	0.777	-0.364
149	High	P62736	Actin, aortic smooth muscle OS=Homo sapiens OX=9606 GN=ACTA2 PE=1 SV=1	5	42	114.8	148.9	0.771	-0.375
150	High	P02790	Hemopexin OS=Homo sapiens OX=9606 GN=HPX PE=1 SV=2	5	51.6	90.3	117.2	0.77	-0.377
151	High	Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens OX=9606	4	103.3	129	167.6	0.77	-0.377

			GN=ITIH4 PE=1 SV=4						
152	High	Q9NQ84	G-protein coupled receptor family C group 5 member C OS=Homo sapiens OX=9606 GN=GPRC5C PE=1 SV=2	2	48.2	83.3	108.2	0.77	-0.377
153	High	P29972	Aquaporin-1 OS=Homo sapiens OX=9606 GN=AQP1 PE=1 SV=3	2	28.5	107.9	141.3	0.763	-0.390
154	High	P27482	Calmodulin-like protein 3 OS=Homo sapiens OX=9606 GN=CALML3 PE=1 SV=2	4	16.9	107.1	141.6	0.756	-0.404
155	High	P62805	Histone H4 OS=Homo sapiens OX=9606 GN=HIST1H4A PE=1 SV=2	3	11.4	81.8	110.4	0.741	-0.432
156	High	P62979	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens OX=9606 GN=RPS27A PE=1 SV=2	3	18	103.3	139.8	0.739	-0.436
157	High	P0C0L5	Complement C4-B OS=Homo sapiens OX=9606 GN=C4B PE=1 SV=2	3	192.6	106.8	144.4	0.739	-0.436
158	High	P12273	Prolactin-inducible protein OS=Homo sapiens OX=9606 GN=PIP PE=1 SV=1	2	16.6	119.3	163.2	0.731	-0.452
159	High	P13473	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens OX=9606 GN=LAMP2 PE=1 SV=2	2	44.9	62.9	86.1	0.73	-0.454
160	High	P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens OX=9606 GN=KRT2 PE=1 SV=2	5	65.4	64.9	96.5	0.672	-0.573
161	High	Q92820	Gamma-glutamyl hydrolase OS=Homo sapiens OX=9606 GN=GGH PE=1 SV=2	3	35.9	115	171.1	0.672	-0.573
162	High	Q01459	Di-N-acetylchitobiase OS=Homo sapiens OX=9606 GN=CTBS PE=1 SV=1	2	43.7	107.7	164.5	0.655	-0.610
163	High	P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens OX=9606 GN=KRT10 PE=1 SV=6	14	58.8	57.4	88.8	0.646	-0.630
164	High	P30086	Phosphatidylethanolamine-binding	3	21	68.1	109.5	0.622	-0.685

			protein 1 OS=Homo sapiens OX=9606 GN=PEBP1 PE=1 SV=3						
165	High	P08637	Low affinity immunoglobulin gamma Fc region receptor III-A OS=Homo sapiens OX=9606 GN=FCGR3A PE=1 SV=2	2	29.1	81.3	136.7	0.595	-0.749
166	High	P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens OX=9606 GN=KRT1 PE=1 SV=6	17	66	56	95.3	0.588	-0.766
167	High	P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens OX=9606 GN=KRT4 PE=1 SV=4	12	57.3	59.2	108.7	0.545	-0.876
168	High	Q9Y646	Carboxypeptidase Q OS=Homo sapiens OX=9606 GN=CPQ PE=1 SV=1	2	51.9	84.6	162.8	0.52	-0.943
169	High	P07911	Uromodulin OS=Homo sapiens OX=9606 GN=UMOD PE=1 SV=1	7	69.7	94.1	186.4	0.505	-0.986
170	High	P05062	Fructose-bisphosphate aldolase B OS=Homo sapiens OX=9606 GN=ALDOB PE=1 SV=2	2	39.4	57.7	116.4	0.496	-1.012
171	High	P09211	Glutathione S-transferase P OS=Homo sapiens OX=9606 GN=GSTP1 PE=1 SV=2	2	23.3	54	124.9	0.433	-1.208
172	High	P07355	Annexin A2 OS=Homo sapiens OX=9606 GN=ANXA2 PE=1 SV=2	8	38.6	78.5	184	0.427	-1.228
173	High	P63104	14-3-3 protein zeta/delta OS=Homo sapiens OX=9606 GN=YWHAZ PE=1 SV=1	5	27.7	37.6	116.9	0.321	-1.639

Data identified at confidence interval (CI): 99 %, FDR < 0.01; Peptides no. ≥ 2.

Table S 3: Dataset of protein expression for IUGR and control at 15 weeks of pregnancy

No.	Protein FDR Confidence: Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of IUGR at 15 W	Abundances of reporter ion of control at 15 W	Abundance ratio (Case / Control) of IUGR at 15 W	Log ₂ fold change
1	High	P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=AZGP1 PE=1 SV=2	13	34.2	271.3	41.6	6.528	2.707
2	High	P02750	Leucine-rich alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=LRG1 PE=1 SV=2	5	38.2	195.9	56.7	3.458	1.790
3	High	P02763	Alpha-1-acid glycoprotein 1 OS=Homo sapiens OX=9606 GN=ORM1 PE=1 SV=1	3	23.5	236.2	68.6	3.442	1.783
4	High	P19652	Alpha-1-acid glycoprotein 2 OS=Homo sapiens OX=9606 GN=ORM2 PE=1 SV=2	2	23.6	193.5	78.5	2.464	1.301
5	High	P02765	Alpha-2-HS-glycoprotein OS=Homo sapiens OX=9606 GN=AHSG PE=1 SV=1	3	39.3	142.1	62.9	2.26	1.176
6	High	P04217	Alpha-1B-glycoprotein OS=Homo sapiens OX=9606 GN=A1BG PE=1 SV=4	3	54.2	146.8	69.5	2.112	1.079
7	High	P01009	Alpha-1-antitrypsin OS=Homo sapiens OX=9606 GN=SERPINA1 PE=1 SV=3	14	46.7	181.5	91.8	1.978	0.984
8	High	P62805	Histone H4 OS=Homo sapiens OX=9606 GN=HIST1H4A PE=1 SV=2	3	11.4	149	79.4	1.876	0.908
9	High	P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens OX=9606 GN=KRT9 PE=1 SV=3	5	62	153.2	81.7	1.875	0.907
10	High	P30086	Phosphatidylethanolamine-binding protein 1 OS=Homo sapiens OX=9606 GN=PEBP1 PE=1 SV=3	3	21	190.9	109	1.751	0.808
11	High	O00187	Mannan-binding lectin serine protease 2 OS=Homo sapiens OX=9606 GN=MASP2 PE=1 SV=4	3	75.7	155.6	89.4	1.741	0.800

12	High	P02788	Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF PE=1 SV=6	3	78.1	102.7	59.4	1.729	0.790
13	High	P02787	Serotransferrin OS=Homo sapiens OX=9606 GN=TF PE=1 SV=3	12	77	153.9	89.8	1.715	0.778
14	High	P0DOY2	Immunoglobulin lambda constant 2 OS=Homo sapiens OX=9606 GN=IGLC2 PE=1 SV=1	3	11.3	161.9	96.6	1.676	0.745
15	High	P07476	Involucrin OS=Homo sapiens OX=9606 GN=IVL PE=1 SV=2	9	68.4	176.3	105.7	1.669	0.739
16	High	Q93088	Betaine--homocysteine S- methyltransferase 1 OS=Homo sapiens OX=9606 GN=BHMT PE=1 SV=2	3	45	167.5	102.3	1.637	0.711
17	High	P01019	Angiotensinogen OS=Homo sapiens OX=9606 GN=AGT PE=1 SV=1	2	53.1	141.5	87.1	1.625	0.700
18	High	P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens OX=9606 GN=KRT5 PE=1 SV=3	11	62.3	196.4	124.7	1.574	0.654
19	High	P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens OX=9606 GN=KRT1 PE=1 SV=6	17	66	182	117.5	1.549	0.631
20	High	P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens OX=9606 GN=PTGDS PE=1 SV=1	3	21	143.5	93.4	1.536	0.619
21	High	P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens OX=9606 GN=KRT6A PE=1 SV=3	19	60	222.3	145.7	1.526	0.610
22	High	P01008	Antithrombin-III OS=Homo sapiens OX=9606 GN=SERPINC1 PE=1 SV=1	4	52.6	178.1	117.1	1.52	0.604
23	High	Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1 OS=Homo sapiens OX=9606 GN=LYVE1 PE=1 SV=2	2	35.2	143.8	94.7	1.519	0.603
24	High	P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens OX=9606 GN=KRT4 PE=1 SV=4	12	57.3	237.6	158.8	1.496	0.581
25	High	P01011	Alpha-1-antichymotrypsin OS=Homo	6	47.6	148.6	100.1	1.484	0.569

			sapiens OX=9606 GN=SERPINA3 PE=1 SV=2						
26	High	P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens OX=9606 GN=KRT6B PE=1 SV=5	20	60	196.3	132.4	1.483	0.569
27	High	P08185	Corticosteroid-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA6 PE=1 SV=1	3	45.1	126.8	85.9	1.475	0.561
28	High	P08571	Monocyte differentiation antigen CD14 OS=Homo sapiens OX=9606 GN=CD14 PE=1 SV=2	4	40.1	157	106.7	1.472	0.558
29	High	P06396	Gelsolin OS=Homo sapiens OX=9606 GN=GSN PE=1 SV=1	6	85.6	147.7	101.1	1.461	0.547
30	High	P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens OX=9606 GN=KRT13 PE=1 SV=4	18	49.6	214.6	147.2	1.459	0.545
31	High	P01871	Immunoglobulin heavy constant mu OS=Homo sapiens OX=9606 GN=IGHM PE=1 SV=4	2	49.4	130.4	93	1.401	0.486
32	High	P04075	Fructose-bisphosphate aldolase A OS=Homo sapiens OX=9606 GN=ALDOA PE=1 SV=2	2	39.4	148.1	106.3	1.393	0.478
33	High	Q03154	Aminoacylase-1 OS=Homo sapiens OX=9606 GN=ACY1 PE=1 SV=1	2	45.9	171.1	125.4	1.364	0.448
34	High	Q6UXB8	Peptidase inhibitor 16 OS=Homo sapiens OX=9606 GN=PI16 PE=1 SV=1	2	49.4	142.6	106.8	1.335	0.417
35	High	P01614	Immunoglobulin kappa variable 2D-40 OS=Homo sapiens OX=9606 GN=IGKV2D-40 PE=1 SV=2	2	13.3	136.6	102.8	1.329	0.410
36	High	Q02487	Desmocollin-2 OS=Homo sapiens OX=9606 GN=DSC2 PE=1 SV=1	3	99.9	140	109.2	1.282	0.358
37	High	P12830	Cadherin-1 OS=Homo sapiens OX=9606	8	97.4	140	110.1	1.271	0.346

			GN=CDH1 PE=1 SV=3						
38	High	P05543	Thyroxine-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA7 PE=1 SV=2	6	46.3	140	110.6	1.266	0.340
39	High	Q6EMK4	Vasorin OS=Homo sapiens OX=9606 GN=VASN PE=1 SV=1	5	71.7	159	127	1.252	0.324
40	High	Q9UBG3	Cornulin OS=Homo sapiens OX=9606 GN=CRNN PE=1 SV=1	4	53.5	166.4	133.1	1.25	0.322
41	High	P00450	Ceruloplasmin OS=Homo sapiens OX=9606 GN=CP PE=1 SV=1	13	122.1	112	89.8	1.247	0.318
42	High	P15586	N-acetylglucosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GNS PE=1 SV=3	3	62	134.1	108	1.241	0.312
43	High	P02760	Protein AMBP OS=Homo sapiens OX=9606 GN=AMBP PE=1 SV=1	5	39	136.7	110.6	1.237	0.307
44	High	P05937	Calbindin OS=Homo sapiens OX=9606 GN=CALB1 PE=1 SV=2	3	30	186	152.3	1.221	0.288
45	High	Q92820	Gamma-glutamyl hydrolase OS=Homo sapiens OX=9606 GN=GGH PE=1 SV=2	3	35.9	133.8	109.8	1.219	0.286
46	High	Q13510	Acid ceramidase OS=Homo sapiens OX=9606 GN=ASAH1 PE=1 SV=5	3	44.6	134	110.5	1.213	0.279
47	High	Q9UBC9	Small proline-rich protein 3 OS=Homo sapiens OX=9606 GN=SPRR3 PE=1 SV=2	5	18.1	162.5	134.3	1.21	0.275
48	High	B9A064	Immunoglobulin lambda-like polypeptide 5 OS=Homo sapiens OX=9606 GN=IGLL5 PE=2 SV=2	3	23	136.4	113.8	1.199	0.262
49	High	P02675	Fibrinogen beta chain OS=Homo sapiens OX=9606 GN=FGB PE=1 SV=2	2	55.9	136.6	113.9	1.199	0.262
50	High	Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens OX=9606 GN=LMAN2 PE=1 SV=1	3	40.2	128.3	108.5	1.183	0.242
51	High	P34059	N-acetylgalactosamine-6-sulfatase	2	58	146.8	124.8	1.176	0.234

			OS=Homo sapiens OX=9606 GN=GALNS PE=1 SV=1						
52	High	P06702	Protein S100-A9 OS=Homo sapiens OX=9606 GN=S100A9 PE=1 SV=1	7	13.2	147	125.6	1.171	0.228
53	High	P01834	Immunoglobulin kappa constant OS=Homo sapiens OX=9606 GN=IGKC PE=1 SV=2	3	11.8	138.4	119.5	1.158	0.212
54	High	O94919	Endonuclease domain-containing 1 protein OS=Homo sapiens OX=9606 GN=ENDOD1 PE=1 SV=2	3	55	122.1	106.7	1.145	0.195
55	High	P29508	Serpin B3 OS=Homo sapiens OX=9606 GN=SERPINB3 PE=1 SV=2	6	44.5	166.2	147	1.13	0.176
56	High	P01859	Immunoglobulin heavy constant gamma 2 OS=Homo sapiens OX=9606 GN=IGHG2 PE=1 SV=2	4	35.9	104.4	93.2	1.119	0.162
57	High	P39059	Collagen alpha-1(XV) chain OS=Homo sapiens OX=9606 GN=COL15A1 PE=1 SV=2	2	141.6	137.7	123.7	1.113	0.154
58	High	P43121	Cell surface glycoprotein MUC18 OS=Homo sapiens OX=9606 GN=MCAM PE=1 SV=2	3	71.6	135.4	123.4	1.098	0.135
59	High	P05109	Protein S100-A8 OS=Homo sapiens OX=9606 GN=S100A8 PE=1 SV=1	5	10.8	141.8	131.2	1.081	0.112
60	High	P09211	Glutathione S-transferase P OS=Homo sapiens OX=9606 GN=GSTP1 PE=1 SV=2	2	23.3	137.3	127.1	1.08	0.111
61	High	P05090	Apolipoprotein D OS=Homo sapiens OX=9606 GN=APOD PE=1 SV=1	3	21.3	134.6	125.8	1.07	0.098
62	High	P60709	Actin, cytoplasmic 1 OS=Homo sapiens OX=9606 GN=ACTB PE=1 SV=1	6	41.7	119	112.5	1.058	0.081
63	High	P10451	Osteopontin OS=Homo sapiens OX=9606 GN=SPP1 PE=1 SV=1	11	35.4	132.4	125.2	1.057	0.080
64	High	P15289	Arylsulfatase A OS=Homo sapiens	6	53.6	129.8	123	1.055	0.077

			OX=9606 GN=ARSA PE=1 SV=3						
65	High	P02751	Fibronectin OS=Homo sapiens OX=9606 GN=FN1 PE=1 SV=4	11	262.5	139.8	132.9	1.052	0.073
66	High	P10599	Thioredoxin OS=Homo sapiens OX=9606 GN=TXN PE=1 SV=3	3	11.7	135.9	129.6	1.048	0.068
67	High	P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens OX=9606 GN=HSPA8 PE=1 SV=1	9	70.9	122.2	117.2	1.043	0.061
68	High	P05155	Plasma protease C1 inhibitor OS=Homo sapiens OX=9606 GN=SERPING1 PE=1 SV=2	6	55.1	147.6	141.6	1.042	0.059
69	High	Q8NFZ8	Cell adhesion molecule 4 OS=Homo sapiens OX=9606 GN=CADM4 PE=1 SV=1	2	42.8	165.1	158.6	1.041	0.058
70	High	P0C0L5	Complement C4-B OS=Homo sapiens OX=9606 GN=C4B PE=1 SV=2	3	192.6	135.9	130.9	1.039	0.055
71	High	P01040	Cystatin-A OS=Homo sapiens OX=9606 GN=CSTA PE=1 SV=1	2	11	155.6	151.4	1.027	0.038
72	High	Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens OX=9606 GN=ITIH4 PE=1 SV=4	4	103.3	140.3	137.3	1.022	0.031
73	High	P06733	Alpha-enolase OS=Homo sapiens OX=9606 GN=ENO1 PE=1 SV=2	4	47.1	119.1	117.3	1.015	0.021
74	High	P19835	Bile salt-activated lipase OS=Homo sapiens OX=9606 GN=CEL PE=1 SV=3	3	79.3	174	173.4	1.003	0.004
75	High	P00924	Enolase 1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ENO1 PE=1 SV=3	14	46.8	100	100	1	0.000
76	High	P15144	Aminopeptidase N OS=Homo sapiens OX=9606 GN=ANPEP PE=1 SV=4	12	109.5	121.1	123.2	0.983	-0.025
77	High	Q16270	Insulin-like growth factor-binding protein 7 OS=Homo sapiens OX=9606	3	29.1	160.9	163.6	0.983	-0.025

			GN=IGFBP7 PE=1 SV=1						
78	High	Q9NZP8	Complement C1r subcomponent-like protein OS=Homo sapiens OX=9606 GN=C1RL PE=1 SV=2	3	53.5	140.8	143.5	0.981	-0.028
79	High	Q9BRK3	Matrix remodeling-associated protein 8 OS=Homo sapiens OX=9606 GN=MXRA8 PE=1 SV=1	4	49.1	141.1	144.2	0.978	-0.032
80	High	P12109	Collagen alpha-1(VI) chain OS=Homo sapiens OX=9606 GN=COL6A1 PE=1 SV=3	10	108.5	142.7	146.1	0.977	-0.034
81	High	P09668	Pro-cathepsin H OS=Homo sapiens OX=9606 GN=CTSH PE=1 SV=4	2	37.4	124.8	127.9	0.976	-0.035
82	High	Q8WZ75	Roundabout homolog 4 OS=Homo sapiens OX=9606 GN=ROBO4 PE=1 SV=1	2	107.4	133.3	137.3	0.971	-0.042
83	High	P78324	Tyrosine-protein phosphatase non-receptor type substrate 1 OS=Homo sapiens OX=9606 GN=SIRPA PE=1 SV=2	3	54.9	126.1	130.6	0.965	-0.051
84	High	A8K2U0	Alpha-2-macroglobulin-like protein 1 OS=Homo sapiens OX=9606 GN=A2ML1 PE=1 SV=3	6	161	130.5	135.5	0.963	-0.054
85	High	P05154	Plasma serine protease inhibitor OS=Homo sapiens OX=9606 GN=SERPINA5 PE=1 SV=3	4	45.6	138.2	144.4	0.957	-0.063
86	High	P05062	Fructose-bisphosphate aldolase B OS=Homo sapiens OX=9606 GN=ALDOB PE=1 SV=2	2	39.4	132.6	138.6	0.956	-0.065
87	High	P55290	Cadherin-13 OS=Homo sapiens OX=9606 GN=CDH13 PE=1 SV=1	3	78.2	121.8	128.5	0.948	-0.077
88	High	Q7KYR7	Butyrophilin subfamily 2 member A1 OS=Homo sapiens OX=9606 GN=BTN2A1 PE=1 SV=3	3	59.6	130.9	139.9	0.936	-0.095

89	High	P05060	Secretogranin-1 OS=Homo sapiens OX=9606 GN=CHGB PE=1 SV=2	2	78.2	130.2	139.2	0.935	-0.097
90	High	P62736	Actin, aortic smooth muscle OS=Homo sapiens OX=9606 GN=ACTA2 PE=1 SV=1	5	42	101.6	108.8	0.934	-0.099
91	High	P62979	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens OX=9606 GN=RPS27A PE=1 SV=2	3	18	137.4	147.1	0.934	-0.099
92	High	P07195	L-lactate dehydrogenase B chain OS=Homo sapiens OX=9606 GN=LDHB PE=1 SV=2	2	36.6	124	132.8	0.933	-0.100
93	High	P01857	Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1 SV=1	7	36.1	128.8	138.4	0.931	-0.103
94	High	Q9BY67	Cell adhesion molecule 1 OS=Homo sapiens OX=9606 GN=CADM1 PE=1 SV=2	3	48.5	138.9	150.4	0.924	-0.114
95	High	P50995	Annexin A11 OS=Homo sapiens OX=9606 GN=ANXA11 PE=1 SV=1	3	54.4	137.6	150.1	0.917	-0.125
96	High	P31947	14-3-3 protein sigma OS=Homo sapiens OX=9606 GN=SFN PE=1 SV=1	5	27.8	107.2	117.4	0.913	-0.131
97	High	Q6UVK1	Chondroitin sulfate proteoglycan 4 OS=Homo sapiens OX=9606 GN=CSPG4 PE=1 SV=2	2	250.4	124.5	138	0.902	-0.149
98	High	P19440	Glutathione hydrolase 1 proenzyme OS=Homo sapiens OX=9606 GN=GGT1 PE=1 SV=2	2	61.4	134.1	148.6	0.902	-0.149
99	High	P60174	Triosephosphate isomerase OS=Homo sapiens OX=9606 GN=TPI1 PE=1 SV=3	2	30.8	140.9	156.8	0.899	-0.154
100	High	Q96NY8	Nectin-4 OS=Homo sapiens OX=9606 GN=NECTIN4 PE=1 SV=1	3	55.4	127.9	142.5	0.898	-0.155
101	High	Q16651	Prostasin OS=Homo sapiens OX=9606 GN=PRSS8 PE=1 SV=1	2	36.4	134.5	150.7	0.893	-0.163

102	High	O00391	Sulfhydryl oxidase 1 OS=Homo sapiens OX=9606 GN=QSOX1 PE=1 SV=3	5	82.5	145.8	163.7	0.891	-0.167
103	High	P01023	Alpha-2-macroglobulin OS=Homo sapiens OX=9606 GN=A2M PE=1 SV=3	2	163.2	88.7	99.6	0.89	-0.168
104	High	P02768	Serum albumin OS=Homo sapiens OX=9606 GN=ALB PE=1 SV=2	19	69.3	117.1	132	0.887	-0.173
105	High	P02790	Hemopexin OS=Homo sapiens OX=9606 GN=HPX PE=1 SV=2	5	51.6	116.8	131.8	0.886	-0.175
106	High	Q14515	SPARC-like protein 1 OS=Homo sapiens OX=9606 GN=SPARCL1 PE=1 SV=2	4	75.2	119.2	134.8	0.884	-0.178
107	High	P26992	Ciliary neurotrophic factor receptor subunit alpha OS=Homo sapiens OX=9606 GN=CNTFR PE=1 SV=2	2	40.6	136	154.3	0.881	-0.183
108	High	P14384	Carboxypeptidase M OS=Homo sapiens OX=9606 GN=CPM PE=1 SV=2	2	50.5	143.4	163.3	0.878	-0.188
109	High	P24855	Deoxyribonuclease-1 OS=Homo sapiens OX=9606 GN=DNASE1 PE=1 SV=1	3	31.4	154	175.8	0.876	-0.191
110	High	Q8NBJ4	Golgi membrane protein 1 OS=Homo sapiens OX=9606 GN=GOLM1 PE=1 SV=1	7	45.3	140.8	160.8	0.875	-0.193
111	High	Q06830	Peroxiredoxin-1 OS=Homo sapiens OX=9606 GN=PRDX1 PE=1 SV=1	2	22.1	111.3	127.6	0.872	-0.198
112	High	P54802	Alpha-N-acetylglucosaminidase OS=Homo sapiens OX=9606 GN=NAGLU PE=1 SV=2	5	82.2	107.5	124.4	0.864	-0.211
113	High	Q08380	Galectin-3-binding protein OS=Homo sapiens OX=9606 GN=LGALS3BP PE=1 SV=1	10	65.3	118.5	137.3	0.863	-0.213
114	High	Q07075	Glutamyl aminopeptidase OS=Homo sapiens OX=9606 GN=ENPEP PE=1 SV=3	3	109.2	106.6	123.6	0.863	-0.213
115	High	Q8TDQ0	Hepatitis A virus cellular receptor 2 OS=Homo sapiens OX=9606	2	33.4	127.3	148.1	0.86	-0.218

			GN=HAVCR2 PE=1 SV=3						
116	High	P10909	Clusterin OS=Homo sapiens OX=9606 GN=CLU PE=1 SV=1	4	52.5	135.6	158.2	0.857	-0.223
117	High	P25815	Protein S100-P OS=Homo sapiens OX=9606 GN=S100P PE=1 SV=2	2	10.4	106.8	125.1	0.853	-0.229
118	High	Q9HCU0	Endosialin OS=Homo sapiens OX=9606 GN=CD248 PE=1 SV=1	2	80.8	127.9	149.9	0.853	-0.229
119	High	Q9P121	Neurotrimin OS=Homo sapiens OX=9606 GN=NTM PE=1 SV=1	2	37.9	144.9	170.2	0.851	-0.233
120	High	P07339	Cathepsin D OS=Homo sapiens OX=9606 GN=CTSD PE=1 SV=1	6	44.5	116.6	137.4	0.849	-0.236
121	High	P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens OX=9606 GN=KRT16 PE=1 SV=4	6	51.2	139	164.6	0.844	-0.245
122	High	P01042	Kininogen-1 OS=Homo sapiens OX=9606 GN=KNG1 PE=1 SV=2	6	71.9	125	149.2	0.838	-0.255
123	High	P63104	14-3-3 protein zeta/delta OS=Homo sapiens OX=9606 GN=YWHAZ PE=1 SV=1	5	27.7	101.8	122.5	0.832	-0.265
124	High	P10153	Non-secretory ribonuclease OS=Homo sapiens OX=9606 GN=RNASE2 PE=1 SV=2	2	18.3	133.5	161.6	0.826	-0.276
125	High	Q08345	Epithelial discoidin domain-containing receptor 1 OS=Homo sapiens OX=9606 GN=DDR1 PE=1 SV=1	4	101.1	107.9	131.5	0.82	-0.286
126	High	P98164	Low-density lipoprotein receptor- related protein 2 OS=Homo sapiens OX=9606 GN=LRP2 PE=1 SV=3	6	521.6	118.1	144.5	0.818	-0.290
127	High	P01133	Pro-epidermal growth factor OS=Homo sapiens OX=9606 GN=EGF PE=1 SV=2	11	133.9	135.4	166.4	0.814	-0.297
128	High	P05164	Myeloperoxidase OS=Homo sapiens OX=9606 GN=MPO PE=1 SV=1	3	83.8	103.6	128.2	0.808	-0.308

129	High	Q99988	Growth/differentiation factor 15 OS=Homo sapiens OX=9606 GN=GDF15 PE=1 SV=3	3	34.1	119.9	149.5	0.802	-0.318
130	High	Q9ULI3	Protein HEG homolog 1 OS=Homo sapiens OX=9606 GN=HEG1 PE=1 SV=3	3	147.4	120.7	151.7	0.795	-0.331
131	High	Q8IYS5	Osteoclast-associated immunoglobulin- like receptor OS=Homo sapiens OX=9606 GN=OSCAR PE=1 SV=3	2	30.5	129.9	164.2	0.791	-0.338
132	High	P98160	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens OX=9606 GN=HSPG2 PE=1 SV=4	12	468.5	106.1	134.6	0.788	-0.344
133	High	O75882	Attractin OS=Homo sapiens OX=9606 GN=ATRN PE=1 SV=2	4	158.4	118.9	152.5	0.779	-0.360
134	High	Q68CJ9	Cyclic AMP-responsive element-binding protein 3-like protein 3 OS=Homo sapiens OX=9606 GN=CREB3L3 PE=1 SV=2	2	49	116.7	149.8	0.779	-0.360
135	High	P07355	Annexin A2 OS=Homo sapiens OX=9606 GN=ANXA2 PE=1 SV=2	8	38.6	114.3	148.2	0.772	-0.373
136	High	P01876	Immunoglobulin heavy constant alpha 1 OS=Homo sapiens OX=9606 GN=IGHA1 PE=1 SV=2	5	37.6	110.5	143.8	0.768	-0.381
137	High	P06870	Kallikrein-1 OS=Homo sapiens OX=9606 GN=KLK1 PE=1 SV=2	3	28.9	138.9	181.2	0.767	-0.383
138	High	P09603	Macrophage colony-stimulating factor 1 OS=Homo sapiens OX=9606 GN=CSF1 PE=1 SV=2	4	60.1	122.1	159.8	0.764	-0.388
139	High	Q9UQ72	Pregnancy-specific beta-1-glycoprotein 11 OS=Homo sapiens OX=9606 GN=PSG11 PE=2 SV=3	7	37.1	87.1	115	0.757	-0.402
140	High	P01833	Polymeric immunoglobulin receptor	7	83.2	112.9	149.5	0.755	-0.405

			OS=Homo sapiens OX=9606 GN=PIGR PE=1 SV=4						
141	High	P02774	Vitamin D-binding protein OS=Homo sapiens OX=9606 GN=GC PE=1 SV=1	5	52.9	117	157.5	0.743	-0.429
142	High	Q9Y646	Carboxypeptidase Q OS=Homo sapiens OX=9606 GN=CPQ PE=1 SV=1	2	51.9	113	153.7	0.735	-0.444
143	High	P15941	Mucin-1 OS=Homo sapiens OX=9606 GN=MUC1 PE=1 SV=3	3	122	118.7	162.9	0.729	-0.456
144	High	P10253	Lysosomal alpha-glucosidase OS=Homo sapiens OX=9606 GN=GAA PE=1 SV=4	4	105.3	125.6	173.8	0.723	-0.468
145	High	P04746	Pancreatic alpha-amylase OS=Homo sapiens OX=9606 GN=AMY2A PE=1 SV=2	7	57.7	109.9	152.3	0.722	-0.470
146	High	P27482	Calmodulin-like protein 3 OS=Homo sapiens OX=9606 GN=CALML3 PE=1 SV=2	4	16.9	128.9	179.1	0.72	-0.474
147	High	Q6GTX8	Leukocyte-associated immunoglobulin-like receptor 1 OS=Homo sapiens OX=9606 GN=LAIR1 PE=1 SV=1	2	31.4	121.2	168.4	0.72	-0.474
148	High	P04083	Annexin A1 OS=Homo sapiens OX=9606 GN=ANXA1 PE=1 SV=2	6	38.7	76.9	107	0.719	-0.476
149	High	P06703	Protein S100-A6 OS=Homo sapiens OX=9606 GN=S100A6 PE=1 SV=1	3	10.2	96.1	134	0.717	-0.480
150	High	O60494	Cubilin OS=Homo sapiens OX=9606 GN=CUBN PE=1 SV=5	3	398.5	99.2	142.7	0.695	-0.525
151	High	Q01459	Di-N-acetylchitinase OS=Homo sapiens OX=9606 GN=CTBS PE=1 SV=1	2	43.7	89.5	129.1	0.694	-0.527
152	High	P01861	Immunoglobulin heavy constant gamma 4 OS=Homo sapiens OX=9606 GN=IGHG4 PE=1 SV=1	3	35.9	109.4	158.3	0.691	-0.533
153	High	P29972	Aquaporin-1 OS=Homo sapiens OX=9606 GN=AQP1 PE=1 SV=3	2	28.5	96.7	140.7	0.687	-0.542

154	High	O43451	Maltase-glucoamylase, intestinal OS=Homo sapiens OX=9606 GN=MGAM PE=1 SV=5	4	209.7	108.6	159.3	0.681	-0.554
155	High	Q9NQ84	G-protein coupled receptor family C group 5 member C OS=Homo sapiens OX=9606 GN=GPRC5C PE=1 SV=2	2	48.2	124.4	183.3	0.679	-0.559
156	High	P80188	Neutrophil gelatinase-associated lipocalin OS=Homo sapiens OX=9606 GN=LCN2 PE=1 SV=2	3	22.6	98.5	145.5	0.677	-0.563
157	High	P12273	Prolactin-inducible protein OS=Homo sapiens OX=9606 GN=PIP PE=1 SV=1	2	16.6	118.8	179.4	0.662	-0.595
158	High	P11021	Endoplasmic reticulum chaperone BiP OS=Homo sapiens OX=9606 GN=HSPA5 PE=1 SV=2	5	72.3	117.9	178.8	0.659	-0.602
159	High	P00738	Haptoglobin OS=Homo sapiens OX=9606 GN=HP PE=1 SV=1	4	45.2	80.7	122.7	0.658	-0.604
160	High	P18510	Interleukin-1 receptor antagonist protein OS=Homo sapiens OX=9606 GN=IL1RN PE=1 SV=1	2	20	101.3	153.8	0.658	-0.604
161	High	P11464	Pregnancy-specific beta-1-glycoprotein 1 OS=Homo sapiens OX=9606 GN=PSG1 PE=1 SV=1	4	47.2	81.2	125.2	0.649	-0.624
162	High	P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens OX=9606 GN=KRT10 PE=1 SV=6	14	58.8	134.4	207.8	0.647	-0.628
163	High	P01591	Immunoglobulin J chain OS=Homo sapiens OX=9606 GN=JCHAIN PE=1 SV=4	2	18.1	91	149.7	0.608	-0.718
164	High	P50895	Basal cell adhesion molecule OS=Homo sapiens OX=9606 GN=BCAM PE=1 SV=2	3	67.4	106.7	185.7	0.575	-0.798
165	High	P62873	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1 OS=Homo	2	37.4	97.5	172.8	0.564	-0.826

			sapiens OX=9606 GN=GNB1 PE=1 SV=3						
166	High	P07911	Uromodulin OS=Homo sapiens OX=9606 GN=UMOD PE=1 SV=1	7	69.7	82.1	148.7	0.552	-0.857
167	High	P16070	CD44 antigen OS=Homo sapiens OX=9606 GN=CD44 PE=1 SV=3	3	81.5	78.6	146.8	0.535	-0.902
168	High	P01877	Immunoglobulin heavy constant alpha 2 OS=Homo sapiens OX=9606 GN=IGHA2 PE=1 SV=4	3	36.6	82.3	155.3	0.53	-0.916
169	High	P08637	Low affinity immunoglobulin gamma Fc region receptor III-A OS=Homo sapiens OX=9606 GN=FCGR3A PE=1 SV=2	2	29.1	110.9	232.1	0.478	-1.065
170	High	P13473	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens OX=9606 GN=LAMP2 PE=1 SV=2	2	44.9	98.8	210.1	0.47	-1.089
171	High	P31944	Caspase-14 OS=Homo sapiens OX=9606 GN=CASP14 PE=1 SV=2	2	27.7	85	189.9	0.448	-1.158
172	High	P13796	Plastin-2 OS=Homo sapiens OX=9606 GN=LCP1 PE=1 SV=6	4	70.2	52.4	161.2	0.325	-1.621
173	High	P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens OX=9606 GN=KRT2 PE=1 SV=2	5	65.4	89.2	277.4	0.321	-1.639

Data identified at confidence interval (CI): 99 %, FDR < 0.01; Peptides no. ≥ 2.

Table S 4: Dataset of protein expression for IUGR and control at 20 weeks of gestation

No.	Protein FDR Confidence: Combined	Accession No.	Protein name	# Unique Peptides	MW [kDa]	Abundances of reporter ion of IUGR at 20 W	Abundances of reporter ion of control at 20 W	Abundance ratio (Case / Control) of IUGR at 20 W	Log ₂ fold change
1	High	P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=AZGP1 PE=1 SV=2	13	34.2	171.4	58.6	2.927	1.549
2	High	P02675	Fibrinogen beta chain OS=Homo sapiens OX=9606 GN=FGB PE=1 SV=2	2	55.9	133.2	50.5	2.638	1.399
3	High	P02750	Leucine-rich alpha-2-glycoprotein OS=Homo sapiens OX=9606 GN=LRG1 PE=1 SV=2	5	38.2	173.2	70.8	2.445	1.290
4	High	P02763	Alpha-1-acid glycoprotein 1 OS=Homo sapiens OX=9606 GN=ORM1 PE=1 SV=1	3	23.5	131.4	62.7	2.097	1.068
5	High	P19652	Alpha-1-acid glycoprotein 2 OS=Homo sapiens OX=9606 GN=ORM2 PE=1 SV=2	2	23.6	127.6	67.3	1.895	0.922
6	High	P16070	CD44 antigen OS=Homo sapiens OX=9606 GN=CD44 PE=1 SV=3	3	81.5	104.1	60.3	1.727	0.788
7	High	Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1 OS=Homo sapiens OX=9606 GN=LYVE1 PE=1 SV=2	2	35.2	110.5	69.2	1.598	0.676
8	High	Q01459	Di-N-acetylchitobiase OS=Homo sapiens OX=9606 GN=CTBS PE=1 SV=1	2	43.7	149.5	96.3	1.552	0.634
9	High	P01008	Antithrombin-III OS=Homo sapiens OX=9606 GN=SERPINC1 PE=1 SV=1	4	52.6	118.3	76.8	1.541	0.624
10	High	P01019	Angiotensinogen OS=Homo sapiens	2	53.1	148.4	98.9	1.501	0.586

			OX=9606 GN=AGT PE=1 SV=1						
11	High	A8K2U0	Alpha-2-macroglobulin-like protein 1 OS=Homo sapiens OX=9606 GN=A2ML1 PE=1 SV=3	6	161	116.8	81.2	1.439	0.525
12	High	P12273	Prolactin-inducible protein OS=Homo sapiens OX=9606 GN=PIP PE=1 SV=1	2	16.6	93	65.7	1.415	0.501
13	High	P63104	14-3-3 protein zeta/delta OS=Homo sapiens OX=9606 GN=YWHAZ PE=1 SV=1	5	27.7	162.3	116.3	1.396	0.481
14	High	Q9UBG3	Cornulin OS=Homo sapiens OX=9606 GN=CRNN PE=1 SV=1	4	53.5	116.2	87.8	1.323	0.404
15	High	Q6UXB8	Peptidase inhibitor 16 OS=Homo sapiens OX=9606 GN=PI16 PE=1 SV=1	2	49.4	116.6	90	1.296	0.374
16	High	P02787	Serotransferrin OS=Homo sapiens OX=9606 GN=TF PE=1 SV=3	12	77	102.4	80	1.28	0.356
17	High	P04217	Alpha-1B-glycoprotein OS=Homo sapiens OX=9606 GN=A1BG PE=1 SV=4	3	54.2	85.3	67.8	1.257	0.330
18	High	P18510	Interleukin-1 receptor antagonist protein OS=Homo sapiens OX=9606 GN=IL1RN PE=1 SV=1	2	20	130.9	104.2	1.256	0.329
19	High	P29508	Serpin B3 OS=Homo sapiens OX=9606 GN=SERPINB3 PE=1 SV=2	6	44.5	92.4	77.3	1.196	0.258
20	High	Q9UBC9	Small proline-rich protein 3 OS=Homo sapiens OX=9606 GN=SPRR3 PE=1 SV=2	5	18.1	102.1	85.9	1.188	0.249
21	High	Q14515	SPARC-like protein 1 OS=Homo sapiens OX=9606 GN=SPARCL1 PE=1 SV=2	4	75.2	124.7	105.2	1.185	0.245
22	High	P07355	Annexin A2 OS=Homo sapiens OX=9606 GN=ANXA2 PE=1 SV=2	8	38.6	108.9	92.3	1.179	0.238
23	High	P13796	Plastin-2 OS=Homo sapiens OX=9606	4	70.2	118.9	101.2	1.174	0.231

			GN=LCP1 PE=1 SV=6						
24	High	P02765	Alpha-2-HS-glycoprotein OS=Homo sapiens OX=9606 GN=AHSG PE=1 SV=1	3	39.3	97	82.8	1.171	0.228
25	High	P06702	Protein S100-A9 OS=Homo sapiens OX=9606 GN=S100A9 PE=1 SV=1	7	13.2	102.9	88.8	1.158	0.212
26	High	P08185	Corticosteroid-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA6 PE=1 SV=1	3	45.1	89.1	77.1	1.155	0.208
27	High	P02760	Protein AMBP OS=Homo sapiens OX=9606 GN=AMBP PE=1 SV=1	5	39	109.3	94.8	1.153	0.205
28	High	P27482	Calmodulin-like protein 3 OS=Homo sapiens OX=9606 GN=CALML3 PE=1 SV=2	4	16.9	97.7	85.1	1.148	0.199
29	High	Q8NFZ8	Cell adhesion molecule 4 OS=Homo sapiens OX=9606 GN=CADM4 PE=1 SV=1	2	42.8	91.3	82	1.113	0.154
30	High	B9A064	Immunoglobulin lambda-like polypeptide 5 OS=Homo sapiens OX=9606 GN=IGLL5 PE=2 SV=2	3	23	103.4	93.6	1.105	0.144
31	High	P10599	Thioredoxin OS=Homo sapiens OX=9606 GN=TXN PE=1 SV=3	3	11.7	84.6	77.3	1.094	0.130
32	High	P01040	Cystatin-A OS=Homo sapiens OX=9606 GN=CSTA PE=1 SV=1	2	11	94.4	86.5	1.091	0.126
33	High	O00187	Mannan-binding lectin serine protease 2 OS=Homo sapiens OX=9606 GN=MASP2 PE=1 SV=4	3	75.7	98.5	90.8	1.085	0.118
34	High	Q6GTX8	Leukocyte-associated immunoglobulin-like receptor 1 OS=Homo sapiens OX=9606 GN=LAIR1 PE=1 SV=1	2	31.4	60.1	55.5	1.081	0.112
35	High	P01614	Immunoglobulin kappa variable 2D-40	2	13.3	118	110.5	1.068	0.095

			OS=Homo sapiens OX=9606 GN=IGKV2D-40 PE=1 SV=2						
36	High	P0DOY2	Immunoglobulin lambda constant 2 OS=Homo sapiens OX=9606 GN=IGLC2 PE=1 SV=1	3	11.3	107.4	101.5	1.059	0.083
37	High	P08571	Monocyte differentiation antigen CD14 OS=Homo sapiens OX=9606 GN=CD14 PE=1 SV=2	4	40.1	93.8	89.6	1.047	0.066
38	High	P55290	Cadherin-13 OS=Homo sapiens OX=9606 GN=CDH13 PE=1 SV=1	3	78.2	95.9	91.6	1.047	0.066
39	High	P07476	Involucrin OS=Homo sapiens OX=9606 GN=IVL PE=1 SV=2	9	68.4	80.1	76.5	1.046	0.065
40	High	P02768	Serum albumin OS=Homo sapiens OX=9606 GN=ALB PE=1 SV=2	19	69.3	85.7	82.1	1.044	0.062
41	High	O94919	Endonuclease domain-containing 1 protein OS=Homo sapiens OX=9606 GN=ENDOD1 PE=1 SV=2	3	55	127.4	122	1.044	0.062
42	High	P05543	Thyroxine-binding globulin OS=Homo sapiens OX=9606 GN=SERPINA7 PE=1 SV=2	6	46.3	95.4	91.8	1.039	0.055
43	High	P01834	Immunoglobulin kappa constant OS=Homo sapiens OX=9606 GN=IGKC PE=1 SV=2	3	11.8	101	98.4	1.027	0.038
44	High	P07911	Uromodulin OS=Homo sapiens OX=9606 GN=UMOD PE=1 SV=1	7	69.7	101.4	98.9	1.025	0.036
45	High	P01011	Alpha-1-antichymotrypsin OS=Homo sapiens OX=9606 GN=SERPINA3 PE=1 SV=2	6	47.6	96.2	94.2	1.022	0.031
46	High	P98160	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens OX=9606 GN=HSPG2 PE=1 SV=4	12	468.5	105.8	103.9	1.018	0.026

47	High	P02790	Hemopexin OS=Homo sapiens OX=9606 GN=HPX PE=1 SV=2	5	51.6	104.2	102.4	1.017	0.024
48	High	P05109	Protein S100-A8 OS=Homo sapiens OX=9606 GN=S100A8 PE=1 SV=1	5	10.8	104.8	104.6	1.002	0.003
49	High	P00924	Enolase 1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ENO1 PE=1 SV=3	14	46.8	100	100	1	0.000
50	High	Q93088	Betaine--homocysteine S- methyltransferase 1 OS=Homo sapiens OX=9606 GN=BHMT PE=1 SV=2	3	45	100.5	100.7	0.998	-0.003
51	High	P06396	Gelsolin OS=Homo sapiens OX=9606 GN=GSN PE=1 SV=1	6	85.6	89.3	90	0.991	-0.013
52	High	P01861	Immunoglobulin heavy constant gamma 4 OS=Homo sapiens OX=9606 GN=IGHG4 PE=1 SV=1	3	35.9	72.3	73.6	0.983	-0.025
53	High	Q9HCU0	Endosialin OS=Homo sapiens OX=9606 GN=CD248 PE=1 SV=1	2	80.8	53.9	55.5	0.97	-0.044
54	High	P39059	Collagen alpha-1(XV) chain OS=Homo sapiens OX=9606 GN=COL15A1 PE=1 SV=2	2	141.6	84.1	88.5	0.95	-0.074
55	High	P50995	Annexin A11 OS=Homo sapiens OX=9606 GN=ANXA11 PE=1 SV=1	3	54.4	108.4	114.7	0.945	-0.082
56	High	P01009	Alpha-1-antitrypsin OS=Homo sapiens OX=9606 GN=SERPINA1 PE=1 SV=3	14	46.7	69.5	74.4	0.934	-0.099
57	High	P09211	Glutathione S-transferase P OS=Homo sapiens OX=9606 GN=GSTP1 PE=1 SV=2	2	23.3	141	151.8	0.929	-0.106
58	High	P01871	Immunoglobulin heavy constant mu OS=Homo sapiens OX=9606 GN=IGHM PE=1 SV=4	2	49.4	73.6	79.5	0.927	-0.109
59	High	P05155	Plasma protease C1 inhibitor	6	55.1	79.6	86.7	0.918	-0.123

			OS=Homo sapiens OX=9606 GN=SERPING1 PE=1 SV=2						
60	High	O43451	Maltase-glucoamylase, intestinal OS=Homo sapiens OX=9606 GN=MGAM PE=1 SV=5	4	209.7	101.7	111.6	0.911	-0.134
61	High	P00450	Ceruloplasmin OS=Homo sapiens OX=9606 GN=CP PE=1 SV=1	13	122.1	75.4	83.5	0.903	-0.147
62	High	P05090	Apolipoprotein D OS=Homo sapiens OX=9606 GN=APOD PE=1 SV=1	3	21.3	76.2	84.9	0.897	-0.157
63	High	P24855	Deoxyribonuclease-1 OS=Homo sapiens OX=9606 GN=DNASE1 PE=1 SV=1	3	31.4	89.8	100.7	0.892	-0.165
64	High	P06703	Protein S100-A6 OS=Homo sapiens OX=9606 GN=S100A6 PE=1 SV=1	3	10.2	104.9	119	0.882	-0.181
65	High	P02774	Vitamin D-binding protein OS=Homo sapiens OX=9606 GN=GC PE=1 SV=1	5	52.9	87.1	100.4	0.868	-0.204
66	High	P98164	Low-density lipoprotein receptor- related protein 2 OS=Homo sapiens OX=9606 GN=LRP2 PE=1 SV=3	6	521.6	88.4	102.5	0.862	-0.214
67	High	Q02487	Desmocollin-2 OS=Homo sapiens OX=9606 GN=DSC2 PE=1 SV=1	3	99.9	92.4	107.3	0.861	-0.216
68	High	P01857	Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1 SV=1	7	36.1	92.2	107.7	0.856	-0.224
69	High	Q13510	Acid ceramidase OS=Homo sapiens OX=9606 GN=ASAH1 PE=1 SV=5	3	44.6	81	94.9	0.854	-0.228
70	High	P80188	Neutrophil gelatinase-associated lipocalin OS=Homo sapiens OX=9606 GN=LCN2 PE=1 SV=2	3	22.6	106.7	126.3	0.845	-0.243
71	High	P15941	Mucin-1 OS=Homo sapiens OX=9606 GN=MUC1 PE=1 SV=3	3	122	69.3	82.3	0.841	-0.250
72	High	P07339	Cathepsin D OS=Homo sapiens	6	44.5	93.4	111.7	0.836	-0.258

			OX=9606 GN=CTSD PE=1 SV=1						
73	High	Q92820	Gamma-glutamyl hydrolase OS=Homo sapiens OX=9606 GN=GGH PE=1 SV=2	3	35.9	76	90.9	0.836	-0.258
74	High	Q9Y646	Carboxypeptidase Q OS=Homo sapiens OX=9606 GN=CPQ PE=1 SV=1	2	51.9	86.2	104.3	0.826	-0.276
75	High	P01833	Polymeric immunoglobulin receptor OS=Homo sapiens OX=9606 GN=PIGR PE=1 SV=4	7	83.2	90.6	110	0.824	-0.279
76	High	Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens OX=9606 GN=ITIH4 PE=1 SV=4	4	103.3	73.9	89.8	0.823	-0.281
77	High	P14384	Carboxypeptidase M OS=Homo sapiens OX=9606 GN=CPM PE=1 SV=2	2	50.5	89.8	109.6	0.819	-0.288
78	High	P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens OX=9606 GN=PTGDS PE=1 SV=1	3	21	78.5	95.9	0.818	-0.290
79	High	Q9ULI3	Protein HEG homolog 1 OS=Homo sapiens OX=9606 GN=HEG1 PE=1 SV=3	3	147.4	82.2	100.6	0.818	-0.290
80	High	P02751	Fibronectin OS=Homo sapiens OX=9606 GN=FN1 PE=1 SV=4	11	262.5	87	107.3	0.811	-0.302
81	High	P12830	Cadherin-1 OS=Homo sapiens OX=9606 GN=CDH1 PE=1 SV=3	8	97.4	92.4	114.9	0.804	-0.315
82	High	Q8TDQ0	Hepatitis A virus cellular receptor 2 OS=Homo sapiens OX=9606 GN=HAVCR2 PE=1 SV=3	2	33.4	92.5	115	0.804	-0.315
83	High	P0C0L5	Complement C4-B OS=Homo sapiens OX=9606 GN=C4B PE=1 SV=2	3	192.6	87.1	108.7	0.801	-0.320
84	High	P00738	Haptoglobin OS=Homo sapiens OX=9606 GN=HP PE=1 SV=1	4	45.2	47.8	60	0.796	-0.329
85	High	P09668	Pro-cathepsin H OS=Homo sapiens OX=9606 GN=CTSH PE=1 SV=4	2	37.4	94.4	118.7	0.795	-0.331

86	High	P06733	Alpha-enolase OS=Homo sapiens OX=9606 GN=ENO1 PE=1 SV=2	4	47.1	93.8	118.2	0.794	-0.333
87	High	Q9NZP8	Complement C1r subcomponent-like protein OS=Homo sapiens OX=9606 GN=C1RL PE=1 SV=2	3	53.5	94.1	120.2	0.783	-0.353
88	High	P12109	Collagen alpha-1(VI) chain OS=Homo sapiens OX=9606 GN=COL6A1 PE=1 SV=3	10	108.5	75.7	96.9	0.782	-0.355
89	High	P43121	Cell surface glycoprotein MUC18 OS=Homo sapiens OX=9606 GN=MCAM PE=1 SV=2	3	71.6	98.1	125.5	0.782	-0.355
90	High	Q8NBJ4	Golgi membrane protein 1 OS=Homo sapiens OX=9606 GN=GOLM1 PE=1 SV=1	7	45.3	81.4	105.7	0.77	-0.377
91	High	P05164	Myeloperoxidase OS=Homo sapiens OX=9606 GN=MPO PE=1 SV=1	3	83.8	95.3	123.8	0.77	-0.377
92	High	Q9BY67	Cell adhesion molecule 1 OS=Homo sapiens OX=9606 GN=CADM1 PE=1 SV=2	3	48.5	76.3	99.4	0.768	-0.381
93	High	P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens OX=9606 GN=KRT10 PE=1 SV=6	14	58.8	80.9	105.7	0.766	-0.385
94	High	P01023	Alpha-2-macroglobulin OS=Homo sapiens OX=9606 GN=A2M PE=1 SV=3	2	163.2	75.1	98	0.766	-0.385
95	High	P11464	Pregnancy-specific beta-1- glycoprotein 1 OS=Homo sapiens OX=9606 GN=PSG1 PE=1 SV=1	4	47.2	91.6	119.7	0.765	-0.386
96	High	P25815	Protein S100-P OS=Homo sapiens OX=9606 GN=S100P PE=1 SV=2	2	10.4	100.2	131.9	0.76	-0.396
97	High	O75882	Attractin OS=Homo sapiens OX=9606 GN=ATRN PE=1 SV=2	4	158.4	87.6	116.3	0.754	-0.407
98	High	P04075	Fructose-bisphosphate aldolase A	2	39.4	98.6	131	0.752	-0.411

			OS=Homo sapiens OX=9606 GN=ALDOA PE=1 SV=2						
99	High	P15144	Aminopeptidase N OS=Homo sapiens OX=9606 GN=ANPEP PE=1 SV=4	12	109.5	84.5	112.4	0.751	-0.413
100	High	P31944	Caspase-14 OS=Homo sapiens OX=9606 GN=CASP14 PE=1 SV=2	2	27.7	93.7	125.2	0.748	-0.419
101	High	P31947	14-3-3 protein sigma OS=Homo sapiens OX=9606 GN=SFN PE=1 SV=1	5	27.8	87.2	117.8	0.74	-0.434
102	High	P78324	Tyrosine-protein phosphatase non- receptor type substrate 1 OS=Homo sapiens OX=9606 GN=SIRPA PE=1 SV=2	3	54.9	93.3	126.2	0.74	-0.434
103	High	P04746	Pancreatic alpha-amylase OS=Homo sapiens OX=9606 GN=AMY2A PE=1 SV=2	7	57.7	76.7	104.3	0.736	-0.442
104	High	P05060	Secretogranin-1 OS=Homo sapiens OX=9606 GN=CHGB PE=1 SV=2	2	78.2	84.3	114.6	0.736	-0.442
105	High	P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens OX=9606 GN=HSPA8 PE=1 SV=1	9	70.9	86.4	117.7	0.734	-0.446
106	High	P09603	Macrophage colony-stimulating factor 1 OS=Homo sapiens OX=9606 GN=CSF1 PE=1 SV=2	4	60.1	90.3	123.6	0.731	-0.452
107	High	P01859	Immunoglobulin heavy constant gamma 2 OS=Homo sapiens OX=9606 GN=IGHG2 PE=1 SV=2	4	35.9	82.6	113.2	0.729	-0.456
108	High	P08637	Low affinity immunoglobulin gamma Fc region receptor III-A OS=Homo sapiens OX=9606 GN=FCGR3A PE=1 SV=2	2	29.1	81.1	111.4	0.728	-0.458
109	High	P62979	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens OX=9606	3	18	76.9	106.1	0.725	-0.464

			GN=RPS27A PE=1 SV=2						
110	High	Q8WZ75	Roundabout homolog 4 OS=Homo sapiens OX=9606 GN=ROBO4 PE=1 SV=1	2	107.4	85.4	118	0.724	-0.466
111	High	Q16651	Prostasin OS=Homo sapiens OX=9606 GN=PRSS8 PE=1 SV=1	2	36.4	91.6	127.2	0.72	-0.474
112	High	P19440	Glutathione hydrolase 1 proenzyme OS=Homo sapiens OX=9606 GN=GGT1 PE=1 SV=2	2	61.4	83.6	116.4	0.718	-0.478
113	High	P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens OX=9606 GN=KRT1 PE=1 SV=6	17	66	87.8	122.8	0.715	-0.484
114	High	P10451	Osteopontin OS=Homo sapiens OX=9606 GN=SPP1 PE=1 SV=1	11	35.4	79.2	111.1	0.713	-0.488
115	High	Q6UVK1	Chondroitin sulfate proteoglycan 4 OS=Homo sapiens OX=9606 GN=CSPG4 PE=1 SV=2	2	250.4	76.9	108	0.712	-0.490
116	High	Q06830	Peroxiredoxin-1 OS=Homo sapiens OX=9606 GN=PRDX1 PE=1 SV=1	2	22.1	95.6	134.3	0.712	-0.490
117	High	P26992	Ciliary neurotrophic factor receptor subunit alpha OS=Homo sapiens OX=9606 GN=CNTFR PE=1 SV=2	2	40.6	84.3	118.5	0.711	-0.492
118	High	P06870	Kallikrein-1 OS=Homo sapiens OX=9606 GN=KLK1 PE=1 SV=2	3	28.9	77.7	110.7	0.702	-0.510
119	High	P50895	Basal cell adhesion molecule OS=Homo sapiens OX=9606 GN=BCAM PE=1 SV=2	3	67.4	84.4	121.1	0.697	-0.521
120	High	P05154	Plasma serine protease inhibitor OS=Homo sapiens OX=9606 GN=SERPINA5 PE=1 SV=3	4	45.6	65	93.5	0.695	-0.525
121	High	Q16270	Insulin-like growth factor-binding protein 7 OS=Homo sapiens OX=9606	3	29.1	65.7	97	0.678	-0.561

			GN=IGFBP7 PE=1 SV=1						
122	High	Q8IYS5	Osteoclast-associated immunoglobulin-like receptor OS=Homo sapiens OX=9606 GN=OSCAR PE=1 SV=3	2	30.5	93.9	139.4	0.674	-0.569
123	High	Q9BRK3	Matrix remodeling-associated protein 8 OS=Homo sapiens OX=9606 GN=MXRA8 PE=1 SV=1	4	49.1	86.8	129	0.673	-0.571
124	High	P10153	Non-secretory ribonuclease OS=Homo sapiens OX=9606 GN=RNASE2 PE=1 SV=2	2	18.3	77.9	116.3	0.67	-0.578
125	High	Q9NQ84	G-protein coupled receptor family C group 5 member C OS=Homo sapiens OX=9606 GN=GPRC5C PE=1 SV=2	2	48.2	84.3	125.9	0.67	-0.578
126	High	P04083	Annexin A1 OS=Homo sapiens OX=9606 GN=ANXA1 PE=1 SV=2	6	38.7	69.9	106.8	0.654	-0.613
127	High	Q96NY8	Nectin-4 OS=Homo sapiens OX=9606 GN=NECTIN4 PE=1 SV=1	3	55.4	72	111.4	0.647	-0.628
128	High	P62873	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1 OS=Homo sapiens OX=9606 GN=GNB1 PE=1 SV=3	2	37.4	73.6	113.9	0.646	-0.630
129	High	P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens OX=9606 GN=KRT9 PE=1 SV=3	5	62	100.3	155.7	0.644	-0.635
130	High	P07195	L-lactate dehydrogenase B chain OS=Homo sapiens OX=9606 GN=LDHB PE=1 SV=2	2	36.6	77.3	121.3	0.637	-0.651
131	High	P15586	N-acetylglucosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GNS PE=1 SV=3	3	62	74.3	118.3	0.628	-0.671
132	High	P01042	Kininogen-1 OS=Homo sapiens	6	71.9	78.3	124.9	0.626	-0.676

			OX=9606 GN=KNG1 PE=1 SV=2						
133	High	P01876	Immunoglobulin heavy constant alpha 1 OS=Homo sapiens OX=9606 GN=IGHA1 PE=1 SV=2	5	37.6	69.2	110.7	0.625	-0.678
134	High	P54802	Alpha-N-acetylglucosaminidase OS=Homo sapiens OX=9606 GN=NAGLU PE=1 SV=2	5	82.2	77.1	123.4	0.625	-0.678
135	High	P19835	Bile salt-activated lipase OS=Homo sapiens OX=9606 GN=CEL PE=1 SV=3	3	79.3	56.5	90.7	0.623	-0.683
136	High	Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens OX=9606 GN=LMAN2 PE=1 SV=1	3	40.2	86.5	139.2	0.622	-0.685
137	High	P60709	Actin, cytoplasmic 1 OS=Homo sapiens OX=9606 GN=ACTB PE=1 SV=1	6	41.7	83.1	134.5	0.618	-0.694
138	High	Q08380	Galectin-3-binding protein OS=Homo sapiens OX=9606 GN=LGALS3BP PE=1 SV=1	10	65.3	79	128.6	0.614	-0.704
139	High	P01877	Immunoglobulin heavy constant alpha 2 OS=Homo sapiens OX=9606 GN=IGHA2 PE=1 SV=4	3	36.6	60.7	99.9	0.608	-0.718
140	High	Q68CJ9	Cyclic AMP-responsive element-binding protein 3-like protein 3 OS=Homo sapiens OX=9606 GN=CREB3L3 PE=1 SV=2	2	49	84.3	139.6	0.604	-0.727
141	High	Q7KYR7	Butyrophilin subfamily 2 member A1 OS=Homo sapiens OX=9606 GN=BTN2A1 PE=1 SV=3	3	59.6	75	124.9	0.601	-0.735
142	High	P01591	Immunoglobulin J chain OS=Homo sapiens OX=9606 GN=JCHAIN PE=1 SV=4	2	18.1	80.2	133.6	0.601	-0.735
143	High	P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens OX=9606	5	65.4	69.9	119.7	0.584	-0.776

			GN=KRT2 PE=1 SV=2						
144	High	P05062	Fructose-bisphosphate aldolase B OS=Homo sapiens OX=9606 GN=ALDOB PE=1 SV=2	2	39.4	88	151.5	0.581	-0.783
145	High	P01133	Pro-epidermal growth factor OS=Homo sapiens OX=9606 GN=EGF PE=1 SV=2	11	133.9	69.2	120.1	0.576	-0.796
146	High	Q03154	Aminoacylase-1 OS=Homo sapiens OX=9606 GN=ACY1 PE=1 SV=1	2	45.9	64.8	113.6	0.571	-0.808
147	High	Q9UQ72	Pregnancy-specific beta-1- glycoprotein 11 OS=Homo sapiens OX=9606 GN=PSG11 PE=2 SV=3	7	37.1	80.9	142.4	0.568	-0.816
148	High	P05937	Calbindin OS=Homo sapiens OX=9606 GN=CALB1 PE=1 SV=2	3	30	57.8	101.7	0.568	-0.816
149	High	O60494	Cubilin OS=Homo sapiens OX=9606 GN=CUBN PE=1 SV=5	3	398.5	76.4	134.8	0.567	-0.819
150	High	P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens OX=9606 GN=KRT4 PE=1 SV=4	12	57.3	57.6	102.4	0.563	-0.829
151	High	O00391	Sulfhydryl oxidase 1 OS=Homo sapiens OX=9606 GN=QSOX1 PE=1 SV=3	5	82.5	66.8	121	0.552	-0.857
152	High	P10909	Clusterin OS=Homo sapiens OX=9606 GN=CLU PE=1 SV=1	4	52.5	65.3	120	0.544	-0.878
153	High	Q07075	Glutamyl aminopeptidase OS=Homo sapiens OX=9606 GN=ENPEP PE=1 SV=3	3	109.2	67	126.9	0.528	-0.921
154	High	Q08345	Epithelial discoidin domain-containing receptor 1 OS=Homo sapiens OX=9606 GN=DDR1 PE=1 SV=1	4	101.1	75.2	145.2	0.518	-0.949
155	High	Q9P121	Neurotrimin OS=Homo sapiens OX=9606 GN=NTM PE=1 SV=1	2	37.9	67.1	130.6	0.513	-0.963

156	High	Q6EMK4	Vasorin OS=Homo sapiens OX=9606 GN=VASN PE=1 SV=1	5	71.7	66.1	131.1	0.504	-0.989
157	High	P15289	Arylsulfatase A OS=Homo sapiens OX=9606 GN=ARSA PE=1 SV=3	6	53.6	71.3	141.9	0.502	-0.994
158	High	P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens OX=9606 GN=KRT5 PE=1 SV=3	11	62.3	60.8	126.9	0.479	-1.062
159	High	P10253	Lysosomal alpha-glucosidase OS=Homo sapiens OX=9606 GN=GAA PE=1 SV=4	4	105.3	62.3	139.3	0.447	-1.162
160	High	P29972	Aquaporin-1 OS=Homo sapiens OX=9606 GN=AQP1 PE=1 SV=3	2	28.5	74.8	171.4	0.436	-1.198
161	High	P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens OX=9606 GN=KRT13 PE=1 SV=4	18	49.6	52	122.3	0.425	-1.234
162	High	Q99988	Growth/differentiation factor 15 OS=Homo sapiens OX=9606 GN=GDF15 PE=1 SV=3	3	34.1	73.6	173.7	0.424	-1.238
163	High	P30086	Phosphatidylethanolamine-binding protein 1 OS=Homo sapiens OX=9606 GN=PEBP1 PE=1 SV=3	3	21	66	156	0.423	-1.241
164	High	P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens OX=9606 GN=KRT6A PE=1 SV=3	19	60	52.2	132.5	0.394	-1.344
165	High	P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens OX=9606 GN=KRT6B PE=1 SV=5	20	60	51.6	139.8	0.369	-1.438
166	High	P34059	N-acetylgalactosamine-6-sulfatase OS=Homo sapiens OX=9606 GN=GALNS PE=1 SV=1	2	58	62.4	175.3	0.356	-1.490
167	High	P60174	Triosephosphate isomerase OS=Homo sapiens OX=9606 GN=TPI1 PE=1 SV=3	2	30.8	39.3	118	0.333	-1.586

168	High	P62736	Actin, aortic smooth muscle OS=Homo sapiens OX=9606 GN=ACTA2 PE=1 SV=1	5	42	65.4	198.7	0.329	-1.604
169	High	P11021	Endoplasmic reticulum chaperone BiP OS=Homo sapiens OX=9606 GN=HSPA5 PE=1 SV=2	5	72.3	59.4	182.3	0.326	-1.617
170	High	P62805	Histone H4 OS=Homo sapiens OX=9606 GN=HIST1H4A PE=1 SV=2	3	11.4	42.8	134.2	0.319	-1.648
171	High	P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens OX=9606 GN=KRT16 PE=1 SV=4	6	51.2	44.8	157.9	0.284	-1.816
172	High	P13473	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens OX=9606 GN=LAMP2 PE=1 SV=2	2	44.9	50.5	218	0.232	-2.108
173	High	P02788	Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF PE=1 SV=6	3	78.1	28.7	156.2	0.184	-2.442

Data identified at confidence interval (CI): 99 %, FDR < 0.01; Peptides no. ≥ 2.

6.4 Appendix 4: Overlay of all FTIR spectra for PE and control at 15 week of pregnancy

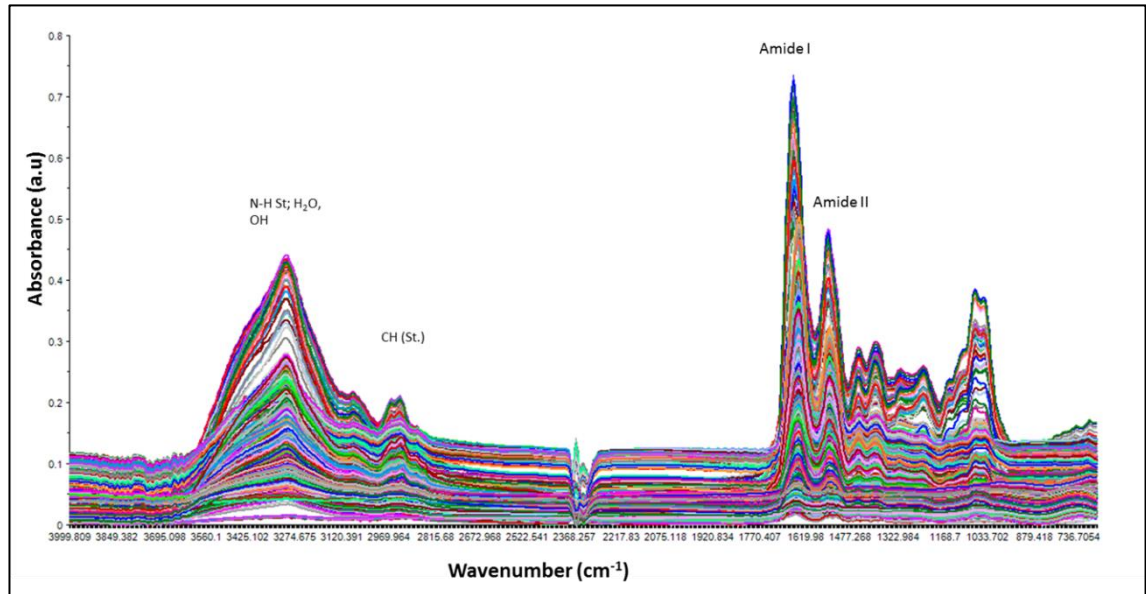


Figure S 1: Overlay of FTIR Spectra for PE and control at 15 week of gestational age at wavenumber range (4000-600) cm⁻¹

6.5 Appendix 5: Overlay of all FTIR spectra for IUGR and control at 15 week gestational age

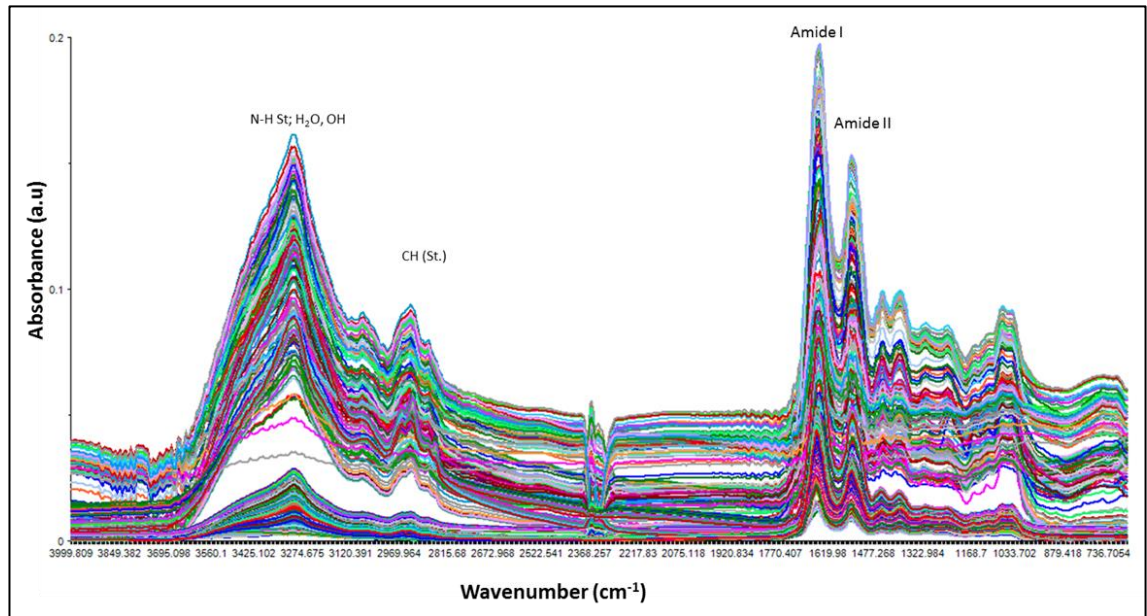


Figure S 2: Overlay of FTIR Spectra for IUGR and control urine samples at 15 weeks' gestation, 4000-600 cm⁻¹